# NEW METHODOLOGY FOR MEASURING SEMANTIC FUNCTIONAL SIMILARITY BASED ON BIDIRECTIONAL INTEGRATION

By

Jong Cheol Jeong

Submitted to the graduate degree program in Bioengineering and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Bo Luo

_____

Co-chairperson Xue-wen Chen

_____

Arvin Agah

_____

Jerzy Grzymala-Busse

_____

Jun Huan

_____

Wonpil Im

Date Defended: April 12, 2013

The Dissertation Committee for Jong Cheol Jeong

certifies that this is the approved version of the following dissertation:

# NEW METHODOLOGY FOR MEASURING SEMANTIC FUNCTIONAL SIMILARITY BASED ON BIDIRECTIONAL INTEGRATION

_____

Chairperson Bo Luo

_____

Co-chairperson Xue-wen Chen

Date approved: April 12, 2013

## Abstract

1.2 billion users in facebook, 17 million articles in Wikipedia, and 190 million tweets per day have demanded significant increase of information processing through Internet in recent years. Similarly life sciences and bioinformatics also have faced issues of processing *Big data* due to the explosion of publicly available genomic information resulted from the Human Genome Project (HGP) and the increasing usage of high throughput technology. HGP was completed in 2003 and resulted in identifying 20,000-25,000 genes in human DNA and determining the sequences of three billion human base pairs. The information requires huge amount of data storage and becomes difficult to process using on-hand database management tools or traditional data processing applications. This thesis introduces new method, Biological and Statistical Mean (BSM) score to calculate functional similarity between gene products (GPs) that can help to extract biologically relevant and statistically robust information from large-scale biomedical, genomic and proteomic data sources. BSM score is defined by 16 different scoring matrices derived from principles of multi-view learning in machine learning algorithm and five different databases including Gene Ontology, UniProt, SCOP, CATH, and KUPS. The proposed method also shows how diverse databases and principles in machine learning theory can be integrated into a simple scoring function, and how the simple concept can give significant impact on the studies in biomedical and human life sciences. The comprehensive evaluations and performance comparisons with other conventional methods show that BSM score clearly outperforms other methods in terms of sensitivity of clustering similarity functional groups and coverage of identifying related genes. As a part of potential applications handling large amount of diverse data sources in medical domain, this thesis introduces similarity-based drug target identification and disease networks using BSM scores. Application of BSM score is freely available through http://www.ittc.ku.edu/chenlab/goal/

## Acknowledgement

I would like to gratefully and sincerely thank my advisor, Professor Xue-wen Chen, for his effective guidance and encouragement that he has given me during my graduate studies at University of Kansas. I thank Prof. Arvin Agah and Jerzy Grzymala-Busse for spending their valuable times to review my dissertation and give great advices. I also appreciate Dr. Bo Luo to help me figuring out the issues during the preparation of my defense. I thank Prof. Jun Huan. He has always given valuable advices whenever I asked him, and no matter what they are. I also thank Prof. Wonpil Im who has helped me opening my eyes in new exciting research area and given excellent advices on finding my new career.

I would especially like to thank my parents, Jaewook Jeong and Sungim Lee. Without their unconditional love and support for my entire life, nothing has been possible. I also appreciate the patience and love from my families including Haeyoung Jeong, Haeja Jeong, and my nephews and nieces.

I also thank my father and mother in law, Jaeyong Kim and Bokja Yang and brother and sister in law, Jongchang Kim, Eunok Kim, and Eunkyung Kim.

Most importantly, I truly give my deepest thanks to my wife Eunmi Kim. During this journey, her support, encouragement, patience and eternal love and devotion are the actual source of my endeavor and endurance. I know it is not enough just say thank to her, but I do know that she understands what I am thinking and what I really want to say to her as she always did.

I thank my two sons, Jayden Geonu Jeong and Joshua Myeongu Jeong although they may need several more years to read what I have written in this page. They are the source of my happiness and the most glorious and precious gift in my life. They make me laugh all the times.

Finally, I appreciate my brother Jongdae Jeong who had been waiting for long time to hear the success of my PhD defense. Although we cannot talk and see each other anymore, I know he must be happy at the beyond the sky and keep watching me with smile.

A portion of this work was supported by the US National Science Foundation Award (IIS-0644366). The opinions, findings, or conclusions in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF EQUATIONS

# 1 Introduction

## 1.1 Motivation

Measuring similarities among genes and gene products (GPs) are neither a new nor special issue in current genomics and proteomics anymore, but this issue could be the one of the issues having longest history in modern computational biology or bioinformatics. Indeed, this issue is continuously evolved in different manners and to give motivation to produce new technologies. For example, the technology of measuring sequence similarity is evolved from pair-wise sequence alignment [1-4] to multiple sequence alignment [5-13]. Methods for measuring secondary and higher structural similarity keeps evolving and improving their algorithms [14-22]. In spite of existing numerous technologies in similarity issue shown above, new aspect of measuring similarities are continuously introduced such as similarity in chemical structure [23-25] or in activity across cell lines [26-30]. By considering this phenomenon, a question naturally arises: "*what makes this issue getting important and having such immortal lifecycle in its evolution?*" The reasons should be varied from study to study, but the one of the most inevitable reasons would be the strong relation to making pipelines among diverse data sources. This has been particularly important after initiating and finishing the human genome project (HGP) and increasing usage of high throughput technology. HGP was completed in 2003 and resulted in identifying 20,000-25,000 genes in human DNA and determining the sequences of three billion human base pairs. The information requires huge amount of data storage and becomes difficult to process using on-hand database management tools or traditional data processing applications. This introduces many issues in diverse areas. For example, biologists

and medical scientists have been faced problems of retrieving relevant data that are required to conduct researches on their purpose such that identifying new discovery is not trivial question any more due do requiring exploring huge amount of data and big scales of data which make difficult to manual analysis (e.g. identifying functional modules from microarray data). Another importance can be found in relations of treating disease or more precisely, designing new drugs or making drugs safer by identifying the relationships of molecular mechanisms or functions in cellular life [26, 31-40]. In fact, reports show that developing a new drug requires average of $800 million and 12 years [41-43] and the drug approvals from Biotech industry are about 45% and 95% of failures are from phase III [44]. Studies on deceased subjects showed that 3.5% of subjects were suspected to have died from fatal adverse drug reactions (ADRs) [45] and more than 7% of patients are suffered from adverse drug reactions [46]. Indeed, the percentage of patients who suffers from ADRs is significantly increased up to 33% with elderly patients [47]. This implies that adverse drug reactions and delay of clinical trials are often caused by the insufficient information about characteristics of proteins or their interactions; therefore, if successfully controlled drug targets or proteins having enriched information obtained from real clinical trials are identified as functional neighbors of unknown or newly discovered disease related genes or GPs in a specific level, both the risk of ADRs and the cost of developing new treatments should be significantly reduced. Although modern sequence similarity-based methods have devoted for improving identification of biomarkers for diagnosis, disease activity assessment or prognosis [48-50], studies show that their devotion toward actual drug discovery is very disappointing [39, 40, 51, 52] due to the lack of relations between drug targets and biological validations and leads bottle neck problem in drug validation [38, 39, 53]. To narrow down the gaps between identified novel drug targets and their validations, alternative approaches

are emerged for helping the discovery of new types of drug targets such that instead of discovering breakthrough drug targets, the ones similar to the well known targets are identified, so that the risk of failure on target/drug validation can be reduced by recycling the information derived from previous information gathered from clinical trials [26, 54-57]. As a consequence of these efforts, studies measuring functional relations among potential drug targets are of interest since defining closeness between proteins can be directly used for identifying and validating the drug targets by comparing the known functions of targets instead of implied from homology-based methods derived from sequence or structure similarity. In other words, this type of approach compares real factors proved by actual experiments including wet-laboratory works but not predictions. Recently more direct approach measuring functional similarity among genes and their products are introduced by using GO graph structure and/or information content (IC) of GO aspect [58-76]. GO [77] is a dictionary of biological terms describing molecular function, biological process and cellular component of gene with controlled vocabulary. Although adequateness and limitations of GO terms are somewhat controversial [78], there are several reasons that the issues of GO-based functional similarity is considered to be more important than other methods. First, GO functional terms are defined by several sources including wet-laboratory works; therefore, by using GO annotation implies integrating multiple data sources and gives more sophisticated and specified information about the relations on protein functions. Second, the multiple data sources can be purified by controlling the hierarchal relations of terms and the evidence of providing functions known as GO evidence code. Third, compared to Enzyme Commission (EC) number, GO provides a unique and most advanced form of a bridge between different biological communities by providing system level interpretations of biological data. Finally, GO database itself provides both data repository of genes and GPs and

attributes across species and databases, so this makes easy to integrate and get the information of genes and GPs from wide range of data sources. These unique and outstanding features/roles of GO itself and increasing demands on development of new tools to make connections between identifying new drug targets and their validations have encouraged me to build a new method measuring functional similarity. Indeed, this study has found that conventional methods often neglect and makes pointless aims on their algorithms and evaluations. It often happens due to the lack of guidelines for developing and evaluating GO-based semantic similarity measurements, regardless of introducing new measurements. Therefore, this paper introduces new semantic functional similarity methods together with proposing novel guidelines on developing and evaluating GO-based methods. At the same time, this study also shows that the proposed method defining functional similarity between genes or GPs can be easily extended to make pipelines used for exploring and retrieving information from diverse sources or existing networks.

# 2 Background studies

## 2.1 Introductions on functional similarity

Functional similarity methods can be briefly categorized into four different groups: i) present/absence of embedding GO graph, ii) the sources of the defining weights, iii) the location of weights (i.e, node/edge-based weights and calculating distance or similarity), and iv) the number of target terms (i.e, a pair of GO terms vs. a group of GO terms).

### 2.1.1 Referring GO graph vs. experimental data

Conventional methods for measuring functional similarity often use the topology of GO graph [63, 67, 74, 79] due to the simplicity of defining problems and broad support of GO functional annotations from wide range of existing databases. By limiting the scope of protein function inside domains of GO topology, this automatically narrows down the problems of measuring semantic similarity into defining the distance between GO terms appeared in GO graph. Indeed GO functional terms of GPs are widely supported by well known databases including UniProt [80] and EMBL [81]; therefore, by adapting GO, the given problems can be much simpler and conveniently reformulated. Compared to GO-based methods, some methods are referring external experimental data such as protein-protein interaction networks, microarray, and sequence or phylogenetic profiles to define similarities or differences of GPs in their own purposes: identifying functional modules [82-84], analyzing experimental results [68, 85-88] and predicting protein functions [30, 76, 89].

### 2.1.2 Weight sources

Methods for measuring GO-based semantic similarity can be grouped into three categories based on the source of defining their weight: the number of GPs associated with a GO term, the hierarchical relations of GO terms and integration of multiple sources. Resnik [90] defines the weight based on IC by considering the number of annotated GPs, and variations in this approach have been applied into several different methods including Lin [91] and Tao *et al*. [79] which normalized Resnik's score, Couto *et al*. [63, 64] which considered multiple ancestor nodes, Schlicker *et al*. [58, 59] which combined Resnik and Lin's method, and IntelliGO [66] which considered the weight of GO evidence code. Although IC-based weighting schemes are one of the most popular approaches to define functional similarity, the principle itself has inevitable defects on specific GO terms having no annotated GPs and/or the root node as their common ancestor. The details on these issues are discussed in the section of *current issues*. Besides these defects, biased [92] annotations and containing errors [93, 94] in GPs are also problematic although GO itself is systemically designed and more consistent among different released versions. Due to the systematic approach to build a graph structure and more conserved domains of GO terms compared to GPs, some methods are measuring the semantic functional similarity by using statistics of GO topology. Wu *et al*. [74] used the number of shared nodes existed in the path of two GO terms as the similarity score. simUI [95] defines GO similarity of two GO terms by using the ratio between shared nodes and total unique number of nodes in GO graph and can be considered as normalized version of Wu *et al*. Other studies integrating both approaches and external data sources have been studied as well. Lerman and Shakhnovich [62] and del Pozo *et al*. [65] defined the weight of edges based on the co-occurrence of domain between a pair of nodes from SCOP [96] and InterPro [97] respectively. Jiang and Conrath [98]

defined the weight of an edge by integrating its depth and local density and the difference of IC between parent and child nodes. XOA [72, 99, 100] compares similarity scores calculated by multiple methods including Resnik, Jiang, Lin and Bodenreider's method and takes the maximum similarity among them.

### 2.1.3 Node vs. Edge based methods

The topology of GO requires two critical elements, nodes and edges to get a final graph structure in which the nodes correspond to the GO terms and edges denote the links between child and parent nodes. Due to the graph structure of considered as a specialized characteristic in GO, methods are often focused on weighting either nodes or edges, and for convenience, it now is called node and edge based methods respectively. The principle of node-based method is that similar GO term pairs are likely to be correlated with their associations or expressions of annotated genes or GPs. Node-based methods often define GO similarity as the average weight of common ancestors of which two GO terms share their ancestor GO terms or the weight of the nearest common ancestor (NCA) which is one of the furthest common ancestor terms from the root node, so this type of method focuses on defining the weight of individual node. Resnik, Lin, Cuoto, simUI, Bodenreider and Schlicker's method can be grouped into node-based methods. Edge based methods define the weight of edges directly connecting two GO nodes. More precisely, this type of methods defined the weight of edges that connect child and parent nodes, so that the weight of edge can be directly used for defining similarity or used for calculating distance or similarity of two GO terms by exploring the path of two GO terms. The idea behind edge-based method is that the child nodes are more important than parent nodes due to the characteristics of hierarchical structure in GO (i.e, the lower terms which are far from the root node give more information than terms located close to root node since the deeper term describes

the more specific functions of GPs), so the edges closer to leaf nodes has higher weights than edges closer to the root node. simLP [95] defines similarity between two GO terms by measuring the depth of the longest shared path from the root node (GO:0003674, 'molecular_function'). Ye *et al*. [101] and Pekar [102] introduced normalized version of simLP. RSS [75] defined similarity by combining relative depth and normalized common path coverage. Wang *et al*. [73] introduced propagated bottom-up weighting scheme to define the weight of edges such that a fixed initial weight of the edge between a leaf node and its direct parent is first defined. The weight of rest edges are calculated by multiplying the fixed initial weight to the weight of an edge associated with directly connected child node of the target edge, so the initial weight is propagated from the edge associated with the leaf node to all the way toward the root node. Some methods can be located in between as integrating both node and edge-based approaches. For example, IntelliGO [66] can be considered as extended version of simLP multiplied by a coefficient calculated from the number of annotated GPs.

### 2.1.4 Similarity from a pair vs. a group

Although the main stream of the methods in measuring functional similarity is focused on measuring semantic similarity based on GO, studies specially give more attentions on two distinct tasks: measuring semantic similarity for a pair of GO terms or GPs. Most of conventional methods are more focused on GO term similarity including the works from Resnik [90], Lin [91], Gentleman (i.e, simLP and simUI) [95], Ye [101], Jiang [98], Bodenreider [84], Schlikcer [59], and Wu [74]. This type of methods does not explicitly define the way to define GP similarity or tend to ignore similarity between GPs. This happens because these methods are often originated from different purposes like natural language processing (NLP) or information retrieval, so these methods are often biased toward solving problems in information theory rather

than the biological targets. Thus identifying functional modules and predicting protein functions in which the real problems are often required dealing with GPs are conventionally ignored. Compared to the methods emphasizing the similarity between GO terms, some methods are relatively more focusing on GP similarity. The simplest method measuring a pair of GP similarity is counting the number of GO term matches [29] or its normalized version [71]. Similarity scores averaged on all against all GO terms and their variations are also introduced [61, 66, 72, 99, 100, 103]. More sophisticate versions of matching score based on Hausdorff distance [104] and its variations is also applied [62, 65, 73, 79]. Some methods directly define GP similarity by identifying a most specific GO term pair having the minimum number of annotated GPs [67] or co-expressed microarray data [29, 87].

## 2.2 Current issues

So far it has been reviewed above, there is no single rule that can categorize one from another. On the other hand, this may imply that current methods are fully specified and diverse enough to be used for clinical trials or practical studies. However, in reality, conventional methods measuring functional similarity is rarely found in practical studies including drug design and its related studies. In fact, approaches using single experiment or dataset are often preferred such as microarray data or PPI networks rather than GO-based methods in which multiple experimental verifications are often required to annotate the functions of GPs. During the research on this topic, only few case studies of GO-based functional similarity measurement has been found: evaluating the similarity of GO or GP itself [105] and analyzing gene expression data [106]. This low usage of conventional methods implies two important criticisms about the method at current stage: i) the conventional methods are not mature enough to be applied yet, and ii) some points in overall design of current methods are not suitable or possibly misled in practical points of view. The first criticism may be too broad and unclear, so let's consider the following facts that the functional annotations of a protein are defined by multiple experiments or screenings. In other words, why the methods relying on the annotations identified with various experiments are considered less confidant or rarely used? One of the reasons can be found from their contradicted performances reported in various studies [61, 64, 68, 70, 87, 106, 107], so this unreliable performances make end users difficult to choose it or less confident on using conventional methods. Therefore, the meaning of 'immature' in the first critique can be redefined as insufficient general performances containing low quality and inconsistent performance of the existing methods. One of assumptions which cause this problem is the lack of guidelines of validation procedures since most evaluations are conducted with the task

independent general approaches; therefore although the validation or testing data are derived from various types of sources and problem domains such as functional predictions [79, 108], human interactive vision analysis [73], common cellular pathways [66, 103, 109], sharing functional domains [63, 64, 66], text mining [84] and sequence similarity [59, 69, 70, 72, 100], these approaches often require over simplified assumptions without clarifying biological relevancies. For example, majority of evaluations is conducted based on the assumption that sequence similarity and functional similarities are proportional. However, the studies showed that identifying the functional similarity between two GPs by using sequence similarity alone is not enough, and the relations between functional similarity and sequences similarities are not necessarily to be proportional but more likely to be correlated [62, 92, 93, 110-112]. By considering the fact that most conventional methods treat the problem of functional similarity without reflecting the ultimate goals or clarifying the definition in evaluation process, most validation methods are not suitable for revealing the task specific true powers. Therefore, through this study, I propose specific guidelines of evaluation process that increase the intuitive interpretation of similarity scores through comprehensive evaluation processes including comparisons of structure, phylogenetic profile and sequence similarity and biological/topological clusters. Although validation processes take important role to decide the performance of the GO-based functional similarity measurements, the fundamental theory consisting of methods would be more crucial to define their performance and biological relevancies. Some conventional methods often derive their theoretical background from natural language processing (NLP) without serious consideration of the differences. There are two distinct differences between biological processes and NLP. First the dictionary used in NLP is believed to be fully discovered and their relations are implicitly assumed to be finite and fully specified. Second,

documents in NLP are assumed to be completely known and time-invariant [113] while annotated GPs are highly biased and time variant. In contrast to NLP, my belief about the role of a gene or GP in cellular life is still accumulating and evolving, so even if GO exists, the terms and relations are neither finite nor fully specified. For example, the majority of methods employ the information content (IC) as their major distributor to define functional similarity [58, 59, 61, 63, 64, 66, 72, 79, 90, 91, 98-100] which is calculated by the ratio between the number of overall GPs and the number of GPs directly annotated by a specific GO term. However, IC itself can be highly biased due to the uncompleted genome and interest of studies toward specific species and genomes. To make matters worse, the biased data does not guarantee to reflect the true distribution of gene or GPs in nature; therefore, although IC-based method works well in NLP, it does not guarantee its promising performances in biological point of view. To check out the tendency of growing annotated GPs, all released versions of gene ontology database have downloaded through GO data repository (http://archive.geneontology.org/), and calculated the statistics of annotated GPs.

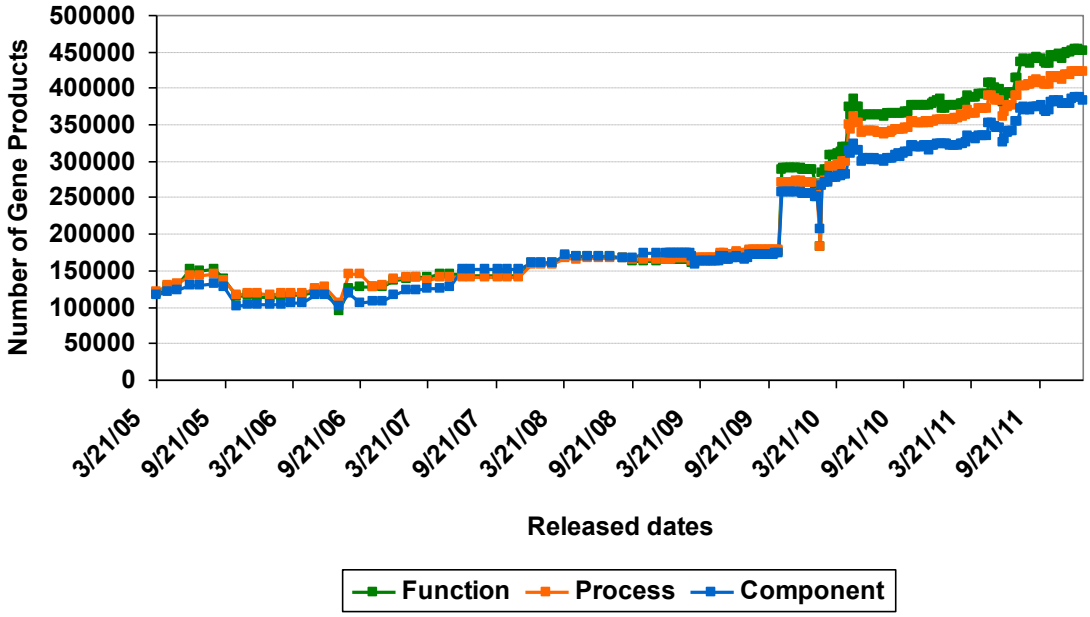



**Figure 2.1** Number of annotated GPs along with the time line of released versions.
This graph shows the number of annotated GPs from March 2005 to January 2012.

Figure 2.1 shows an interesting point that even if by considering annotated GPs alone, the distributions of GPs belonging to timelines are varied widely. Indeed, by comparing Figure 2.1 and 2, the significant biased distributions in GPs are observed such that overall number of GPs in Figure 2.1 is dramatically increased at around September 2009 but Fly and Yeast do not show significant changes at this period. More interestingly, overall annotated GPs in Figure 2.1 are significantly decreased at around March 2010 but in Figure 2.2, Bacteria only have similar pattern observed in overall annotated GPs at Figure 2.1.

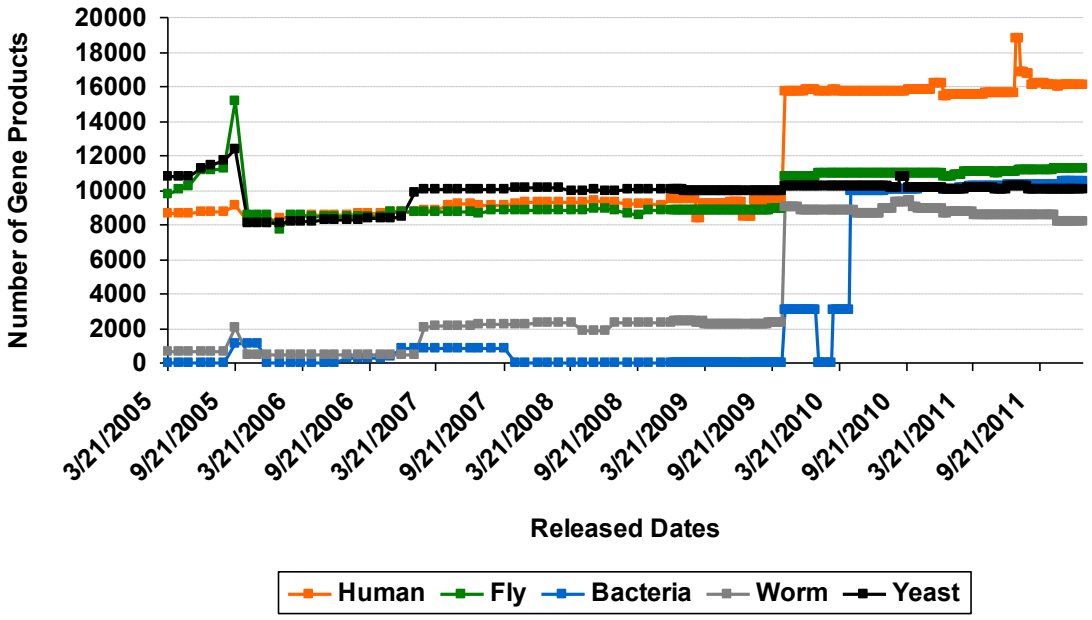



**Figure 2.2** Number of annotated GPs with five model organisms.
The model organisms consist of seven species: Homo sapiens (Human), Drosophila melanogaster (Fruit fly), Escherichia coli (Bacteria), Caenorhabditis briggsae (Worm), Caenorhabditis elegans (Worm), Saccharomyces cerevisiae (Baker's yeast), and Schizosaccharomyces prombe (Fission yeast).

These biased and contradicted distributions of GPs shows that IC-based method is not suitable for defining the weight of individual node since the biased distributions are directly affecting the frequency of GO annotations which actually is the most critical information to define functional similarity in most cases. However, there is no such evidence that these biased distributions actually reflect the true nature of functional relations, especially by considering the gaps between the numbers of known sequences and annotated GPs which can be referred by Figure 2.3; therefore, the lack of biological connections between the results from biased IC and actual functional similarity increase the uncertainty of reflecting true nature of functional relations. Indeed, this fact also contradict the commonly used assumptions in NLP, documents are completely known and time-invariant [113].

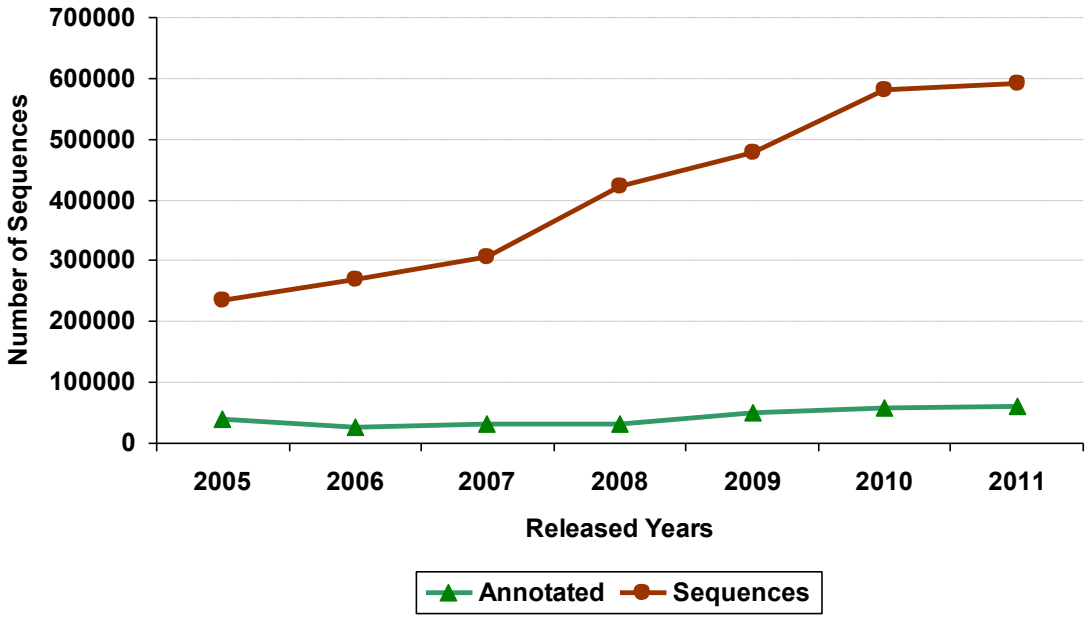



**Figure 2.3** Difference in the number of annotated GPs

The five model organisms and known sequences are retrieved from SwissProt (ftp://ftp.uniprot.org/pub/databases/uniprot/)

Besides biological uncertainty of IC and time-variant characteristics of GO, two more problems should be considered when information theory is considered as significant contributor to define functional similarity: i) GPs with no annotation and ii) root shared GO terms. When this study started, GO database released at April 2011 was first used as the master database of this research. The database shows that total 9833 functional GO terms exist, and among them 820 GO terms are obsolete, so 9013 GO terms are actually used. Since most methods derived from information theory use IC as a part of their critical information to define functional similarity, it is interesting to see the distributions of GPs at individual GO terms. Surprisingly any related studies on this issue have not been found. Although most studies use IC as a part of critical scoring functions to define functional similarity, the fundamental sources of this score, (i.e. the distributions of GPs) have not been reported yet. Simple statistics shows that 2479 GO terms out of 9013 GO terms do

not have any associated GPs which is about 28% of total GO terms, so 6534 out of 9013 GO terms are actually meaningful terms in IC-based methods. 153 out of 2479 GO terms are not leaf nodes having at least a child node. These non-leaf nodes without any annotated GPs are corresponding to 6% of all GO terms having no annotated GPs. In other words, 28% of GO terms cannot calculate their functional similarity. This is a serious defect on IC-based methods but no study has been claimed on this issue. These GO terms with no annotated GPs give inevitable restrictions on related studies including but not limited in protein function predictions and identifying functional modules. This may be considered as minor problem in most cases since no annotation means that there is no existing reference data, so why GO terms having no GPs have to be considered? They do not exist, so there is no reason to calculate functional similarity for these terms. However, this is not true, since none existing annotation does not necessarily mean that there is no gene product exists, since the GO terms and annotated GPs are not simultaneously updated and reflected into the database, so these terms are still needed especially for discovering new functions of already annotated GPs. Therefore, the inability of calculating IC should be considered as the limitations of IC-based methods. Figure 2.4 shows another interesting point such that 3,679 GO terms are associated with less than 50 GPs which are 56% of entire meaningful GO terms (i.e. 6534 = 9,013 – 2,479) depicted as a green diagonal bar. By considering less than five annotated GPs which are associated with a green diagonal bar, 26% of entire meaningful GO terms are falling into this range. In other words, more than half of GO terms control functional similarity with less than 50 GPs and majority of them are distributed to measuring functional similarity with less than five GPs. Is it enough to represent the nature of functional similarity with less than five GPs? As discussed previously, what if the biased distributions of GPs are considered together with this small sample issue? The uncertainty of the

relationship between IC and functional similarity must be increased significantly. The other critical problem in IC-based method can be found when the nearest common ancestors (NCA) are considered together [59, 79, 90, 91]. A *shadow problem* is one of the typical issues caused in conventional methods such that the specificity (i.e, how specific they are – the GO terms located in lower/deeper graph is considered to have more specific information about the function of a GP) of individual GO terms is not contributed to the calculation of GO term similarity by ignoring the influence of terms located below the NCA [64, 91]. In contrast to shadow problem, a *shallow problem* is also reported as significant defects in Jiang [98] and Lin's [91] method. For example, if two sets of GO term pairs having same depth are considered, it is possible that a pair of which the NCA is located near the root node has higher similarity score than the other pair of which the NCA of this pair is located far away from the root node due to the biased distributions of annotated GPs [73, 87] although based on human perspectives, NCA located near the root of the ontology (more general term) should be more significantly penalized than the pair with NCA located far away from the root of the ontology (more specific term).

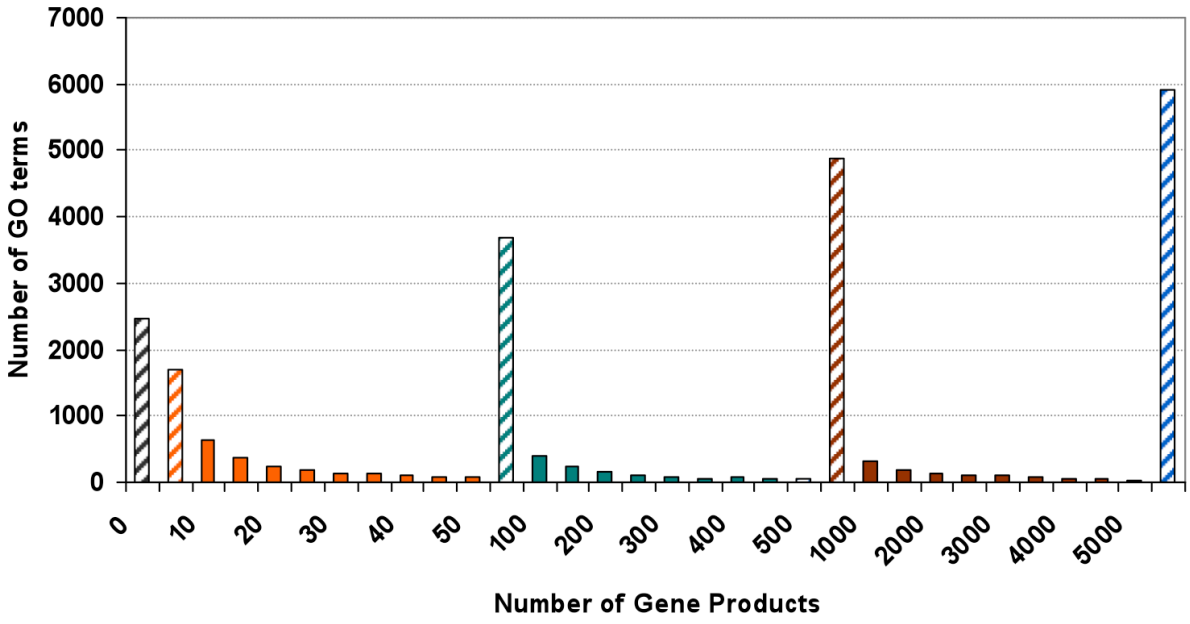



**Figure 2.4** GO term associations toward GPs.

Bars with diagonal lines show the number of accumulated gene product. 5927/5000 denotes that there are 5927 GO terms which have less than 5000 annotated GPs. Bars having solid colors show the corresponding number of GO terms based on the given range of GPs. GO terms with no annotation (2479/0) are excluded from other statistics in this graph (i.e. they are not accumulated to calculate bars with diagonal stripes.

Besides these known problems, more general and critical problems of IC-based methods are also found such that IC is calculated by the ratio between the number of overall GPs and the number of GPs annotated by a specific GO term; therefore, if the nearest common ancestor (NCA) is considered, the functional similarity of two GO terms having NCA with root node always has nil similarity which means there is no functional similarity between two GO terms. By considering 9,013 GO terms, all possible pairs among them are 40,621,591 pairs, and all possible pairs have been searched to find their NCAs by identifying the common paths through cluster machine with 264 nodes for 12 days and the results show that 20,986,023 pairs which are 52% of all possible

GO pairs have their NCA as a root node. This result shows that methods derived from IC and NCA-based methods cannot distinct the differences of GO terms or can mislead the results for more than half of all possible GO term pairs. Indeed, these pairs with nil similarity are highly problematic since they lead biologically irrelevant statistical point of view by treating all root-shared pairs as having same similarity or distance. In other words, more than half of GO-term pairs are treated as exactly same. Therefore, consequently, these equally scored root-shared terms are propagated toward measuring the functional similarity/distance between two GPs as proportionally penalizing the similarity score based on the number of nil GO similarity pairs in two GPs. However, just by considering a simple well known biological fact of fusion genes in which two different genes having different functions can develop new functions, nil similarity of root shared GO terms biologically does not make any sense. To check out the biological relevance of using IC and NCA-based method, protein-protein interaction (PPI) pairs associated with root shared GO terms are investigated from KUPS, the PPI database which contains manually curated PPIs retrieved from IntAct, MINT, and HPRD [114]. To enforce the significance of contributing root-shared terms toward PPIs, individual GPs containing a pair of root-shared GO term pairs are randomly retrieved. Once GPs annotated with at least a pair of root-shared GO terms are retrieved, the relationships of nominated GPs are figured out through PPI networks in KUPS. This ensures that root-shared same GO annotation pairs associated with retrieved PPIs are significantly contributed toward trigging protein bindings. Some of findings from these retrieved GPs are of interest. For example, ribonuclease III activity (GO:0004525) and metal ion binding (GO:0046872) are shared by root-node molecular function (GO:0003674) and retrieves all 53 GPs from UniProt which contain both GO:0004525 and GO0046872). These GPs are all associated with ‘dicer-like endoribonuclease’ proteins which cleave the

phosphodiester bond within a polynucleotide chain and lead double-stranded RNA (dsRNA) and pre-microRNA (miRNA) into short double-stranded RNA fragments called small interfering RNA (siRNA). As expected from their names, these enzymes share EC 3.1.26.- which corresponds to 'ribonuclease'. Indeed, they also share a cofactor which is 'magnesium or manganese'. By considering the significant functional similarity and large amount of retrieved GPs, it is stunning that the root-shared GO term GO:0004525 and GO:0046872 are only terms shared by 53 GPs consisted with 516 GO terms. In short, co-occurrence of GO terms, GO:0004525 and GO:0046872, should not be neglected or equally treated with other root-shared terms since ribonuclease III activity (GO:0004525) and metal ion binding (GO:0046872) are necessary GO terms which describe the functions and cofactors of the groups of nominated GPs. Indeed, they are the only GO terms shared by all retrieved GPs which are all dicer-like endoribonuclease and have same or similar biological functions together with cofactors. Therefore, IC and NCA-based methods which always treat root-shared GO terms equally are not suitable to identify groups of GPs having strong agreement on root-shared GO terms. For more details about statistics of root-shared GO terms, biological functions, cofactors and others please refer supplement material [S-1,2]. The second example is more interesting such that ribose phosphate diphosphokinase activity (GO:0004749) and Identical protein binding (GO:0042802) are root-shared GO terms commonly found in four ribose-phosphate pyrophosphokinase proteins and a phosphoribosyl pyrophosphate synthase-associated protein. Similar to the previous example, these GPs share EC 2.7.6.1 except Phosphoribosyl pyrophosphate synthase-associated protein (Q14558). As shown in protein names and EC, their biological functions are same such that they are involved in nucleotide, histidine, and tryptophan biosynthesis and active in heteromultimeric complexes with other 5-phosphoribose 1-diphosphate synthesis. Q14558 plays

a negative regulatory role in 5-phosphoribose 1-diphosphate synthesis; therefore, all retrieved GPs are related with regulatory of 5-phosphoribose 1-diphosphate synthesis. In contrast to previous example, all retrieved GPs consist of PPI networks including isolated self-binding of Q14558. These functional similarity and binding affinity are also considered as important facts that root-shared GO terms should be carefully treated when GO or GP similarities are measured (for more details about biological functions, cofactors and others please refer supplement material [S-3,4]). The third example is even more interesting compared to previous two examples. In this example, instead of considering a pair of root-shared GO terms, triplet root-shared GO terms are considered. The fact that increasing the number of root-shared GO terms can increase the significance of GP similarity influenced by root-shared GO terms shows that equally treating or ignoring the root-shared GO terms in calculating GO or GP similarity is biologically not relevant. In this example, total four GPs are retrieved from UniProt of which each GP is annotated with three GO terms and others. Three GO terms, which is the triplet of root-shared GO terms associated with retrieved GPs, are GO:0004716, GO:0046982, and GO:0005524 which are 'receptor signaling protein tyrosine kinase activity', 'Protein heterodimerization activity', and 'ATP binding' respectively. Same as previous two examples, retrieved GPs share EC number and protein name which are 2.7.10.1 and 'receptor tyrosine-protein kinase' correspondingly. All pairs are root-shared relations except GO:0046982 and GO:000552 which is shared by ' binding' (GO:0005488) which is direct child of root term and often neglected by IC and NCA-based method as well. By considering the fact that all retrieved GPs are protein tyrosine kinase, they are part of several cell surface receptor complexes and need a coreceptor for ligand binding which activates several signaling cascades to trigger appropriate cellular responses. Besides sharing EC number, biological function and catalytic

activity, the similarity of quaternary structure, homo-and/or heterodimerization of EGFR, ERBB2, ERBB3 or ERBB4 and significant PPIs existing among retrieved GPs also shows that the root-shared triplet GO terms are essential for defining a group of biologically similar GPs (for more details about biological functions, cofactors and others please refer supplement material [S-5,6]). More statistics on root-shared GO terms and associated GPs and PPIs can be found in [S-1~6].

In contras to node-based methods, edge based methods often take advantage on several issues especially when the topology of GO graph is applied rather than relying on annotated GPs: i) stabilized or completed GO terms which deliver positive impact to the assumptions of time-invariant and completeness of data, ii) relaxed conditions of GO terms associated with no annotated GPs and root-shared terms, and iii) reduced risk of biased sources. Figure 2.5 shows the changes of GO terms based on released versions. Compared to Figure 2.1, the graph is more stabilized although the terms associated with biological process are continuously increased. Indeed, edge-based methods often rely on GO term itself, so issues related to annotated GPs can be ignored. However, edge-based methods often assume that similarity of unit edges (i.e, between child and parent node) is same. In other words, relations between any pair of parent and child node have a same distribution or proportional to the depth of the node or edge, so no matter what GO terms are given, they would have exactly same similarity if they have same depth.

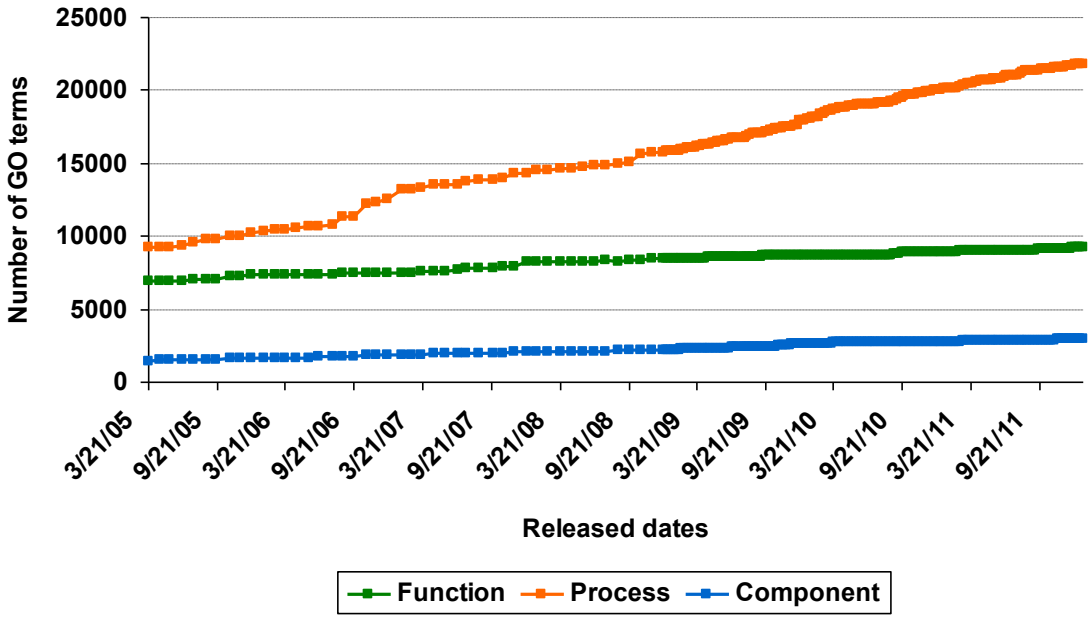



**Figure 2.5** Statistics of GO terms in GO categories based on the released dates.
Function (green dots in the middle) denotes molecular function, Process (orange dots on the top) denotes biological process, and Component (blue dots at the bottom) denotes cellular component in gene ontology category.

However, this assumption ignores fidelity of cellular life in which some proteins or functions have distinct associations compared to others as having high specificity and low affinity of their relations [115-117]. Besides the fidelity, specificity of GO terms should be considered as well. Figure 2.6 shows the distributions of GO terms at each level or depth. In the context the depth and level of a GO term are treated to be same. Level 1 or depth 1 denotes the root node, the most general term. As the depth is getting deeper and the levels are getting higher by increasing the number, they are further away from the root node and have more specific information about the function of a GP. In the graph, each level consists of two different types of GO terms: leaf nodes and others. Leaf nodes are the most specific GO term of a current GO version, and others are less specific GO terms than leaf nodes.

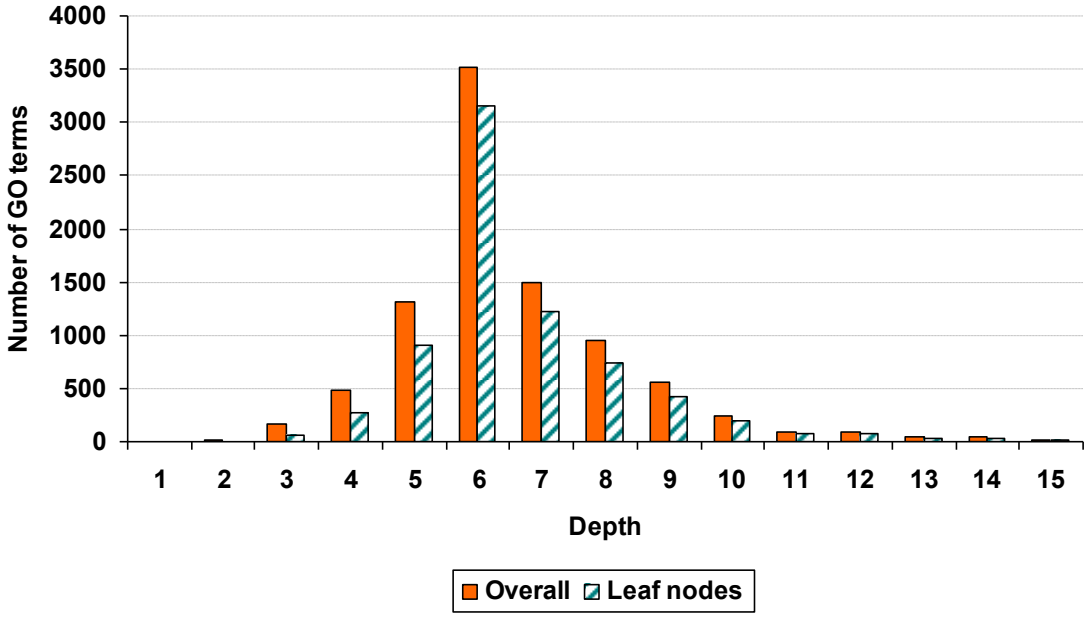



**Figure 2.6** Distributions of GO term and leaf nodes at different depth levels.
The 'overall' denoted as orange bars consist of all GO terms which pass the corresponding depth (i.e, horizontal axis) to get their maximum depth (i.e, leaf node). Leaf nodes consist of GO terms having no child nodes. The overall statistics shows that more than 72% of GO terms consist of leaf nodes.

This shows that the depth of GO term is not enough to reflect the significance of information at the given GO term. For example, let us consider that the distance/depth between two GO term pairs are same, and one pair is from depth 6 but both of them are not the leaf node. In contrast to this pair, the other pair is from depth 6 but both of them are leaf nodes. Intuitively, the later case should receive higher similarity score if the specificity/detail of GO term is considered. In summary, there are several issues being considered to design methods of measuring GO-based functional similarity: i) biased distributions of annotated GPs, ii) complete and time-invariant assumptions of GPs, iii) shadow & shallow problem, iv) nil similarity on root-shared GO terms, v) validations upon lack of specific aims, vi) considering similarity score from GO terms having no annotated GPs and vii) the assumption of uniform distance between parent and child nodes.

Besides these critical issues, the range of the score also has to be considered. For example, non-normalized score in which the score is ranged from 0 to ∞ makes difficult to identify the significance of scores. So far it has been discussed that conventional methods have several pitfalls in their theoretical background and algorithms including intuitively and biologically irrelevant assumptions and validation methods. To overcome these pitfalls, new method to measure functional similarity based on gene ontology functional terms are proposed. The proposed method has several distinct features compared to conventional methods: i) having a broad definition of functional similarity by integrating biologically and statistically specified relations among GPs including co-occurrence in PPIs, structural similarity in functional modules, and geometrical distance in GO instead of focusing only on the relations of GO terms in GO graph, ii) reducing the noise effects caused by biased datasets, iii) evaluating the method with strictly designed procedure by considering both GO and GP similarity and iv) validating the methods with comprehensive datasets and biological and biomedical points of view. The details of proposed method can be found in method section.

## 2.3 Conventional methods

For the evaluation purposes, six conventional methods are compared to the proposed method. The brief introduction of each method is described below.

Resnik's [90] method uses Information Content (IC) of Nearest Common Ancestor (NCA), which is the closest node from leaf nodes or the farthest node from the root node shared by two GO terms, as the functional similarity of two GO terms. The final score is negative logarithm of the ratio between the number of annotated GPs corresponding to NCA of two GO terms, $go_i$ and $go_j$, and total number of GPs associated with molecular function. The equation is defined by

$$\boldsymbol{simResnik}(\boldsymbol{go_i}, \boldsymbol{go_j}) = -\boldsymbol{log}(\boldsymbol{p}(\boldsymbol{NCA}))$$
$$\boldsymbol{where}\ \boldsymbol{p}(\boldsymbol{NCA}) = \frac{\boldsymbol{\#\ of\ GPs\ associated\ with\ NCA\ of\ go_i\ and\ go_j}}{\boldsymbol{\#\ of\ GPs\ associated\ with\ root\ node}} \quad \textbf{(Equation 2-1)}$$

The range of Resnik's similarity, *sim*Resnik(·) is a value greater than and equal to zero. Due to the unlimited range of maximum score, defining the relative significance of similarity score directly from the provided scoring scale is not trivial; therefore, the score is normalized between 0 and 1 by using maximum and minimum value of the observation. The equation of normalization is following

$$GO_{norm}(go_i, go_j) = \frac{GO_{sim}(go_i, go_j) - min(GO_{sim}(\cdot))}{max(GO_{sim}(\cdot)) - min(GO_{sim}(\cdot))} \quad \textbf{(Equation 2-2)}$$

$GO_{norm}$ denotes normalized GO term similarity for two GO terms, $go_i$ and $go_j$. $GO_{sim}(\cdot)$ denotes all observed GO term similarity calculated by conventional methods. $GO_{sim}(go_i, go_j)$ calculates GO term similarity between $go_i$ and $go_j$. In short, the given normalization function transforms

original value to relative values ranged from 0 to 1 which corresponds to the minimum and maximum value of all observed GO term similarity scores.

To overcome the difficulty of interpreting significance of GO term similarity resulted in non-normalized Resnik's score, Lin introduced normalized version of Resnik's method while avoiding the shadow problems also caused in Resnik's method [91]. The *shadow problem* is caused by directly using NCA-based scoring scheme. In Resnik's method, any pairs of GO term similarity can be same if their NCAs are the same. In other words, by focusing on NCA only, Resnik's method ignores the specificity or details of GO terms that can be interpreted as the distance between NCA and leaf nodes. The equation of Lin's method is defined below.

$$simLin(go_i, go_j) = \frac{2 \times \sum_{k=root}^{NCA} w_k}{\sum_{n=root}^{go_i} w_n + \sum_{m=root}^{go_j} w_m} \quad \textbf{(Equation 2-3)}$$

In this equation the weight of each node is defined by IC similar to Resnik's method.

$$w = -log(p(c))$$
$$where\ p(c) = \frac{\#\ of\ GPs\ associated\ with\ current\ GO\ term}{\#\ of\ GPs\ associated\ with\ root\ node} \quad \textbf{(Equation 2-4)}$$

The equation shows that Lin's method can be defined as ratio between the summation of ICs belonging to the path from root node to NCA and the summation of ICs belonging to the path from root node to target nodes. Although the *shadow problem* can be avoided as considering both the specificity of NCA and leaf nodes, a new problem, *shallow problem* arises in this case. There are two factors which directly affect this problem: score scheme focusing on the ratios only and the biased distribution of annotated GPs. Ratio is not necessary to reflect true

specificity of GO terms since this only considers relative difference of two distances. Indeed, the annotated GPs are not necessarily to be evenly distributed to individual functional term. Therefore, these two factors can cause such situation that a pair of GO terms having both NCA and leaf nodes closer to root node has higher similarity score than a GO term pair having both NCA and leaf nodes far away from the root node.

Schlicker *et al*. [58, 59] introduced the combination of Resnik and Lin's method. The equation is defined by

$$simSchlicker(go_i, go_j) = \frac{2 \cdot log\ p(NCA)}{log\ p(go_i) + log\ p(go_j)} \cdot (1 - p(NCA)) \qquad \textbf{(Equation 2-5)}$$

In the equation, $p(go_i)$ and $p(go_j)$ are the probability of two GO terms, $go_i$ and $go_j$ calculated by the number of annotated GPs, and $p$(NCA) is the probability of their NCA. Therefore, the left and right term corresponds to Lin and Resnik's method respectively. This score is designed to be decreased when p(NCA) is increased. In other words, the specificity of NCA defines the final GO similarity score by rewarding or penalizing the score made at NCA.

As an edge-based method, Jiang method [98] is also used for comparison. Jiang's method is different from others in two categories. First instead of focusing on nodes, Jiang's method uses weighted edges by using the density of parent terms in GO graph. Second this method calculates semantic distance rather than similarity by measuring distance of two GO terms based on the space of GO graph. The equation of this method is defined below.

$$diffJiang(go_i, go_j) = \sum_{i \in min[path(go_i, go_j)]} w_i$$

$$w_i = w(c,p) = \left(\beta + (1-\beta)\frac{\bar{E}}{E(p)}\right) \cdot \left(\frac{d(p)+1}{d(p)}\right)^{\alpha} \cdot (IC(c) - IC(p)) \cdot T(c,p) \quad \textbf{(Equation 2-6)}$$

To make comprehensive comparisons, Gentleman's method [95] which is also known as simUI are also included in performance evaluation. In this equation, $w_i$ denotes the weight of $i^{th}$ edge which is one of edges connecting two GO terms $go_i$ and $go_j$ with shortest distance, $i \in min|path(go_i, go_j)|$. The weight of each edges is defined between a child node and its direct parent node, $w$(c,p). E(p) denotes the number of edges from a parent node to its direct child nodes, and for convenience, it now is called a density of parent node. $\bar{E}$ denotes average density in entire GO graph. d(p) denotes depth of parent node and IC(c) and IC(p) are the information content (IC) at child node and parent node respectively and can be calculated same as Resnik's method. T(c,p) is a link relation factor (i.e. 'is-a' and 'part-of' relation) and set to 1 for all relations. α and β are parameters to control the significance between depth and density factors. The optimal value 0.5 and 0.3 suggested in original paper are used as the default value of α and β respectively. The range of Jiang's distance is greater and equal to zero, so the observed values are normalized by Equation 2-2. simUI uses the number of GO terms to calculate the similarity of GO terms instead of using IC based on the number of GPs. The equation is defined in Equation 2-7.

$$simUI(go_i, go_j) = \frac{2 \times |path(root, NCA)|}{|path(root, go_i) \cup path(root, go_j)|} \quad \textbf{(Equation 2-7)}$$

The similarity score is defined by ratio between the number of GO terms belonging to the path from root node to NCA $|path(root, NCA)|$ and the number of GO terms resulted in the union of all nodes belonging to the path from root node to each of GO terms, $go_i$ and $go_j$,$|path(root, go_i) \cup path(root, go_j)|$. The equation looks similar to Lin's method. Indeed, due to the lack of considering GO term specificity, the *shallow problem* observed in Lin's method is also detected in Gentleman's method. The range of similarity score is between 0 and 1.

Ye's method [101] is also considered in the experiments. Compared to other conventional method this method more focuses on the specificity of GO terms by directly adapting the relative depth of NCA into its similarity function.

$$simYe(go_i, go_j) = \frac{depth(NCA) - min[depth(\cdot)]}{max[depth(\cdot)] - min[depth(\cdot)]} \quad \textbf{(Equation 2-8)}$$

The final value of the Equation 2-8 is normalized depth of NCA, *depth*(NCA) defined by two GO terms, $go_i$ and $go_j$. The normalization is derived from maximum and minimum depth of entire functional GO terms in GO graph which corresponds to max[*depth*(·)] and min[*depth*(·)] respectively; therefore, the range of similarity score is between 0 and 1.

# 3 Methodology

## 3.1 Principals of proposed method

Biological and Statistical Mean (BSM) score is designed to integrate biological facts derived from protein domains and PPI networks and statistical significance derived from GO graph structure. To derive biologically relevant scores, two commonly believed biological facts are used: i) if GPs and genes share functional domains then they have similar functions and ii) interacting protein pairs tend to have similar function. To maximize the statistical significance of GO graph, two principals are derived from the hierarchical structure of GO term relations and information theory such that the score is intentionally designed to give higher score for GO terms located closer to leaf node or further away from root nodes, and the score is penalized for the terms have many siblings or edges between parent child nodes. The definition of functional similarity applied in the proposed method also can be explained as an integration of two different approaches toward identifying functional relations of GPs. Briefly speaking, conventional methods can be categorized into two groups either top-down or bottom-up approaches. However, both of them have intrinsic and reciprocal drawbacks. Top-down methods are favorable for evaluating or calculating GP similarity by directly analyzing the consequences of various genetic tests [29, 76, 87] or interactions between GPs [118]. These top-down approaches only consider direct consequences of an experiment or association observed from GPs having multiple functions. However, gathering data for targeted genes or GPs are neither always possible nor reliable due to limited amount of available sources and biased/inconsistent datasets [93, 110-112, 119]. Although some of top-down methods use the GO based scoring scheme [62, 65] before directly calculating the similarity of specific GPs, this type of method has two significant

drawbacks: i) it requires additional datasets, and it is not always possible to get relevant datasets for the given task ii) acquired dataset can be biased toward species and case histories; therefore, this can mislead and/or confuse the human interpretations toward relationships of functional terms (e.g. parent and child relations in GO term may not be observed or can be under/over estimated). In contrast to top-down approach, bottom-up approaches evaluate functional similarity by comparing predefined individual functions of GPs. Indeed, most of GO based methods are categorized into bottom-up approach, and this type of methods can bypass the pitfalls in top-down approach by referring multiple evidences projected into GO terms. However, biological relevance toward measuring GP similarity is somewhat questionable due to unrevealed relations of different functional branches in GO topology. In order words, GO itself does not define any functional priority or precedence among functional terms to bring the final consequences or reactions from sets of GO functional terms in a GP or a group of GPs; therefore, top-down and bottom-up approaches have reciprocal relations on their drawbacks and advantages. To get the advantages of both methods while minimizing drawbacks, a bidirectional method is proposed. The proposed method integrates both top-down and bottom-up approaches. To reduce and minimize the noise effects in top-down approach, the principles of multi view learning [120, 121] is applied by projecting independent observations into an integrated probabilistic distribution.

## 3.2 Biological and statistical mean score

The definition of functional similarity governed by the proposed method is from the consideration of biological relations identified from experimentally supported evidences and the topological similarity of GO terms referred from graphical structure of GO database. This new paradigm of functional similarity is refined and integrated in a new scoring function defined in the Equation 3-1. In the equation, $GO_{sim}$ consists of two distinct terms: reference score $R_{scr}$ and gene ontology graph-based score $G_{scr}$.

$$BSM_{sim}(t_1, t_2) = \alpha \cdot R_{scr} + \beta \cdot G_{scr}$$
$$= \alpha \cdot \left(M_{scr}(t_1, t_2)\right) + \beta \cdot exp\left(-\frac{1}{1-\frac{Ep}{E_{max}}}\right)^{\left(1-\frac{d_{NCA}}{d_{max}}\right)} \quad \textbf{(Equation 3-1)}$$

The function $GO_{sim}$ requires two parameters, $t_1$ and $t_2$ which are the pair of targeted gene ontology (GO) functional terms of which their semantic similarity is expected to be measured. $R_{scr}$ and $G_{scr}$ denote reference score. $R_{scr}$ is derived from external data sources, and $G_{scr}$ is derived from GO graph-based score in which the score is calculated from the topology of GO. $\alpha$ and $\beta$ are real values by satisfying $\alpha+\beta=1$ to define a weight for reference and GO graph based score correspondingly. For the experiments equal contribution $\alpha=0.5$ and $\beta=0.5$ are used, so the contribution of $R_{scr}$ and $G_{scr}$ toward final functional similarity score is equal. As shown in Equation 3-1, $R_{scr}$ is directly calculated from $M_{scr}$ which is a score matrix derived from SCOP and CATH database. The details of creating score matrix, $M_{scr}$ are described in following section. In contrast to $R_{scr}$, $G_{scr}$ is consisting of four variables: $d_{NCA}$ denotes the depth of NCA, and *Ep* denotes the number of edges between the parents of nearest common ancestor (NCA) of two GO

functional terms and their corresponding child nodes; therefore *Ep* can be easily calculated by counting the number of siblings of NCA. $d_{max}$ and $E_{max}$ are the maximum depth and number of edges in the given problem respectively. The values of these variables can be calculated from two different manners by considering either entire GO graph or local path only. To help better understanding the equations, variables used in the graph-based score are graphically visualized in Figure 3.1.

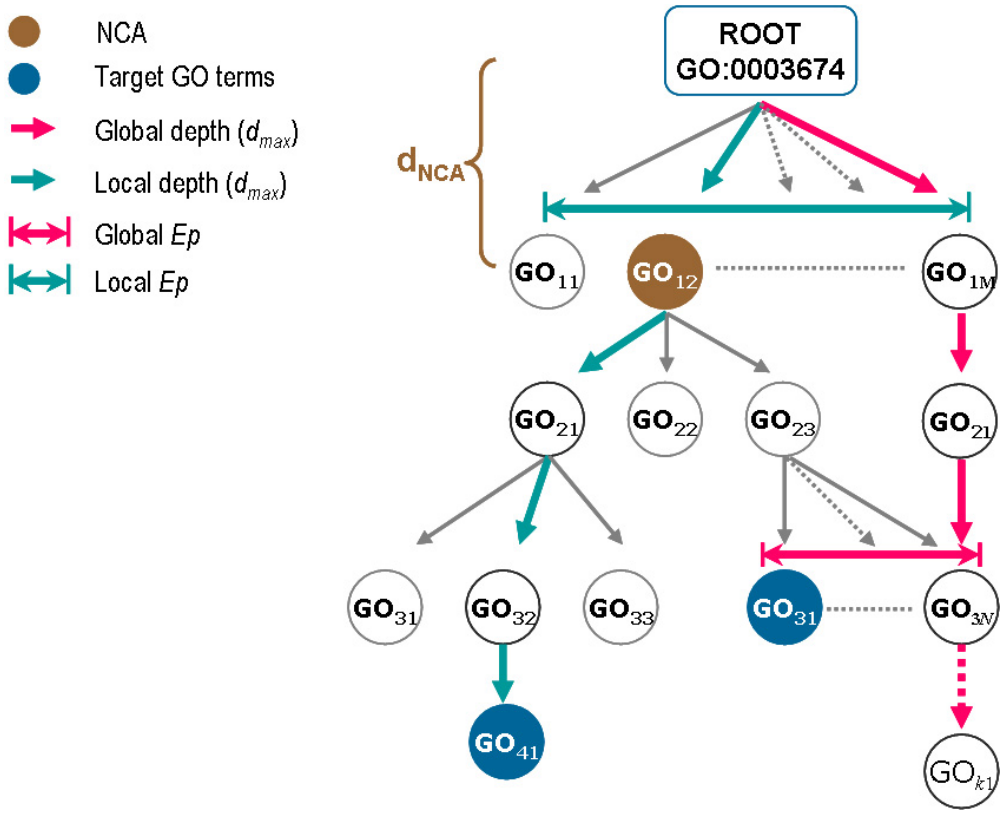



**Figure 3.1** Visualization of variables in $G_{scr}$.

Global and local are defined as maximum values found by considering entire GO graph or local path toward target nodes respectively. In this graph, the first digit in GO index denotes the depth of a GO term and the second digit denotes the order of edges belonging to a parent node. For example, $GO_{12}$ is nearest common ancestor of two target GO terms, $GO_{41}$ and $GO_{31}$. This node is located at depth 1 and second child (i.e, next to $GO_{11}$) of ROOT node. In practice, there is no such an order, so this index is only used for counting total number of child nodes of the target node. With this example, the *global* $d_{max}$=k and *local* $d_{max}$ = 4. By using same approach both *global* and *local* $E_{max}$ and Ep are N and M respectively where N > M.

In short, the rationale of the given equation is that the score of GO functional similarity should reflect both affinity of biological interactions among GPs and systemic significance of functional relations, thus such scores can provide a standard tool for analyzing associations of GPs under functional similarities and associations. Consequently this method can help elucidating and translating the relations of cellular functions in high throughput way without human interaction. The details of each score term will be further described below together with more detailed principles and algorithms.

### 3.2.1 The reference score, $R_{scr}$

The reference score described here shows a general procedure to calculate $R_{scr}$ from diverse databases, and it is not necessarily limited on specific databases. $R_{scr}$ requires at least two different databases which contain information about functions of domains, genes or GPs since the principles of multi view learning show that a high consensus of independent hypotheses or observations results in a low generalization error [120, 121]. In order to meet specified requirements, two different protein structure classification databases, SCOP [122] and CATH [123] are considered as two independent views. Creating $R_{scr}$ or more precisely $M_{scr}$ in Equation 3-1 requires several steps of procedures.

*Step one, downloading and parsing databases*: SCOP 1.75 (http://scop.mrc-lmb.cam.ac.uk/scop/parse/) and CATH 3.4 (http://release.cathdb.info/v3.4.0/) are downloaded and then each database is parsed to extract required information. Domain and homologous super family are retrieved from SCOP and CATH respectively to extract a group of proteins having high conservation on evolutionary and functional relationships. From this

filtering step, 10,214 domains and 2,549 homologous super families are retrieved from SCOP and CATH respectively.

*Step two, identifying associated proteins and mapping identification*: $R_{scr}$ is based on the assumption that closely related functions share the patterns of domains being present or absent on genes or GPs. Studies show that this assumption is biologically relevant. For example, Bodenreider *et al*. [84] used the same assumption to identify dependence relations among three different groups of GO, molecular function, cellular component, and biological process by projecting annotations from five different databases into a vector space, Yu *et al*. [76] annotated target genes by selecting GO terms associated with highly correlated reference genes with target gene in microarray data, and del Pozo *et al*. [65] defined GO similarity by using co-occurrence of GO terms upon the annotations of InterPro databases [124]. Once domains are retrieved upon this assumption, associated GPs are retrieved by referring PDB [125] names in each database. To get the annotations of PDB structure, the retrieved PDB IDs are mapped into UniProt [80] accession number.

*Step three, generating score matrix ($M_{scr}$) from the occurrences of GO annotation by projecting them into multiple views*: Annotations of proteins containing domains defined by SCOP and CATH are retrieved from UniProt. The occurrences of annotations are accumulated into the array of score matrix consisting with all domains and functional GO terms. The matrix filled with the occurrence of exact annotations is now called Original Direct Annotation (ODA). To make multiple views, diverse observations are applied. First, Hierarchical topology of GO database is applied such that an annotation is tagged in the score matrix then all ancestor nodes belonging to the target nodes (i.e. from the target node to root node) are also counted as current occurrences. This hierarchical annotation is called Original Hierarchical Annotation (OHA). To

generate more generalized view, protein-protein interaction (PPI) networks are considered. The applied PPIs are derived from KUPS [114] which is public databases consisting with manually curated PPIs derived from three PPI data bases, IntAct, MINT, and HPRD. Annotations of both target proteins associated with a domain and their direct PPIs are simultaneously tagged into a score matrix since it is well known that two proteins tend to have same function if they interact each other. This now is called PPI-based Direct Annotation (PDA). By using same approach creating OHA, PPI-based Hierarchical Annotation (PHA) is also produced by including direct interacting pairs in PPI network as current occurrences. Therefore, each protein structure classification database can produce four score matrix, ODA, OHA, PDA, and PHA, and this is resulted in total eight score matrix from both SCOP and CATH. Finally, more views are introduced by using binary projection of all existing views. Instead of counting the occurrences, these views only checks whether a GO term is appeared or not. In short, above all eight score matrix is converted into simple binary value such that if an array has equal or greater than an occurrence then the associated occurrence is projected into 1 in binary views and otherwise 0. Therefore, total 16 score matrices are created and they are all used for calculating $R_{scr}$.

*Final step, calculating GO similarity, $R_{scr}$ from 16 score matrices*: The final $R_{scr}$ is the average of cosine similarities calculated from 16 scoring matrices. The equation of cosine similarity is defined below together with graphical descriptions of score matrix at Figure 3.2. $GO_1$ and $GO_2$ are targeted two functional GO terms for which their semantic functional similarity is calculated. $\mathbf{GO}_1$ and $\mathbf{GO}_2$ are two column vectors in a scoring matrix associated with occurrence of domains. Due to the property of the cosine function, the values are ranged from 0 to 1 in which 0 means no similarity and 1 means exact match.

$$GO_{sim}(GO_1, GO_2) = \cos\theta(GO_1, GO_2) = \frac{GO_1 \cdot GO_2}{\|GO_1\| \times \|GO_2\|} \quad \textbf{(Equation 3-2)}$$

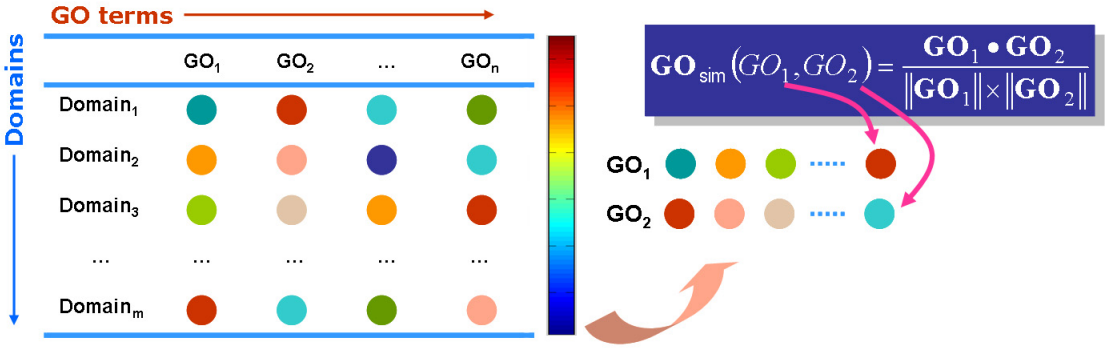


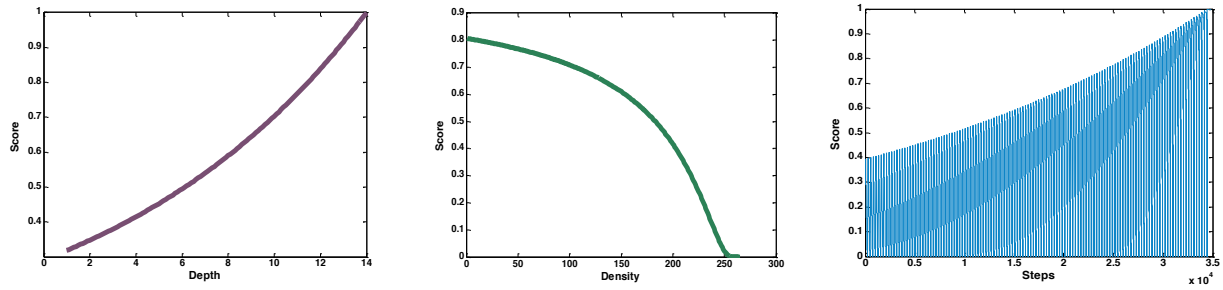

**Figure 3.2** Visualization of a score matrix and its similarity equation.
Each row consists of domains defined by SCOP or CATH and each column consists of all known functional GO terms. Colored circle represents the frequency of occurrences corresponding to domains and GO terms. Assume that the intensity of color is proportional to the frequency of occurrences (i.e. blue is lowest and red is highest). Each column vector of GO terms is fed into cosine similarity function, and the result is defined as semantic similarity between two GO terms.

### 3.2.2 The graph-based score, $G_{scr}$

The principle of graph-based score is to reflect the specificity of GO terms into their similarity measurements. To accomplish this aim, the specificity of a GO term pair is scaled by projecting both the depth of a node and certainty of a path into an equation. Four variables, Ep, $d_{NCA}$, $E_{max}$ and $d_{max}$ are introduced and the complete equation is defined below.

$$G_{scr} = exp\left(-\frac{1}{1-\frac{Ep}{E_{max}}}\right)^{\left(1-\frac{d_{NCA}}{d_{max}}\right)} \quad \textbf{(Equation 3-3)}$$

In Equation 3-3, Ep and $d_{NCA}$ are representing number of siblings and depth of NCA and used for defining certainty and depth respectively. The weight of certainty and depth toward the specificity is relative to their maximum values, $E_{max}$ and $d_{max}$. Thus these values can be defined in two different scopes, local and global. Local maximum indicates the maximum number of siblings or depth that belongings to the paths from root node to the leaf node that contains the target node. Global maximum results from exploring all possible paths associated with molecular functions in GO graph. The equations are designed to return higher score when the path is simpler and the depth is deeper which are corresponding to more distinctive and specific functional terms respectively. To see the effects of each variable, let's separate the equation into two terms: certainty term containing Ep and specificity term containing $d_{NCA}$. To get a high score from the given equation, the value of certainty term needs to be increased and specificity term should be decreased. Therefore, in order to increase the certainty term, the number of siblings of NCA should be small. In contrast to certainty term, the depth of NCA should be increased to increase the specificity term.




(a) (b) (c)

**Figure 3.3** Tendencies of $G_{scr}$.
(a) the tendency of $G_{scr}$ by increasing the depth of $d_{NCA}$ (b) the tendency of $G_{scr}$ by increasing the density of Ep (c) overall tendency of $G_{scr}$ which is normalized score between 0 and 1

In short, the equation results high score when the number of siblings of NCA, Ep is small and the depth of NCA, $d_{NCA}$ is deeper or getting closer to the leaf node. The visualization of $G_{scr}$ tendency is shown in Figure 3.3 Two major scoring schemes, $R_{scr}$ and $G_{scr}$, used in the proposed method have been introduced, and these scores represent top-down and bottom-up methods respectively; therefore, through weighted integration by using two more variable $\alpha$ and $\beta$, the proposed method, bidirectional integration is completed.

### 3.2.3 Calculating the similarity between Gene Products

In reality, calculating similarity score between individual GO terms is practically not much useful since most functions in GPs are described with more than a single GO term; therefore, additional metric is required for measuring similarity between groups of GO terms or GPs. To calculate the similarity of GPs, enhanced version of Hausdorff metric is used in this study by averaging the normalized Hausdorff metric. Similar approaches have been used for other studies as well [61, 62, 65, 73]. The equation is defined by

$$Hd(GP_1, GP_2) = \frac{1}{2}\left(\sqrt{\frac{1}{|\mathbf{GP_1}|}\sum_{x_i \in \mathbf{GP_1}} [GO_{sim}(x_i, \mathbf{GP}_2)]^2} + \sqrt{\frac{1}{|\mathbf{GP_2}|}\sum_{x_i \in \mathbf{GP_2}} [GO_{sim}(x_i, \mathbf{GP}_1)]^2}\right)$$

$$where\ GO_{sim}(x, \mathbf{GP}_i) = \max_{y \in \mathbf{GP}_i} GO_{sim}(x, y), i = 1,2 \qquad \textbf{(Equation 3-4)}$$

The Hausdorff metric between two GPs, $GP_1$ and $GP_2$ are defined by $Hd(GP_1, GP_2)$. This function averages two normalized similarity scores derived from reciprocal directions, from $GP_1$ to $GP_2$ and vice versa. $|\mathbf{GP}_1|$ and $|\mathbf{GP}_2|$ denotes the number of GO terms associated with two GPs,

$GP_1$ and $GP_2$ respectively. $GO_{sim}(\cdot)$ refers any method that can measure the similarity between two GO terms. However, it should be careful when the final score is calculated since the final score is purely depends on the $GO_{sim}$ whether it measures similarity or distance. If $GO_{sim}$ measures similarity then the final score should be based on the maximum score, but if distance is measured then it should be based on minimum score at each GO term. To give better understand on this metric, details are described with graphical representation.

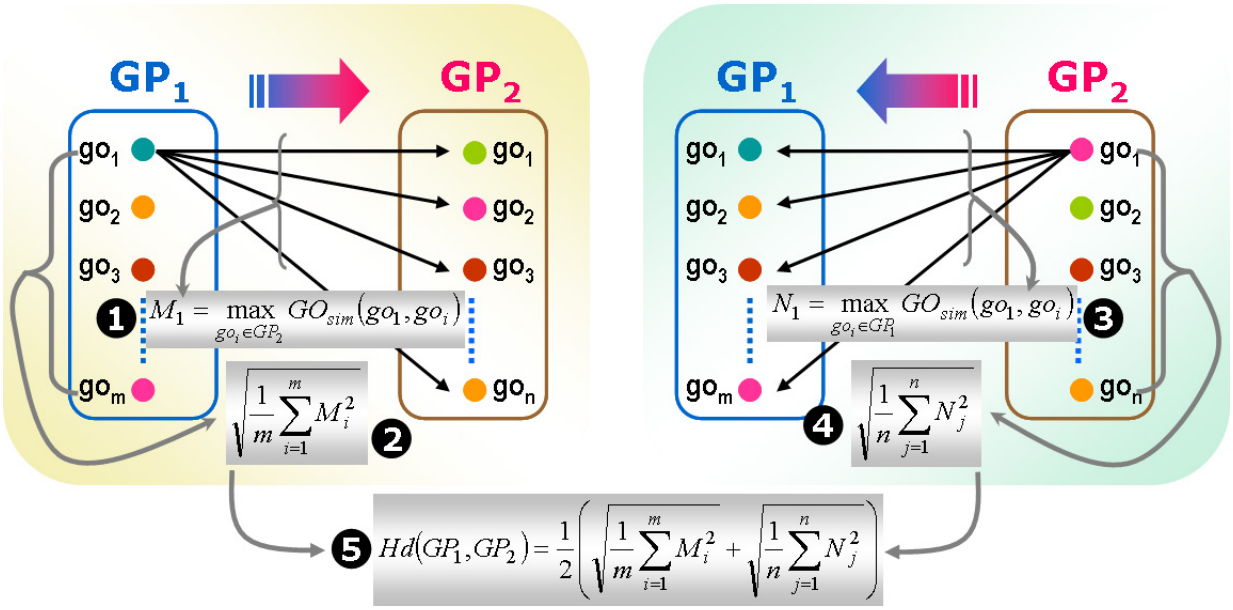



**Figure 3.4** Visualization of calculating similarity of GPs.
1) Identifying maximum (minimum) similarity (distance) between one of the GO terms in $GP_1$ against all GO terms in $GP_2$. 2) normalizing maximum similarity. 3) and 4) reproducing 1) and 2) with opposite direction respectively. 5) the final Hausdorff score is the average of normalized similarity.

So far, overall procedures of the proposed method from measuring a pair of GO term to GP similarity have been introduced. In summary the proposed method integrates five different databases, GO, UniProt, KUPS, CATH and SCOP to reflect broad range of biological factors into functional similarity.

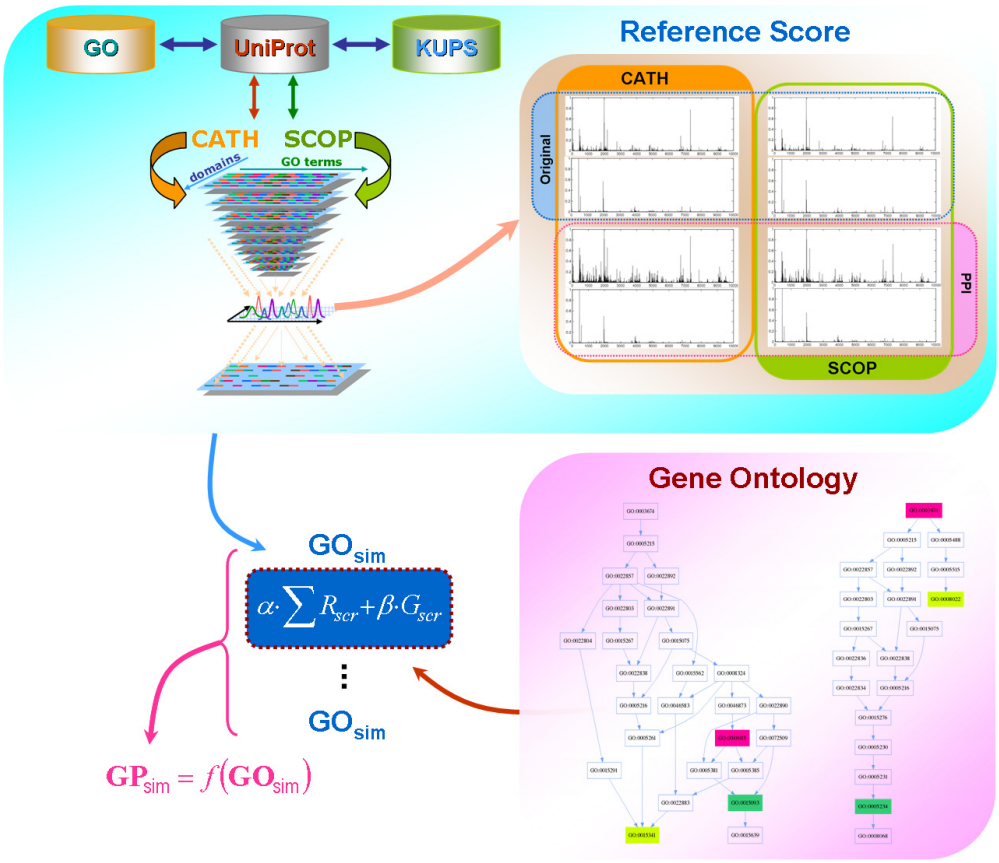



**Figure 3.5** Visualization of proposed method.

To the clockwise, first image shows the way to create $R_{scr}$ which integrates five databases by mapping and retrieving information among GO, Uniprot, KUPS, CATH and SCOP. The initial score matrix is composed of annotation occurrences upon domains defined by CATH and SCOP. As producing multiple views, diverse score matrices are derived such as referring annotations from PPIs, hierarchical annotations and binary projections. The second Figure shows the distribution of annotations in different score matrix: left (orange) and right (green) show two databases, CATH and SCOP respectively, and top (blue) and bottom (PPI) show the distribution of scores whether PPI pairs are considered or not. These distributions only show the occurrences of annotation, but actual score matrices contains more views by projecting all scores into binary conditions based on appearance and absence of GP occurrences; therefore, this makes total 16 different score matrices. The third Figure shows examples of GO graphs in which the light green and green are a pair of target GO terms, and pink node denotes their NCA. This information is fed into equation 3-3 and resulted in $G_{scr}$. The last Figure shows that the final score of semantic similarity between GO terms are made by the summation of weighted $R_{scr}$ and $G_{scr}$. Finally the GP similarity is defined by a function of individual GO similarity which is defined in equation 3-4

To reduce the generalization error or bias effects, the principles of multi-view learning is adapted such that observed factors are projected into diverse spectrums of data distribution as multiple views. Both integration in wide range of biological factors and systematically designed algorithm, which helps to reduce generalization error, guarantee that the proposed method can effectively handle the problems of measuring functional similarity. The visualization of proposed method is shown in Figure 3.5.

### 3.2.4 Theoretical relevance of $R_{scr}$

The reference score, $R_{scr}$ has two main purposes: giving a metric of GO terms shared by the root node and reducing general errors caused by single dataset. As mentioned above, 52% of GO term pairs out of all possible pairs of currently existing GO terms are shared by root node. This means conventional methods defining similarity based on the location of nearest common ancestor (NCA) will give nil similarity for these pairs or treat them as having exactly same similarity. Defining them as either nil similarity or same similarity does not seem like biologically reasonable, especially by considering the facts that that more then half of GO term pairs are shared by root node and most of GPs have multiple GO terms. The semantic similarity of proposed method calculates root-shared GO terms with following assumption. Two GO terms are close if two GO terms tend to be co-annotated by GPs having same functional domains and/or their interacting partners since GO terms are designed for describing the function of GPs; therefore, if the function of two GPs are similar each other then their functional terms should be similar as well. In fact, it is well known that if two proteins interact each other then they tend to have similar functions [126-128].

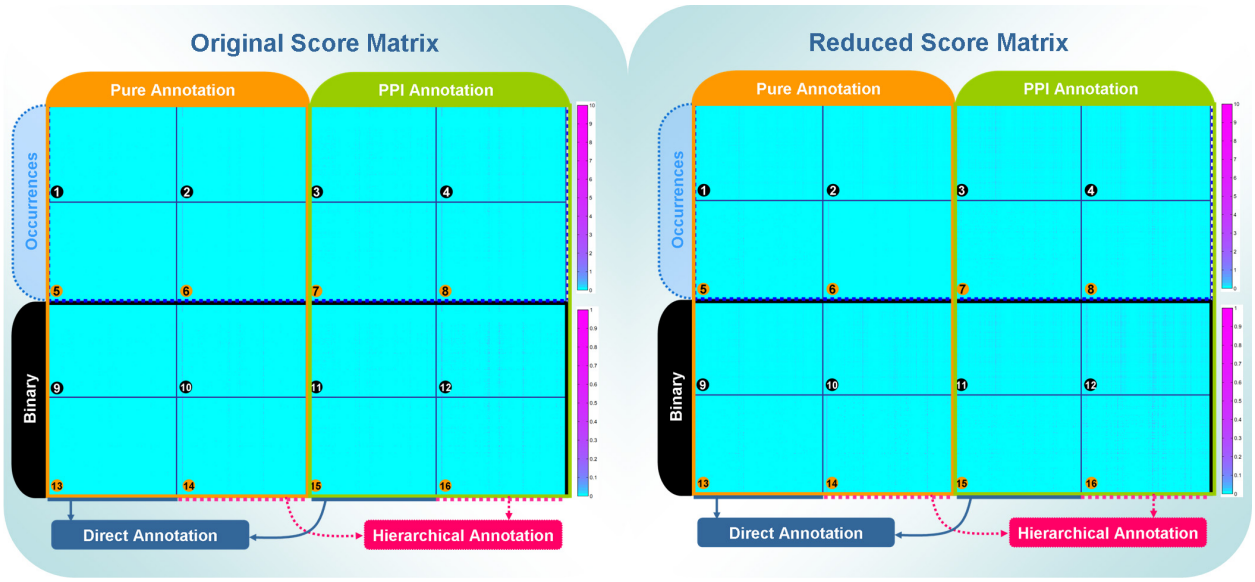



**Figure 3.6** Heat map of 16 score matrices.

The left image, 'original score matrix' is the actual matrix that used for calculating $R_{scr}$ and the right image, 'reduced score matrix' is heat maps to get better resolution of correlation significance by eliminating columns and rows with all zero elements. The black circle numberings depicted in the each of left bottom of score matrices indicate score matrices derived from CATH and the orange circle numberings depicted in the left bottom of rest score matrices are corresponding to score matrices derived from SCOP. Upper two rows denoted as "occurrences" in 16 score matrices are created by occurrences of GO terms at each domain. Bottom two rows denoted as "Binary" in 16 score matrices are obtained by projecting occurrences into binary value (i.e, one or zero) based on their presence and absence respectively. First two columns denoted as "Pure Annotation" of 16 score matrices are calculated by relying on the annotations of GPs considering corresponding functional domains only. Rest two columns denoted as "PPI Annotation" among 16 score matrices are extended versions of the first two columns by considering corresponding functional domains and all GPs containing the domains while they appear in integrated PPI database (i.e, KUPS). First and third column denoted as "Direct Annotation" counts the occurrence of annotations with exact GO terms. Second and fourth column denoted as "Hierarchical Annotation" counts the occurrence of annotations by including all ancestors of the target GO term derived from hierarchical structure of GO.

To incorporate this assumption into the proposed similarity score, multiple score matrices are designed which are derived from two domain structure databases SCOP [129] and CATH [123], and manually curated comprehensive protein-protein interaction networks, KUPS [130]. The final score matrices are completed by projecting the initial score matrix into statistically different

views. The final score measuring similarity between two GO terms is made by integrating these multiple views or observations; thus the final score represents the high consensus of multiple observations. In fact, similar approach is already well known in machine learning area as a multi-view learning which is know to reduce generalization error [120, 121]. The actual distributions and relationships among 16 score matrices are depicted from Figure 3.6~9. In Figure 3.6, figures show the heat map of the actual 16 matrices. Each column of the matrices corresponds to all functional GO terms in GO database and each row corresponds to the domains. The numbers in black and orange circles denote CATH and SCOP database respectively. The domains are from two different databases. Thus, these score matrices are created independently. Since the original score matrix is too sparse to visualize differences in a small figure, reduced score matrix is also shown in the right image at Figure 3.6 Reduced score matrices are drawn by eliminating columns and rows of which elements of either columns or rows do not have any associated GPs. These figures show that the patterns of score matrices derived from two different databases are similar each other. In other words, the pattern of GO terms annotated by GPs having similar domains is conserved even if it is derived from different domain databases. This implies that the similarity of patterns between two GO terms can be used for defining the functional relations between two GO terms since their patterns are derived from functional domains. To check out the distributions of domains upon GO terms, simple statistics are conducted such that the correlations among the normalized distribution of annotated GPs along with GO terms are calculated. Let's assume that a score matrix is composed of $m$ by $n$ matrix. Here $m$ corresponds to the number of domains, and $n$ corresponds to the number of GO terms in GO database and then the normalized distributions of GO terms are defined by $\mathbf{G}$ =

$[f(G_1), f(G_2), \ldots, f(G_n)]$ in which $f(G_i)$ denotes the normalized value of $G_i$ ($i$=1,…$n$). The equation is defined below.

$$G_i = \sum_{d=1}^{m} g_{d,i} \quad i = 1 \cdots n$$

$$f(G_i) = \frac{G_i - min(\mathbf{G})}{max(\mathbf{G}) - min(\mathbf{G})}$$

**(Equation 3-5)**

Here $g_{d,i}$ denotes the value corresponding to $d^{th}$ row and $i^{th}$ column of a score matrix; therefore, this equation shows the relative distributions of domains belonging to GO terms at the given matrix. To check out relations within and between score matrices, correlations of normalized distributions are calculated. For the convenience, each of 16 score matrices is named as following: original direct annotation (ODA), original hierarchical annotation (OHA), PPI direct annotation (PDA), PPI hierarchical annotation (PHA), binary original direct annotation (bODA), binary original hierarchical annotation (bOHA), binary PPI direct annotation (bPDA), and binary PPI hierarchical annotation (bPHA) in which *original* and *PPI* denotes that GO term occurrences are counted by using GPs having target domains only or including direct interacting pairs of the GPs respectively. *Direct* and *Hierarchical* annotation denotes whether the occurrences of GO terms are counted by using an exact match or including all ancestor nodes respectively. More precisely, in Figure 3.6, the name of individual matrix is given as: ODA, OHA, PDA, PHA, bODA, bOHA, bPDA, bPHA of which the name is corresponding to the order of numbers from 1 to 4 and from 9 to 12 for CATH and from 5 to 8 and 13-16 for SCOP.

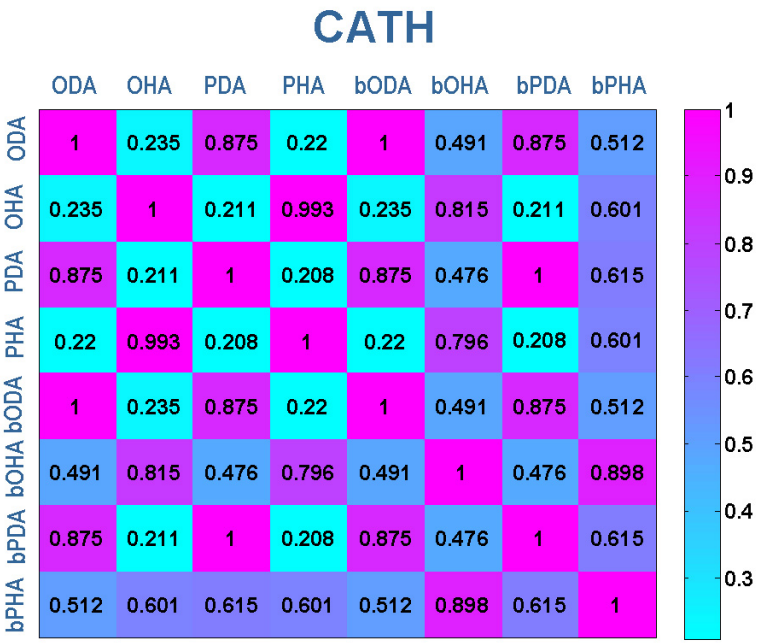



**Figure 3.7** The visualization of correlation coefficient derived from score matrices within CATH.

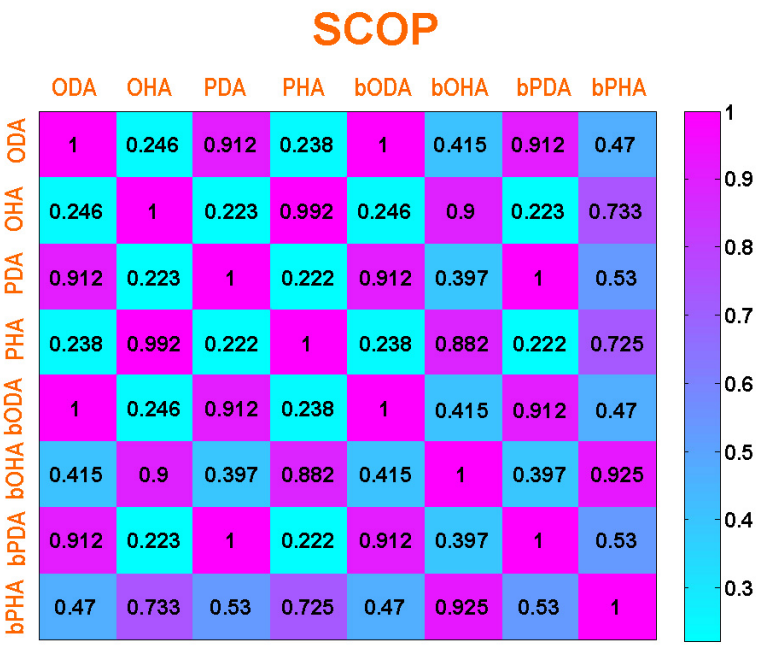



**Figure 3.8** The visualization of correlation coefficient derived from score matrices within SCOP

Figure 3.7 and 3.8 show the correlations within score matrices derived from CATH and SCOP respectively. Each of heat maps clearly shows that homologous scoring schemes conserve high correlation. For example, no matter what databases are referred, direct annotations are highly correlated with any types of direct annotations and similarly the patterns of hierarchical annotations are also high correlated each other. Interestingly the binary-based score matrices also conserve similar patterns with the ones in occurrence-based score matrices. However, compared to the occurrence-based score matrices, their correlations are rather smoothly distributed by increasing the correlation between homologous annotations and decreasing the correlation between heterologous annotations such that the correlation within direct annotations or hierarchical annotations are increased but the correlation between direct and hierarchical annotation are decreased. Thus these two different scoring schemes give positive effects on reducing the noises or biased observations by increasing the number of views. The correlations between two independent databases, CATH and SCOP are also calculated, and the results are shown in Figure 3.9. The patterns of correlated relations between CATH and SCOP are very similar to the ones calculated from within the individual databases. The overall statistics show that the distributions of GPs associated with individual GO terms in multiple matrices are highly conserved, so this should results in reductions of biased and noise effect; thus this implies that the power of co-annotations categorized at same functional domains can be used for robust similarity measurement of two GO terms.

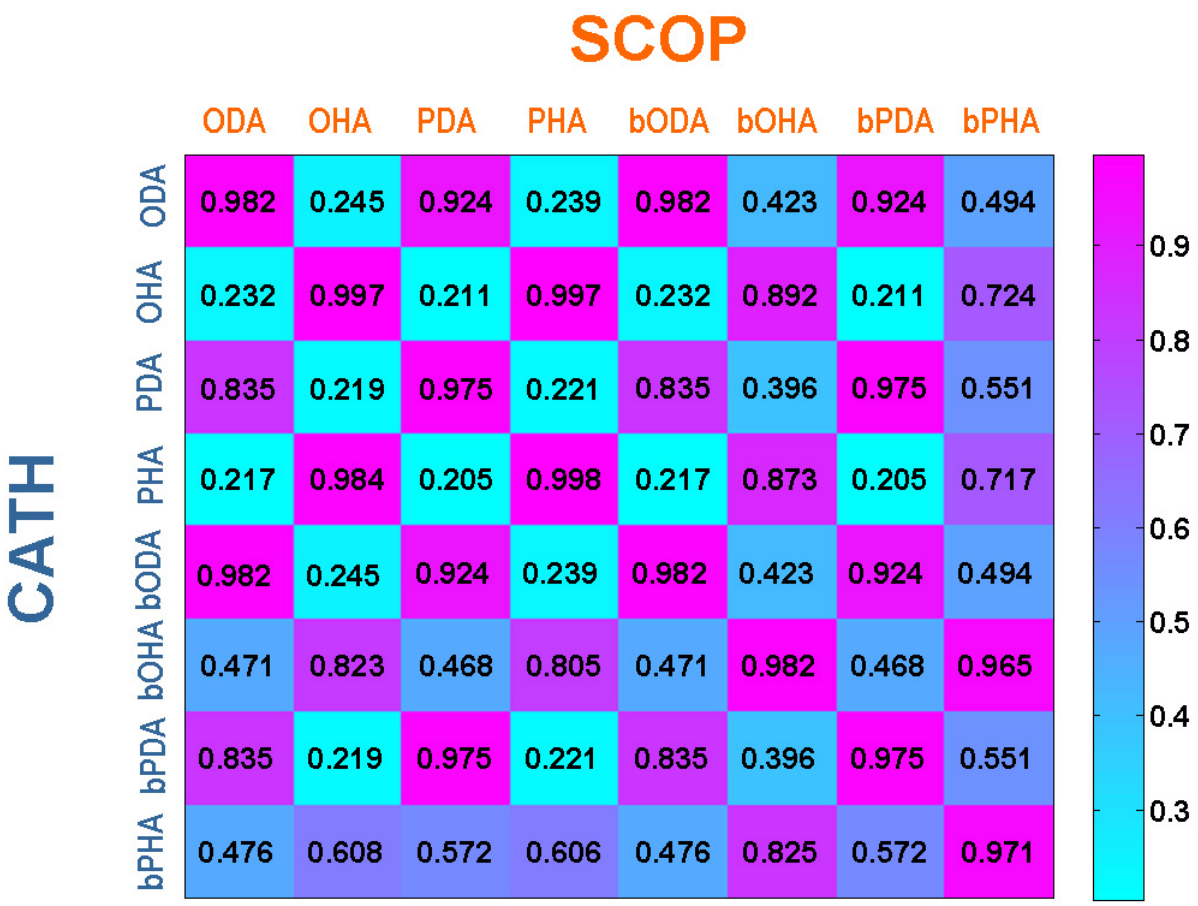
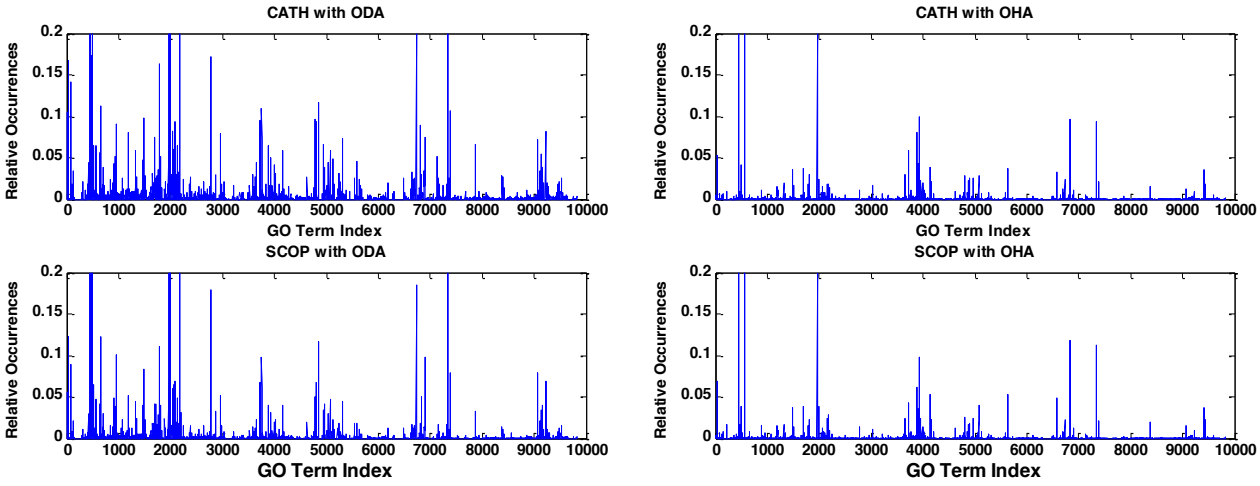



**Figure 3.9** The visualization of correlation coefficient between CATH and SCOP.

For the convenience, the names of 16 score matrices are given as following in the rest of this article: original direct annotation (ODA), original hierarchical annotation (OHA), PPI direct annotation (PDA), and PPI hierarchical annotation (PHA) from CATH and SCOP. In addition, the occurrences in all eight score matrices are projected into binary space which checks whether each array of score matrices are present or absent. At each term, *Original* and *PPI* denotes that GO term occurrences are counted by using GPs having target domains only or including direct interacting pairs of the GPs respectively. *Direct* and *Hierarchical* annotation denotes whether either using an exact match or including all ancestor nodes respectively to count the occurrences of GO terms. More precisely, numbers from 1 to 4 and from 9 to12 are derived from CATH, and numbers from 5 to 8 and from 13 to 16 are derived from SCOP. The numbers in each group (i.e. CATH and SCOP) can be denoted with original direct annotation (ODA), original hierarchical annotation (OHA), PPI direct annotation (PDA), PPI hierarchical annotation (PHA), binary original direct annotation (bODA), binary original hierarchical annotation (bOHA), binary PPI direct annotation (bPDA), and binary PPI hierarchical annotation (bPHA) as corresponding to the order of numbers.

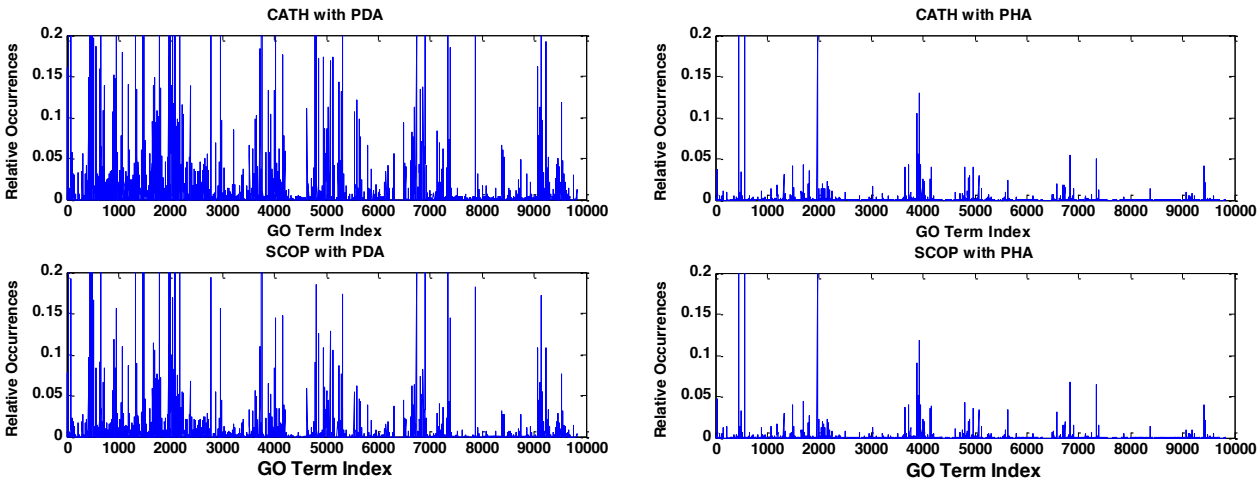



(a) (b)

**Figure 3.10** Relative occurrences of GO term with original accumulated array.

The number of occurrences is sequentially increased if a new occurrence is detected at the previously counted array corresponding to same domains and GO terms. Relative occurrences are calculated by the ratio between each corresponding occurrence and the maximum occurrence at the score matrix. In order to emphasize the differences between score matrices, the relative occurrences smaller than 0.2 are magnified (i.e, the relative occurrences, greater than 0.2 are not fully depicted (a) First row, the histogram of original direct annotation (ODA) from CATH and second row, the histogram of ODA from SCOP. (b) First row, the histogram of original hierarchical annotation (OHA) from CATH and second row, the histogram OHA from SCOP.

Figure 3.10~13 are provided for visualizing comparisons among normalized distributions of individual score matrices derived from occurrences and binary. Due to the significant difference between maximum and minimum number of GP occurrences, the figures used cut-off value of 0.2 In the graphs the horizontal axis denotes the index of total 9,833 GO functional terms, and vertical axis denotes *relative occurrences* calculated by the ratio between each corresponding occurrence and the maximum occurrence at the score matrix.




(a) (b)

**Figure 3.11** Relative occurrences of GO term with PPI-based accumulated array.
(a) First row, the histogram of PPI direct annotation (PDA) from CATH and second row, the histogram PDA from SCOP. (b) First row, the histogram of PPI hierarchical annotation (PHA) from CATH and second row, the histogram PHA from SCOP. *Original* and *PPI* denotes that GO term occurrences are counted by using GPs having target domains only or including direct interacting pairs of the GPs respectively. *Direct* and *Hierarchical* annotation denotes whether the occurrences of GO terms are counted by using an exact match or including all ancestor nodes of the target GO term respectively.

Figure 3.10 (a) compared relative occurrences between CATH and SCOP upon ODA which correspond to the top and bottom of the figure. Although there are subtle differences, for example, the picks around 7,000 and 9,000 have different numbers and shapes, overall patterns look very similar each other; thus this results in high correlation as shown in Figure 3.9 (i.e. 0.982). Rest of figures, Figure 3.10 (b) and Figure 3.11 (a) and (b) also shows similar patterns. Low correlation or dissimilar patterns can be identified by comparing heterogeneous matrices. Comparing PDA from CATH and PHA from SCOP, which are located in Figure 3.11 (a) and (b) respectively, could be one of the examples to figure out the differences between two figures such that the density of picks are clearly noticeable to shows the differences. In fact, their correlations

are denoted relatively low and can be found in Figure 3.9 (i.e. 0.205). OHA from CATH and PDA from SCOP which are located in the Figure 3.10 (b) and Figure 3.11 (a) respectively also show the clear difference of patterns between two matrices. As expected, their correlation is relatively low and can be found in Figure 3-9 (i.e. 0.211). Binary version of Figure 3.10 and 11 are depicted in Figure 3.12 and 13. Indeed, compared to direct annotation, the patterns of the hierarchical annotation between binary and occurrence-based scoring matrix show significant difference. As expected and described before, the patterns are more evenly distributed in hierarchical annotations in binary-based score matrix. This implies that binary-based scoring scheme helps to make uniform distribution by emphasizing the fidelity or specific functions in cellular life while fading the weight of general GO terms located in higher level (i.e, closer to root GO term). In other words, under the consideration of hierarchical annotations, broader terms have higher occurrences than specific term since the general terms are counted redundantly while their descendent terms are counted. However, binary based-scoring matrix does not rely on the occurrence of annotation, but it only considers the status of appearance or absence. The consequence of binary-based scoring matrix can be easily found in Figure 3.7~9 where correlations in heterogeneous binary-based score (e.g, correlation between bODA and bOHA is 0.491, 0.415 and 0.423 in Figure 3.7, 8 and 9 respectively) is higher than correlations in occurrence-based heterogeneous matrices (e.g, correlation between ODA and OHA is 0.235, 0.246 and 0.245 in Figure 3.7, 8 and 9 respectively). It is of interest to analyze the relations between the final score of proposed method and $R_{scr}$ only. For the analysis, 50 GO terms having $R_{scr}$ score greater than 0.5 are randomly selected to ensure that the final score is derived from both $R_{scr}$ and $G_{scr}$. The second column in table 1 shows that $R_{scr}$ is likely to give higher score for '*tight*' relations of which NCA and two GO terms have parent and child relationship.

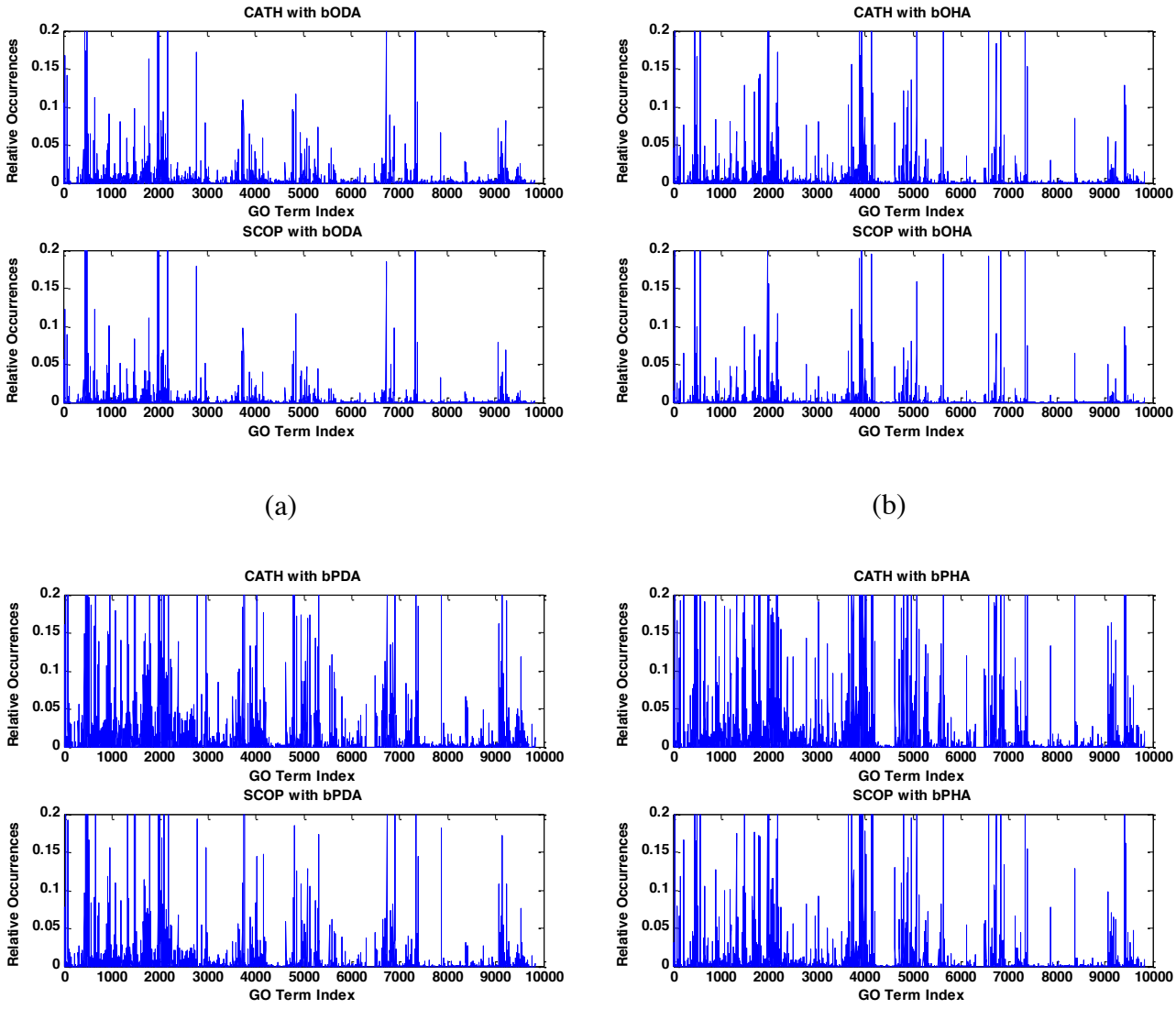



**Figure 3.12** Relative occurrences of GO term with binary array.

The number of occurrences are denoted as 1 or zero if corresponding array is present or absent respectively. In order to emphasize the differences between score matrices, the relative occurrences smaller than 0.2 are magnified (i.e, the maximum value shown in the graph is limited to less than and equal to 0.2). (a) First row, the histogram of binary original direct annotation (bODA) from CATH and second row, the histogram of bODA from SCOP. (b) First row, the histogram of binary original hierarchical annotation (bOHA) from CATH and second row, the histogram bOHA from SCOP. (c) First row, the histogram of binary PPI direct annotation (bPDA) from CATH and second row, the histogram bPDA from SCOP. (d) First row, the histogram of binary PPI hierarchical annotation (bPHA) from CATH and second row, the histogram bPHA from SCOP.

This seems very reasonable since they actually have same parent and indeed, they are the most specific term in their path. To see the details of the results please refer supplement material [S-7]. To evaluate biological relevance of $R_{scr}$, simple statistics of annotated GPs in UniProt is conducted, and the results are shown in Table 3.1. In this table 1st column shows the average score of GO term pairs in a category. 2nd and 4th columns show average number of GPs annotated by 'either' or 'both' of GO term pairs in a category respectively. 3rd and 5th columns show the average number of GPs having an interaction partner among the GPs in 2nd and 4th column respectively. As an intuitive comparison, 1st column shows that $R_{scr}$ is somewhat proportional to the geometric similarity: the closer distance the higher score. 2nd column shows the difference between $R_{scr}$ and IC-based score since IC-based method penalizes GO terms associated higher number of GPs. Indeed, the problems of penalty policy in IC-based method can be also observed from 1st, 2nd and 4th columns since the order of intuitive score and IC-based score are contradicted such that the direction of scores goes opposition direction. In other words, 'tight' should have higher score than 'root' under the intuitive comparison, but by considering the characteristics of IC-based method in which the higher number of GPs has lower score, IC-based methods relatively give higher score to 'root' and lower score to 'tight' relations. This intuitive contradiction of IC-based methods is also reported by Wang et al. [73]. The 3rd and 5th columns show very interesting point and support that the root-shared pairs having relatively high score are not from random selection. As shown in the 3rd and 5th columns at the table, the higher score should be from the effect of considering interacting partners in PPI networks since their annotated GPs are relatively high affinity to be interact. In short, this brief statistic experiment shows that $R_{scr}$ tends to reflect both intuitive and biological relations into its scoring scheme by giving positive feedback toward GO terms having high number of GPs co-annotated and

involved in PPIs. This also shows the significant different between the proposed method and conventional IC-based methods in which all root-shared GO terms are considered as either same or nil similarity.

**Table 3.1** Comparing scores and biological significance of GO pairs
❶ Average reference score ($R_{scr}$) ❷ Average number of GPs annotated by a GO term in each category for which total number of GPs annotated by GO terms in the given category is divided by total number of GO terms in the given category ❸ Average number of GPs including direct interaction partners of GPs annotated by one of GO terms in the target GO term pairs at the given category ❹ average number of GPs annotated by both GO terms consisting a pair of GO terms at each category ❺ average number of GPs which is interacting and annotated by both GO terms consisting a target GO term pair at each category. In category, *Tight* denotes direct parent-child relations between both GO terms and their NCA. *Root* denotes NCA of a GO pair is the root node. *Loose* denotes all GO pairs except *Tight* and *Root* relations

| Category | ❶ | ❷ | ❸ | ❹ | ❺ |
|---|---|---|---|---|---|
| **Tight (17)** | 0.69 | 56.94 | 0.70 | 10.94 | 0.11 |
| **Loose (26)** | 0.68 | 29.92 | 0.76 | 2.42 | 0.23 |
| **Root (7)** | 0.66 | 23.85 | 0.85 | 2.14 | 0.42 |
| **Overall (50)** | 0.68 | 38.26 | 0.76 | 5.28 | 0.22 |

# 4 Experiments and results

## 4.1 Experiment I: measuring functional correlation

### 4.1.1 Dataset

To evaluate the performances of the proposed method and other conventional methods, similarity scores calculated from SCOP domain similarities are compared. To get the annotations of each domain in SCOP, high-quality truly domain-centric GO annotations are downloaded from Superfamily DB (http://supfam.cs.bris.ac.uk) [131] and parsed and integrated into our local database to retrieve the information later. Based on the annotation of SCOP domain at family level, total 2,690 unique domains are retrieved. 44 out of 2,690 domains are annotated with non-existing GO terms in current GO DB, so these domains are eliminated from the list. Finally total 2,646 domains are obtained from SCOP; therefore, total 3,501,981 all possible pairs of SCOP domains are used for experiments. The list of SCOP domains used in the experiments are attached in supplementary material [S-8].

### 4.1.2 Similarity measurements

To evaluate methods, four different similarity scores are compared to functional similarities calculated by the proposed method and six conventional methods. Four different similarity score is derived from three categories: sequence, phylogenetic profile and structure similarity. Once all similarity scores are calculated, all scores are grouped into 50 bins based on target similarity. For example, similarity scores calculated by BSM score are grouped into 50 bins from 0 to 1, and then the scores, which are calculated from sequence, phylogenetic profile and structure similarity, corresponding to domain pairs associated with each bin is averaged.

***Sequence similarity*:** sequence similarity is one of the most popular methods identifying the inheritance of homology [132] and homologous proteins are likely to share their functions although several limitations have been reported [94, 111, 133, 134]; therefore this leads the assumption that sequence similarity and functional similarity tend to be correlated. To get the sequence similarity, all SCOP domain sequences are downloaded from ASTRAL [135] (http://astral.berkeley.edu/), and sequence similarity of all possible pairs of domains are calculated by BLAST program with default parameters [136](http://blast.ncbi.nlm.nih.gov).

***Phylogenetic profile similarity*:** the rational behind phylogenetic profiles are from the hypothesis that functionally linked proteins are evolved in a correlation fashion; therefore the homologs of functionally related proteins can be found in the same subset of organisms [137]. The phylogenetic profiles are created by checking whether the homologs of proteins are present or absent in a same organism. To create phylogenetic profiles, psiBLAST searches homologs of target proteins against proteins from 25 fully sequenced genomes (see supplementary material [S-9]) with e-value less than $10^{-10}$. The similarity of phylogenetic profiles between domains is calculated by mutual information (MI) as following:

$$MI(\mathbf{x},\mathbf{y}) = \sum_{x=0}^{1}\sum_{y=0}^{1} p(x,y)\cdot \log\left[\frac{p(x,y)}{p(x)\cdot p(y)}\right] \qquad \textbf{(Equation 4-1)}$$

*where* $x \in \mathbf{x}$ *and* $y \in \mathbf{y}$

Here, **x** and **y** denote two vectors of phylogenetic profiles associated with any pairs of two domains. $x$ and $y$ are elements of two vectors correspondingly. $p(x,y)$ is a joint probability of two domains derived from four possible states of occurrences in which 0 states absence and 1 denotes

presence. $p(x)$ and $p(y)$ describe the probability of absence and presence of corresponding domains.

***Structural similarity*:** the knowledge of protein structures often provides a crucial insight of protein function. Indeed, protein structure is known as more conserved than sequence in evolutionary process[138]; therefore, this implies that the structural similarity can represent the functional relations between two proteins. To increase the relevance of measuring structural similarity, two different structural alignments are used: TM-align [22] and Dali [139]. Corresponding programs are downloaded from http://zhanglab.ccmb.med.umich.edu/TM-align/ and http://ekhidna.biocenter.helsinki.fi respectively. The similarity scores are calculated by 3,501,981 all possible pairs of domains on cluster machine.

### 4.1.3 Results

To compare the performance between the proposed method and other conventional methods, the tendency of domain similarity is analyzed. The hypothesis in this experiment is that if genes or GPs are functionally similar, then their structure, sequence, and phylogenetic profiles tend to be similar each other. For example, if a pair of GP similarity is close to 1 which corresponds to exact match or closely related each other, then structure, sequence, and phylogenetic profile could be identical or very similar each other as well. This hypothesis is commonly used in many studies although the hypothesis has not been comprehensively verified yet [58, 59, 61, 62, 65, 69, 70, 72, 99, 100]. Therefore, this hypothesis should be verified before further experiments proceed. To do this, similarity relations among three categories are compared, and the results are shown in Figure 4.1.

The details about experimental procedures are described in method section. Figure 4.1 shows tendency of domain similarity based on the score calculated by TM-align (TM), Dali (DALI), phylogenetic profile (PHY), and sequence alignment (SEQ). To help the visual comparison, fitted graph is also drawn through each method, and prefix '*fit*' is added in front of each method to identify fitted graph from the graphs derived from actual similarity scores. Graphs show overall similarity scores derived from three different categories are likely to be correlated.

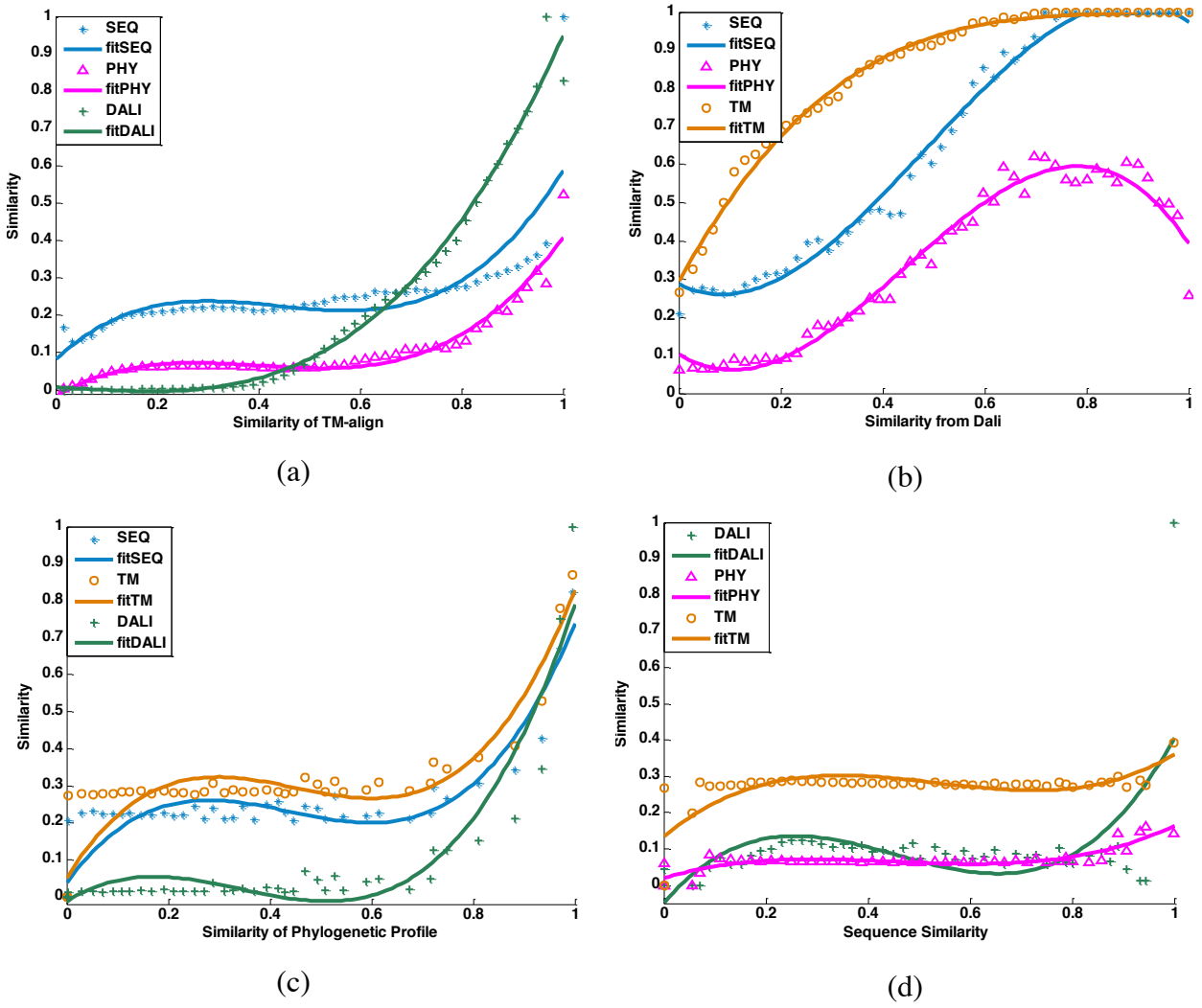



**Figure 4.1** The tendency of domain similarity among three categories.
Structure(TM-align and DALI), sequence (SEQ), and phylogenetic (PHY) profile. Horizontal axis denotes similarity score based on one of four methods (i.e, TM-align, DALI, SEQ, and PHY), and vertical axis denotes similarity score for rest of methods. (a) and (b) the tendency of domain similarity based on structure similarity calculated by TM-align and Dali respectively. (c) the tendency of domain similarity based on phylogenetic profile. (d) the tendency of domain similarity based on sequence similarity.

However, as observed in Figure 4.1, their correlations are not exactly proportional to each other. More precisely, in Figure 4.1 (a), there is stationary area from 0.2 to 0.5 in TM similarity, and strong correlation is observed at higher than structure similarity 0.6. In Figure 4.1 (b), Dali shows very clear correlation between 0.2 and 0.8. However, when zero similarity considering the

most left bottom of each graph at Dali (i.e, Figure 4.1 (b)) is compared to same part of all others, Dali's initial graphs show higher similarity score than those of others (i.e. SEQ and PHY stays around 0.3 and others stay around 0.1). This shows lower Dali score is somewhat chaotic and unreliable. Indeed, when Dali reaches more than structural similarity 0.8, TM-align and sequence similarities are convergent, and phylogenetic profile similarity is dropped significantly. This mismatched tendency shows that higher similarity score of Dali does not guarantee to make agreement on different characteristics of similarity. The tendency of graphs derived from phylogenetic profile in Figure 4.1 (c) and from TM-align in Figure 4.1 (a) show very similar patterns. In contrast to others, Figure 4.1 (d) shows sequence similarity requires very high similarity to observe similarities such that the correlations among three categories are only observed at high similarity degree (i.e. greater than 0.8). However, this result is not that surprising and can be found in previous studies such that the results show finding similar GPs often requires high sequence similarity [49, 50, 140]. Indeed, unreliable performances of sequence similarity toward identifying functional similarity also have been reported [39, 40, 51, 52]. The unreliable sequence similarity itself is not a major concern in this study. The actual problem is that most conventional methods identifying functional similarity have been verified their methods by using sequence similarity only or heavily focusing on the results from sequence similarity to emphasize the performance of proposed methods [59, 61, 65, 69, 71, 72]. Besides showing the relations of sequence similarity toward other similarity scores, as a guidance of evaluation, it is strongly recommended to use multiple measure measurements to verify the rational of the similarity measurement methods. Overall results show that correlations do exist among three categories of similarity measurements although there exist some specific tendencies of which they are not exactly proportional between similarity scores and functional similarity,

but with certain degree or range they do show clear correlations. Indeed, due to the existence of difference tendencies along with different range of similarity score, it is challenge to find such method that makes agreement on correlations among three different categories. In other words, the tendency of all agreement on correlation can be used for one of evaluation criteria to judge similarity methods as a statistical point of view. By following the proposed evaluation criteria and upon the verification of correlation tendency among three categories, conventional methods and proposed methods are compared, and results are shown in Figure 4.2~5. Horizontal and vertical axis in Figure 4.2~5 is similarity score derived from the given method and criteria measurements derived from three different categories respectively. Figure 4.2 and 4.3 are the proposed method, BSM score and Ye's method respectively and show similar patterns of similarity graph along with different critera. The differences between two graphs are that Ye's method is somewhat looks like shrunken version of the proposed method. In fact, the range of score in Ye's methods is between 0 and 0.5. This happens due to the effect of normalization process in which the depth of NCA is normalized by maximum depth of entire graph. In contrast to Ye's method, the proposed method has score range between 0 and 1. More interestingly, the graphs shown in the proposed method (Figure 4.2) are very similar to those of TM-align in Figure 4.1 (a). By considering the fact that no structural information is used, it is stunning that *the proposed method can derive the result as similar as the ones derived from structural information*

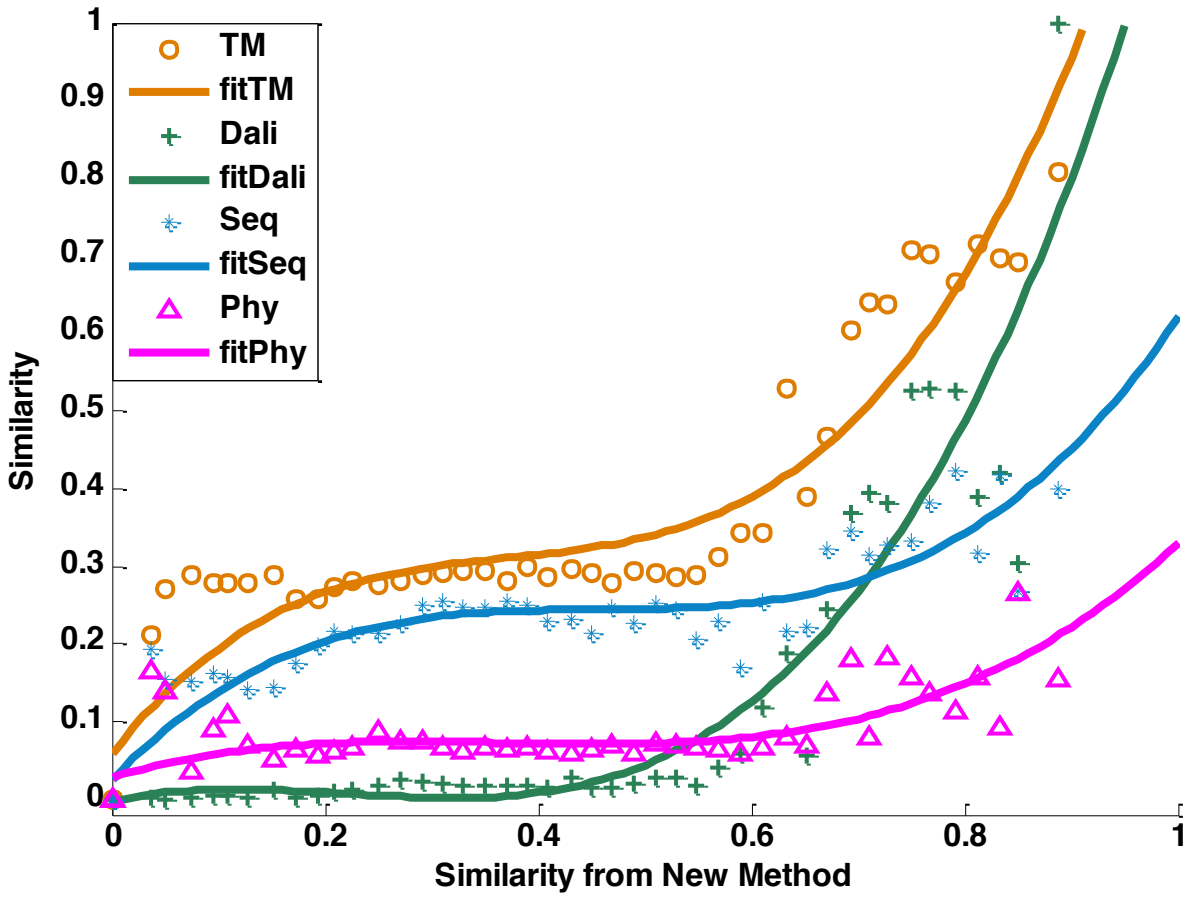
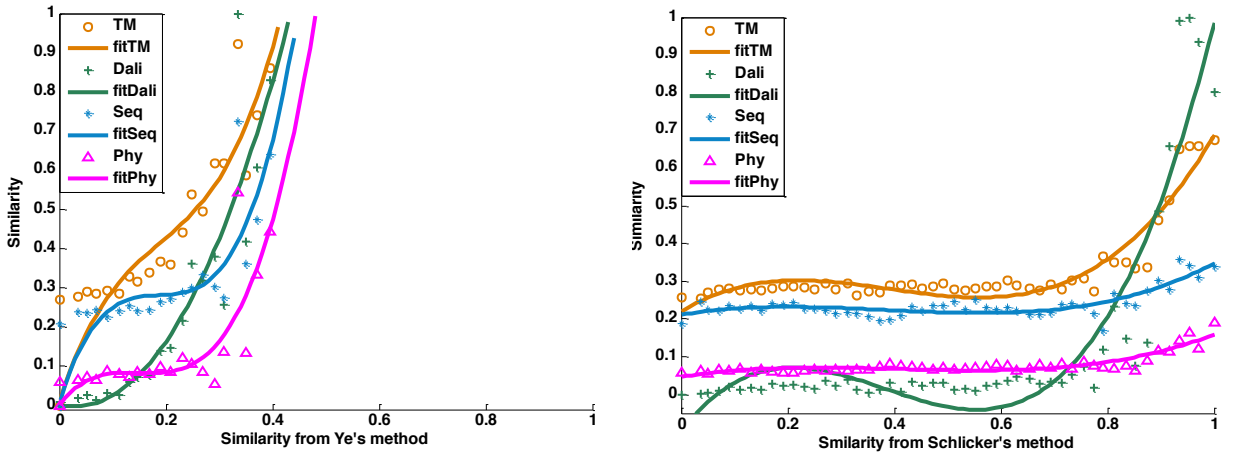



**Figure 4.2** Correlation measurements based on BSM score.

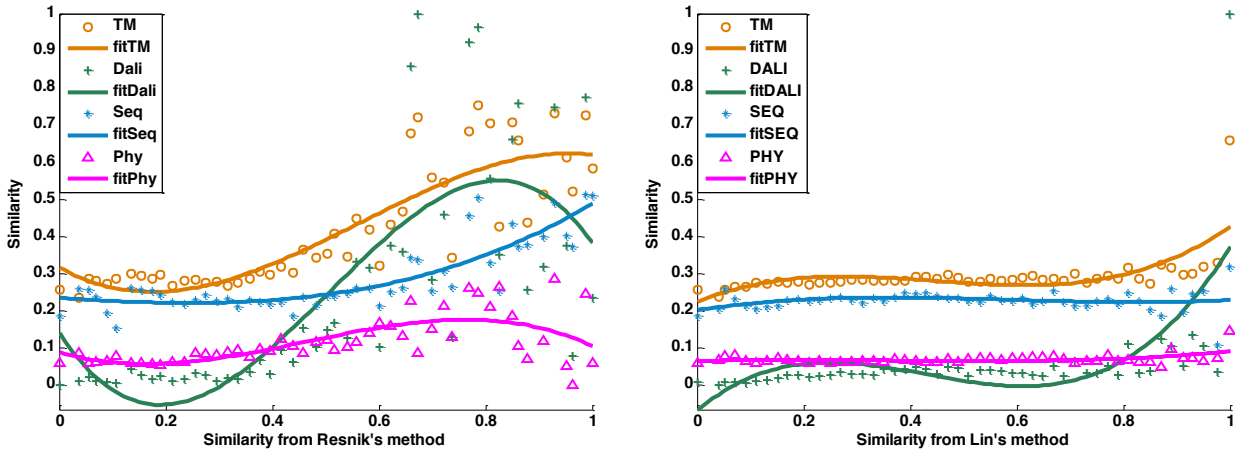
Comparing tendency of similarity score based on structural, sequence, and phylogenetic similarity based on BSM score. To get the structural similarity two well-known alignments are used TM-align and Dali which are denoted as TM and Dali respectively. For the sequence similarity, pair-wise alignments are conducted and denoted as Seq. Phylogenetic profile similarity is denoted as Phy. To help reading the graph, fitted graph is also drawn and identified by adding prefix 'fit' in front of each method. Horizontal axis denotes functional similarity calculated by corresponding methods, and vertical axis denotes similarity score derived from three categories (i.e. structure, sequence, and phylogenetic profile)

Especially, by considering the facts that the number of annotated GPs without having structure information is much higher than the number of GPs with known structures [141], the potential usage of this method and its performance is very significant in the current stage of in-silico method for identifying protein functions.




(a) (b)

**Figure 4.3** Correlation measurements with Ye and Schlicker's method.
Comparing tendency of similarity score based on structural, sequence, and phylogenetic similarity based on (a) Ye and (b) Schlicker's method.

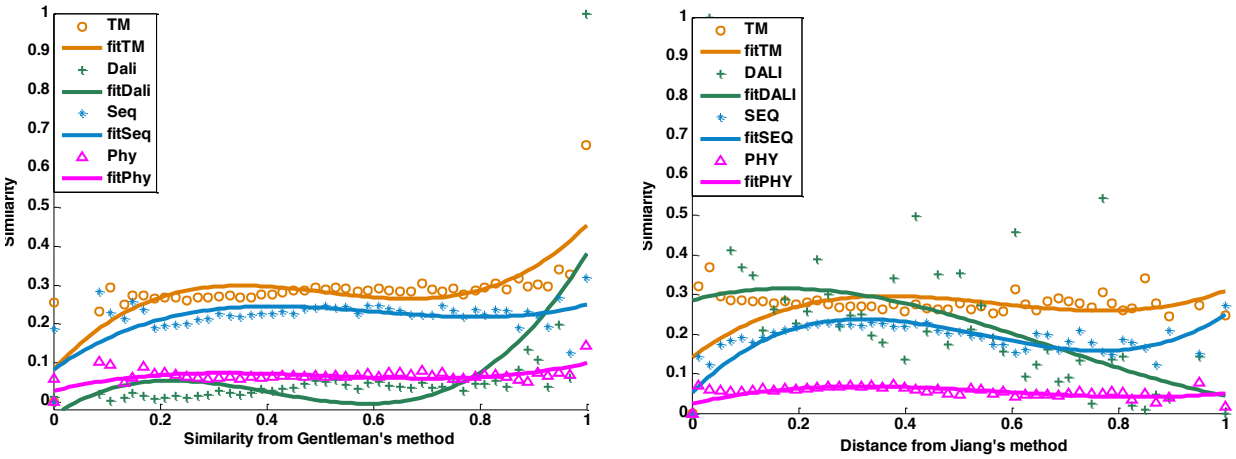



(a) (b)

**Figure 4.4** Correlation measurements with Resnik and Lin's method.
Comparing tendency of similarity score based on structural, sequence, and phylogenetic similarity based on (a) Resnik and (b) Lin's method.

In other words, this result shows the high possibility that the proposed method can devote identifying similar genes or GPs as accurate as structure based methods such as TM-align and Dali do. In fact, based on proposed criteria that the better similarity measurement tools should

bring the high agreement on correlations among three different similarity measurements, the evaluation shows that TM-align and the proposed method clearly outperform Dali.




(a) (b)

**Figure 4.5** Correlation measurements with Gentleman and Jiang's method.
Comparing tendency of similarity score based on structural, sequence, and phylogenetic similarity based on (a) Gentleman and (b) Jiang's method.

Except the proposed method and Ye's method, all other methods are failed to make agreement on the tendency of similarity with others. Interestingly, violations of Dali on similarity tendency are observed across all methods except the proposed method and Ye's method such that in figures 4.3~5, the score range from 0.2 to 0.6, Dali score is anti-proportional to the scores derived from Schlicker, Lin, and Gentleman's methods. Resnik's method in Figure 4.4 (a) show two anti-proportional area: from 0 to 0.2 and from 0.8 to 1.0. Lin, Gentleman and Jiang's method which are Figure 4.4 (b), 4.5 (a), and 4.5 (b) respectively show similar tendency. Interestingly, their patterns are more likely to be the patterns derived from sequence similarity shown in Figure 4.1-(d) in which the similarity is mostly flat along with criteria similarities.

## 4.2 Experiment II: measuring clustering performances

### 4.2.1 Dataset

It is well known that proteins in SCOP classification at superfamily level tend to have common evolutionary origins and functional features. Indeed, the family level which is more specific level than superfamily has even more significant relations in terms of protein functions [62, 96]. Therefore, the given assumption is that proteins in same family category should be out expressed in similarity networks based on the score of functional similarity calculated by proposed method and other conventional methods since proteins in same SCOP family level has more functionally similar than protein in other family level. In similarity network each node corresponds to SCOP domain and edges correspond to the score of functional similarity. To emphasize the expression level or distinction between different family levels, threshold value is introduced such that the threshold value is considered as functional similarity score that can isolate one family group from others. In other words, threshold value is defined by changing the similarity score until the network reaches maximum number of clusters or highest precision depending on evaluation criteria. The maximum number of clusters is defined by counting total number of clusters in the network with a threshold value that produces the highest number of isolated clusters. The precision is defined as below.

$$precision = \frac{\#\ of\ misclustered\ edges}{total\ \#\ of\ edges\ in\ a\ network} \qquad \textbf{(Equation 4-2)}$$

Mis-clustered edges are defined as minority edges in a cluster. In other words, the majority edges belonging to a certain SCOP family is considered as a member of the true cluster.

### 4.2.2 Results

Upon the results of correlation comparisons, the BSM score and Ye's method do not show clear difference. In fact, both to them work very well, especially compared to the method requiring structure information (i.e. TM-align). To analyze the difference between the proposed method and others, the second evaluations are designed and conducted. The hypothesis of this evaluation is that if groups of functionally similar domains are known, then good similarity analyzer could group them together with certain range of similarity. This type of method is well known in machine learning area as clustering methods or unsupervised learning, but the difference between conventional clustering methods and this experiment is that in the given problem, the clusters are already known, so by comparing the results of clusters made by different methods and known clusters, performances and their differences among different algorithms can be compared. To compare the clustering performances of conventional methods and proposed method, all 2646 domains in SCOP databases are used as testing data such that the functional similarity of all possible combinations of SCOP domains (i.e. 3,501,517 domain pairs) are calculated by each similarity method including the proposed method. Once all similarity score is calculated, the threshold of similarity score for each method is adjusted until it reaches the best performance to which either it has maximum number of clusters or highest average precision. Figure 4.6 shows the maximum precision of each methods. In this result, both proposed method (NEW) and Ye's method reached maximum precision 1.0. However, by considering the similarity score 0.75 (75%) and 1.0 (100%), the proposed method has reached maximum precision in earlier stage.

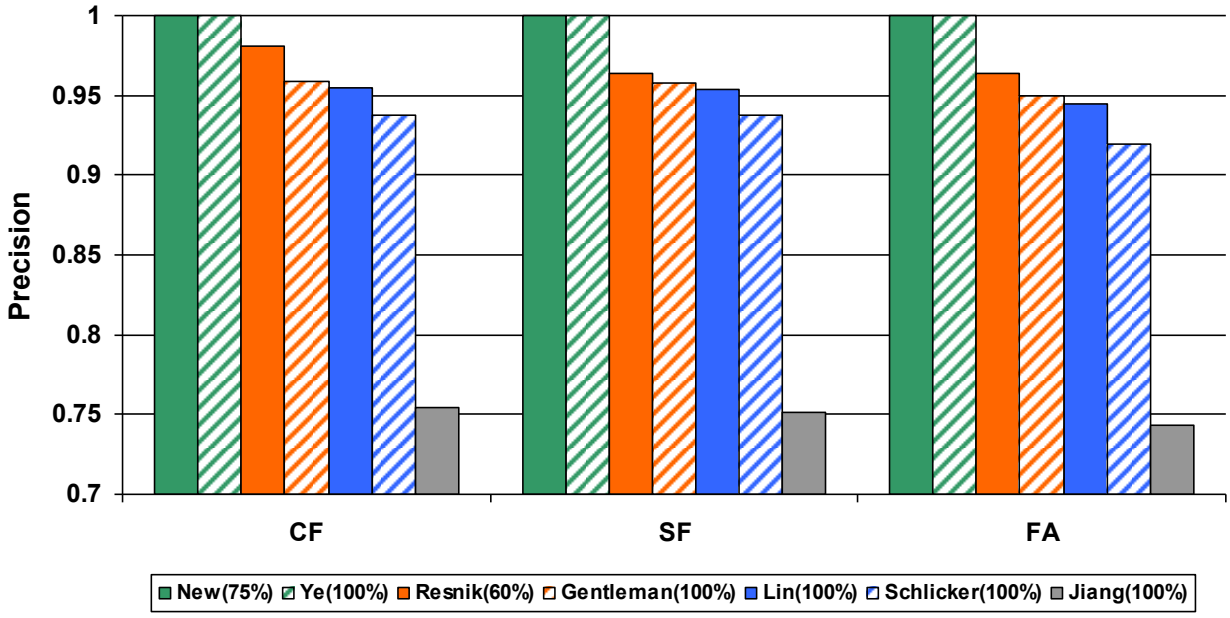



**Figure 4.6** Comparing precisions among clusters.

The threshold to functional similarity score is changed until individual method reaches the best clustering performance at each criterion. Maximum precision of each method according to three different levels of domain family in SCOP classification: Class family (CF), Superfamily (SF) and Family (FA) in which the order corresponds to the broad classification to the most specific classification in SCOP definition. The values in parenthesis denote threshold rate.

Figure 4.7 shows the number of mixed domains. As expected, the proposed method and Ye's method do not have any mixed domains since both of them were reached 100% precision as shown in Figure 4.6. By considering the number of clusters between the proposed method and Ye's method as shown in Figure 4.8, the difference between the proposed method and Ye's methods are very clear such that the proposed method outperforms Ye's method more than 20 folds with this criterion. Although Gentleman's method and Lin's method have better or similar performance upon the consideration of number of clusters, by considering the precision and number of mixed domains shown in Figure 4.6 and 4.7, their errors are far exceeded those of

BSM score. Therefore, by considering all three criteria, BSM score clearly outperforms all other conventional methods.

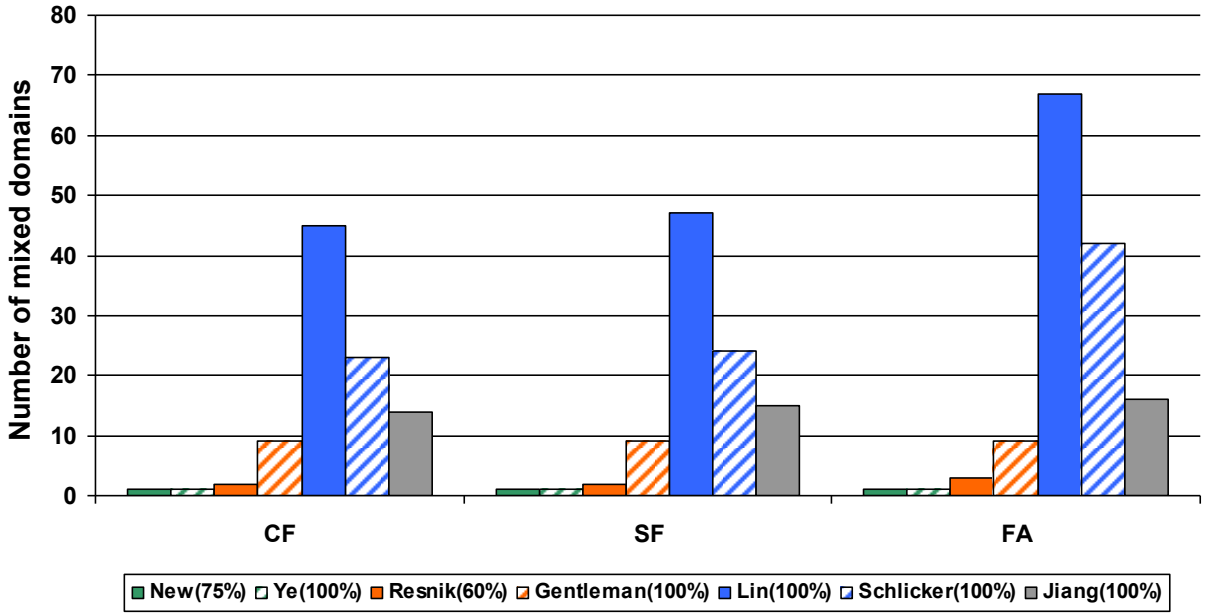



**Figure 4.7** Comparing false clusters.

The threshold to functional similarity score is changed until individual method reaches the best clustering performance at each criterion. The average number of mixed domain families along with three different levels in SCOP classification: Class family (CF), Superfamily (SF) and Family (FA). The values in parenthesis denotes threshold rate.

Figure 4.8 shows the difference of the average number of clusters made by certain threshold among the methods such that with 70% of similarity rate, the proposed method clearly outperforms others, and with average above 70% of similarity rate, the proposed method, Gentleman, Lin, and Schlicker's method are comparative. This inconsistent result shows very interesting and critical point of view to define the balance between human cognition and the significance of score reflecting GP relations. Problems originated from normalized and

non-normalized scores are already discussed [64, 91], and although they share the same concept of human cognition at the level of interpreting the score, they show clear difference by considering the gaps between the score significance at algorithm level and user's expectation. In other words, what does the meaning of 70% similarity in each method? Are they can be considered as significant in terms of their functional relations? If the human cognition and the score significance were involved, the evaluation criteria would be more complicated. Unfortunately no study has been conducted on this issue, so as best of my knowledge believe this is the first time to claim the significance of human cognition at reported scores as important evaluation criteria.

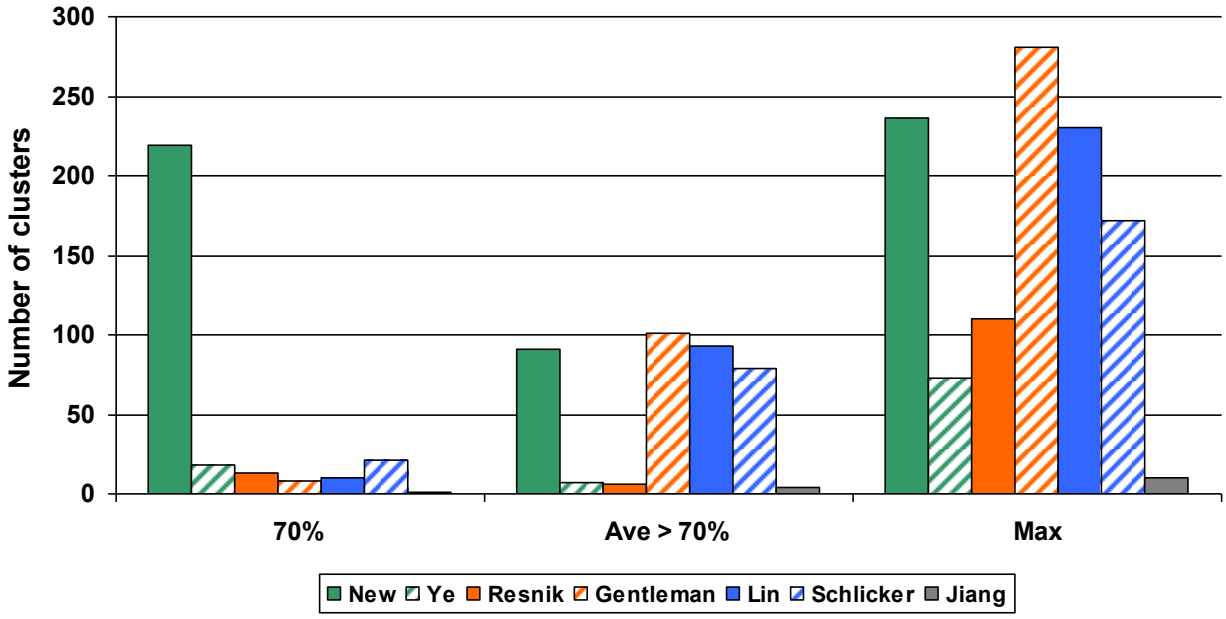
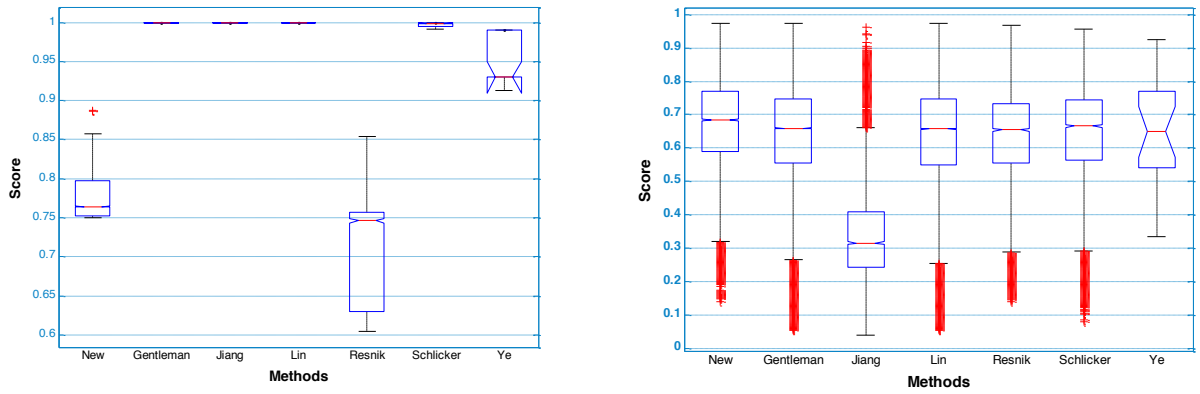



**Figure 4.8** Comparing number of clusters.
The threshold to functional similarity score is changed until individual method reaches the best clustering performance at each criterion. Number of clusters identified by the given method with specific threshold range such that the first group of graphs is drawn with 70% of threshold, next group is drawn by averaging the identified number of clusters greater than 70% and the final group is shown to compare the maximum number of identified clusters.

Let's come back to the Figure 4.8, as shown in this figure, the proposed method, Ye and Resinik's methods shows smaller number of clusters by increasing the number similarity rate. In fact, the proposed method reached the maximum number of clusters on 65% similarity rate. Compared to these methods, Gentleman, Lin, Schlicker, and Jiang's method increase the number of cluster while increasing the similarity rate. Indeed, they reached the maximum number of clusters with 100% similarity. Two problems can be found in this result: i) they define the GPs in each cluster with exactly same similarity rate (i.e, 100% similarity) and ii) some of clusters consist of more than two domain families which show that even if GPs in a cluster have 100% similarities, this does not mean they function exact same although user's expectations are often considering 100% similarities are exact same or extremely similar. Indeed, the second group (i.e, Gentleman, Lin, Schlicker, and Jiang's method) would fail to further distinct GPs in a same cluster in such case they cannot narrow down potential neighboring gene or GPs to identify drug targets, so their contributions on this purpose could significantly reduced or limited. The more significant problem is that the question of how the score should be interpreted? Although 100% similarity rate is somewhat reliable as shown in the result, by considering the similarity rate 100% in general or as definition of their maximum similarity, the score does not match the results or user's expectation, which should be exact same genes or GPs. More directly, what does meaning of less than 100% similarity, and what is the significance of these scores that contribute identifying genes or GP similarity as a human cognition point of view? Although the question is very important since this can directly affects interpreting the results and can be propagated to making final decision of the essay or application, the answer is not easy due to the lack of guidelines in this issue. In order to answer this question, this issue is applied to one of criteria in performance evaluation. As shown in the results of precision, the results of proposed method

could be reliable if it reaches higher than 70% similarity. Indeed, the 100% similarity score will not be produced unless two GPs are consisted with exactly same GO terms. Therefore, as a human cognition point of view, the proposed method is only satisfied since Ye and Resnik's method have relatively high chance to produce maximum similarity score for GPs having different GO terms with same depth or the same number of annotated GPs.



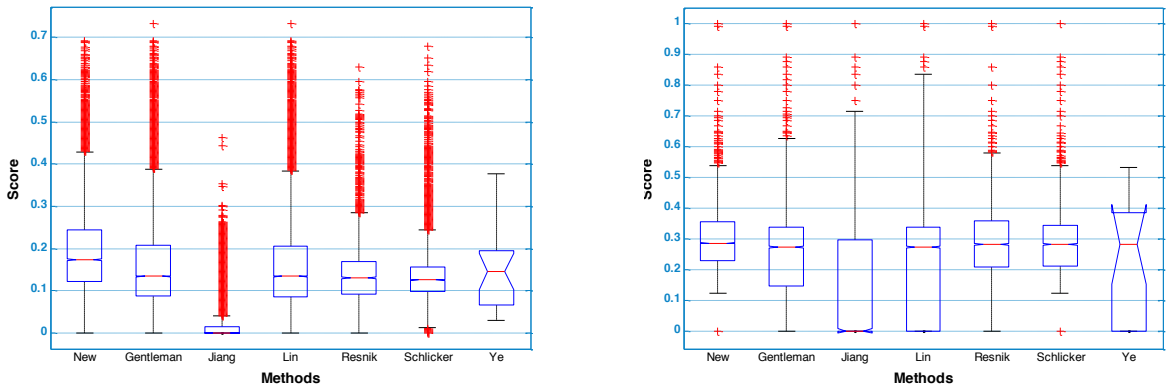

(a) (b)

**Figure 4.9** Analysis of statistical distributions.

Analyzing statistical distributions of similarity score based on three categories (i.e. structure, sequence, phylogenetic profile). The distributions are calculated from each clusters made by SCOP domains with maximum precision. Corresponding similarity rates are NEW 75%, Ye 100%, Resnik 60%, Gentleman 100%, Lin 100%, Schlicker 100% and Jiang 100%. (a) the distribution of GP similarity scores at the highest precision (b) the distribution of structural similarity calculated by TM-align at highest precision.

To analyze the clustering tendency, the distributions of five different base similarity measurements inside clusters are considered. Figure 4.9 (a) shows the distributions of GP similarity at the highest precision, and the proposed method and Resnik's method have lower similarity score than other methods. Interestingly the range of GP similarity at the proposed method is relatively larger than Gentleman, Jiang, Lin, and Schlicker's method and smaller than Resnik and Ye's method. It is important to have lower score with reaching the highest precision

since this gives more opportunity to further investigate the relations of inside clusters. In other words, BSM score is capable of identifying more specific relations within a cluster while maintaining stabilizing the range of human cognition. Therefore, based on two criteria, possibility of further investigation and the significance of similarity score at human cognition, BSM score clearly outperforms other methods.




(a) (b)

**Figure 4.10** Analysis of statistical distributions.

Analyzing statistical distributions of similarity score based on three categories (i.e. structure, sequence, phylogenetic profile). The distributions are calculated from each clusters made by SCOP domains with maximum precision. Corresponding similarity rates are NEW 75%, Ye 100%, Resnik 60%, Gentleman 100%, Lin 100%, Schlicker 100% and Jiang 100%. (a) the distribution of structural similarity calculated by DALI at highest precision. (b) the distribution of sequence similarity at highest precision.

Figure 4.9 (b) and 4.10 (a) show the distribution of structural similarity at the highest precision calculated by TM-align and Dali respectively. With TM-align, although all methods tend to have high structural similarity, the proposed method and Ye's method slightly have higher structural similarity than others. The results from Dali make better discrimination the proposed method and others although their similarity score is lower than TM-align. The lower structural similarity is also observed in Figure 4.1 (b) and 4.4 (a). More importantly, the difference between proposed

method and Ye's method is much clear. The sequence similarity in Figure 4.10 (b) shows that the proposed method, Resnik and Schlicker's method have relatively higher and denser sequence similarity than others.

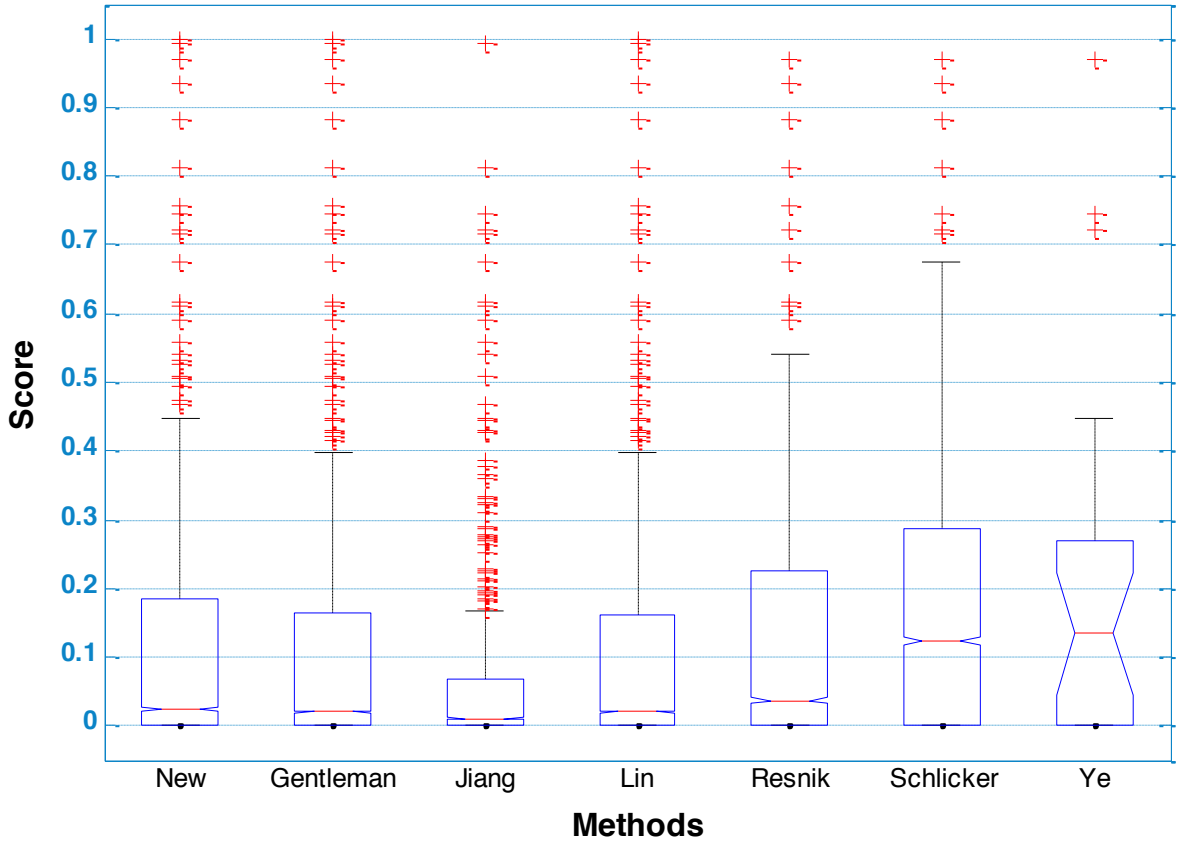


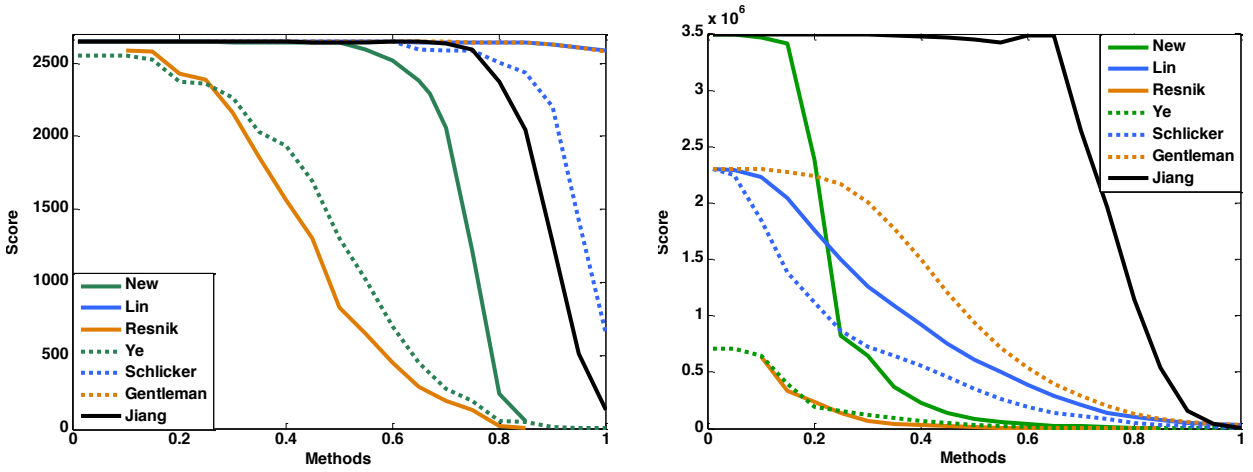

**Figure 4.11** Analysis of statistical distributions.

Analyzing statistical distributions of similarity score based on three categories (i.e. structure, sequence, phylogenetic profile). The distributions are calculated from each clusters made by SCOP domains with maximum precision. Corresponding similarity rates are NEW 75%, Ye 100%, Resnik 60%, Gentleman 100%, Lin 100%, Schlicker 100% and Jiang 100%. the distribution of phylogenetic profile similarity at the highest precision.

Although the results of phylogenetic profile similarity shown in Figure 4.11 shows that Schlicker and Ye's method have slightly higher profile similarity score than others, by considering outliers and the upper and lower quartile, the difference of phylogenetic profile similarity distributions is

not significant. In short, the proposed method outperforms other methods at several criteria: the ability of discriminating specific relations with high precision, providing stabilized significance of similarity score at human cognition level, and relatively higher structural and sequence similarity.

To get the broad spectrum of clustering tendency, the changes of clusters are further analyzed by continuously increasing the threshold of similarity score of each method.

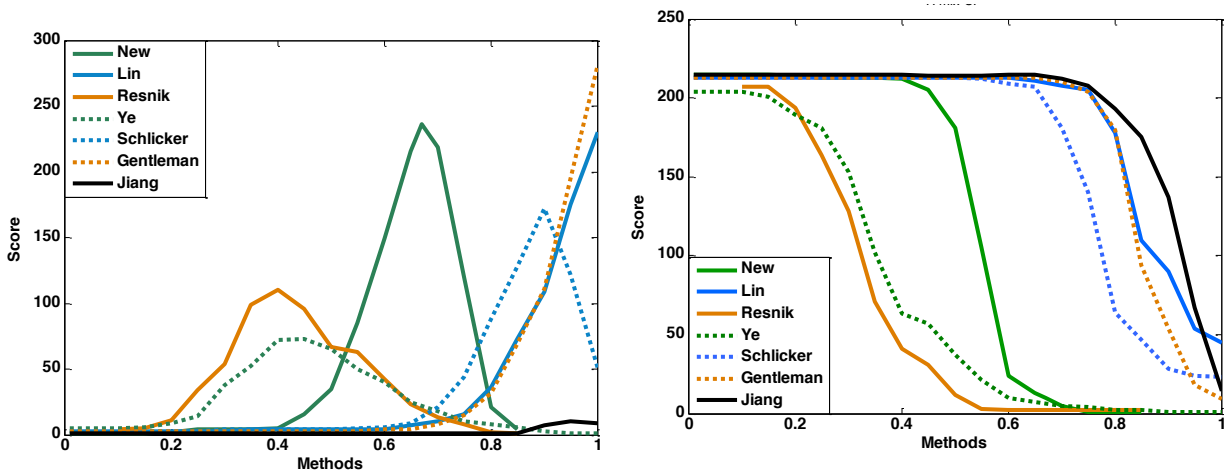



(a) Nodes (b) Edges

**Figure 4.12** Tendency of clusters upon nodes and edges.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) *Nodes* are the individual domains in clusters at the given threshold. (b) *Edges* are the relations connecting two nodes in clusters at the given threshold. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

The graph *Nodes* in Figure 4.12 (a) shows that all methods dramatically decrease the number of nodes at a certain threshold except Lin and gentleman's method which do not show significant changes in the number of nodes. For example, the proposed method has significant decrease of the number of nodes at the threshold value around 0.7, but Lin's and Gentleman's method stays

at same number as the proposed method has at similarity score around 0.7. In contrast to node changes, the patterns of *edge* changes in Lin and Gentleman's method are similar to others.



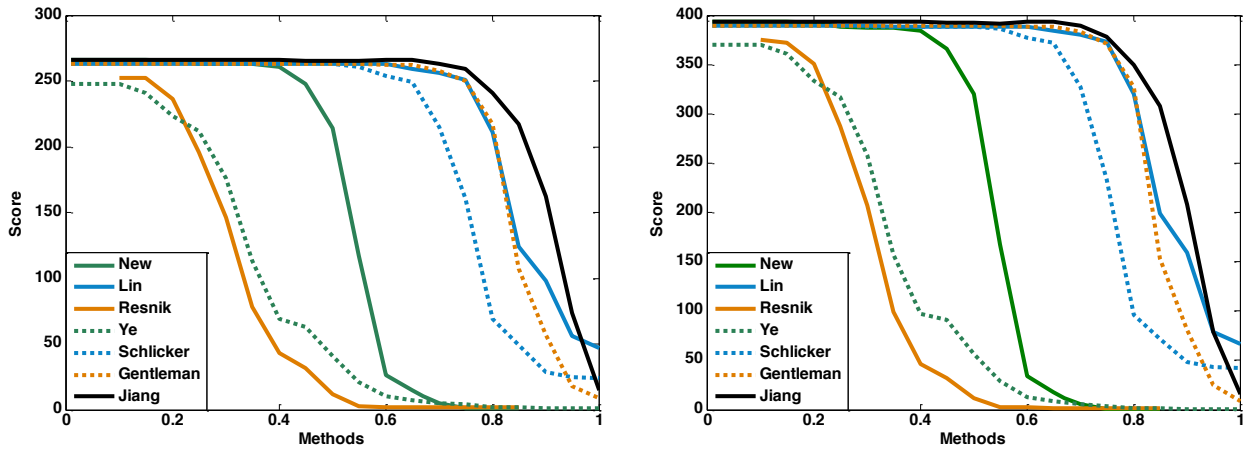

(a) Clusters

(b) HmixCF

**Figure 4.13** Tendency of clusters upon clusters and highest false class family clusters.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) Clusters are the total number of isolated groups at the given threshold. (b) HmixCF denotes the highest number of misclassified edges among all clusters at the level of 'class-family' in SCOP classification. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

The patterns of *cluster* changes show Lin and Gentleman's reaches the highest number of clusters with the proposed method, but their patterns are different from others such that they only have the ascent of graph and descent does not exist. Indeed, overall graphs from *Nodes*, *Edges*, and *Clusters* show that clusters made by Lin and Gentleman's method have relatively sparse relations than others. These characteristics of Lin and Gentleman's method imply that their methods are not favorable for identifying specific relations at higher similarity score. Indeed, this can be problematic when identifying the threshold of trust level is required.

Figure 4.13 (b), 4.14, 4.15, and 4.16 (a) are measuring the highest number of edges in mixed clusters and precision at three different classification levels: class-family, super-family, and family which are Figure 4.13 (b) HmixCF, 4.14 (a) HmixSF, 4.14 (b) HmixFA, 4.15 (a) precCF, 4.15 (b) precSF, and 4.16 (a) precFA respectively.

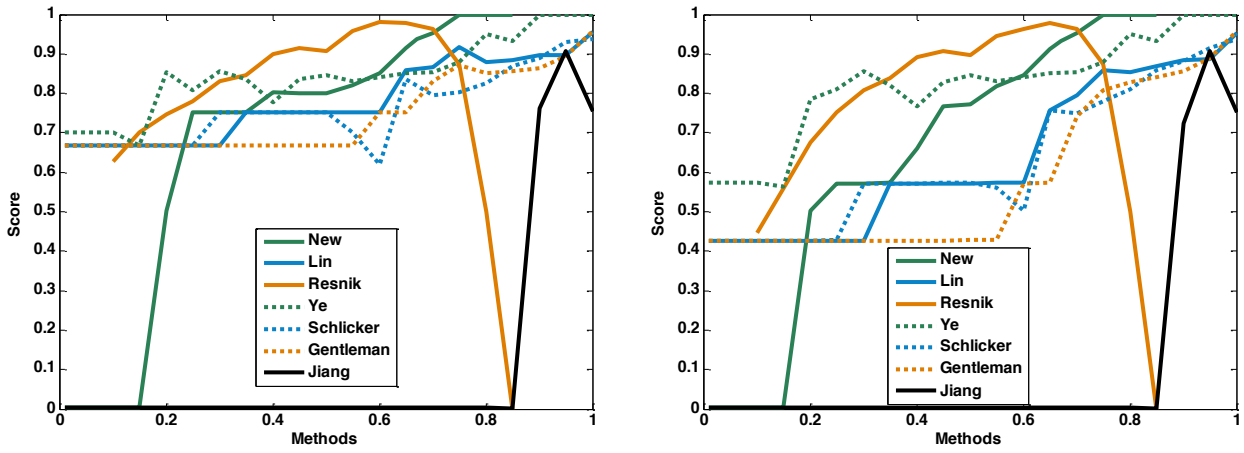



(a) HmixSF (b) HmixFA

**Figure 4.14** Tendency of clusters upon highest false super family clusters and family clusters.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) *HmixSF* denotes the highest number of misclassified edges among all clusters at the level of 'super-family' in SCOP classification. (b) *HmixFA* denotes the highest number of misclassified edges among all clusters at the level of 'family' in SCOP classification. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

Ideal similarity analyzer should have high correlation between similarity score and corresponding precisions, and inverse correlation between similarity score and the number of misclassified GP pairs. More precisely, if the similarity analyzer is practically useful then it should give clear relations between significance of scores and functional similarity such that if

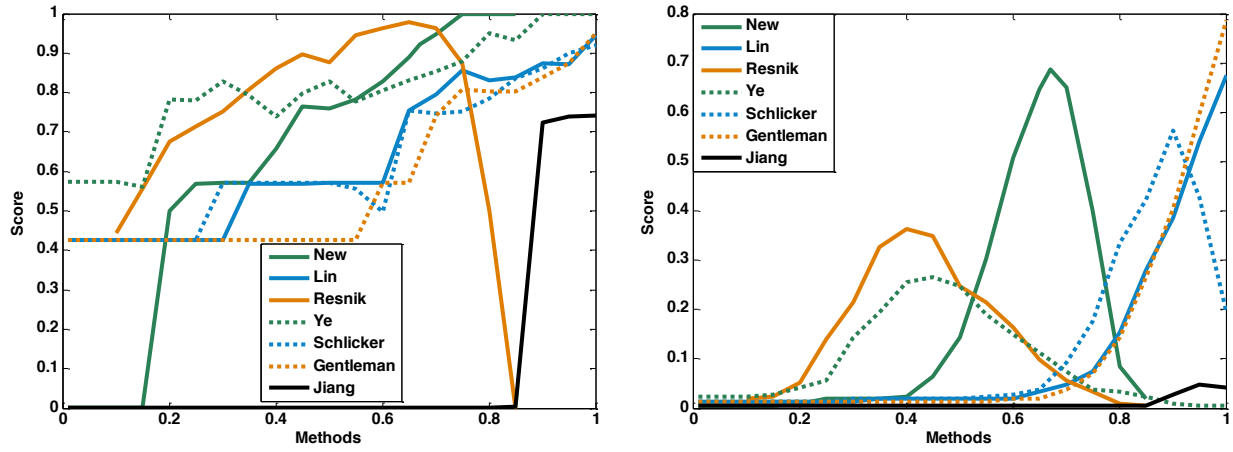
the score is getting higher then the precision should be same or getting higher by eliminating unrelated gene or GPs, and the number of misclassified GP pairs should be decreased.




(a) precCF (b) precSF

**Figure 4.15** Tendency of precision upon class family and super family.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) *precCF* denotes an averaged precision from all clusters at the level of 'class-family' in SCOP classification. (b) *precSF* denotes an averaged precision from all clusters at the level of super-family in SCOP classification. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

Figures based on precision which are Figure 4.15 and 4.16 (a) show that all methods decrease the number of misclassified GPs while increasing the similarity score, but Lin, Gentleman, Schlicker, and Jiang's methods do not reached to get complete clusters of which all GP pairs are correctly clustered. Interestingly the patterns of mixed domain families are almost conserved along with different levels of domain families although the number of misclassified GP pairs is increased as narrowing down to domain families. This implies that each method has its own confidence level for similarity score, and the confidence level must be conserved for all different levels of domain families while preserving smallest number of misclassified GP pairs.

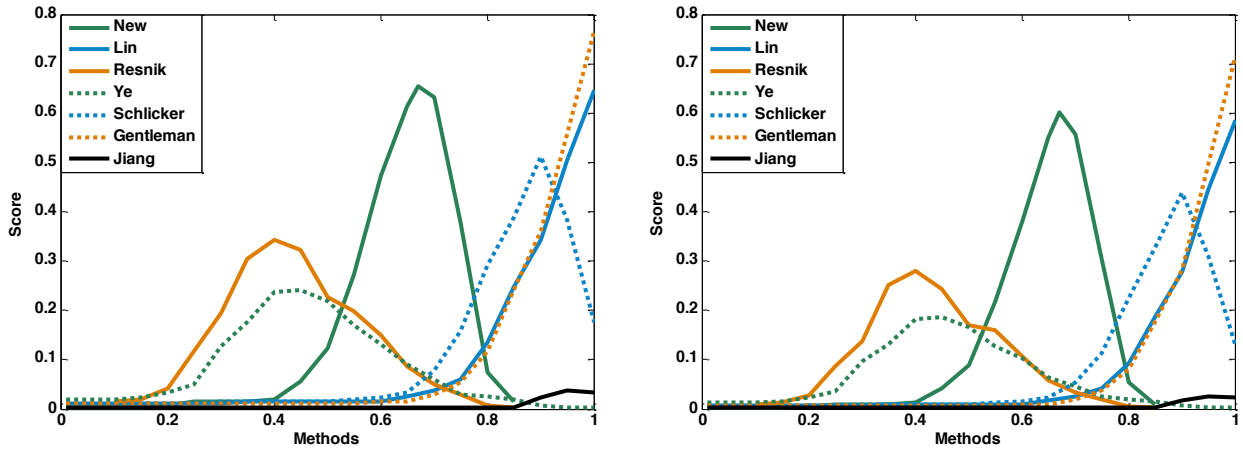



(a) precFA (b) covCF

**Figure 4.16** Tendency of precision upon family and of coverage upon class family.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) *precFA* denotes an averaged precision from all clusters at the level of 'family' in SCOP classification. (b) *covCF* denotes an averaged coverage from all clusters at the level of class-family in SCOP classification. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

To make a better comparison and get clearer picture of confidence range, let's look closer into the precision graphs. These precision graphs show that the proposed method only has strong correlations along with precision score. For example, Lin and Gentleman's method decrease their precision by increasing the similarity score from 0.7 to 0.8. Resnik's method shows dramatic changes of precision after the threshold similarity 0.7; this happens since the clusters made with higher similarity scores are composed with small number of GP pairs and almost all of them are mixed with more than two domain families. Rest of methods, Ye, Schlicker, and Jiang's method, are even more fluctuated. These results show that the proposed method is well characterized with

its confidence rate and similarity score by preserving consistent correlation between precision and similarity score.




(a) covSF (b) covFA

**Figure 4.17** Tendency of coverage upon super family and family.
Analyzing the tendency of clusters while changing the threshold of similarity score defined by individual methods. Horizontal axis denotes the threshold of similarity score, and vertical axis denotes score of each category. (a) *covSF* denotes an averaged coverage of clusters from all clusters at the level of 'super-family' in SCOP classification. (b) *covFA* denotes an averaged coverage from all clusters at the level of 'family' in SCOP classification. In SCOP classification, class-family, super-family, and family is ordered list from broader to the more specific classification level.

Figure 4.16 (b) and 4.17 show mean coverage of clusters defined by each method. Every method shows same tendency that broader definition of domain family has the higher coverage. BSM score has obtained one of the highest coverage along with Lin and Gentleman's method.

So far, the results show that the proposed method outperforms conventional methods, and some of methods are compatible to the proposed method. However, the details of differences are still unclear. To further explore on this issue, the direction of investigation is switched to approaching opposite way, such that instead of analyzing successfully classified clusters, mixed clusters are analyzed to understand the behavior of similarity methods. In order to get mixed

clusters, the threshold of the BSM score is set to 0.7 which is corresponding to the second highest precision of the method (i.e, the score 0.75 reached 100% precision). Ye's method has too small number of GPs at 95% similarity score which corresponds to the second highest precision of this method.

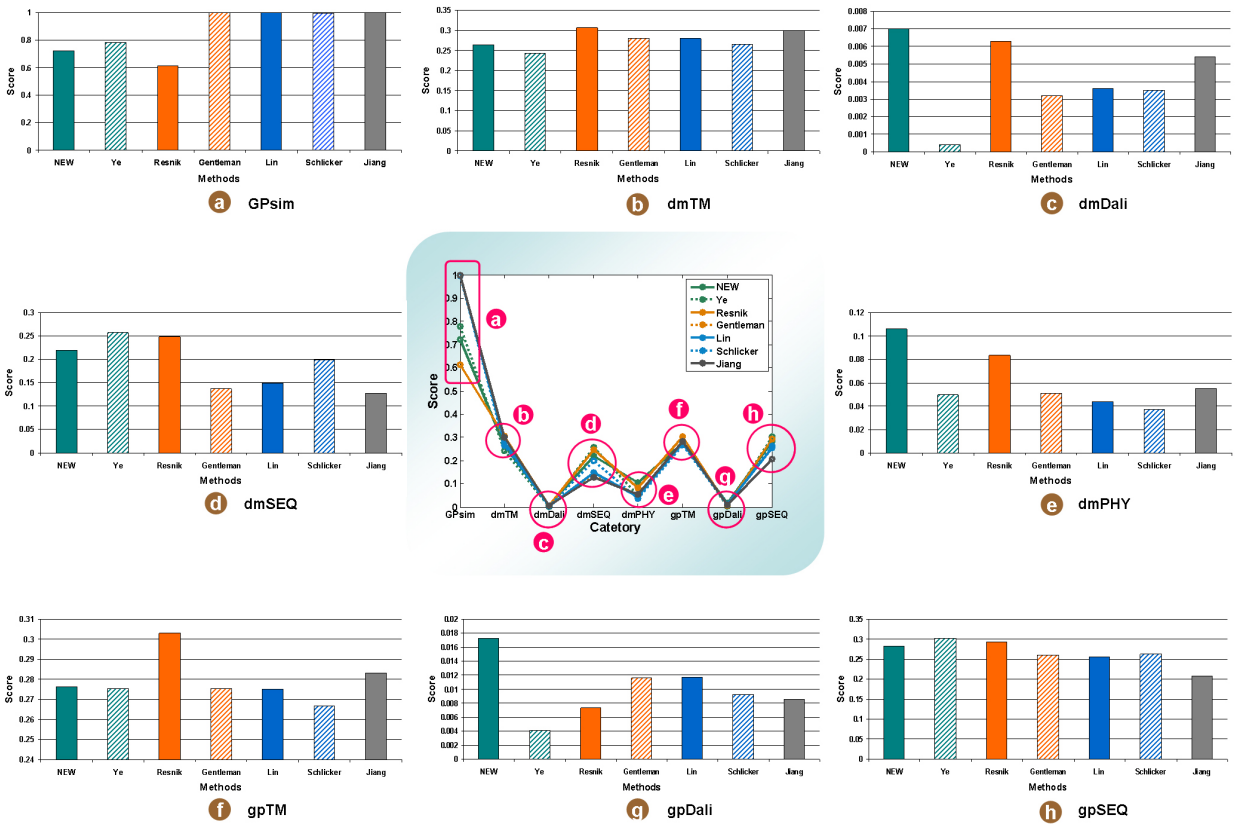



**Figure 4.18** Distributions of score in mixed clusters.

The graphs are drawn by analyzing clusters consisting with mixed domain families. Horizontal and vertical axis of each graph show the name of methods involved in this experiments and corresponding scores respectively. The graph in the center highlighted with green background shows the overall patterns of measurement. Each dot in this Figure is enlarged corresponding to their alphabetic index. (a) GPsim: Averaged similarity score calculated by each method (b) dmTM: averaged domain similarity calculated by TM-align (c) dmDali: averaged domain similarity calculated by Dali (d) dmSEQ: averaged sequence similarity of SCOP domain (e) dmPHY: averaged phylogenetic profile similarity of SCOP domain (f) gpTM: averaged GP similarity calculated by TM-align (g) gpDali: averaged GP similarity calculated by Dali (h) gpSEQ: averaged sequence similarity of GPs.

Therefore, to get statistically meaningful number of GPs, the threshold is reduced to 75% similarity. All other methods are obtained with their highest precision since they did not reach the 100% precision. As discussed previously, Figure 4.18 (a) shows that Gentleman, Lin, schlicker and Jiang are not suitable for identifying closely related domains from others since the similarity score of mixed clusters already reaches its maximum score. In Figure 4.18 (b) and (c), although all methods show that misclassified GP pairs have relatively lower structure similarity than majority ones which are shown in Figure 4.9 (b) and 4.10 (a), except Jiang's method which relatively conserved than others. BSM score and Resink's methods are relatively higher similarity on both TM and DALI. Figure 4.18 (d) shows that New, Ye and Resnik's methods tend to conserver the sequence similarity similar with majority ones sown in Figure 4.10 (b). Figure 4.18 (e) shows that the phylogenetic profile similarity in mixed clusters have maintained slightly higher score than average of majority ones, especially for BSM score and Resinik's method. Overall scores of mixed clusters show that domains clustered by proposed method and Resnik's method are closely related by the means of three similarity criteria. To understand relations between domain and GPs, GPs containing each domain are retrieved from SCOP, and corresponding structure and sequence information are downloaded from PDB. The structure similarity of these GPs calculated by TM-align shows that Resnik and Jiang's method obtain higher score than others while the proposed method remains as similar as others. However, Dali shows different results such that BSM score has relatively higher score than others. The sequence similarity shown in Figure 4.18 (h) show that New, Ye, and Resnik's method get relatively higher score. Overall results from examining mixed clusters, BSM score and Resnik's method constantly show that the mixed class has strong relations of the remaining clusters.

However, by considering all comparisons including correlations of similarity score against to structure, sequence, and phylogenetic profile scores, precision and coverage of clusters, and statistics on pure and mixed clusters, the proposed method constantly outperforms all other methods although some methods show compatible performance at certain experiment. More importantly, analyzing the results obtained from mixed clusters shows that misclassified domains in a mixed cluster has somewhat strong relations to other dominant domains since their similarities cannot be distinct from majority domains by structure, sequence and phylogenetic profile similarity. Therefore, it is of great interest to figure out the biological relations between minor and majority domains in a cluster defined by the proposed method.

***Analyzing clusters with mixed SCOP families***: As continuous analysis on mixed clusters defined by the proposed methods, the relations between heterogeneous SCOP families defined by proposed method are investigated. To get the mixed clusters, similarity score 0.7 is used as similarity threshold. Figure 4.19 shows a cluster consisting with three different families: *Ferritin* (47241), *2Fe-2S ferredoxin-related* (54293), and *purple acid phosphatase-like* (56301) in which the numbers in parenthesis correspond to *sunid* in SCOP classification. *Ferritin* is one of the major non-heme iron storage proteins in animals, plants, and microorganisms [142, 143]. It consists of a mineral core of hydrated ferric oxide, and a multi-subunit protein shell which encloses the former and assures its solubility in an aqueous environment. Purple acid phosphatases (PAPs) are ubiquitous binuclear metal-containing acid hydrolases characterized by their acidic pH optima and their intense purple color due to a TyrO-to-FeIII charge-transfer transition [143].

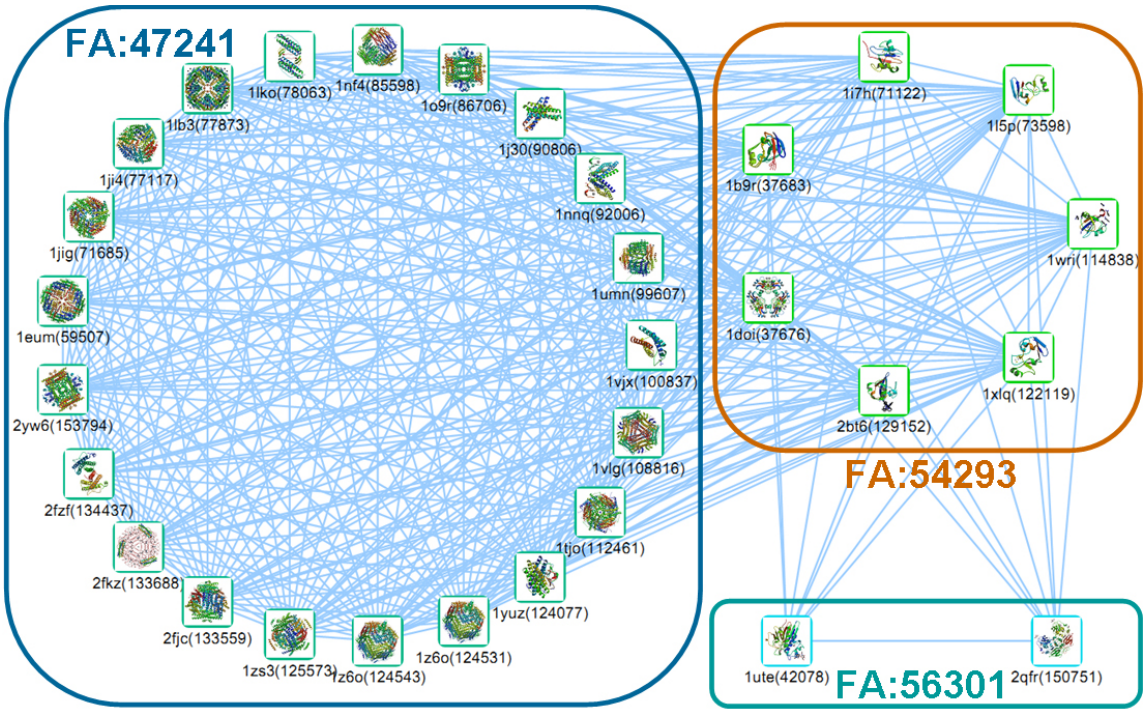



**Figure 4.19** Analysis of mixed families of FA:47241, FA:54293, and FA:56301.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. The numbers in parenthesis correspond to sunid in SCOP classification. 1UTE and 2QFR are associated with FA:56301 which is Purple acid phosphatase-like. 1B9R, 1I7H, 1L5P, 1WRI, 1XLQ, 2BT6, and 1DOI are associated with FA:54293 which is 2Fe-2S ferredoxin-related. All others are associated with FA:47241 which is Ferritin.

Therefore, it seems 2Fe-2S ferredoxin makes strong relation between Ferritin and Purple acid phosphatases by the roles of electron-transfer carriers and donors since 2Fe-2S ferredoxins function as electron carriers in the photosynthetic electron transport chain and as electron donors to various cellular proteins [144]. In hydroxylating bacterial dioxygenase systems, they serve as intermediate electron-transfer carriers between reductase flavoproteins and oxygenase [145].

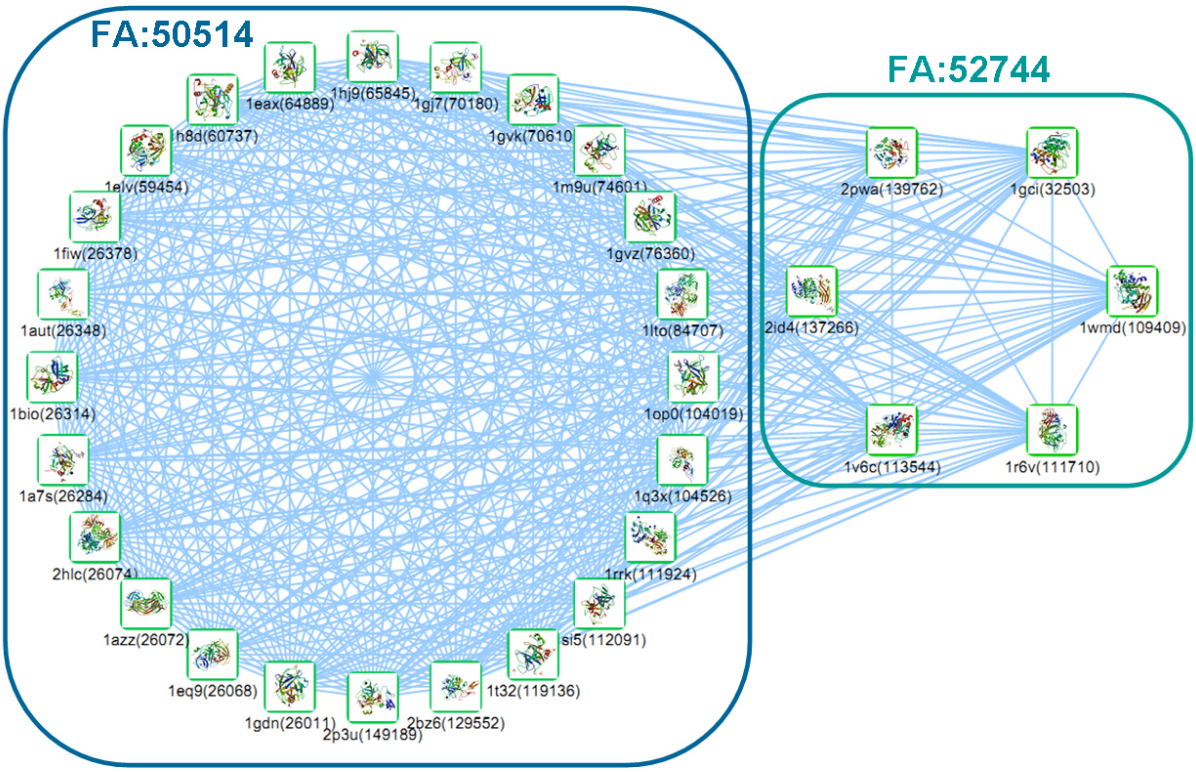



**Figure 4.20** Analysis of mixed families of FA:50514 and FA:52744.

Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. The numbers in parenthesis correspond to sunid in SCOP classification. 2PWA, 1GCI, 1WMD, 1R6V, 1V6C, and 2ID4 are associated with FA:52744 which is Subtilases family and all others are associated with FA:50514 which is Eukaryotic proteases family.

Figure 4.20 shows a cluster consisting with two different families: *Eukaryotic protease* (50514) and *Subtilases* (52744). Two families are clearly related to their functional roles such that *Subtilases* are a family of subtilisin-like serine proteases and *Eukaryotic protease* is also protease found specially in *Eukaryotic*. This means both of them are enzymes related to hydrolysis of the peptide bonds in proteins although serine is serving as the nucleophilic amino acid at the active site [146] and more widespread so that this can be found in eubacteria, archaebacteria, eukaryotes and viruses [147].

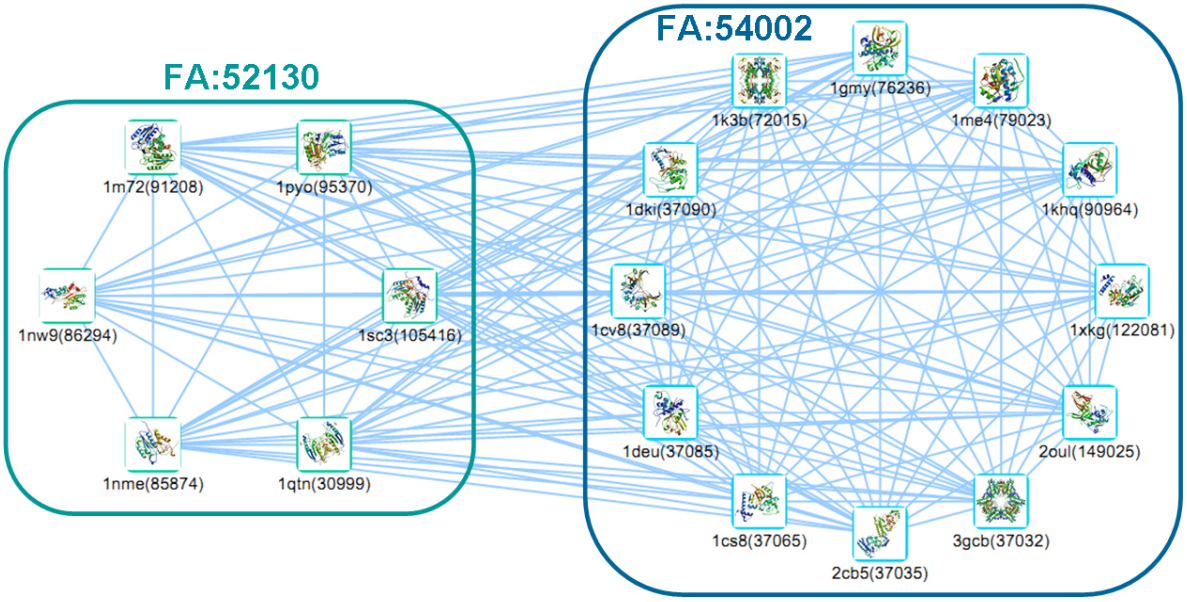



**Figure 4.21** Analysis of mixed families of FA:52130 and FA:54002.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. The numbers in parenthesis correspond to sunid in SCOP classification. 1M72, 1PYO, 1SC3, 1QTN, 1NME, and 1NW9 are associated with FA:52130 which is Caspase catalytic domain family and all others are associated with FA:54002 which is Papain-like family.

Figure 4.21 shows a cluster consisting with two different SCOP families: *Caspase catalytic domain* (52130) and *Papain-like* (54002). These two families are closely related to apoptosis which is known as programmed cell death and immunization such that Caspases are a family of cysteine proteases that play essential roles in apoptosis, necrosis, and inflammation [148], and Papain is involved in deprotonation of Cys-25 by His-159 and known to cleave the Fc which is crystallisable portion of antibodies from the Fab which is antigen-binding portion [149].

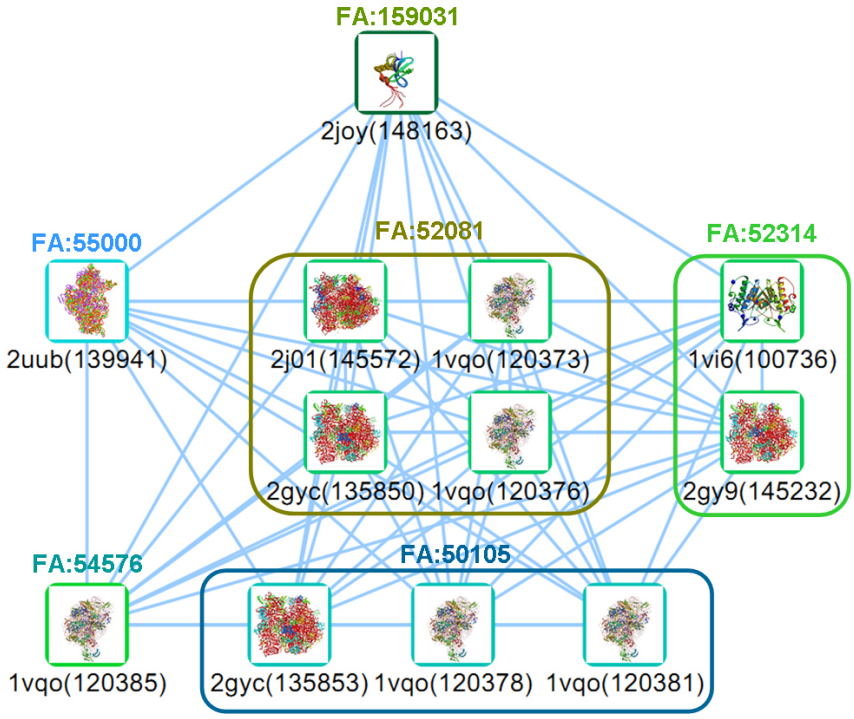



**Figure 4.22** Analysis of mixed families of FA:159031, 55000, 52081, 52314, 54576, and 50105.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. The numbers in parenthesis correspond to sunid in SCOP classification. 2JOY is associated with FA:159031 which is Ribosomal protein L14e. 2UUB is associated with FA:55000 which is Ribosomal protein S10. 1VQO (suid:120385) is associated with FA:54576 which is Ribosomal protein L31e. 1VI6 and 2GY9 are associated with FA52314 which is Ribosomal protein S2. 2GYC and two 1VQOs consisting of sunid 120378 and 120381 are associated with FA:50105 which is Ribosomal proteins L24p and L21e. 2J01, 1VQO (sunid: 120373, 120376), and 2GYC are associated with FA:52081 which is Ribosomal proteins L15p and L18e.

Figure 4.22 shows a cluster consisting with six different families: Ribosomal protein L14e (159031), Ribosomal proteins L24p (50105), Ribosomal protein S10 (55000), Ribosomal protein L31e (54576), Ribosomal protein S2 (52314), and Ribosomal proteins L15p and L18e (52081). As appeared in their names, they are the ribosomal proteins that make up ribosomal subunits involved in cellular process of translation. Indeed, Ribosomal protein L14e (159031) and Ribosomal proteins L24p (50105) are the subgroup of super-family, translation proteins SH3-like domain (50104).

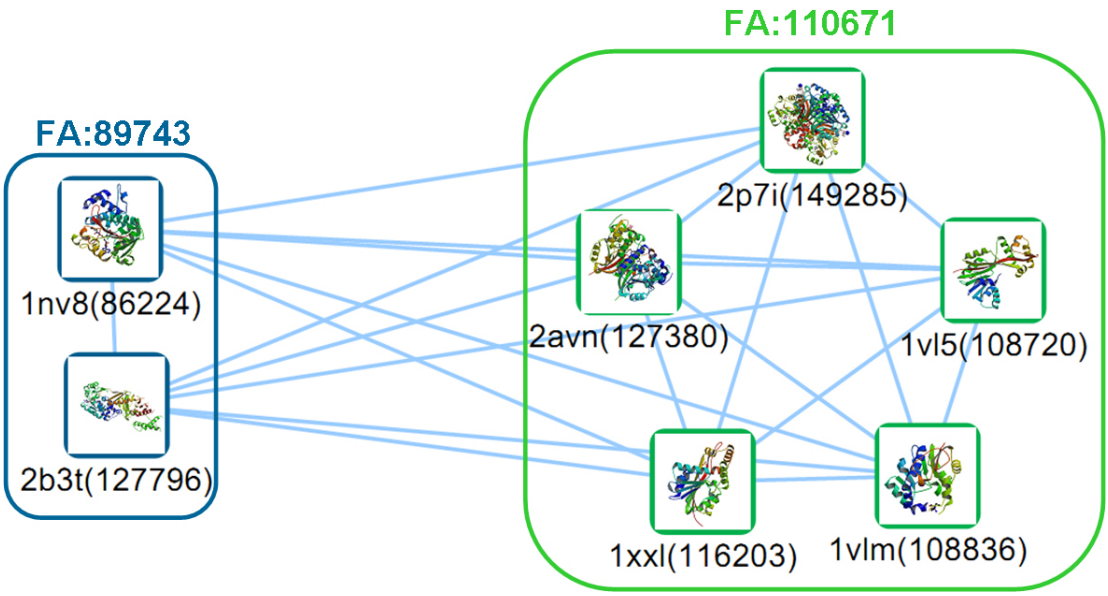



**Figure 4.23** Analysis of mixed families of FA:89743 and FA:110671.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. 1NV8 and 2B3T are associated with FA:89743 which is N5-glutamine methyltransferase, HemK family and all others are associated with FA:110671 which is UbiE/COQ5 methyltransferase family.

Figure 4.23 shows a cluster consisting with two different families: *N5-glutamine methyltransferase, HemK* (89743) and *UbiE/COQ5 methyltransferase* (110671). As shown in their names, both of them are a *methyltransferase* which also known as a methylase and transfers a methyl group from a donor to an acceptor [150]. Indeed, they are the subgroup of superfamily *S-adenosyl-L-methionine-dependent methyltransferases* (53335) which is and found primarily in bacteria.

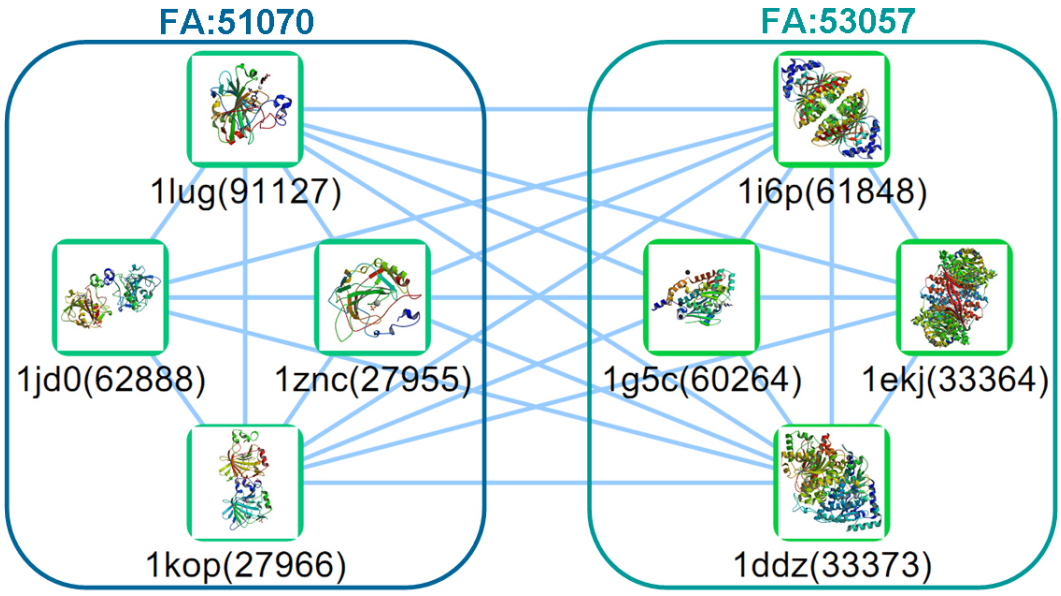



**Figure 4.24** Analysis of mixed families of FA:51070 and FA:53057.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. 1JD0, 1KOP, 1LUG, and 1ZNC are associated with FA:51070 which is Carbonic anhydrase family and all others are associated with FA:53057 which is beta-carbonic anhydrase, cab family.

Figure 4.24 shows a cluster consisting with two different families: *Carbonic anhydrase* (51070) and *beta-carbonic anhydrase, cab* (53057). As shown in their names, both of them are the *carbonic anhydrases* which form a family of enzymes that catalyze the rapid interconversion of carbon dioxide and water to protons and bicarbonate which is also known as hydrogen carbonate, an intermediate form in the deprotonation of carbonic acid [151].

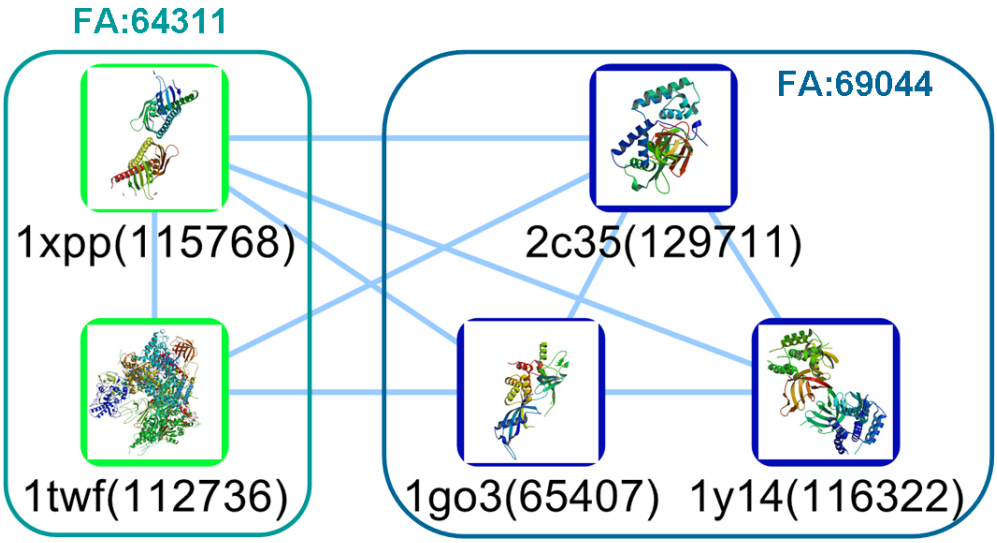



**Figure 4.25** Analysis of mixed families of FA:64311 and FA:69044.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. 1XPP and 1TWF are associated with FA:64311 which is RBP11/RpoL and all others are associated with FA:69044 which is RNA polymerase II subunit RBP4 (RpoF).

Figure 4.25 shows a cluster consisting with two different families: *RBP11/RpoL* (64311) and *RNA polymerase II subunit RBP4 (RpoF)* (69044). These two different families are closely related as the subunit of retinol-binding proteins (*RBP*) which are carrier proteins that bind retinol, and retinol plays a crucial role in the growth and differentiation of various body tissues [152]. Indeed, they also share the functions of RNA polymerases as subunit of DNA-directed RNA polymerases which are responsible for the polymerisation of ribonucleotides that produces RNA [153].

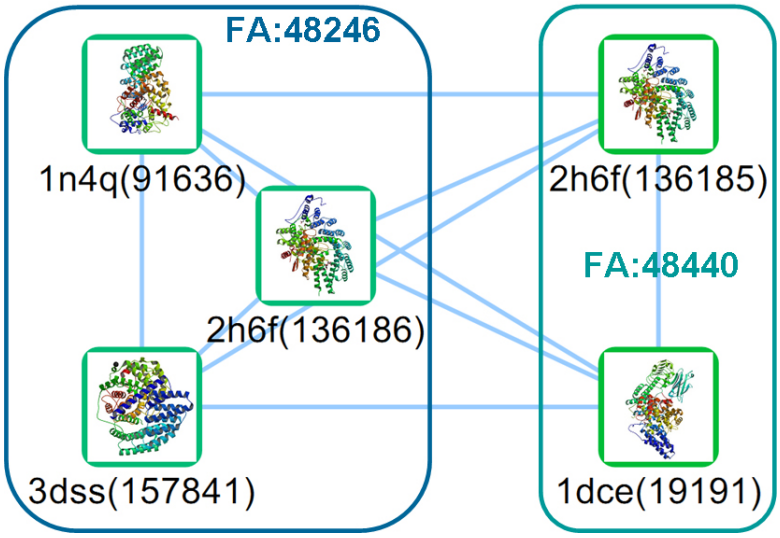



**Figure 4.26** Analysis of mixed families of FA:48246 and FA:48440.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. 2H6F and 1DCE are associated with FA:48440 which is Protein prenylyltransferase and all others are associated with FA:48246 which is Protein prenyltransferases.

Figure 4.26 shows a cluster consisting with two different families: *protein prenylyltransferase* (48440) and *protein prenyltransferases* (48246). As shown in their names, both families function as prenyltransferases which are a class of enzymes that transfer allylic prenyl groups to acceptor molecules [154].

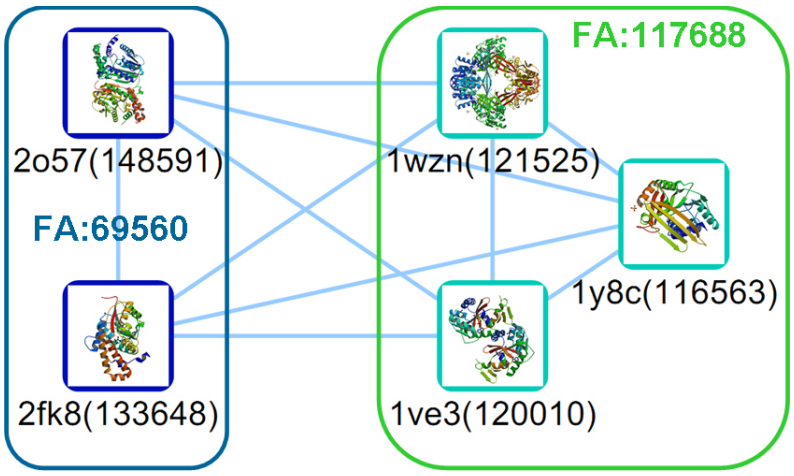



**Figure 4.27** Analysis of mixed families of FA:69560 and FA:117688.
Analyzing relations of mixed SCOP families defined by the proposed method with 70% similarity. 2O57 and 2FK8 are associated with FA:69560 which is Mycolic acid cyclopropane synthase and all others are associated with FA:117688 which is CAC2371-like family.

Figure 4.27 shows a cluster consisting with two different families *Mycolic acid cyclopropane synthase* (69560) and *CAC2371-like* (117688). These two families are subgroup of super-family, *S-adenosyl-L-methionine-dependent methyltransferasess* (53335). Indeed, *Mycolic acid cyclopropane synthase* corresponds to Enzyme Commission (EC) 2.1.1.79 which is a *cyclopropane-fatty-acyl-phospholipid synthase*, the family of transferases that catalyzes the transfer of a methyl group from the donor to Delta9-olefinic acyl chain in phosphatidylethanolamine, phosphatidylglycerol or phosphatidylinositol [155], and CAC2371 corresponds to EC 2.1.1.- which is Methyltransferase. Putative methyltransferase including *Hydroxymycolate synthase* (2fk8) which transfers a methyl group from the S-adenosylmethionine (SAM) cofactor and, subsequently or simultaneously, a water molecule onto the double bound of ethylene substrates, leading to the formation of the hydroxylated product at the distal position [156]. By considering the facts that SCOP classification is based on

structural similarity, and $R_{scr}$ is derived from SCOP classification, the expected scores are supposed to highly correlated to structural similarity. However, the final integrated scores (i.e, both $R_{scr}$ and $G_{scr}$) show that overall scores tend to give higher scores to the mixed families having similar roles or functions than structural similarity. For example, in Figure 4.22 and 4.24~6, their functional roles are very much the same as expected from their names, but they do not share any common classification level until it reaches the top level structure classification. This observation is very important to explain the results showing lower structural similarity in cluster consisting of mixed families than clusters consisting of same family. Indeed, this analysis shows that the proposed methods are neither dominant at structural similarity nor other specific similarities, but the method more importantly considers the similarity in roles and functions in GPs. In fact, this result clearly shows that the proposed bidirectional integration method well fit the goal of this study, identifying functional similarity.

It is possible to argue that the previous experiments are not fair to compare the performance between the proposed method and others since SCOP classification is already integrated in the proposed method as one of the features making $R_{scr}$. However, by considering the fact that the SCOP itself is considered as one of features to define final score, and $R_{scr}$ itself cannot reconstruct SCOP since the SCOP classification is reinterpreted to get statistical significance of GO relations. Nevertheless, it is still interesting to evaluate and compare the proposed method in different ways. More detailed analysis about each cluster is attached to [S-10] at supplementary section.

## 4.3 Experiment III: identifying new disease related genes

### 4.3.1 Dataset

To identify disease related genes, OMIM dataset is derived from previous studies of Schlicker et al. [60]. OMIM provides information about genetic disorders and related human genes, and the dataset used in experiments contains total 99 disease and 413 genes. OMIM dataset used in the experiments are listed in supplementary material [S-11]. To identify the relations between genes and specific disease, all possible pairs derived from 413 genes are involved to get their similarity scores from BSM score and other conventional methods. Based on the score, OMIM disease network is constructed, and their relations are emphasized by using the threshold of functional similarity. In order to compare the performances, the threshold is changed until all methods involved in this experiment get reasonable results.

### 4.3.2 Results

Online Mendelian Inheritance in Man (OMIM) database [157] is used to collect disease related genes. OMIM contains information about human gene and genetic disorders. For the experiments total 99 disease and 413 genes are retrieved from OMIM. Functional similarities are calculated from all possible pairs of genes. The results from all methods are drawn in Figure 4.28. The Figure located in top left of Figure 4.28 is drawn with 70% similarity as a threshold. As previously discussed as human cognition issue, the score showing greater than 50% similarity should give some meaningful information if the similarity range from 0 to 1.0 is considered. However, the results show that none of them give any meaningful results. It may imply that the scoring scheme applied in conventional method is somewhat different from human cognition, or at least it requires extra efforts or processes to determine problem dependent threshold value.

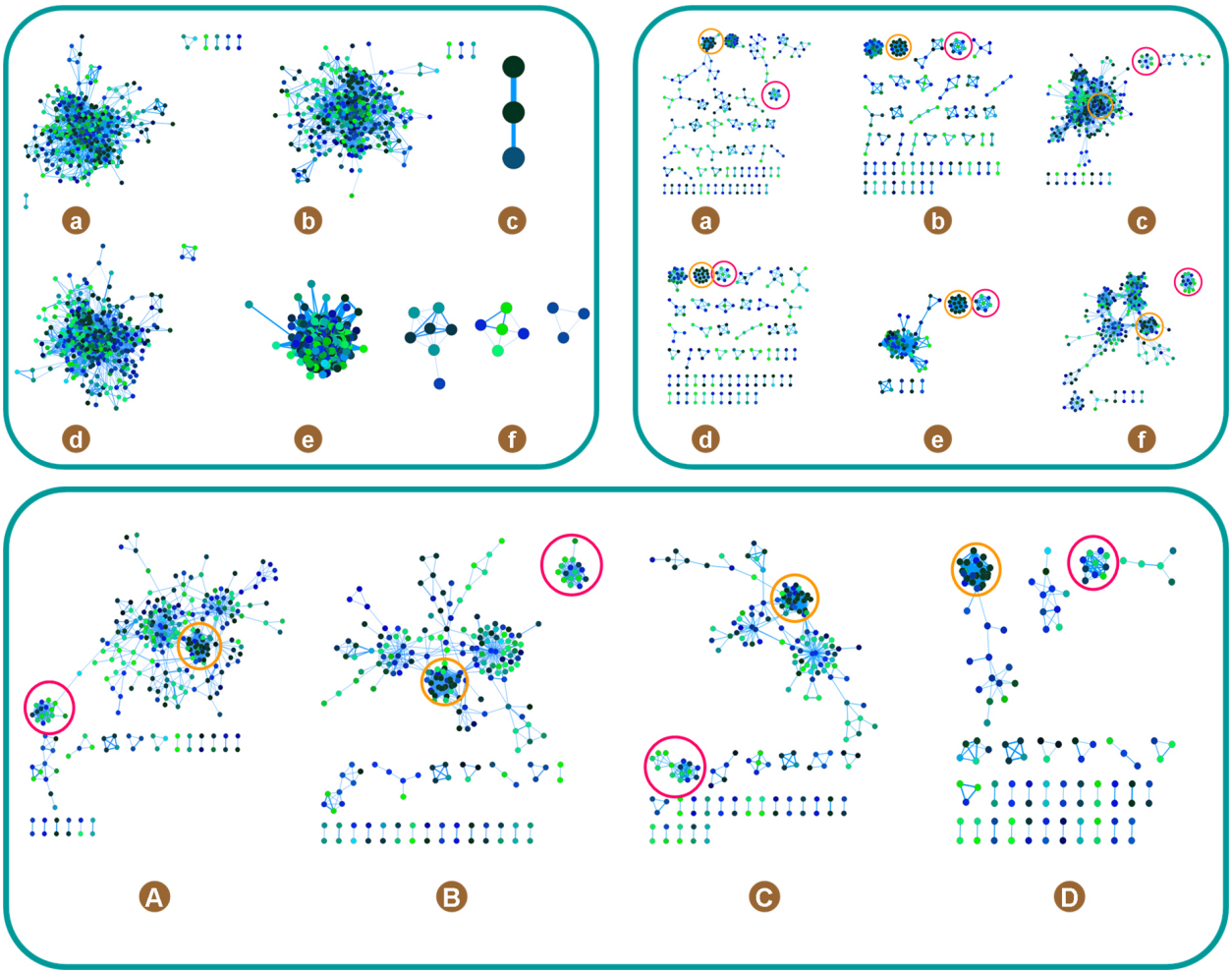



**Figure 4.28** OMIM network generated by different methods.
OMIM network with threshold: **1)** top-left: threshold with 70% GP similarity. **2)** top-right: clusters identified with highest precision measured by SCOP definition. Two distinct clusters are identified by all methods which are emphasized by yellow and red circles. **3)** Bottom – the variations of network upon changing threshold of new methods. Lower case a~f indicates Schlicker (1.0), Lin (1.0), Resnik (0.6), Gentleman (1.0), Jaing (1.0), and Ye's method (1.0) respectively in which the number in parenthesis indicates threshold value. Upper cases A~D in the bottom Figure indicate GP similarity greater than 0.6, 0.65, 0.7, and 0.75 respectively.

In fact, to find optimized values for following experiments, threshold values with finding maximum number of target genes are used such that Schlicker with 0.85, Lin with 0.9, Resnik with 0.3, Gentleman with 0.9, Jaing with 0.95, and Ye with 0.35 in which the order corresponds

to lower case of alphabet from '*a*' to '*f*' in both top-left and top-right of Figure 4.28. We note that the optimal threshold values used in following experiments and previous experiments are different as using the threshold corresponding to maximum number of clusters and the highest precision respectively. As shown above, the threshold value of certain methods is dramatically changed based on the given problem. For example, Ye and Resnik's method were 0.35 and 0.3 in current experiment respectively, and 1.0 and 0.6 chosen by producing highest precision on SCOP clustering problem respectively. In contrast to these cases, the proposed method is very stable on choosing threshold. In fact this varied between 0.65 and 0.75. Again, this experiment also shows the importance of human cognition between the calculated score and functional similarity. To illustrate the changes of network based on similarity scores, Figure 4.28 is drawn as changing the threshold of the proposed method. As expected, the structure of network is more intuitive than other conventional methods as showing the different levels of information while changing the threshold value. Interestingly no matter what methods were applied, in Figure 4.28, all methods have two distinct disease groups: mitochondrial deficiency related disease which is enclosed by a red circle and diseases dominated by protocadherin beta which is enclosed by a yellow circle.

To further explorer the differences of similarity analyzers, mitochondrial deficiency related disease are analyzed since at the current stage, resources in this group have more specific information than the other group. In fact, literatures show that some diseases are strongly associated with mitochondrial deficiency related disease which is the target cluster. *Leber optic atrophy* is a mitochondrially inherited degeneration of retinal ganglion cells and their axons that leads to an acute or subacute loss of central vision; this affects predominantly young adult males [158].

**Table 4.1** Disease-related genes identified by BSM score.

| UniProt ID | DISEASE | EC | GENE | Name of Protein | OMIM |
|---|---|---|---|---|---|
| P03891 | LEBER OPTIC ATROPHY | 1.6.5.3 | MT-ND2 | NADH-ubiquinone oxidoreductase chain 2 | 535000 |
| P03901 | LEBER OPTIC ATROPHY | 1.6.5.3 | MT-ND4L | NADH-ubiquinone oxidoreductase chain 4L | 535000 |
| P03915 | LEBER OPTIC ATROPHY | 1.6.5.3 | MT-ND5 | NADH-ubiquinone oxidoreductase chain 5 | 535000 |
| O75251 | LEIGH SYNDROME | 1.6.5.3 | NDUFS7 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial | 256000 |
| O00217 | LEIGH SYNDROME | 1.6.5.3 | NDUFS8 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial | 256000 |
| P49821 | LEIGH SYNDROME | 1.6.5.3 | NDUFV1 | NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial | 256000 |
| P03897 | LEIGH SYNDROME | 1.6.5.3 | MT-ND3 | NADH-ubiquinone oxidoreductase chain 3 | 256000 |
| P03923 | MELAS SYNDROME | 1.6.5.3 | MT-ND6 | NADH-ubiquinone oxidoreductase chain 6 | 540000 |
| P03905 | MELAS SYNDROME | 1.6.5.3 | MT-ND4 | NADH-ubiquinone oxidoreductase chain 4 | 540000 |
| P03886 | MELAS SYNDROME | 1.6.5.3 | MT-ND1 | NADH-ubiquinone oxidoreductase chain 1 | 540000 |
| O15239 | MITOCHONDRIAL COMPLEX I DEFICIENCY | N/A | NDUFA1 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 1 | 252010 |
| P28331 | MITOCHONDRIAL COMPLEX I DEFICIENCY | 1.6.5.3 | NDUFS1 | NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial | 252010 |
| O43181 | MITOCHONDRIAL COMPLEX I DEFICIENCY | N/A | NDUFS4 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial | 252010 |
| Q8N183 | MITOCHONDRIAL COMPLEX I DEFICIENCY | N/A | NDUFAF2 | Mimitin, mitochondrial | 252010 |
| O75306 | MITOCHONDRIAL COMPLEX I DEFICIENCY | 1.6.5.3 | NDUFS2 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial | 252010 |
| P21912 | PHEOCHROMOCYTOMA | 1.3.5.1 | SDHB | Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial | 171300 |

*Leigh syndrome* is a rare neurometabolic disorder caused by mutations in mitochondrial DNA and leads to loss of head control and motor skills by affecting the central nervous system [159]. *Melas syndrome* is one of the families of mitochondrial cytopathies caused by defects in the mitochondrial genome which is inherited purely from the female parent and leads to vision loss, problems with movement, and a loss of intellectual function [160]. *Pheochromocytoma* is a rare tumor of adrenal gland tissue caused by germline and a mitochondrial complex II gene mutation [161]. It results in the release of too much epinephrine and norepinephrine, hormones that control heart rate, metabolism, and blood pressure [162]. *Mitochondrial complex I deficiency* is the most common enzymatic defect of the oxidative phosphorylation disorders and causes a wide range of clinical disorders, ranging from lethal neonatal disease to adult-onset neurodegenerative

disorders [163]. As shown above, all diseases are connected through neurodegenerative disorders and their signs and symptoms. More precisely they are all caused by mitochondrial mutations or deficiency. Indeed 12 out of 16 genes have exact same EC number, EC 1.6.5.3 which is NADH:ubiquinone reductase (H(+)-translocating) presented in mitochondria and aerobic bacteria. More information about identified disease related genes are shown in table 4.1. Based on these evidences showing strong connecting between genes and diseases, the power of identifying related genes are analyzed by counting the number of genes in table 4.1.

**Table 4.2** Disease-related genes identified by conventional methods.

| Method | Threshold | Previous threshold | Identified genes |
|---|---|---|---|
| Schlicker | 0.85 | 1.0 | MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6, NDUFA1, NDUFS4 |
| Lin | 0.90 | 1.0 | MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6, NDUFA1, NDUFS4 |
| Resnik | 0.30 | 0.6 | MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6, NDUFA1, NDUFS4, DLD |
| Gentleman | 0.90 | 1.0 | MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6, NDUFA1, NDUFS4 |
| Jiang | 0.95 | 1.0 | MT-ND2, MT-ND3, MT-ND4, MT-ND5, MT-ND6, NDUFA1, NDUFS4 |
| Ye | 0.35 | 1.0 | MT-ND1, MT-ND4L, NDUFAF2, NDUFS1, NDUFS2,NDUFS7, NDUFS8, NDUFV1 |

The results shown in table 4.2 show that Schlicker, Lin, Gentleman, and Jiang group exactly same genes. Resnik's method found a gene, DLD, which is related to *Maple syrup urine disease*. However, this disease is caused by homozygous or compound heterozygous mutation in specific genes, so they are different from mitochondrial mutations or deficiency. Ye's method detect 15 genes out of 16 genes, so this method is only compatible to the proposed method. However, by considering their threshold value, 0.35, as Ye's method point of view, this cluster may not be considered as sharing strong functional similarity when the range of minimum score 0 and maximum score 1 is considered. Indeed, the optimal threshold value used in this experiment was defined by sequentially increasing the threshold value until it creates a cluster with the maximum

number of related genes. Therefore, in real case, these genes may not be found by Ye's method due to the dramatic difference of threshold values defined upon precision and number of clusters. Overall results clearly show that the proposed methods outperform conventional methods upon both statistical analysis and biological relevancies. Therefore, the possible application derived from the proposed method is of interested to promise the success in real world problems.

To test the proposed method into more realistic problems, the proposed method is applied to finding new disease related genes such that the previous mitochondrial mutation and deficiency disease related genes are used as template genes to discover new disease related genes. To identify related genes, proteins in UniProt are used as a pool of potentially related genes. To get both benefits of improving the relationships and reducing searching space, the pool of potential genes are screened by keeping genes with at least one exact same cellular component GO term with the template genes since functionally related genes often share their cellular component. To make the balance reducing search space and increasing the detection rate of functionally related genes, threshold value 0.6 is used. In fact, this value is slightly lower than the recommended range of threshold value (0.65~0.75). The newly identified disease related genes are depicted in Figure 20. As expected, the survey of the newly discovered genes shows that some of them are highly related to mitochondrial mutation or deficiency. For example, FDX1L is a mitochondrial iron-sulfur (Fe/S) protein that transfers electrons from NADPH via FDXR and functions in Fe/S protein biosynthesis, and strongly connected to activities of mitochondrial complex I [164]. NDUFA2 gene encodes one of the accessory subunits of ubiquinone oxidoreductase, and a homozygous splice site mutation. More importantly NDUFA2 can cause Leigh syndrome [165].

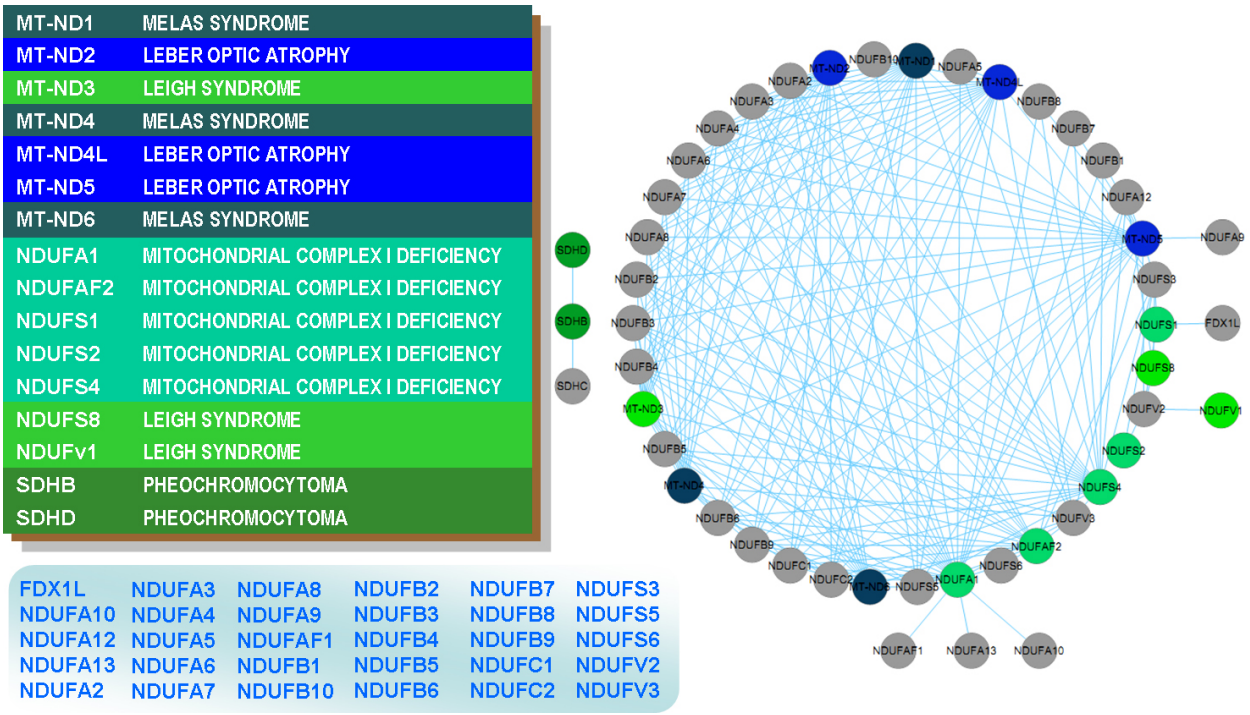



**Figure 4.29** New disease related genes identified by BSM score.
Mitochondrial dysfunction related genes retrieved from OMIM and newly discovered genes: left-top shows the list of target genes used for searching UniProt database. The right shows the association between target genes and newly discovered genes. Bottom-left shows the 30 newly discovered genes with threshold value 0.6.

Loeffen et al. [166] found that NDUFV1, NDUFV2, and NDUFV3 are the subunits of NADH:ubiquinone oxidoreductase which is enzyme complex in the electron transport chain of mitochondria. The authors also thought that HP fraction proteins NDUFA7, NDUFA10, NDUFB2, NDUFB4, and NDUFB10 participated in the transfer of protons from the mitochondrial matrix to the intermembrane space and suggested that HP fraction contribute to communicate with mitochondrial matrix or intermembrane space. Bugianie et al. found that variants of NDUFA8 can cause mitochondrial complex I deficiency [167]. Osterquaard et al. [168] found that NDUFA12 is is required either at a late step in the assembly of complex I, or in the stability of complex I. Lu and Cao [169] show that NDUFA13 which is also known as

GRIM-19 is essential for electron transfer activity of mitochondrial complex I. Puissegur et al. [170] found that NDUFA4 is a target of miR-210 and strongly related to a loss of mitochondrial membrane potential and the apparition of an aberrant mitochondrial phenotype. Marui et al. [171] show that NDUFA5 is strongly related to mitochondria dysfunction by analyzing the relations between SNPs of NDUFA5 and Autism which is caused by central nervous system dysfunction commonly associated with mitochondrial disorders. Melick et al. [172] show that NDUFA6 is associated with mitochondrial complex I deficiency by showing SNPs of NDUFA6 can lead Parkinson's disease which is known to be affected by mitochondrial DNA mutation [173, 174]. Most recently van den Boschet al. identified a homozygous mutation in the NDUFA9 gene from a patient with fatal neonatal Leigh syndrome due to mitochondrial complex I deficiency [175]. The NDUFAF1 gene encodes the human homolog of a Neurospora crassa Cia30 protein involved in the assembly of mitochondrial respiratory complex I [176]. Fassone et al. [177] identified compound heterozygosity for 2 mutations in the NDUFAF1 from the patient with fatal infantile hypertrophic cardiomyopathy which requires diagnosing mitochondrial cardiomyopathy caused by mitochondrial dysfunction. Calvo et al. identified a homozygous mutation in the NDUFB3 from a patient with severe mitochondrial complex I deficiency [178]. Most recently Haack et al. [179] found that mutations of NDUFB9 could cause mitochondrial complex I deficiency by screening the reduction of NDUFB9 and both amount and activity of complex I. Szklarczyk et al. identified that NDUFB7 protein lacks a mitochondrial targeting signal and transmembrane domains [180]. Loublier et al. [181] proved that NDUFB6 is essential for mitochondrial complex I activity by showing the shRNA expression resulted in extinction of additional genes. Ugalde et al. [182] shows that mutations of NDUFS3, NDUFS5, and NDUFA9 severally affect the assembly and/or stability of mitochondrial complex I. From two unrelated

families with mitochondrial complex I deficiency, Kirby et al. [183] identified homozygous mutations in the NDUFS6 gene, causing a splicing abnormality and a large deletion, respectively. Although some of newly discovered genes do not show direct evidences for relations with mitochondrial deficiency, their relations can be deduced from related studies. For example, Granata et al [184] found that based on microarray and pathway analysis, certain genes including NDUFB1, NDUFA6, and NDUFS5 could effectively discriminate chronic kidney disease (CKD) and hemodialysis patients, and many of them were involved in mitochondrial oxidative phosphorylation system. By considering the fact that mitochondria are involved in calcium and iron hemostasis, it is possible that mutations of NDUFB1, NDUFA6, and NDUFS5 can lead mitochondrial dysregulation. The study from Su et al. [185] showed that NDUFB5 gene was the one of the highest number of dysregulated genes appeared in the mitochondrial dysfunction and oxidative phosphorylation pathways. Tucker et al. [186] found that NDUFB8 reflected instability of mitochondrial complex I in the absence of mtDNA-encoded protein found in fibroblast patients. In summary, 24 genes out of 30 newly discovered genes had direct evidences of relations toward mitochondrial complex I deficiency. Studies also showed that three out of rest six genes including NDUFB1, NDUFB5, and NDUFB8 had relatively strong relations to complex I deficiency. Unfortunately no literature has been found that showed the relation between complex I deficiency and three newly discovered genes including NDUFA3, NDUFC1, and NDUFC2. As a result, based on the literature survey, the proposed method shows more than 80% overall accuracy. More importantly, most of literatures shown above have not been added to the most recent OMIM, so the reported literatures in this study should be valuable information to explorer or expand current knowledge in mitochondrial complex I deficiency.

## 4.4 Experiment IV: drawing disease network

### 4.4.1 Dataset

Disease similarity network is drawn based on the average GP similarity measured by BSM score and OMIM database which contains 93 diseases out of 99 OMIM diseases used in this experiments. Disease and associated genes are attached in supplementary section. It is important to notice that this network is different from conventional disease networks by containing functional similarity information among them. To calculate similarity between two diseases, all GP similarity scores associated with a pair of diseases are averaged, and the averaged scores are projected proportionally into the thickness of edges and the intensity between green and red color that correspond to minimum and maximum disease similarity respectively. The degree of each node is depicted proportionally to the radius of nodes. Average similarity score 0.25 is used as a threshold value to avoid making complete connections among nodes in the network which is unfavorable conditions to analyze network with Cytoscape program [187] since this makes circular shaped fully connected network. Upon the proposed drawing strategy and threshold cutoff, existing similarity scores in Figure 4.30 are ranged between 0.25 and 0.85.

### 4.4.2 Results

As shown in Figure 4.30, majority diseases which are depicted in the middle of the network with relevantly higher degree (i.e. the big circles in the middle) are densely connected to each other, and this is not trivial task to analyze each individual edge even if BSM-based network provides similarity score instead of just showing their connections. Besides these high degree disease nodes, four interesting groups have been found that they are relatively loose but show clear and strong relations among them. Each group has been analyzed starting from top right to top left with clockwise direction.

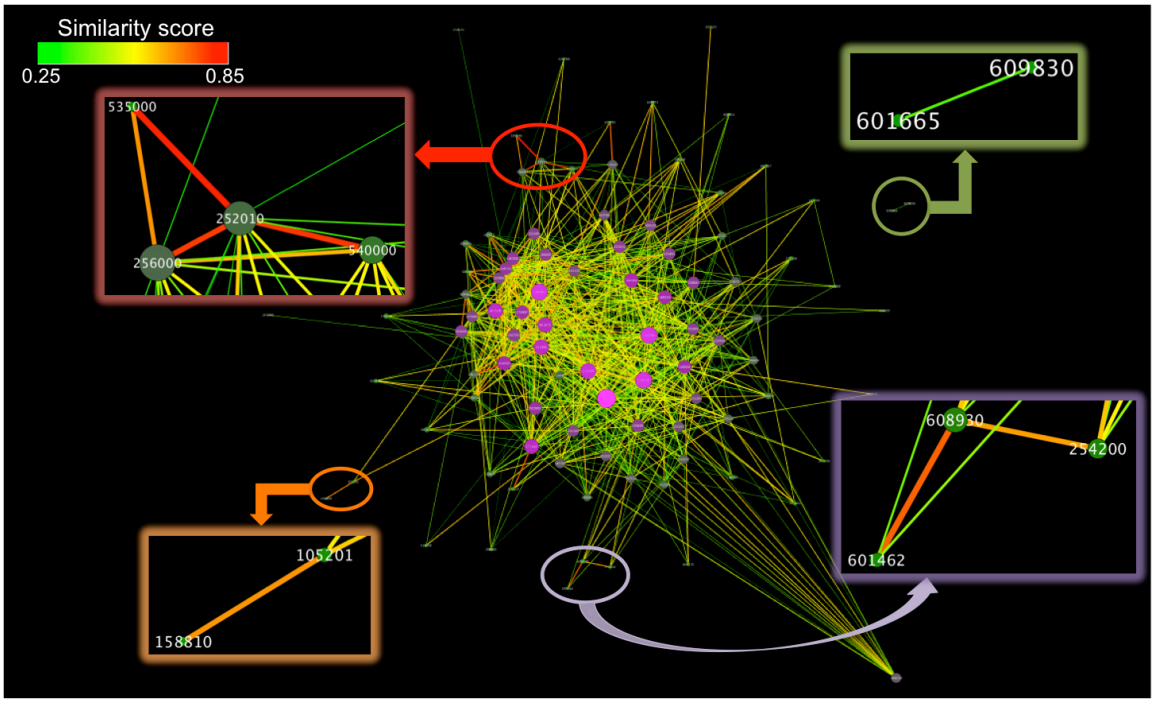



**Figure 4.30** Disease Similarity Network created by BSM score.
The size of nodes corresponds to degree of edges. The thickness of edges is corresponds to the similarity. The colors of edges are ranged from green to red that are the minimum and maximum similarity respectively. Four interesting groups of disease are magnified, and the numbers belonging to each node correspond to OMIM ID. The group located in top right corner consists of 601665 and 609830 which are *obesity* and *abdominal body fat distribution* respectively. The group located in bottom right corner consists of 254200, 601462, and 608930 which are *myasthenia gravis* (MG), *slow-channel congenital myasthenic syndrome* (SCCMS), and *fast-channel congenital myasthenic syndrome* respectively. The group located in bottom left norner consists of 105201 which now is 105210 and 158810 which are *amyloidosis* and bethlem myopathy respectively. The group located in the top left corner consists of 252010, 256000, 535000, 540000 which are *mitochondrial complex I deficiency*, *leigh syndrome*, *leber optic atrophy*, and *melas syndrome* respectively.

The group consisting of 601665 and 609830 are *obesity* and *abdominal body fat distribution* respectively. These diseases are the only group that is completely isolated from other diseases. It is interesting that although obesity is well known disease related to many diseases, the proposed network does not show significant relations to others except this pair. Indeed their similarity is 2.8, so they are not significantly related to each other, although their names are somewhat

related. There are two reasons why this happens. First, well known obesity related diseases, for example, caused by heterozygous mutations in the melanocortin-4 receptor gene (MC4R) such as *leptin deficiency* [188], *leptin receptor deficiency* [189], *prohormone convertase-1 deficiency* [190], and *proopiomelanocortin deficiency* [191] do not appear in applied datasets. Indeed, although the direct causal relations between obesity and many diseases are suspected, they are still unclear [192-195]. Second, the similarity score between *obesity* and *abdominal body fat distribution* are low due to high numbers of the diverse functional terms. In fact, *obesity* consists of total 27 unique functional terms and *abdominal body fat distribution* consists of total 18 unique functional terms; therefore, these diverse functional terms could affect to decrease similarity between two diseases since BSM-score is based on average similarity score, so the higher number of lower similarity scores could be dominant to make final similarity. In fact, abdominal body fat distribution has been removed from OMIM database due to unclear distinction of phenotype from that in another entry and increasing number of related genes and their variants [196].

The second group consists of 254200, 601462, and 608930 which are *myasthenia gravis* (MG), *slow-channel congenital myasthenic syndrome* (SCCMS), and *fast-channel congenital myasthenic syndrome* (FCCMS) respectively. This groups shows relatively higher similarity among three diseases. The similarity score between MG and FCCMS is 0.52 and between SCCMS and FCCMS is 0.63. Interestingly, MG and SCCMS do not show direct, although these three diseases are linked to *generalized epilepsy with febrile seizures plus* which corresponds to OMIM 604233 which is a syndromic autosomal dominant disorder where afflicted individuals can exhibit numerous epilepsy phenotypes which are diverse set of chronic neurological disorders characterized by seizures [197, 198]. MG is a neuromuscular disorder caused by

producing antibodies to block the muscle cells receiving messages from the nerve cell [199]. Researches [200, 201] showed that CHRNA1 gene encodes the alpha subunit of the muscle acetylcholine receptor (AChR), which is the main target of pathogenic autoantibodies in autoimmune MG. FCCMS is a genetic disorder of the neuromuscular junction that can be classified by the site of the transmission defect. Studies [202-205] showed that FCCMS is caused by heterozygous mutations within the AChR subunit genes, for example, CHRNA1, CHRND, CHRNE genes that affect AChR kinetics and results in infrequent and brief ion channel activations. SCCMS is a genetic disorder of the neuromuscular junction and an autosomal dominant form of postsynaptic CMS, which can be caused by mutation in the subunits of the AChR that can be resulted in kinetic abnormalities of the AChR. Studies [206-209] show that SCCMS is strongly related to mutations of CHRNA1, CHRNB1, CHRND, or CHRNE gene. The relations shown in the network are meaningful such that SCCMS and FCCMS are strongly related to each other by sharing majority of causing genes, CHRNA1, CHRND, and CHRNE genes. The reason of absent links between MG and SCCMS can be explained that other unrelated genes are dominant to defining final similarities. It is important that the absent links in the BSM-based network do not necessary to be showing absence of actual relations since the proposed networks are drawn based on similarity cut-off. In other words, the relation between MG and SCCMS still exist since they share one gene, but by considering their similarity, their similarity is not strong enough because SCCMS consists of much more functions than MG, and the definition of similarity between a pair of GPs or diseases in proposed method is equilibrium on their functions with specific functional terms. Therefore, the lack of functional equilibrium between MG and SCCMS leads the less similarity in this case.

The third group consists of 105210 and 158810 which are *amyloidosis* and *bethlem myopathy* respectively. Hereditary *amyloidoses* are a clinically and genetically heterogeneous group of autosomal dominantly inherited diseases characterized by the deposit of unsoluble protein fibrils in the extracellular matrix due to the mutations in transthyretin (TTR) gene [210]. Studies show TTR aggregations are significantly related to clinical syndromes, for example, familial amyloid polyneuropathy (FAP) and cardiomyopathy are caused by mutant TTR protein aggregation in peripheral and autonomic nerves and heart, respectively. [211]. *Bethlem myopathy* is a skeletal muscular disease of which the muscle fibers do not function for any one of many reasons, resulting in muscular weakness. People with this condition experience progressive muscle weakness and develop joint stiffness (contractures) in their fingers, wrists, elbows, and ankles that can restrict movement. Many studies show significant relations between *amyloidosis* and *myopathy* including causal relations (i.e. amyloidosis causes myopathy) [212-217]. Indeed, myopathy is diagnosed with muscle biopsy by identifying *amyloidosis*. For example, *amyloid myopathy* (AM) is a manifestation of primary systemic amyloidosis and diagnosed by identifying *amyloidosis* in the gastric mucosal biopsy [218].

The group located in the top left corner consists of 252010, 256000, 535000, 540000 which are *mitochondrial complex I deficiency*, *leigh syndrome*, *leber optic atrophy*, and *melas syndrome* respectively. For this group, the details of similarities upon related genes are already discussed in the section 'Identifying disease related genes' (i.e. please refer the context associated with Figure 20). In short, this group shows strong relations among them such that all these diseases are caused by mutations in mitochondrial genomes and related to neurodegenerative disorders. Indeed, their signs and symptoms are very similar each other including lost of vision and motor skills. Results show that BSM score can be applied to

constructing disease similarity network which gives more information than simple diseases network by providing semantic confidences of disease similarity. Indeed, the successful identifications of similar diseases show that the proposed method can be applied to real world problems for screening similar and outlier diseases in high throughput manner.

# 5 Gene Ontology Analysis Layer (GOAL)

## 5.1 Introduction

One of the most fundamental and important tasks of computational biology or bioinformatics is providing solutions to help understanding inter-relationship between genes and GPs. Although significant progress in the conventional methods based on sequence and structure similarity has been made, their roles in defining functional similarity that has significant impact on revealing inter-cellular relationship are still limited [111, 219-221]. Due to their limitations and computational complexity many studies have been conducted by directly targeting measuring functional similarities among genes or GPs [59, 62, 90, 91, 95, 98, 101]. Identifying functionally similar or closely related genes and GPs has especially significant impacts on clinical studies and trials by supporting the prediction of metastasis or development of treatments for certain diseases since the identified similar genes and GPs can give critical information to understand molecular mechanisms or functions in cellular life of unknown or newly discovered diseases by referring the gene or GPs having similar functions [26, 54-57, 222, 223]. Therefore, it is not surprising that the interest of measuring the semantic similarity among genes and GPs have been significantly increased. Indeed, identifying functional similarities are particularly useful in validating drugs and targets by making possible to recycle the information obtained from target/drug validations and clinical trials [26]. The current vest amount of GPs with known functions stored in various databases, and systematically defined functional annotations like Gene Ontology (GO) [77] are of promising to develop biologically and statistically more relevant method toward measuring functional similarities among GPs. Although several methods have been introduced, their services are limited on their interests and do not provide stand-alone

outputs that can be used for further analysis of the results. More details of existing services and GOAL are compared and summarized in table 5.1. Four web-services are compared to GOAL: FuSSiMeG [73], FunSimMat [58], ProteInOn [61], and G-SESAME [73].

**Table 5.1** Comparing GOAL and other existing services

| Contents | FuSSiMeG | FunSimMat | ProteInOn | G-SESAME | GOAL |
|---|---|---|---|---|---|
| Source DB (1=GO, 2=UniProt) | 1 | 1,2 | 1 | 1 | 1,2 |
| # of method for GO similarity | 3 | 1 | 6 | 1 | 7 |
| # of method for GP similarity | 3 | N/A | 1 | N/A | 5 |
| # of all possible scores for GP | 6 | N/A | 6 | N/A | 42 |
| GO similarity between groups | No | No | No | No | Yes |
| Supporting direct input | Yes | Yes | Yes | Yes | Yes |
| Supporting file upload | No | No | No | No | Yes |
| Supporting batch process | No | Yes | Yes | No | Yes |
| Navigating 2 nodes | No | No | No | No | Yes |
| Navigating 1 node | No | No | No | No | Yes |
| Handling root-shared GO term | No | No | No | No | Yes |
| Handling GPs without annotation | No | No | No | No | Yes |
| Providing root shared term pairs | No | No | No | No | Yes |
| Providing GO statistics | No | No | No | No | Yes |
| Continuing Service | discontinue | continue | continue | continue | continue |

In addition to limited accessibility, two significant drawbacks have been found in current existing methods: *i) the nil similarity resulted from GPs with no annotation* and *ii) the nil similarity from root-shared GO terms*. Most of conventional methods are based on two distinct features of GO graph: nearest common ancestor (NCA) and information contents-based score. NCA is the common ancestor of two GO terms located furthest from the root node of GO graph, and IC-based scores are calculated using the number of annotated GPs. Due to the incompleteness of GO annotation for existing GPs and unsynchronized update between GO terms and GP annotations in GO database, there exists GO terms having no annotated GPs. NCA-based scores implicitly allows zero similarity score of which NCA of two GO terms is the root node of GO graph. All possible combinations of GO term pairs among 9,013 GO terms have been calculated in current version of GOAL. The results shows that 20,986,023 pairs out of

40,621,591 pairs have root-shared relations which are about 52% of total possible GO terms pairs (i.e. GOAL provides the list of root-shared GO terms as a submenu). The scores are not only statistically significant toward the number of available GPs for measuring their similarities, but it also shows that this kind of scoring scheme under estimates biological relations. For example, the NCA of GO:0004525 and GO:0046872 which are *ribonuclease III activity* and *metal ion binding* respectively is the root node of GO graph (i.e. GO:0003674, molecular function), but annotated GPs have significant relations in their functional roles such that all corresponding GPs are associated with *dicer-like endoribonuclease* protein, match enzyme commission number (EC) EC3.1.26.- which is *ribonuclease*, and even share cofactors *magnesium* and *manganese* (please refer supplementary table S-1 and S-2 for details and more examples). GOAL is designed to minimize the risk of resulting nil similarity by integrating biological characteristics of GPs and statistical significance of GO term relations in GO databases and provides user interfaces to get maximum 42 different scores from embedded methods that can be used for determining the confidence of resulted scores.

## 5.2 Input

GOAL provides user-friendly graphic user interface (GUI) for both measuring GO term and GP similarity. For GO term similarity GOAL provides three different types of interfaces for measuring semantic similarity between two GO terms: i) direct input of two GO terms, ii) list of GO terms pairs through input text box, and iii) list of GO terms through file upload. GP similarity provides two types of GUI: i) direct inputs from input text boxes, and ii) file uploads. Compared to GO similarity GP similarity requires two inputs: GP contents and GP pairs. GP contents define unique GP name and corresponding GO annotations. GP pairs are the list of GP pairs appeared in GP contents. Besides common inputs, Jiang's method requires extra parameters, and they can be easily defined through input forms.

## 5.3 Output

The results of GOAL can be further used and analyzed by researchers, and the size of results can be very large since GO and GP similarities are often resulted in the permutations of seed lists; therefore, GOAL provides tab delimited text file outputs. BSM provides four different scores GLOBAL_MUL, GLOBAL_AVE, LOCAL_MUL, and LOCAL_AVE in which GLOBAL and LOCAL are defined by normalization factor $E_{max}$ and $d_{max}$ shown in equation (4). MUL and AVE denote whether BSM is defined by multiply of $B_{scr}$ and $S_{scr}$ or average of them. Besides BSM, GOAL also provides four different scores for six conventional methods: DIR_GO, ANC_GO, DIR_UniProt, and ANC_UniProt. At the name, GO and UniProt [224] are the two different sources to get annotated GPs. DIR and ANC denotes whether considering annotated GO term only or including ancestors of the GO terms toward calculating GO and GP similarity. The results of six different methods are also provided as downloadable text files.

## 5.4 Submenus

Besides calculating GO and GP similarity, GOAL provides comprehensive GO analysis platform including GO Navigator, GP annotation writer, and daily GO statistics. GO Navigator visualize either single GO term or GO term pairs based on user preference. Writing the input of calculating GP similarity is often cumbersome since it often requires retrieving annotated GO terms of target GPs. To improve user accessibility, GOAL provides automatic input writer to get annotated GPs from the list of UniProt IDs or gene names. The daily updated GO statistics is one of unique service in GOAL. Statistics menu provides statistics of GO terms and annotated GPs from 2005 to present. In fact, from August 2011 to present, GOAL provides daily updated GO statistics. To help the people developing new method for measuring GP similarity, GOAL also provides GUI for retrieving root-shared GO terms. At the current version GOAL contains total 20,986,023 root-shared GO term pairs. The results provided by GOAL are all downloadable, so that users can easily redistribute and analyze the results. Overall schematic diagram of GOAL is shown in Figure 5.1.

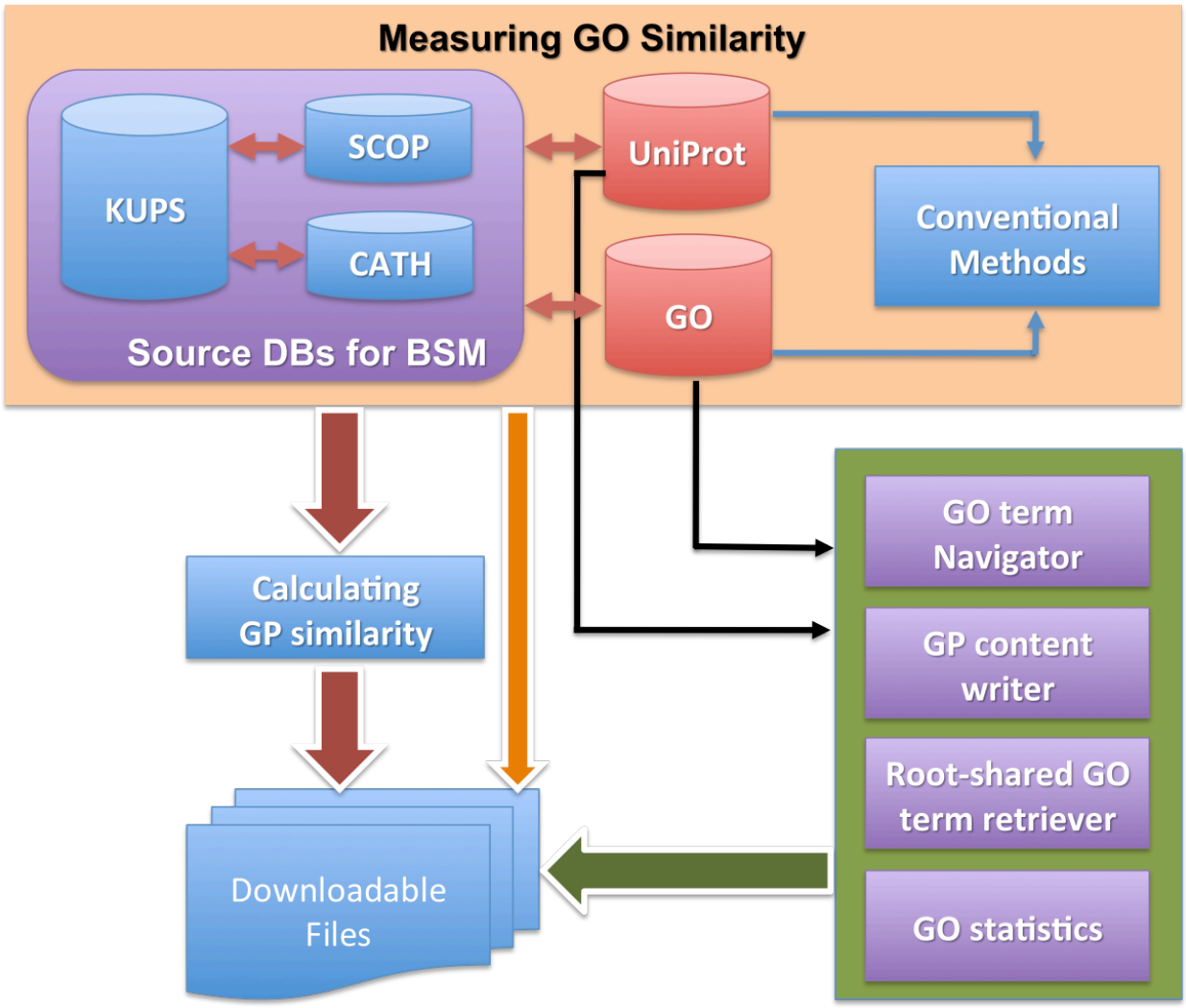



**Figure 5.1** GOAL schematic diagram.

GOAL contains two major semantic similarity scores: BSM and six conventional methods. BSM is derived from five DBs: KUPS, SCOP, CATH, UniProt, and GO. Six conventional methods simultaneously calculate GO similarity scores based on known GP annotations from UniProt and GO which is one of unique features in GOAL. Upon user's request GOAL can calculate GP similarity from five different methods. Thus compared to other existing services, GOAL provides multiple scores resulted from two different DBs and five different GP scoring methods so that users can determine the confidence of resulted scores by comparing consistence of the scores. Besides measuring functional similarity of GO terms and GPs, GOAL also provides various tools to support researchers analyzing resulted scores and developing new methods to define semantic similarity between GPs. To help user's accessibility and post analysis, GOAL also provides all results through downloadable files and web-view including images and text files.

# 6 Conclusion and future works

The significance of the proposed method can be found in several aspects: 1) solving problems of the nil similarity and root-shared terms: the most critical defects in current methods are resulted in nil similarity or giving same similarity scores to any pairs of root-shared GO terms which are resulted from the absence of annotated GPs and root-shared GO terms respectively. Proposed method prevents the nil similarity by calculating functional similarity through the topology of GO graph instead of using the statistics of annotated GPs. To minimize the effects resulted from root-shared GO terms, reference score matrix derived from GO term distributions among known domains is introduced. 2) reflecting specificity and fidelity of GO terms into final score: methods relying on the rationale of nearest common ancestor (NCA) often ignore the specificity and fidelity of GO terms due to shadow problem in which the specificity of target GO term located below the NCA GO term is ignored by considering NCA only. To handle this problem, density and depth of GO terms are integrated in proposed score function such that the deeper and simpler path from root to NCA will reward higher similarity score. 3) combining top-down and bottom-up approaches: by solving the problems stated above, integration of top-down and bottom-up approaches are proposed. As an initial effort, the weight of each approaches are evenly distributed; therefore, the final score will have unbiased results from both graph search-based method and statistical results of GO term distribution upon known functional domains. 4) Narrowing down gaps between theoretical background and biological relevancies: conventional methods dealing with functional similarity are often derived form theoretical background of information theory, but there exist certain gaps between the background theory and real clinical trials as discussed in preliminary study. The proposal is designed for minimizing

the gaps between theoretical approaches and practical usages by comprehensively evaluating the proposed method in multiple steps. 5) Providing guidelines on evaluation process: conventional methods are often focused on evaluations on statistical point of view (e.g. evaluating the similarity score upon sequence and structure similarity). However, statistical evaluations are not enough to elucidate specific characteristics of the methods. The proposed method is evaluated through multiple steps including statistical evaluations with multiple comparisons on conventional methods, and biological analysis based on literature review. It is also emphasized that the human cognition on reported scores by showing unfavorable differences between reported score and user expectation or interpretation on that score should be considered. To show the potentials of the proposed method to be used in real problems, disease related genes are identified. The results show that the proposed methods are very promising to discover disease related genes as showing overall accuracy greater than 80%. This thesis shows new way to define functional similarity between GPs by using terms associated with molecular function in GO database. GO contains three different categories: molecular function, biological process, and cellular component. In future, I will complete similarity scores in other categories and combine them to define more rigorous relations among GPs. Finally I will apply them as a pipeline to connect diverse data sources, so that users can easily explore *Big data*.

# Supplementary materials

## [S-1] Statistics of root shared GO terms and GPs

**300 randomly selected Gene Products (GPs) annotated at least a pair of root-shared terms**

- # GPs: number of GPs annotated at least a pair of root-shared terms (i.e. GO term1 and GO term2)
- # PPIs: number of protein-protein interactions among # of GPs selected above step

| GO term1 | GO term2 | # of GPs | # of PPIs | GO term1 | GO term2 | # of GPs | # of PPIs |
|---|---|---|---|---|---|---|---|
| GO:0004674 | GO:0005524 | 2517 | 372 | GO:0004748 | GO:0046914 | 99 | 9 |
| GO:0004715 | GO:0005524 | 186 | 86 | GO:0004702 | GO:0005524 | 10 | 9 |
| GO:0004842 | GO:0008270 | 400 | 77 | GO:0004872 | GO:0005525 | 35 | 9 |
| GO:0016563 | GO:0043565 | 295 | 52 | GO:0004842 | GO:0043130 | 13 | 9 |
| GO:0004674 | GO:0042802 | 35 | 38 | GO:0030374 | GO:0046966 | 12 | 9 |
| GO:0016563 | GO:0046982 | 58 | 36 | GO:0004672 | GO:0042802 | 6 | 9 |
| GO:0004674 | GO:0046872 | 435 | 34 | **GO:0004716** | **GO:0046982** | **4** | **9** |
| GO:0004709 | GO:0005524 | 58 | 32 | GO:0004674 | GO:0004872 | 172 | 8 |
| GO:0004714 | GO:0005524 | 134 | 28 | GO:0004697 | GO:0046872 | 33 | 8 |
| GO:0004707 | GO:0005524 | 178 | 27 | GO:0004715 | GO:0046872 | 20 | 8 |
| GO:0004842 | GO:0005524 | 264 | 25 | GO:0004708 | GO:0005524 | 26 | 8 |
| GO:0004713 | GO:0005524 | 251 | 24 | GO:0004579 | GO:0042802 | 4 | 8 |
| GO:0004749 | GO:0005524 | 200 | 24 | GO:0016563 | GO:0019901 | 18 | 8 |
| GO:0016563 | GO:0042803 | 51 | 24 | GO:0004639 | GO:0005524 | 539 | 7 |
| GO:0004693 | GO:0005524 | 264 | 23 | GO:0004523 | GO:0008270 | 112 | 7 |
| GO:0004871 | GO:0046982 | 26 | 22 | GO:0004523 | GO:0005198 | 76 | 7 |
| GO:0004872 | GO:0005524 | 267 | 21 | GO:0004748 | GO:0005524 | 72 | 7 |
| GO:0004672 | GO:0005524 | 245 | 19 | GO:0004533 | GO:0008270 | 75 | 7 |
| GO:0004716 | GO:0005524 | 22 | 19 | GO:0004533 | GO:0005198 | 61 | 7 |
| GO:0004872 | GO:0042802 | 15 | 19 | GO:0005083 | GO:0005524 | 30 | 7 |
| GO:0016769 | GO:0030170 | 417 | 18 | GO:0004872 | GO:0046966 | 15 | 7 |
| GO:0004550 | GO:0005524 | 751 | 17 | GO:0004872 | GO:0030374 | 10 | 7 |
| GO:0004697 | GO:0005524 | 71 | 17 | GO:0004872 | GO:0042809 | 10 | 7 |
| GO:0005024 | GO:0005524 | 7 | 17 | GO:0004526 | GO:0042802 | 3 | 7 |
| GO:0004889 | GO:0015464 | 25 | 17 | GO:0004871 | GO:0016563 | 29 | 7 |
| GO:0005024 | GO:0046872 | 6 | 17 | GO:0004523 | GO:0046872 | 1203 | 6 |
| GO:0004722 | GO:0046872 | 186 | 16 | GO:0004519 | GO:0046872 | 553 | 6 |
| GO:0004871 | GO:0005524 | 87 | 16 | GO:0004824 | GO:0046872 | 452 | 6 |
| GO:0004826 | GO:0005524 | 1028 | 15 | GO:0004849 | GO:0005524 | 235 | 6 |
| GO:0016847 | GO:0030170 | 126 | 15 | GO:0004674 | GO:0005083 | 27 | 6 |
| GO:0004842 | GO:0042802 | 21 | 15 | GO:0004702 | GO:0046872 | 24 | 6 |
| **GO:0004749** | **GO:0042802** | **5** | **15** | GO:0004871 | GO:0008134 | 20 | 6 |
| GO:0004871 | GO:0005096 | 56 | 14 | GO:0004674 | GO:0004871 | 15 | 6 |
| GO:0004872 | GO:0046872 | 93 | 13 | GO:0004871 | GO:0046872 | 14 | 6 |
| GO:0004691 | GO:0005524 | 29 | 13 | GO:0004674 | GO:0042803 | 13 | 6 |
| GO:0004871 | GO:0019901 | 25 | 13 | GO:0004871 | GO:0043565 | 15 | 6 |
| GO:0004519 | GO:0005524 | 180 | 12 | GO:0005009 | GO:0043559 | 7 | 6 |
| GO:0016566 | GO:0043565 | 46 | 12 | GO:0005009 | GO:0043560 | 7 | 6 |
| GO:0016702 | GO:0031418 | 292 | 11 | GO:0005009 | GO:0046872 | 7 | 6 |
| GO:0004683 | GO:0005524 | 69 | 11 | GO:0004704 | GO:0005524 | 7 | 6 |
| GO:0004871 | GO:0042802 | 18 | 11 | GO:0005009 | GO:0005524 | 7 | 6 |
| GO:0004824 | GO:0005524 | 592 | 10 | GO:0005024 | GO:0046332 | 6 | 6 |
| GO:0004871 | GO:0005525 | 204 | 10 | GO:0004888 | GO:0042802 | 6 | 6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GO:0004871 | GO:0008270 | 44 | 10 | GO:0004815 | GO:0005524 | 1026 | 5 |
| GO:0004683 | GO:0005516 | 66 | 10 | GO:0004813 | GO:0005524 | 728 | 5 |
| GO:0005003 | GO:0005524 | 44 | 10 | GO:0004775 | GO:0005524 | 564 | 5 |
| GO:0004871 | GO:0005109 | 21 | 10 | GO:0004519 | GO:0008270 | 81 | 5 |
| GO:0004871 | GO:0005110 | 14 | 10 | GO:0004743 | GO:0005524 | 102 | 5 |
| GO:0004872 | GO:0051740 | 5 | 10 | GO:0004743 | GO:0030955 | 103 | 5 |
| GO:0004550 | GO:0046872 | 733 | 9 | GO:0004518 | GO:0005524 | 81 | 5 |
| GO:0004802 | GO:0046872 | 73 | 5 | GO:0004740 | GO:0005524 | 19 | 4 |
| GO:0004890 | GO:0005254 | 68 | 5 | GO:0005096 | GO:0005524 | 12 | 4 |
| GO:0004970 | GO:0005234 | 67 | 5 | GO:0004758 | GO:0030170 | 18 | 4 |
| GO:0016740 | GO:0030976 | 54 | 5 | GO:0004709 | GO:0046872 | 16 | 4 |
| GO:0004535 | GO:0046872 | 50 | 5 | GO:0004842 | GO:0004871 | 14 | 4 |
| GO:0004525 | GO:0005524 | 42 | 5 | GO:0004705 | GO:0005524 | 12 | 4 |
| GO:0004843 | GO:0008270 | 38 | 5 | GO:0004888 | GO:0005529 | 12 | 4 |
| GO:0005021 | GO:0005524 | 31 | 5 | GO:0004709 | GO:0042803 | 10 | 4 |
| GO:0004645 | GO:0030170 | 29 | 5 | GO:0004872 | GO:0005047 | 9 | 4 |
| GO:0004872 | GO:0019864 | 20 | 5 | GO:0004871 | GO:0005125 | 10 | 4 |
| GO:0004842 | GO:0042803 | 18 | 5 | GO:0004674 | GO:0043621 | 4 | 4 |
| GO:0004674 | GO:0019901 | 13 | 5 | GO:0004842 | GO:0043621 | 4 | 4 |
| GO:0004842 | GO:0031625 | 16 | 5 | GO:0004674 | GO:0046982 | 7 | 4 |
| GO:0005096 | GO:0005525 | 9 | 5 | GO:0005009 | GO:0042169 | 6 | 4 |
| GO:0004871 | GO:0042803 | 15 | 5 | GO:0005009 | GO:0043548 | 6 | 4 |
| GO:0016563 | GO:0019904 | 23 | 5 | GO:0005009 | GO:0051425 | 6 | 4 |
| GO:0016566 | GO:0042803 | 20 | 5 | GO:0005096 | GO:0042802 | 4 | 4 |
| GO:0004972 | GO:0005234 | 11 | 5 | GO:0004888 | GO:0042803 | 3 | 4 |
| GO:0004715 | GO:0042169 | 8 | 5 | GO:0005006 | GO:0005524 | 4 | 4 |
| GO:0004871 | GO:0005070 | 6 | 5 | GO:0004639 | GO:0042802 | 4 | 4 |
| GO:0004715 | GO:0005131 | 5 | 5 | GO:0004714 | GO:0046982 | 4 | 4 |
| GO:0016566 | GO:0019901 | 9 | 5 | GO:0004674 | GO:0047485 | 3 | 4 |
| GO:0004715 | GO:0030145 | 9 | 5 | GO:0000978 | GO:0003704 | 34 | 4 |
| GO:0004972 | GO:0005262 | 3 | 5 | GO:0004872 | GO:0005161 | 2 | 4 |
| GO:0017124 | GO:0017128 | 3 | 5 | GO:0016564 | GO:0019901 | 2 | 4 |
| GO:0004748 | GO:0042802 | 3 | 5 | GO:0004716 | GO:0019903 | 2 | 4 |
| GO:0016564 | GO:0031625 | 6 | 5 | GO:0005006 | GO:0019903 | 2 | 4 |
| GO:0016564 | GO:0042803 | 26 | 5 | GO:0030617 | GO:0031625 | 2 | 4 |
| GO:0004821 | GO:0005524 | 934 | 4 | GO:0030617 | GO:0034713 | 2 | 4 |
| GO:0004820 | GO:0005524 | 859 | 4 | GO:0004617 | GO:0042802 | 2 | 4 |
| GO:0004618 | GO:0005524 | 719 | 4 | GO:0004643 | GO:0042802 | 2 | 4 |
| GO:0004831 | GO:0005524 | 544 | 4 | GO:0004716 | GO:0042802 | 2 | 4 |
| GO:0004825 | GO:0005524 | 496 | 4 | GO:0004849 | GO:0042802 | 2 | 4 |
| GO:0004812 | GO:0005524 | 358 | 4 | GO:0004867 | GO:0042802 | 2 | 4 |
| GO:0004872 | GO:0005529 | 108 | 4 | GO:0005006 | GO:0042802 | 2 | 4 |
| GO:0004830 | GO:0005524 | 166 | 4 | GO:0004704 | GO:0042803 | 2 | 4 |
| GO:0004890 | GO:0005230 | 63 | 4 | GO:0004872 | GO:0043184 | 2 | 4 |
| GO:0004735 | GO:0005488 | 56 | 4 | GO:0004704 | GO:0046982 | 2 | 4 |
| GO:0004672 | GO:0004872 | 42 | 4 | GO:0005006 | GO:0046982 | 2 | 4 |
| GO:0004523 | GO:0005524 | 45 | 4 | GO:0004871 | GO:0070411 | 4 | 4 |
| GO:0004638 | GO:0005524 | 45 | 4 | GO:0030617 | GO:0070411 | 2 | 4 |
| GO:0004784 | GO:0008270 | 18 | 4 | GO:0004813 | GO:0046872 | 705 | 4 |
| GO:0004703 | GO:0005524 | 30 | 4 | GO:0004798 | GO:0005524 | 609 | 3 |
| GO:0004703 | GO:0004871 | 30 | 4 | GO:0004642 | GO:0005524 | 522 | 3 |
| GO:0005089 | GO:0046872 | 23 | 4 | GO:0004527 | GO:0046872 | 133 | 3 |
| GO:0004672 | GO:0005525 | 29 | 4 | GO:0004797 | GO:0005524 | 245 | 3 |
| GO:0004697 | GO:0008270 | 29 | 4 | GO:0004601 | GO:0020037 | 187 | 3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GO:0005096 | GO:0046872 | 31 | 4 | GO:0004819 | GO:0005524 | 153 | 3 |
| GO:0000978 | GO:0003700 | 21 | 4 | GO:0004591 | GO:0030976 | 109 | 3 |
| GO:0004617 | GO:0051287 | 27 | 4 | GO:0004791 | GO:0050660 | 84 | 3 |
| GO:0016740 | GO:0030170 | 87 | 3 | GO:0019899 | GO:0030374 | 4 | 3 |
| GO:0004616 | GO:0050661 | 57 | 3 | GO:0015421 | GO:0042605 | 3 | 3 |
| GO:0004674 | GO:0005516 | 53 | 3 | GO:0005083 | GO:0042802 | 4 | 3 |
| GO:0004887 | GO:0008270 | 35 | 3 | GO:0004724 | GO:0046872 | 4 | 3 |
| GO:0016829 | GO:0030170 | 55 | 3 | GO:0004724 | GO:0048037 | 4 | 3 |
| GO:0017147 | GO:0042813 | 11 | 3 | GO:0005100 | GO:0017048 | 4 | 3 |
| GO:0030165 | GO:0042813 | 3 | 3 | GO:0004888 | GO:0019901 | 3 | 3 |
| GO:0016538 | GO:0019901 | 6 | 3 | GO:0004714 | GO:0033612 | 3 | 3 |
| GO:0004517 | GO:0005516 | 31 | 3 | GO:0004638 | GO:0042802 | 3 | 3 |
| GO:0004930 | GO:0017147 | 19 | 3 | GO:0005034 | GO:0043424 | 3 | 3 |
| GO:0004867 | GO:0008201 | 26 | 3 | GO:0004842 | GO:0046625 | 3 | 3 |
| GO:0004930 | GO:0030165 | 6 | 3 | GO:0004871 | GO:0050681 | 2 | 3 |
| GO:0004517 | GO:0020037 | 25 | 3 | GO:0005010 | GO:0005158 | 2 | 3 |
| GO:0004517 | GO:0010181 | 23 | 3 | GO:0004674 | GO:0008307 | 3 | 3 |
| GO:0004776 | GO:0005525 | 28 | 3 | GO:0004713 | GO:0008307 | 3 | 3 |
| GO:0004725 | GO:0046872 | 15 | 3 | GO:0016564 | GO:0019904 | 4 | 3 |
| GO:0004517 | GO:0050661 | 21 | 3 | GO:0015433 | GO:0042605 | 2 | 3 |
| GO:0004517 | GO:0050660 | 20 | 3 | GO:0004818 | GO:0005524 | 886 | 2 |
| GO:0004494 | GO:0031419 | 20 | 3 | GO:0004828 | GO:0005524 | 844 | 2 |
| GO:0004494 | GO:0046872 | 20 | 3 | GO:0004827 | GO:0005524 | 833 | 2 |
| GO:0005089 | GO:0005524 | 19 | 3 | GO:0004817 | GO:0005524 | 813 | 2 |
| GO:0005096 | GO:0017124 | 20 | 3 | GO:0004829 | GO:0005524 | 698 | 2 |
| GO:0016301 | GO:0042803 | 17 | 3 | GO:0004594 | GO:0005524 | 640 | 2 |
| GO:0005057 | GO:0005524 | 16 | 3 | GO:0004826 | GO:0046872 | 635 | 2 |
| GO:0000739 | GO:0003700 | 16 | 3 | GO:0004822 | GO:0005524 | 579 | 2 |
| GO:0004674 | GO:0005089 | 14 | 3 | GO:0004592 | GO:0005524 | 563 | 2 |
| GO:0004871 | GO:0008083 | 13 | 3 | GO:0004641 | GO:0005524 | 486 | 2 |
| GO:0005125 | GO:0042056 | 13 | 3 | GO:0004781 | GO:0005524 | 450 | 2 |
| GO:0004872 | GO:0005178 | 14 | 3 | GO:0004834 | GO:0030170 | 455 | 2 |
| GO:0016566 | GO:0046982 | 16 | 3 | GO:0004825 | GO:0046872 | 431 | 2 |
| GO:0005096 | GO:0042803 | 6 | 3 | GO:0004784 | GO:0046872 | 378 | 2 |
| GO:0004706 | GO:0005524 | 8 | 3 | GO:0004810 | GO:0005524 | 365 | 2 |
| GO:0004871 | GO:0008013 | 9 | 3 | GO:0004636 | GO:0005524 | 346 | 2 |
| GO:0005021 | GO:0019838 | 12 | 3 | GO:0004832 | GO:0005524 | 339 | 2 |
| GO:0004872 | GO:0019901 | 8 | 3 | GO:0004789 | GO:0046872 | 334 | 2 |
| GO:0004872 | GO:0005518 | 8 | 3 | GO:0004721 | GO:0046872 | 229 | 2 |
| GO:0004871 | GO:0008009 | 8 | 3 | GO:0016706 | GO:0031418 | 9 | 2 |
| GO:0004675 | GO:0005524 | 4 | 3 | GO:0004721 | GO:0005524 | 149 | 2 |
| GO:0005010 | GO:0043548 | 6 | 3 | GO:0016811 | GO:0070403 | 139 | 2 |
| GO:0004674 | GO:0005262 | 5 | 3 | GO:0004637 | GO:0005524 | 114 | 2 |
| GO:0005034 | GO:0005524 | 5 | 3 | GO:0004637 | GO:0046872 | 110 | 2 |
| GO:0004972 | GO:0016594 | 5 | 3 | GO:0004652 | GO:0005524 | 71 | 2 |
| GO:0004550 | GO:0042802 | 3 | 3 | GO:0004635 | GO:0005524 | 65 | 2 |
| GO:0004871 | GO:0005201 | 5 | 3 | **GO:0004525** | **GO:0046872** | **53** | **2** |
| GO:0004871 | GO:0008022 | 5 | 3 | GO:0004794 | GO:0030170 | 52 | 2 |
| GO:0004842 | GO:0019899 | 2 | 3 | GO:0004887 | GO:0043565 | 30 | 2 |
| GO:0005001 | GO:0019901 | 4 | 3 | GO:0004587 | GO:0030170 | 50 | 2 |
| GO:0004722 | GO:0030165 | 4 | 3 | GO:0004655 | GO:0046872 | 45 | 2 |
| GO:0004725 | GO:0030165 | 4 | 3 | GO:0004519 | GO:0051539 | 43 | 2 |
| GO:0016314 | GO:0030165 | 4 | 3 | GO:0004872 | GO:0042803 | 32 | 2 |

# [S-2] GPs containing root shared GO terms I

- GO:0004525: Ribonuclease III activity
- GO:0046872: Metal ion binding

| | |
|---|---|
| **Accession ID** | A0MQH0 (DICER_CRIGR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.- |
| **Gene Names** | Name:DICER1 Synonyms:DICER |
| **Organism** | Cricetulus griseus (Chinese hamster) (Cricetulus barabensis griseus) |
| **Taxonomic identifier** | 10029 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Within the RLC/miRLC, DICER1 and TARBP2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto EIF2C2/AGO2. EIF2C2/AGO2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from DICER1 and TARBP2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | A1C9M6 (DCL2_ASPCL) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2<br>ORF Names: ACLA_055980 |
| **Organism** | Aspergillus clavatus (strain ATCC 1007 / CBS 513.65 / DSM 816 / NCTC 3887 / NRRL 1) [Complete proteome] |
| **Taxonomic identifier** | 344612 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactore** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A1CBC9 (DCL1_ASPCL) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: ACLA_014840 |
| **Organism** | Aspergillus clavatus (strain ATCC 1007 / CBS 513.65 / DSM 816 / NCTC 3887 / NRRL 1) |

| | |
|---|---|
| | [Complete proteome] |
| **Taxonomic identifier** | 344612 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | |
| **Accession ID** | A1D9Z6 (DCL2_NEOFI) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2<br>ORF Names: NFIA_030600 |
| **Organism** | Neosartorya fischeri (strain ATCC 1020 / DSM 3700 / FGSC A1164 / NRRL 181) (Aspergillus fischerianus) [Complete proteome] |
| **Taxonomic identifier** | 331117 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A1DE13 (DCL1_NEOFI) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: NFIA_075500 |
| **Organism** | Neosartorya fischeri (strain ATCC 1020 / DSM 3700 / FGSC A1164 / NRRL 181) (Aspergillus fischerianus) [Complete proteome] |
| **Taxonomic identifier** | 331117 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A2R345 (DCL21_ASPNC) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2-1<br>Including the following 2 domains:<br>Endoribonuclease dcl2-1 |

| | |
|---|---|
| | EC=3.1.26.-<br>ATP-dependent helicase dcl2-1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2-1<br>ORF Names: An14g03000 |
| **Organism** | Aspergillus niger (strain CBS 513.88 / FGSC A1513) [Complete proteome] |
| **Taxonomic identifier** | 425011 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A2RAF3 (DCL1_ASPNC) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: An18g02950 |
| **Organism** | Aspergillus niger (strain CBS 513.88 / FGSC A1513) [Complete proteome] |
| **Taxonomic identifier** | 425011 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A4RHU9 (DCL2_MAGO7) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease DCL2<br>EC=3.1.26.-<br>ATP-dependent helicase DCL2<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL2<br>Synonyms: MDL2<br>ORF Names: MGG_12357 |
| **Organism** | Magnaporthe oryzae (strain 70-15 / ATCC MYA-4617 / FGSC 8958) (Rice blast fungus) (Pyricularia oryzae) [Complete proteome] |
| **Taxonomic identifier** | 242507 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |

| | |
|---|---|
| **Accession ID** | A4RKC3 (DCL1_MAGO7) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease DCL1<br>EC=3.1.26.-<br>ATP-dependent helicase DCL1<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL1<br>Synonyms: MDL1<br>ORF Names: MGG_01541 |
| **Organism** | Magnaporthe oryzae (strain 70-15 / ATCC MYA-4617 / FGSC 8958) (Rice blast fungus) (Pyricularia oryzae) [Complete proteome] |
| **Taxonomic identifier** | 242507 [NCBI] |
| **Function** | Dicer-like endonuclease which seems not to be involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway, contrary to its DCL2 counterpart. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | A7LFZ6 (DCL4_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 4<br>EC=3.1.26.-<br>Alternative name(s):<br>Dicer-like protein 4<br>Short name=OsDCL4<br>Protein SHOOT ORGANIZATION 1 |
| **Gene Names** | Name: DCL4<br>Synonyms: SHO1<br>Ordered Locus Names: Os04g0509300, LOC_Os04g43050<br>ORF Names: OSJNBb0065L13.5 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Involved in the RNA silencing pathway. Cleaves double-stranded RNA to produce small interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs. Required for the production of 21 nucleotide siRNAs. Regulates shoot apical meristem (SAM) initiation and maintenance, leaf polarization and lemma polarity through the trans-acting siRNAS (ta-siRNAs) pathway, which probably modulate the expression of the ARF2, ARF3, ARF4, ARF14 and ARF15 genes. Can process endogenous 21 nucleotide siRNAs derived from an imperfect inverted repeat. May not be involved in microRNAs (miRNAs) production. |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Subcellular location** | Nucleus |
| **Accession ID** | A8BQJ3 (DCL_GIAIC) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer-like<br>EC=3.1.26.- |
| **Gene Names** | ORF Names: GL50803_103887 |
| **Organism** | Giardia intestinalis (strain ATCC 50803 / WB clone C6) (Giardia lamblia) |
| **Taxonomic identifier** | 184922 [NCBI] |
| **Function** | Involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs. |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |

| | |
|---|---|
| **Subunit structure** | Subunit structure |
| **Accession ID** | B3DLA6 (DICER_XENTR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.- |
| **Gene Names** | Name: dicer1 |
| **Organism** | Xenopus tropicalis (Western clawed frog) (Silurana tropicalis) |
| **Taxonomic identifier** | 8364 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of dicer1, eif2c2/ago2 and tarbp2. Within the RLC/miRLC, dicer1 and tarbp2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto eif2c2/ago2. eif2c2/ago2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from dicer1 and tarbp2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of dicer1, eif2c2/ago2 and tarbp2. Note that the trimeric RLC/miRLC is also referred to as RISC |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | O01326 (RNC_CAEEL) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3<br>EC=3.1.26.3<br>Alternative name(s):<br>Protein drosha<br>Ribonuclease III<br>Short name=RNase III |
| **Gene Names** | Name: drsh-1<br>ORF Names: F26E4.10 |
| **Organism** | Caenorhabditis elegans |
| **Taxonomic identifier** | 6239 [NCBI] |
| **Function** | Executes the initial step of microRNA (miRNA) processing in the nucleus, that is the cleavage of pri-miRNA to release pre-miRNA. Involved in pre-rRNA processing. Cleaves double-strand RNA and does not cleave single-strand RNA. Involved in fertility. Required for the function or synthesis of the let-7 miRNA. |
| **Catalytic activity** | Endonucleolytic cleavage to 5'-phosphomonoester. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | Nucleus |
| **Accession ID** | P0C5H7 (DCL2_EMENI) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2<br>ORF Names: AN10380 |
| **Organism** | Emericella nidulans (strain FGSC A4 / ATCC 38163 / CBS 112.46 / NRRL 194 / M139) (Aspergillus nidulans) |
| **Taxonomic identifier** | 227321 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference |

| | |
|---|---|
| | (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | P22192 (PAC1_SCHPO) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Double-strand-specific pac1 ribonuclease<br>EC=3.1.26.3<br>Alternative name(s):<br>Protein hcs |
| **Gene Names** | Name: pac1<br>Synonyms: hcs<br>ORF Names: SPBC119.11c |
| **Organism** | Schizosaccharomyces pombe (strain 972 / ATCC 24843) (Fission yeast) |
| **Taxonomic identifier** | 284812 [NCBI] |
| **Function** | Digests double-stranded RNA. Converts long double-stranded RNAs into short oligonucleotides, leaving 5'-phosphates on their cleavage products. Probably inhibits mating and meiosis by degrading a specific mRNA required for sexual development. |
| **Catalytic activity** | Endonucleolytic cleavage to 5'-phosphomonoester. |
| **Cofactor** | Magnesium |
| **Subcellular location** | N/A |
| **Accession ID** | P34529 (DCR1_CAEEL) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease dcr-1<br>EC=3.1.26.- |
| **Gene Names** | Name: dcr-1<br>ORF Names: K12H4.8 |
| **Organism** | Caenorhabditis elegans |
| **Taxonomic identifier** | 6239 [NCBI] |
| **Function** | Involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs. Seems to process the precursor of the small temporal RNA let-7 which is involved in developmental timing. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | |
| **Accession ID** | P84634 (DCL4_ARATH) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 4<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL4<br>Ordered Locus Names: At5g20320<br>ORF Names: F5O24.210 |
| **Organism** | Arabidopsis thaliana (Mouse-ear cress) |
| **Taxonomic identifier** | 3702 [NCBI] |
| **Function** | Ribonuclease (RNase) III involved in RNA-mediated post-transcriptional gene silencing (PTGS). Functions in the biogenesis of trans-acting small interfering RNAs (ta-siRNAs, derived from the TAS1, TAS2 or TAS3 endogenous transcripts) by cleaving small dsRNAs into 21-24 nucleotide ta-siRNAs. Functions with the dsRNA-binding protein DRB4 in ta-siRNAs processing. Acts in the RDR6/SGS3/DCL4/AGO7 ta-siRNA pathway involved in leaf developmental timing. Plays a role in transitive silencing of transgenes by processing secondary siRNAs. This pathway, which requires DCL2 and RDR6, amplifies silencing by using the target RNA as substrate to generate secondary siRNAs, providing an efficient |

| | |
|---|---|
| | mechanism for long-distance silencing. Required for the production of the 30-40 nucleotide bacterial-induced long siRNAs (lsiRNA). May participate with DCL3 in the production of 24 nucleotide repeat-associated siRNAs (ra-siRNAs) which derive from heterochromatin and DNA repeats such as transposons. Plays an important role in antiviral RNA silencing. Involved in the production of viral siRNAs derived from the cucumber mosaic virus (CMV), turnip crinkle virus (TCV) and tobacco rattle virus (TRV). Targeted by the viral silencing suppressor (VSR) protein 2b of the cucumber mosaic virus (CMV) that inactivates DCL4 function in RNA silencing. Does not seem to be involved in microRNAs (miRNAs) processing. |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | Interacts with DRB4. |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q09884 (DCR1_SCHPO) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Protein Dicer<br>Alternative name(s):<br>Cell cycle control protein dcr1<br>RNA interference pathway protein dcr1<br>Including the following 2 domains:<br>Endoribonuclease dcr1<br>EC=3.1.26.-<br>ATP-dependent helicase dcr1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcr1<br>ORF Names: SPCC188.13c, SPCC584.10c |
| **Organism** | Schizosaccharomyces pombe (strain 972 / ATCC 24843) (Fission yeast) |
| **Taxonomic identifier** | 284812 [NCBI] |
| **Function** | Required for G1 arrest and mating in response to nitrogen starvation. Ago1 regulation of cytokinesis and cell cycle checkpoints occurs downstream of dcr1. Required, indirectly, for regulated hyperphosphorylation of cdc2.<br>Has a role in the RNA interference (RNAi) pathway which is important for heterochromatin formation, accurate chromosome segregation, centromere cohesion and telomere function during mitosis and meiosis. Digests double-stranded RNA (dsRNA) producing 21 to 23 bp dsRNAs, so-called interfering RNAs (siRNA). Required for both post-transcriptional and transcriptional gene silencing. Required for silencing at the centromeres and for initiation of transcriptionally silent heterochromatin at the mating type locus. Promotes histone H3 'Lys-10' methylation necessary for centromere function. Required for recruitment of swi6 and cohesin to an ectopic dg repeat. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | Cytoplasm. Nucleus |
| **Accession ID** | Q0CEI2 (DCL2_ASPTN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2<br>ORF Names: ATEG_07902 |
| **Organism** | Aspergillus terreus (strain NIH 2624 / FGSC A1156) [Complete proteome] |
| **Taxonomic identifier** | 341663 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective |

| | |
|---|---|
| | destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q0CW42 (DCL1_ASPTN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: ATEG_02092 |
| **Organism** | Aspergillus terreus (strain NIH 2624 / FGSC A1156) [Complete proteome] |
| **Taxonomic identifier** | 341663 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q0UI93 (DCL1_PHANO) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease DCL1<br>EC=3.1.26.-<br>ATP-dependent helicase DCL1<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL1<br>ORF Names: SNOG_08521 |
| **Organism** | Phaeosphaeria nodorum (strain SN15 / ATCC MYA-4574 / FGSC 10173) (Glume blotch fungus) (Septoria nodorum) |
| **Taxonomic identifier** | 321614 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q0UL22 (DCL2_PHANO) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease DCL2<br>EC=3.1.26.-<br>ATP-dependent helicase DCL2<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL2<br>ORF Names: SNOG_07542 |

| | |
|---|---|
| **Organism** | Phaeosphaeria nodorum (strain SN15 / ATCC MYA-4574 / FGSC 10173) (Glume blotch fungus) (Septoria nodorum) |
| **Taxonomic identifier** | 321614 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q10HL3 (DCL2A_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 2a<br>Alternative name(s):<br>Dicer-like protein 2a<br>Short name=OsDCL2a<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL2A<br>Ordered Locus Names: Os03g0583900, LOC_Os03g38740<br>ORF Names: OJ1785_A05.30 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Probably involved in the RNA silencing pathway. May cleave double-stranded RNA to produce short 21-24 nucleotides (nt) RNAs which target the selective destruction of complementary RNAs |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q1DKI1 (DCL1_COCIM) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease DCL1<br>EC=3.1.26.-<br>ATP-dependent helicase DCL1<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL1<br>ORF Names: CIMG_09182 |
| **Organism** | Coccidioides immitis (strain RS) (Valley fever fungus) |
| **Taxonomic identifier** | 246410 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q1DW80 (DCL2_COCIM) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease DCL2<br>EC=3.1.26.-<br>ATP-dependent helicase DCL2 |

| | |
|---|---|
| | EC=3.6.4.- |
| **Gene Names** | Name: DCL2<br>ORF Names: CIMG_05433 |
| **Organism** | Coccidioides immitis (strain RS) (Valley fever fungus) |
| **Taxonomic identifier** | 246410 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q25BN1 (DICER_CHICK) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.- |
| **Gene Names** | Name: DICER1<br>Synonyms: DICER |
| **Organism** | Gallus gallus (Chicken) |
| **Taxonomic identifier** | 9031 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Within the RLC/miRLC, DICER1 and TARBP2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto EIF2C2/AGO2. EIF2C2/AGO2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from DICER1 and TARBP2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Note that the trimeric RLC/miRLC is also referred to as RISC |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q2H0G2 (DCL1_CHAGB) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease DCL1<br>EC=3.1.26.-<br>ATP-dependent helicase DCL1<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL1<br>ORF Names: CHGG_04734 |
| **Organism** | Chaetomium globosum (strain ATCC 6205 / CBS 148.51 / DSM 1962 / NBRC 6347 / NRRL 1970) (Soil fungus) [Complete proteome] |
| **Taxonomic identifier** | 306901 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q2U6C4 (DCL1_ASPOR) Reviewed, UniProtKB/Swiss-Prot |

| | |
|---|---|
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: AO090120000297 |
| **Organism** | Aspergillus oryzae (strain ATCC 42149 / RIB 40) (Yellow koji mold) [Complete proteome] |
| **Taxonomic identifier** | 510516 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q2UNX5 (DCL2_ASPOR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl2<br>ORF Names: AO090001000193 |
| **Organism** | Aspergillus oryzae (strain ATCC 42149 / RIB 40) (Yellow koji mold) [Complete proteome] |
| **Taxonomic identifier** | 510516 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q2VF18 (DCL2_CRYPA) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease DCL-2<br>EC=3.1.26.-<br>ATP-dependent helicase DCL-2<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL-2 |
| **Organism** | Cryphonectria parasitica (Chesnut blight fungus) (Endothia parasitica) |
| **Taxonomic identifier** | 5116 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense, DCL-2 is involved in antiviral defense against mycoviruses like the hypovirus CHV1-EP713 and the reovirus MyRV1-Cp9B21. |

| | |
|---|---|
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q2VF19 (DCL1_CRYPA) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease DCL-1<br>EC=3.1.26.-<br>ATP-dependent helicase DCL-1<br>EC=3.6.4.- |
| **Gene Names** | Name: DCL-1 |
| **Organism** | Cryphonectria parasitica (Chesnut blight fungus) (Endothia parasitica) |
| **Taxonomic identifier** | 5116 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q3EBC8 (DCL2_ARATH) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 2<br>EC=3.1.26.-<br>Alternative name(s):<br>Dicer-like protein 2<br>Short name=AtDCL2 |
| **Gene Names** | Ordered Locus Names: At3g03300<br>ORF Names: T17B22.1 |
| **Organism** | Arabidopsis thaliana (Mouse-ear cress) |
| **Taxonomic identifier** | 3702 [NCBI] |
| **Function** | Ribonuclease (RNase) III involved in RNA-mediated post-transcriptional gene silencing (PTGS). Involved in the processing of natural small interfering RNAs (nat-siRNAs, derived from cis-natural antisense transcripts) by cleaving small dsRNAs into 24 nucleotide nat-siRNAs. Plays an essential role in transitive silencing of transgenes by processing secondary siRNAs. This pathway, which requires DCL4 and RDR6, amplifies silencing by using the target RNA as substrate to generate secondary siRNAs, providing an efficient mechanism for long-distance silencing. May participate with DCL3 in the production of 24 nucleotide repeat-associated siRNAs (ra-siRNAs) which derive from heterochromatin and DNA repeats such as transposons. Plays a role in antiviral RNA silencing. Involved in the production of viral siRNAs derived from the turnip crinkle virus (TCV) and tobacco rattle virus (TRV). Targeted by the viral silencing suppressor (VSR) protein 2b of the cucumber mosaic virus (CMV) that inactivates DCL2 function in RNA silencing. Does not seem to be involved in microRNAs (miRNAs) processing. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | Nucleus. Cytoplasm |
| **Accession ID** | Q4WA22 (DCL2_ASPFU) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl2<br>EC=3.6.4.- |

| | |
|---|---|
| **Gene Names** | Name: dcl2<br>ORF Names: AFUA_4G02930 |
| **Organism** | Neosartorya fumigata (strain ATCC MYA-4609 / Af293 / CBS 101355 / FGSC A1100) (Aspergillus fumigatus) |
| **Taxonomic identifier** | 330879 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q4WVE3 (DCL1_ASPFU) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl1<br>ORF Names: AFUA_5G11790 |
| **Organism** | Neosartorya fumigata (strain ATCC MYA-4609 / Af293 / CBS 101355 / FGSC A1100) (Aspergillus fumigatus) |
| **Taxonomic identifier** | 330879 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q5HZJ0 (RNC_MOUSE) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3<br>EC=3.1.26.3<br>Alternative name(s):<br>Protein Drosha<br>Ribonuclease III<br>Short name=RNase III |
| **Gene Names** | Name: Drosha<br>Synonyms: Etohi2, Rn3, Rnasen |
| **Organism** | Mus musculus (Mouse) |
| **Taxonomic identifier** | 10090 [NCBI] |
| **Function** | Ribonuclease III double-stranded (ds) RNA-specific endoribonuclease that is involved in the initial step of microRNA (miRNA) biogenesis. Component of the microprocessor complex that is required to process primary miRNA transcripts (pri-miRNAs) to release precursor miRNA (pre-miRNA) in the nucleus. Within the microprocessor complex, DROSHA cleaves the 3' and 5' strands of a stem-loop in pri-miRNAs (processing center 11 bp from the dsRNA-ssRNA junction) to release hairpin-shaped pre-miRNAs that are subsequently cut by the cytoplasmic DICER to generate mature miRNAs. Involved also in pre-rRNA processing. Cleaves double-strand RNA and does not cleave single-strand RNA. Involved in the formation of GW bodies |
| **Catalytic activity** | Endonucleolytic cleavage to 5'-phosphomonoester. |

| | |
|---|---|
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | Component of the microprocessor complex, or pri-miRNA processing protein complex, which is composed of DROSHA and DGCR8. The microprocessor complex may contain multiple subunit of DGCR8 and DROSHA. Interacts with DGCR8, SP1 and SNIP1<br>Interacts with SRRT/ARS2 |
| **Subcellular location** | Nucleus By similarity.<br>Nucleus › nucleolus By similarity.<br>Note: A fraction is translocated to the nucleolus during the S phase of the cell cycle.<br>Localized in GW bodies (GWBs), also known as P-bodies |
| **Accession ID** | Q5JK90 (RTL1_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3-like protein 1<br>EC=3.1.26.-<br>Alternative name(s):<br>Ribonuclease III-like protein 1<br>Short name=RNase III-like protein 1 |
| **Gene Names** | Ordered Locus Names: Os01g0551100, LOC_Os01g37050<br>ORF Names: OSJNBa0024F24.33, OSJNBa0066C06.2 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Cleaves double-stranded RNA (dsRNA) |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q5N870 (DCL3A_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 3a<br>Alternative name(s):<br>Dicer-like protein 3a<br>Short name=OsDCL3a<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL3A<br>Ordered Locus Names: Os01g0909200, LOC_Os01g68120<br>ORF Names: P0456E05.21 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Probably involved in the RNA silencing pathway. May cleave double-stranded RNA to produce short 21-24 nucleotides (nt) RNAs which target the selective destruction of complementary RNAs |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q61XX9 (RNC_CAEBR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3<br>EC=3.1.26.3<br>Alternative name(s):<br>Protein drosha<br>Ribonuclease III<br>Short name=RNase III |
| **Gene Names** | Name: drsh-1<br>ORF Names: CBG03798 |
| **Organism** | Caenorhabditis briggsae |
| **Taxonomic identifier** | 6238 [NCBI] |
| **Function** | Executes the initial step of microRNA (miRNA) processing in the nucleus, that is the |

| | |
|---|---|
| | cleavage of pri-miRNA to release pre-miRNA. Involved in pre-rRNA processing. Cleaves double-strand RNA and does not cleave single-strand RNA. Involved in fertility. Required for the function or synthesis of the let-7 miRNA |
| **Catalytic activity** | Endonucleolytic cleavage to 5'-phosphomonoester. |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q69KJ0 (RTL3_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3-like protein 3<br>EC=3.1.26.-<br>Alternative name(s):<br>Ribonuclease III-like protein 3<br>Short name=RNase III-like protein 3 |
| **Gene Names** | Ordered Locus Names: Os06g0358800, LOC_Os06g25250<br>ORF Names: OsJ_21299, OSJNBb0027A16.25 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Cleaves double-stranded RNA (dsRNA) |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q69LX2 (DCL2B_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 2b<br>Alternative name(s):<br>Dicer-like protein 2b<br>Short name=OsDCL2b<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL2B<br>Ordered Locus Names: Os09g0315100, LOC_Os09g14610<br>ORF Names: OSJNBa0057D11.8-1, OSJNBb0079G12.35-1 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Probably involved in the RNA silencing pathway. May cleave double-stranded RNA to produce short 21-24 nucleotides (nt) RNAs which target the selective destruction of complementary RNAs |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q6ATG6 (RTL2_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3-like protein 2<br>EC=3.1.26.-<br>Alternative name(s):<br>Ribonuclease III-like protein 2<br>Short name=RNase III-like protein 2 |
| **Gene Names** | Ordered Locus Names: Os05g0271300, LOC_Os05g18850<br>ORF Names: OJ1653_D06.9, OSJNBb0061M13.3 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Cleaves double-stranded RNA (dsRNA) |
| **Cofactor** | Magnesium or manganese |
| **Subcellular location** | N/A |
| **Accession ID** | Q6TUI4 (DICER_BOVIN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name: |

| | |
|---|---|
| | Endoribonuclease Dicer<br>EC=3.1.26.- |
| **Gene Names** | Name: DICER1<br>Synonyms: DICER |
| **Organism** | Bos taurus (Bovine) |
| **Taxonomic identifier** | 9913 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Within the RLC/miRLC, DICER1 and TARBP2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto EIF2C2/AGO2. EIF2C2/AGO2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from DICER1 and TARBP2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Note that the trimeric RLC/miRLC is also referred to as RISC. Interacts with DHX9, EIF2C1, PIWIL1 and PRKRA. Associates with the 60S ribosome |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q6TV19 (DICER_DANRE) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.- |
| **Gene Names** | Name: dicer1 |
| **Organism** | Danio rerio (Zebrafish) (Brachydanio rerio) |
| **Taxonomic identifier** | 7955 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of dicer1, eif2c2/ago2 and tarbp2. Within the RLC/miRLC, dicer1 and tarbp2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto eif2c2/ago2. eif2c2/ago2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from dicer1 and tarbp2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of dicer1, eif2c2/ago2 and tarbp2. Note that the trimeric RLC/miRLC is also referred to as RISC |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q7S8J7 (DCL1_NEUCR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 1<br>Including the following 2 domains:<br>Endoribonuclease dcl-1<br>EC=3.1.26.-<br>ATP-dependent helicase dcl-1<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl-1<br>ORF Names: NCU08270 |
| **Organism** | Neurospora crassa (strain ATCC 24698 / 74-OR23-1A / CBS 708.71 / DSM 1257 / FGSC 987) |
| **Taxonomic identifier** | 367110 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference |

| | |
|---|---|
| | (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons. Controls the expression of the non-LTR retrotransposon Tad in the African strain, Adiomopoume. |
| **Cofactor** | Magnesium or manganese |
| **Induction** | By double-stranded RNA (dsRNA) |
| **Accession ID** | Q7SCC1 (DCL2_NEUCR) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Dicer-like protein 2<br>Including the following 2 domains:<br>Endoribonuclease dcl-2<br>EC=3.1.26.-<br>ATP-dependent helicase dcl-2<br>EC=3.6.4.- |
| **Gene Names** | Name: dcl-2<br>ORF Names: B14D6.490, NCU06766 |
| **Organism** | Neurospora crassa (strain ATCC 24698 / 74-OR23-1A / CBS 708.71 / DSM 1257 / FGSC 987) |
| **Taxonomic identifier** | 367110 [NCBI] |
| **Function** | Dicer-like endonuclease involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. Produces 21 to 25 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs leading to sequence-specific suppression of gene expression, called post-transcriptional gene silencing (PTGS). Part of a broad host defense response against viral infection and transposons. Controls the expression of the non-LTR retrotransposon Tad in the African strain, Adiomopoume |
| **Cofactor** | Magnesium or manganese |
| **Induction** | By double-stranded RNA (dsRNA). |
| **Accession ID** | Q7XD96 (DCL3B_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 3b<br>Alternative name(s):<br>Dicer-like protein 3b<br>Short name=OsDCL3b<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL3B<br>Ordered Locus Names: Os10g0485600, LOC_Os10g34430<br>ORF Names: OSJNBa0029C15.23 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | Probably involved in the RNA silencing pathway. May cleave double-stranded RNA to produce short 21-24 nucleotides (nt) RNAs which target the selective destruction of complementary RNAs |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q8LMR2 (DCL1_ORYSJ) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 1<br>Alternative name(s):<br>Dicer-like protein 1<br>Short name=OsDCL1<br>EC=3.1.26.- |
| **Gene Names** | Name: DCL1 |

| | |
|---|---|
| | Ordered Locus Names: Os03g0121800, LOC_Os03g02970<br>ORF Names: OJ1705B08.11, OsJ_09217 |
| **Organism** | Oryza sativa subsp. japonica (Rice) [Complete proteome] |
| **Taxonomic identifier** | 39947 [NCBI] |
| **Function** | nvolved in the RNA silencing pathway. Cleaves double-stranded RNA to produce microRNAs (miRNAs) of 21-24 nucleotides which target the selective destruction of complementary RNAs. Regulates by this way the development of the plant. May not be involved in small interfering RNAs (siRNAs) production |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | May interact with ARGONAUTE1 or PINHEAD through their common PAZ domains |
| **Subcellular location** | Nucleus |
| **Accession ID** | Q8R418 (DICER_MOUSE) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.-<br>Alternative name(s):<br>Double-strand-specific ribonuclease mDCR-1 |
| **Gene Names** | Name: Dicer1<br>Synonyms: Dicer, Mdcr |
| **Organism** | Mus musculus (Mouse) |
| **Taxonomic identifier** | 10090 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Within the RLC/miRLC, DICER1 and TARBP2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto EIF2C2/AGO2. EIF2C2/AGO2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from DICER1 and TARBP2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs<br>Involved in cleaving double-stranded RNA in the RNA interference (RNAi) pathway. It produces 21 to 23 bp dsRNAs (siRNAs) which target the selective destruction of homologous RNAs. |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Note that the trimeric RLC/miRLC is also referred to as RISC. Interacts with DHX9, EIF2C1, PIWIL1 and PRKRA. Associates with the 60S ribosome |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q9LTQ0 (RTL2_ARATH) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3-like protein 2<br>EC=3.1.26.-<br>Alternative name(s):<br>Ribonuclease III-like protein 2<br>Short name=RNase III-like protein 2<br>Ribonuclease three-like protein 2 |
| **Gene Names** | Name: RTL2<br>Ordered Locus Names: At3g20420<br>ORF Names: MQC12.21 |
| **Organism** | Arabidopsis thaliana (Mouse-ear cress) |
| **Taxonomic identifier** | 3702 [NCBI] |
| **Function** | Ribonuclease that cleaves double-stranded RNA (dsRNA). Required for 3'-external transcribed spacer (ETS) cleavage of the pre-rRNA precursors. May promote the production of 21 nucleotide small interfering RNA (siRNA) during post-transcriptional gene silencing |

| | |
|---|---|
| | (PTGS). |
| **Cofactor** | Binds magnesium or manganese |
| **Subunit structure** | Homodimer; disulfide-linked |
| **Subcellular location** | Nucleus. Cytoplasm |
| **Accession ID** | Q9LXW7 (DCL3_ARATH) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 3<br>EC=3.1.26.-<br>Alternative name(s):<br>Dicer-like protein 3<br>Short name=AtDCL3 |
| **Gene Names** | Name: DCL3<br>Ordered Locus Names: At3g43920<br>ORF Names: T15B3.60 |
| **Organism** | Arabidopsis thaliana (Mouse-ear cress) |
| **Taxonomic identifier** | 3702 [NCBI] |
| **Function** | Ribonuclease (RNase) III involved in RNA-mediated post-transcriptional gene silencing (PTGS). Involved in the processing of repeat-associated small interfering RNAs (ra-siRNAs, derived from heterochromatin and DNA repeats such as transposons) by cleaving small dsRNAs into 24 nucleotide ra-siRNAs. Plays a role in antiviral RNA silencing. Involved in the production of viral siRNAs derived from the cabbage leaf curl virus (CaLCuV) and tobacco rattle virus (TRV). Targeted by the viral silencing suppressor (VSR) protein 2b of the cucumber mosaic virus (CMV) that inactivates DCL3 function in RNA silencing. Acts redundantly with DICER-LIKE 1 (DCL1) to promote flowering via repression of FLOWERING LOCUS C (FLC). Does not seem to be involved in microRNAs (miRNAs) processing. |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | Interacts with DRB2 and DRB5 |
| **Subcellular location** | Nucleus › nucleolus |
| **Accession ID** | Q9NRR4 (RNC_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribonuclease 3<br>EC=3.1.26.3<br>Alternative name(s):<br>Protein Drosha<br>Ribonuclease III<br>Short name=RNase III<br>p241 |
| **Gene Names** | Name: DROSHA<br>Synonyms: RN3, RNASE3L, RNASEN |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Ribonuclease III double-stranded (ds) RNA-specific endoribonuclease that is involved in the initial step of microRNA (miRNA) biogenesis. Component of the microprocessor complex that is required to process primary miRNA transcripts (pri-miRNAs) to release precursor miRNA (pre-miRNA) in the nucleus. Within the microprocessor complex, DROSHA cleaves the 3' and 5' strands of a stem-loop in pri-miRNAs (processing center 11 bp from the dsRNA-ssRNA junction) to release hairpin-shaped pre-miRNAs that are subsequently cut by the cytoplasmic DICER to generate mature miRNAs. Involved also in pre-rRNA processing. Cleaves double-strand RNA and does not cleave single-strand RNA. Involved in the formation of GW bodies |
| **Catalytic activity** | Endonucleolytic cleavage to 5'-phosphomonoester. |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | Component of the microprocessor complex, or pri-miRNA processing protein complex, |

| | |
|---|---|
| | which is composed of DROSHA and DGCR8. The microprocessor complex may contain multiple subunit of DGCR8 and DROSHA. Interacts with DGCR8, SP1 and SNIP1. Interacts with SRRT/ARS2. |
| **Subcellular location** | Nucleus. Nucleus › nucleolus. Note: A fraction is translocated to the nucleolus during the S phase of the cell cycle. Localized in GW bodies (GWBs), also known as P-bodies. |
| **Accession ID** | Q9SP32 (DCL1_ARATH) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer homolog 1<br>EC=3.1.26.-<br>Alternative name(s):<br>Dicer-like protein 1<br>Short name=AtDCL1<br>Protein ABNORMAL SUSPENSOR 1<br>Protein CARPEL FACTORY<br>Protein SHORT INTEGUMENTS 1<br>Protein SUSPENSOR 1 |
| **Gene Names** | Name: DCL1<br>Synonyms: ASU1, CAF SIN1, SUS1<br>Ordered Locus Names: At1g01040<br>ORF Names: T25K16.4 |
| **Organism** | Arabidopsis thaliana (Mouse-ear cress) |
| **Taxonomic identifier** | 3702 [NCBI] |
| **Function** | Ribonuclease (RNase) III involved in RNA-mediated post-transcriptional gene silencing (PTGS). Functions in the microRNAs (miRNAs) biogenesis pathway by cleaving primary miRNAs (pri-miRNAs) and precursor miRNAs (pre-miRNAs). Functions with DRB1/HYL1 and SERRATE proteins for accurate pri-miRNAs to miRNAs processing. Indirectly involved in the production of trans-acting small interfering RNAs (ta-siRNAs) derived from the TAS1, TAS2 or TAS3 endogenous transcripts by participating in the production of their initiating miRNAs. Involved in the processing of natural siRNAs (nat-siRNAs, derived from cis-natural antisense transcripts) by cleaving 24 nucleotide nat-siRNAs into 21 nucleotide nat-siRNAs. Can produce RDR6-dependent endogenous ta-siRNAs derived from TAS1 and TAS2. Required for the production of 30-40 nucleotide bacterial-induced long siRNAs (lsiRNA). Acts redundantly with DICER-LIKE 3 (DCL3) to promote flowering via repression of FLOWERING LOCUS C (FLC). Represses antiviral RNA silencing through negative regulation of the expression of DCL4 and DCL3. |
| **Cofactor** | Magnesium or manganese |
| **Subunit structure** | Interacts (via N-terminus) with DDL. Interacts (via DRBM domains) with DRB1, DRB2 and DRB5. May interact with AGO1 or AGO10 through their common PAZ domains |
| **Subcellular location** | Nucleus. Note: Localizes to nuclear dicing body (also named D body), a nuclear body distributed throughout the nucleoplasm and involved in miRNA processing. |
| **Accession ID** | Q9UPY3 (DICER_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Endoribonuclease Dicer<br>EC=3.1.26.-<br>Alternative name(s):<br>Helicase with RNase motif<br>Short name=Helicase MOI |
| **Gene Names** | Name: DICER1<br>Synonyms: DICER, HERNA, KIAA0928 |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Required for formation of the RNA induced silencing complex (RISC). Component of the RISC loading complex (RLC), also known as the micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Within the |

| | |
|---|---|
| | RLC/miRLC, DICER1 and TARBP2 are required to process precursor miRNAs (pre-miRNAs) to mature miRNAs and then load them onto EIF2C2/AGO2. EIF2C2/AGO2 bound to the mature miRNA constitutes the minimal RISC and may subsequently dissociate from DICER1 and TARBP2. Also cleaves double-stranded RNA to produce short interfering RNAs (siRNAs) which target the selective destruction of complementary RNAs |
| **Cofactor** | Binds 2 magnesium or manganese ions per subunit |
| **Subunit structure** | Component of the RISC loading complex (RLC), or micro-RNA (miRNA) loading complex (miRLC), which is composed of DICER1, EIF2C2/AGO2 and TARBP2. Note that the trimeric RLC/miRLC is also referred to as RISC. Interacts with DHX9, EIF2C1, PIWIL1 and PRKRA. Associates with the 60S ribosome. |
| **Subcellular location** | Cytoplasm |

**PPI pairs having root-shared GO terms (i.e. GO:0004525 & GO:0046872)**

| PPI1 | PPI2 | Source DB | Taxonomy1 | Taxonomy2 | type | method |
|---|---|---|---|---|---|---|
| Q3EBC8 | Q9LXW7 | IntAct | 3702 | 3702 | direct interaction | far western blotting |
| Q3EBC8 | Q9SP32 | IntAct | 3702 | 3702 | direct interaction | far western blotting |

# [S-3] GPs containing root shared GO terms II

- GO:0004749: Ribose phosphate diphosphokinase activity
- GO:0042802: Identical protein binding
- GO:0000287: Magnesium ion binding
- Root shared except GO:0042802 – GO:0000287 which are shared by GO:0005488 (binding – 2nd direct child of root term, GO:0003674)

| | |
|---|---|
| **Accession ID** | P32895 (KPR1_YEAST) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribose-phosphate pyrophosphokinase 1<br>EC=2.7.6.1<br>Alternative name(s):<br>Phosphoribosyl pyrophosphate synthase 1 |
| **Gene Names** | Name: PRS1<br>Synonyms: PPS1, PRP1, PRPS, PRPS1<br>Ordered Locus Names: YKL181W |
| **Organism** | Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast) |
| **Taxonomic identifier** | 559292 [NCBI] |
| **Function** | 5-phosphoribose 1-diphosphate synthase involved in nucleotide, histidine, and tryptophan biosynthesis. Active in heteromultimeric complexes with other 5-phosphoribose 1-diphosphate synthases (PRS2, PRS3, PRS4 and PRS5). |
| **Catalytic activity** | ATP + D-ribose 5-phosphate = AMP + 5-phospho-alpha-D-ribose 1-diphosphate. |
| **Pathway** | Metabolic intermediate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate from D-ribose 5-phosphate (route I): step 1/1. |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | P38620 (KPR2_YEAST) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribose-phosphate pyrophosphokinase 2<br>EC=2.7.6.1<br>Alternative name(s):<br>Phosphoribosyl pyrophosphate synthase 2 |
| **Gene Names** | Name: PRS2<br>Synonyms: PRPS2, PRS<br>Ordered Locus Names: YER099C |
| **Organism** | Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast) |
| **Taxonomic identifier** | 559292 [NCBI] |
| **Function** | 5-phosphoribose 1-diphosphate synthase involved in nucleotide, histidine, and tryptophan biosynthesis. Active in heteromultimeric complexes with other 5-phosphoribose 1-diphosphate synthases (PRS2, PRS3, PRS4 and PRS5). |
| **Catalytic activity** | ATP + D-ribose 5-phosphate = AMP + 5-phospho-alpha-D-ribose 1-diphosphate. |
| **Pathway** | Metabolic intermediate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate from D-ribose 5-phosphate (route I): step 1/1. |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | P38689 (KPR3_YEAST) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribose-phosphate pyrophosphokinase 3<br>EC=2.7.6.1<br>Alternative name(s):<br>Phosphoribosyl pyrophosphate synthase 3 |
| **Gene Names** | Name: PRS3<br>Synonyms: PRPS3<br>Ordered Locus Names: YHL011C |
| **Organism** | Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast) |

| | |
|---|---|
| **Taxonomic identifier** | 559292 [NCBI] |
| **Function** | 5-phosphoribose 1-diphosphate synthase involved in nucleotide, histidine, and tryptophan biosynthesis. Active in heteromultimeric complexes with other 5-phosphoribose 1-diphosphate synthases (PRS2, PRS3, PRS4 and PRS5). |
| **Catalytic activity** | ATP + D-ribose 5-phosphate = AMP + 5-phospho-alpha-D-ribose 1-diphosphate. |
| **Pathway** | Metabolic intermediate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate from D-ribose 5-phosphate (route I): step 1/1. |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q12265 (KPR5_YEAST) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Ribose-phosphate pyrophosphokinase 5<br>EC=2.7.6.1<br>Alternative name(s):<br>Phosphoribosyl pyrophosphate synthase 5 |
| **Gene Names** | Name: PRS5<br>Synonyms: PRPS5<br>Ordered Locus Names: YOL061W<br>ORF Names: O1213 |
| **Organism** | Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast) |
| **Taxonomic identifier** | 559292 [NCBI] |
| **Function** | 5-phosphoribose 1-diphosphate synthase involved in nucleotide, histidine, and tryptophan biosynthesis. Active in heteromultimeric complexes with other 5-phosphoribose 1-diphosphate synthases (PRS2, PRS3, PRS4 and PRS5). |
| **Catalytic activity** | ATP + D-ribose 5-phosphate = AMP + 5-phospho-alpha-D-ribose 1-diphosphate. |
| **Pathway** | Metabolic intermediate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate biosynthesis; 5-phospho-alpha-D-ribose 1-diphosphate from D-ribose 5-phosphate (route I): step 1/1. |
| **Subcellular location** | Cytoplasm |
| **Accession ID** | Q14558 (KPRA_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Phosphoribosyl pyrophosphate synthase-associated protein 1<br>Short name=PRPP synthase-associated protein 1<br>Alternative name(s):<br>39 kDa phosphoribosypyrophosphate synthase-associated protein<br>Short name=PAP39 |
| **Gene Names** | Name: PRPSAP1 |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Seems to play a negative regulatory role in 5-phosphoribose 1-diphosphate synthesis. |
| **Subunit structure** | Binds to PRPS1 and PRPS2. |
| **Subcellular location** | N/A |

# [S-4] PPI pairs having root-shared GOterms

**PPI pairs having root-shared GO terms (i.e. GO:0004749 & GO:0042802)**

| PPI1 | PPI2 | Source DB | Taxonomy1 | Taxonomy2 | type | method |
|---|---|---|---|---|---|---|
| P32895 | P32895 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P32895 | P38620 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P32895 | P38689 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P32895 | Q12265 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P32895 | P38620 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P38620 | P38620 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P38620 | Q12265 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P32895 | P38689 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P38689 | P38689 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P38689 | Q12265 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P32895 | Q12265 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| P38620 | Q12265 | mint | 4932 | 4932 | physical association | two hybrid pooling |
| P38689 | Q12265 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| Q12265 | Q12265 | mint | 4932 | 4932 | physical association | dhfr reconstruction |
| Q14558 | Q14558 | mint | 9606 | 9606 | physical association | two hybrid |

# [S-5] GPs containing root shared GO terms III

- GO:0004716: Receptor signaling protein tyrosine kinase activity
- GO:0046982: Protein heterodimerization activity
- GO:0005524: ATP binding
- Root shared except GO:0046982 – GO:0005524 which are shared by GO:0005488 (binding – 2nd direct child of root term, GO:0003674)

| | |
|---|---|
| **Accession ID** | P00533 (EGFR_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Epidermal growth factor receptor<br>EC=2.7.10.1<br>Alternative name(s):<br>Proto-oncogene c-ErbB-1<br>Receptor tyrosine-protein kinase erbB-1 |
| **Gene Names** | Name: EGFR<br>Synonyms: ERBB, ERBB1, HER1 |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Receptor tyrosine kinase binding ligands of the EGF family and activating several signaling cascades to convert extracellular cues into appropriate cellular responses. Known ligands include EGF, TGFA/TGF-alpha, amphiregulin, epigen/EPGN, BTC/betacellulin, epiregulin/EREG and HBEGF/heparin-binding EGF. Ligand binding triggers receptor homo- and/or heterodimerization and autophosphorylation on key cytoplasmic residues. The phosphorylated receptor recruits adapter proteins like GRB2 which in turn activates complex downstream signaling cascades. Activates at least 4 major downstream signaling cascades including the RAS-RAF-MEK-ERK, PI3 kinase-AKT, PLCgamma-PKC and STATs modules. May also activate the NF-kappa-B signaling cascade. Also directly phosphorylates other proteins like RGS16, activating its GTPase activity and probably coupling the EGF receptor signaling to the G protein-coupled receptor signaling. Also phosphorylates MUC1 and increases its interaction with SRC and CTNNB1/beta-catenin.<br>Isoform 2 may act as an antagonist of EGF action. |
| **Catalytic activity** | ATP + a [protein]-L-tyrosine = ADP + a [protein]-L-tyrosine phosphate. |
| **Enzyme regulation** | Endocytosis and inhibition of the activated EGFR by phosphatases like PTPRJ and PTPRK constitute immediate regulatory mechanisms. Upon EGF-binding phosphorylates EPS15 that regulates EGFR endocytosis and activity. Moreover, inducible feedback inhibitors including LRIG1, SOCS4, SOCS5 and ERRFI1 constitute alternative regulatory mechanisms for the EGFR signaling. |
| **Subunit structure** | Binding of the ligand triggers homo- and/or heterodimerization of the receptor triggering its autophosphorylation. Heterodimer with ERBB2. Interacts with ERRFI1; inhibits dimerization of the kinase domain and autophosphorylation. Part of a complex with ERBB2 and either PIK3C2A or PIK3C2B. Interacts with GRB2; an adapter protein coupling the receptor to downstream signaling pathways. Interacts with GAB2; involved in signaling downstream of EGFR. Interacts with STAT3; mediates EGFR downstream signaling in cell proliferation. Interacts with RIPK1; involved in NF-kappa-B activation. Interacts (autophosphorylated) with CBL; involved in EGFR ubiquitination and regulation. Interacts with SOCS5; regulates EGFR degradation through TCEB1- and TCEB2-mediated ubiquitination and proteasomal degradation. Interacts with PRMT5; methylates EGFR and enhances interaction with PTPN6. Interacts (phosphorylated) with PTPN6; inhibits EGFR-dependent activation of MAPK/ERK. Interacts with COPG1; essential for regulation of EGF-dependent nuclear transport of EGFR by retrograde trafficking from the Golgi to the ER. Interacts with TNK2; this interaction is dependent on EGF stimulation and kinase activity of EGFR. Interacts with PCNA; positively regulates PCNA. Interacts with PELP1. Interacts with MUC1. Interacts with AP2M1. Interacts with FER. May interact with EPS8; mediates EPS8 phosphorylation. Interacts (via SH2 domains) with GRB2, NCK1 and |

| | |
|---|---|
| | NCK2. Interacts with ATX2. |
| **Subcellular location** | Cell membrane; Single-pass type I membrane protein. Endoplasmic reticulum membrane; Single-pass type I membrane protein. Golgi apparatus membrane; Single-pass type I membrane protein. Nucleus membrane; Single-pass type I membrane protein. Endosome. Endosome membrane. Note: In response to EGF, translocated from the cell membrane to the nucleus via Golgi and ER. Endocytosed upon activation by ligand.<br>Isoform 2: Secreted |
| **Accession ID** | P04626 (ERBB2_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Receptor tyrosine-protein kinase erbB-2<br>EC=2.7.10.1<br>Alternative name(s):<br>Metastatic lymph node gene 19 protein<br>Short name=MLN 19<br>Proto-oncogene Neu<br>Proto-oncogene c-ErbB-2<br>Tyrosine kinase-type cell surface receptor HER2<br>p185erbB2<br>CD_antigen=CD340 |
| **Gene Names** | Name: ERBB2<br>Synonyms: HER2, MLN19, NEU, NGL |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Protein tyrosine kinase that is part of several cell surface receptor complexes, but that apparently needs a coreceptor for ligand binding. Essential component of a neuregulin-receptor complex, although neuregulins do not interact with it alone. GP30 is a potential ligand for this receptor. Regulates outgrowth and stabilization of peripheral microtubules (MTs). Upon ERBB2 activation, the MEMO1-RHOA-DIAPH1 signaling pathway elicits the phosphorylation and thus the inhibition of GSK3B at cell membrane. This prevents the phosphorylation of APC and CLASP2, allowing its association with the cell membrane. In turn, membrane-bound APC allows the localization of MACF1 to the cell membrane, which is required for microtubule capture and stabilization.<br>In the nucleus is involved in transcriptional regulation. Associates with the 5'-TCAAATTC-3' sequence in the PTGS2/COX-2 promoter and activates its transcription. Implicated in transcriptional activation of CDKN1A; the function involves STAT3 and SRC. Involved in the transcription of rRNA genes by RNA Pol I and enhances protein synthesis and cell growth. |
| **Catalytic activity** | ATP + a [protein]-L-tyrosine = ADP + a [protein]-L-tyrosine phosphate. |
| **Enzyme regulation** | Activated by dimerization. Not activated by EGF, TGF-alpha and amphiregulin. Interaction with PTK6 increases its intrinsic kinase activity. |
| **Subunit structure** | Homodimer. Heterodimer with EGFR, ERBB3 and ERBB4. Part of a complex with EGFR and either PIK3C2A or PIK3C2B. May interact with PIK3C2B when phosphorylated on Tyr-1196. Interacts with PRKCABP and PLXNB1. Interacts (when phosphorylated on Tyr-1248) with MEMO1. Interacts with MUC1; the interaction is enhanced by heregulin (HRG). Interacts (when phosphorylated on Tyr-1139) with GRB7 (via SH2 domain). Interacts (when phosphorylated on Tyr-1248) with ERBB2IP. Interacts with KPNB1, RANBP2, EEA1, CRM1, CLTC, PTK6, RPA94 and ACTB. Interacts with SRC |
| **Involvement in disease** | Defects in ERBB2 are a cause of hereditary diffuse gastric cancer (HDGC) [MIM:137215]. A cancer predisposition syndrome with increased susceptibility to diffuse gastric cancer. Diffuse gastric cancer is a malignant disease characterized by poorly differentiated infiltrating lesions resulting in thickening of the stomach. Malignant tumors start in the stomach, can spread to the esophagus or the small intestine, and can extend through the stomach wall to nearby lymph nodes and organs. It also can metastasize to other parts of the body. |

| | |
|---|---|
| | Defects in ERBB2 are involved in the development of glioma (GLM) [MIM:137800]. Gliomas are central nervous system neoplasms derived from glial cells and comprise astrocytomas, glioblastoma multiforme, oligodendrogliomas, and ependymomas.<br>Defects in ERBB2 are a cause of susceptibility to ovarian cancer (OC) [MIM:167000]. Ovarian cancer common malignancy originating from ovarian tissue. Although many histologic types of ovarian neoplasms have been described, epithelial ovarian carcinoma is the most common form. Ovarian cancers are often asymptomatic and the recognized signs and symptoms, even of late-stage disease, are vague. Consequently, most patients are diagnosed with advanced disease.<br>Defects in ERBB2 may be a cause of lung cancer (LNCR) [MIM:211980]. LNCR is a common malignancy affecting tissues of the lung. The most common form of lung cancer is non-small cell lung cancer (NSCLC) that can be divided into 3 major histologic subtypes: squamous cell carcinoma, adenocarcinoma, and large cell lung cancer. NSCLC is often diagnosed at an advanced stage and has a poor prognosis.<br>Defects in ERBB2 are a cause of gastric cancer (GASC) [MIM:613659]. A malignant disease which starts in the stomach, can spread to the esophagus or the small intestine, and can extend through the stomach wall to nearby lymph nodes and organs. It also can metastasize to other parts of the body. The term gastric cancer or gastric carcinoma refers to adenocarcinoma of the stomach that accounts for most of all gastric malignant tumors. Two main histologic types are recognized, diffuse type and intestinal type carcinomas. Diffuse tumors are poorly differentiated infiltrating lesions resulting in thickening of the stomach. In contrast, intestinal tumors are usually exophytic, often ulcerating, and associated with intestinal metaplasia of the stomach, most often observed in sporadic disease.<br>Note=Chromosomal aberrations involving ERBB2 may be a cause gastric cancer. Deletions within 17q12 region producing fusion transcripts with CDK12, leading to CDK12-ERBB2 fusion leading to trunctated CDK12 protein not in-frame with ERBB2. |
| **Subcellular location** | Isoform 1: Cell membrane; Single-pass type I membrane protein. Cytoplasm › perinuclear region. Nucleus. Note: Translocation to the nucleus requires endocytosis, probably endosomal sorting and is mediated by importin beta-1/KPNB1.<br>Isoform 2: Cytoplasm. Nucleus<br>Isoform 3: Cytoplasm. Nucleus |
| **Accession ID** | P06494 (ERBB2_RAT) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Receptor tyrosine-protein kinase erbB-2<br>EC=2.7.10.1<br>Alternative name(s):<br>Epidermal growth factor receptor-related protein<br>Proto-oncogene Neu<br>Proto-oncogene c-ErbB-2<br>p185erbB2<br>p185neu<br>CD_antigen=CD340 |
| **Gene Names** | Name: Erbb2<br>Synonyms: Neu |
| **Organism** | Rattus norvegicus (Rat) |
| **Taxonomic identifier** | 10116 [NCBI] |
| **Function** | Protein tyrosine kinase that is part of several cell surface receptor complexes, but that apparently needs a coreceptor for ligand binding. Essential component of a neuregulin-receptor complex, although neuregulins do not interact with it alone. GP30 is a potential ligand for this receptor. Regulates outgrowth and stabilization of peripheral microtubules (MTs). Upon ERBB2 activation, the MEMO1-RHOA-DIAPH1 signaling pathway elicits the phosphorylation and thus the inhibition of GSK3B at cell membrane. This prevents the phosphorylation of APC and CLASP2, allowing its association with the cell membrane. In turn, membrane-bound APC allows the localization of MACF1 to the cell |

| | |
|---|---|
| | membrane, which is required for microtubule capture and stabilization<br>In the nucleus is involved in transcriptional regulation. Associates with the 5'-TCAAATTC-3' sequence in the PTGS2/COX-2 promoter and activates its transcription. Implicated in transcriptional activation of CDKN1A; the function involves STAT3 and SRC. Involved in the transcription of rRNA genes by RNA Pol I and enhances protein synthesis and cell growth |
| **Catalytic activity** | ATP + a [protein]-L-tyrosine = ADP + a [protein]-L-tyrosine phosphate. |
| **Subunit structure** | Homodimer. Heterodimer with EGFR, ERBB3 and ERBB4. Part of a complex with EGFR and either PIK3C2A or PIK3C2B. May interact with PIK3C2B when phosphorylated on Tyr-1198. Interacts with PRKCABP and PLXNB1. Interacts (when phosphorylated on Tyr-1250) with MEMO1. Interacts with MUC1. Interacts (when phosphorylated on Tyr-1141) with GRB7 (via SH2 domain). Interacts (when phosphorylated on Tyr-1250) with ERBB2IP Interacts with SRC, KPNB1, RANBP2, EEA1, CRM1, CLTC, PTK6, RPA94 and ACTB |
| **Post-translational modification** | Autophosphorylated. Ligand-binding increases phosphorylation on tyrosine residues. Autophosphorylation occurs in trans, i.e. one subunit of the dimeric receptor phosphorylates tyrosine residues on the other subunit. Signaling via SEMA4C promotes phosphorylation at Tyr-1250 |
| **Subcellular location** | Cell membrane; Single-pass type I membrane protein<br>Cytoplasm › perinuclear region<br>Nucleus |
| **Accession ID** | P21860 (ERBB3_HUMAN) Reviewed, UniProtKB/Swiss-Prot |
| **Protein Names** | Recommended name:<br>Receptor tyrosine-protein kinase erbB-3<br>EC=2.7.10.1<br>Alternative name(s):<br>Proto-oncogene-like protein c-ErbB-3<br>Tyrosine kinase-type cell surface receptor HER3 |
| **Gene Names** | Name: ERBB3<br>Synonyms: HER3 |
| **Organism** | Homo sapiens (Human) |
| **Taxonomic identifier** | 9606 [NCBI] |
| **Function** | Binds and is activated by neuregulins and NTAK. |
| **Catalytic activity** | ATP + a [protein]-L-tyrosine = ADP + a [protein]-L-tyrosine phosphate. |
| **Subunit structure** | Monomer and homodimer. Heterodimer with each of the other ERBB receptors<br>Interacts with CSPG5, PA2G4, GRB7 and MUC1. |
| **Post-translational modification** | Ligand-binding increases phosphorylation on tyrosine residues and promotes its association with the p85 subunit of phosphatidylinositol 3-kinase<br>Subject to autophosphorylation |
| **Involvement in disease** | Defects in ERBB3 are the cause of lethal congenital contracture syndrome type 2 (LCCS2) [MIM:607598]; also called Israeli Bedouin multiple contracture syndrome type A. LCCS2 is an autosomal recessive neurogenic form of a neonatally lethal arthrogryposis that is associated with atrophy of the anterior horn of the spinal cord. The LCCS2 syndrome is characterized by multiple joint contractures, anterior horn atrophy in the spinal cord, and a unique feature of a markedly distended urinary bladder. The phenotype suggests a spinal cord neuropathic etiology. |
| **Subcellular location** | Isoform 1: Cell membrane; Single-pass type I membrane protein.<br>Isoform 2: Secreted. |

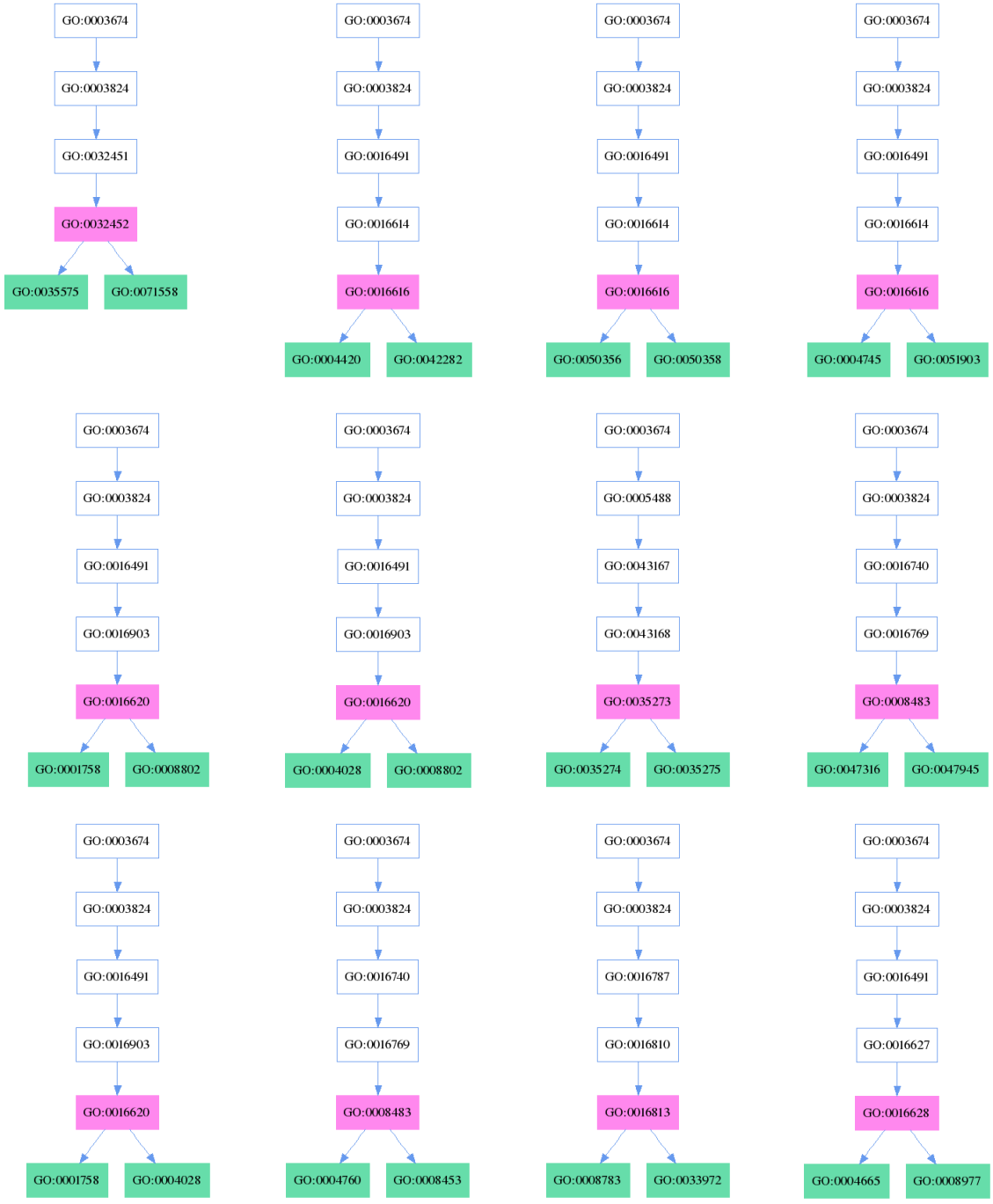
# [S-6] PPI pairs having root-shared GO terms

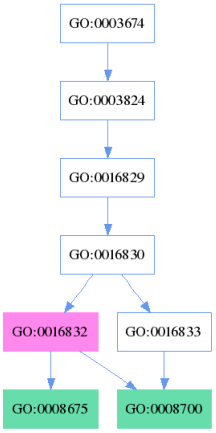
**PPI pairs having root-shared GO terms (i.e. GO:0004716 & GO:0046982)**

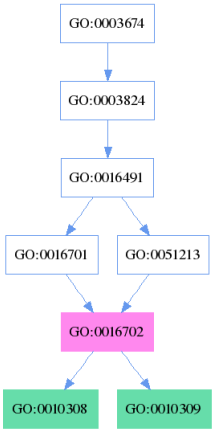
| PPI1 | PPI2 | Source DB | Taxonomy1 | Taxonomy2 | type | method |
|---|---|---|---|---|---|---|
| P00533 | P00533 | mint | 9606 | 9606 | colocalization | fluorescence imaging |
| P00533 | P04626 | mint | 9606 | 9606 | physical association | anti bait coip |
| P00533 | P21860 | intact | 9606 | 9606 | physical association | anti bait coip |
| P00533 | P04626 | mint | 9606 | 9606 | physical association | anti bait coip |
| P04626 | P04626 | intact | 9606 | 9606 | physical association | protein kinase assay |
| P04626 | P21860 | intact | 9606 | 9606 | physical association | anti bait coip |
| P00533 | P21860 | intact | 9606 | 9606 | physical association | anti bait coip |
| P04626 | P21860 | intact | 9606 | 9606 | physical association | anti bait coip |
| P21860 | P21860 | intact | 9606 | 9606 | direct interaction | NMR |

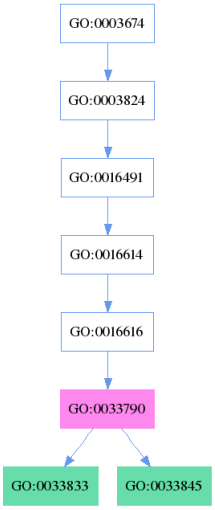
# [S-7] Statistical significance of BSM score

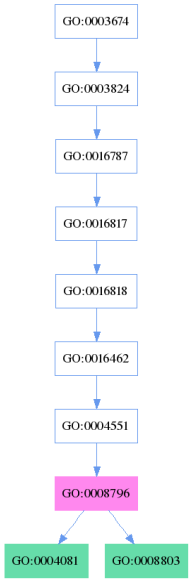
**Visualization of randomly selected 50 GO term similarity scores:** Green colors denote GO pairs and pink colors denote the nearest common ancestor of GO pairs.

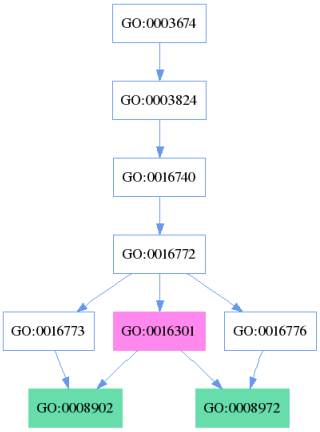


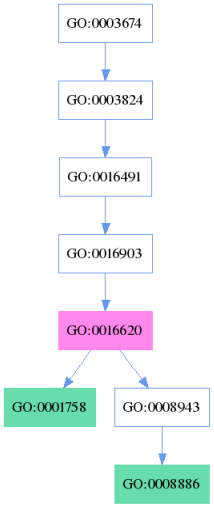

GO:0003674
GO:0003824
GO:0016829
GO:0016830
GO:0016832
GO:0016833
GO:0008675
GO:0008700

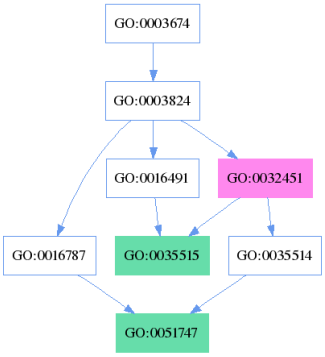

GO:0003674
GO:0003824
GO:0016491
GO:0016701
GO:0051213
GO:0016702
GO:0010308
GO:0010309

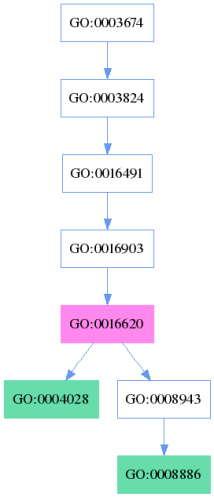

GO:0003674
GO:0003824
GO:0016491
GO:0016614
GO:0016616
GO:0033790
GO:0033833
GO:0033845

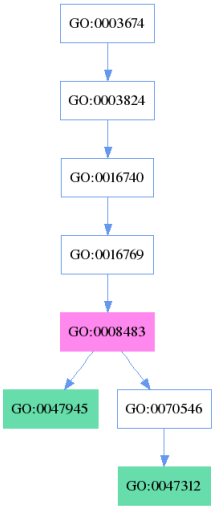

GO:0003674
GO:0003824
GO:0016787
GO:0016817
GO:0016818
GO:0016462
GO:0004551
GO:0008796
GO:0004081
GO:0008803

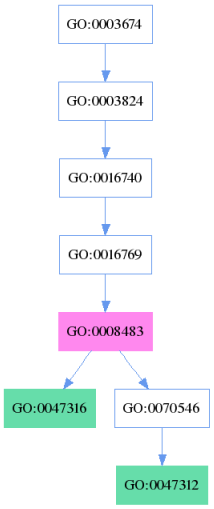

GO:0003674
GO:0003824
GO:0016740
GO:0016772
GO:0016773
GO:0016301
GO:0016776
GO:0008902
GO:0008972

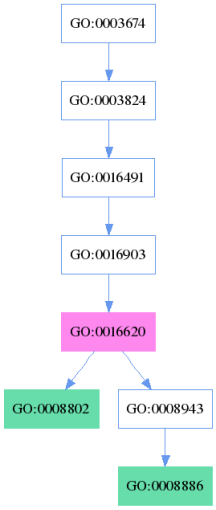

GO:0003674
GO:0003824
GO:0016491
GO:0016903
GO:0016620
GO:0001758
GO:0008943
GO:0008886

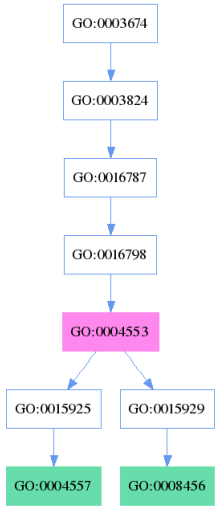

GO:0003674
GO:0003824
GO:0016491
GO:0032451
GO:0016787
GO:0035515
GO:0035514
GO:0051747

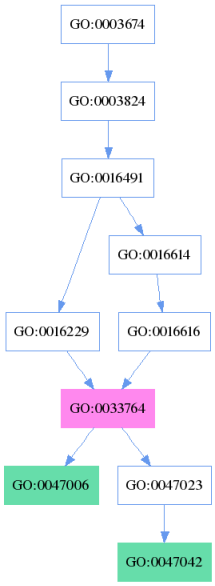

GO:0003674
GO:0003824
GO:0016491
GO:0016903
GO:0016620
GO:0004028
GO:0008943
GO:0008886

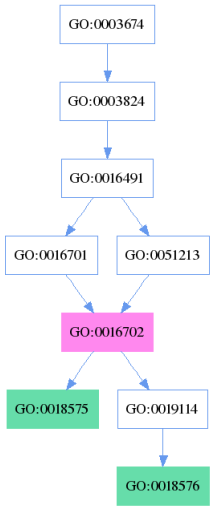

GO:0003674
GO:0003824
GO:0016740
GO:0016769
GO:0008483
GO:0047945
GO:0070546
GO:0047312

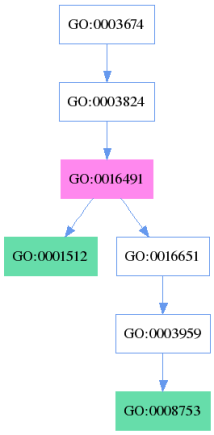

GO:0003674
GO:0003824
GO:0016740
GO:0016769
GO:0008483
GO:0047316
GO:0070546
GO:0047312

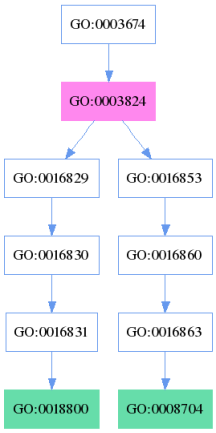

GO:0003674
GO:0003824
GO:0016491
GO:0016903
GO:0016620
GO:0008802
GO:0008943
GO:0008886

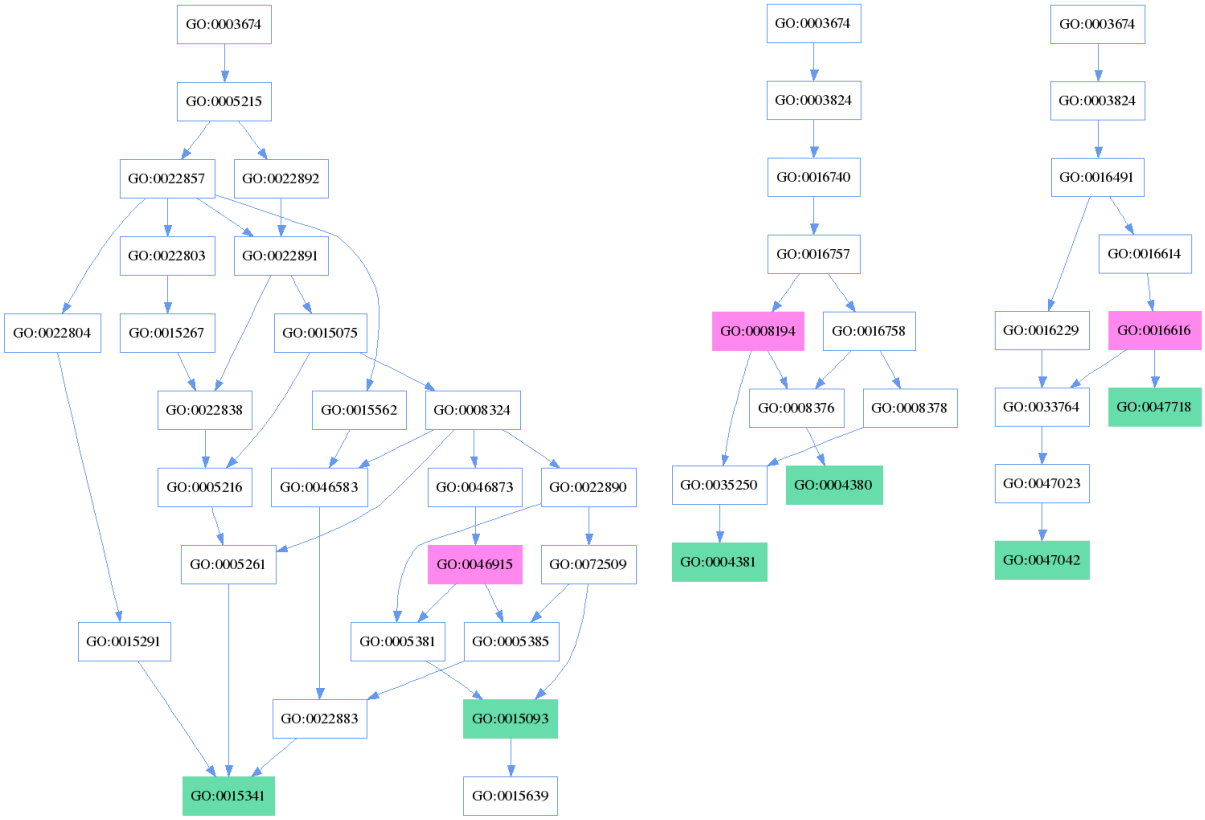

GO:0003674
GO:0003824
GO:0016787
GO:0016798
GO:0004553
GO:0015925
GO:0015929
GO:0004557
GO:0008456

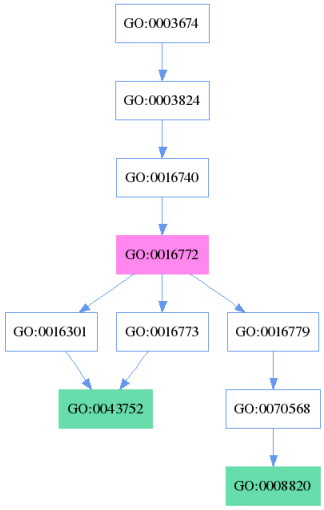

GO:0003674
GO:0003824
GO:0016491
GO:0016614
GO:0016229
GO:0016616
GO:0033764
GO:0047006
GO:0047023
GO:0047042

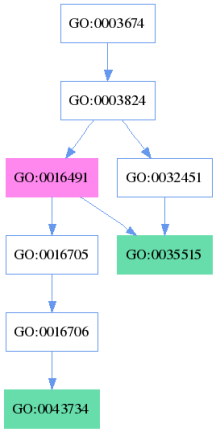

GO:0003674
GO:0003824
GO:0016491
GO:0016701
GO:0051213
GO:0016702
GO:0018575
GO:0019114
GO:0018576

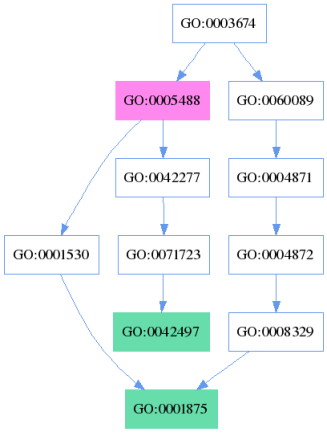

GO:0003674
GO:0003824
GO:0016491
GO:0001512
GO:0016651
GO:0003959
GO:0008753

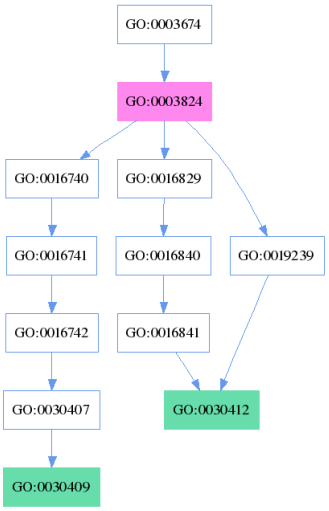

GO:0003674
GO:0003824
GO:0016829
GO:0016853
GO:0016830
GO:0016860
GO:0016831
GO:0016863
GO:0018800
GO:0008704

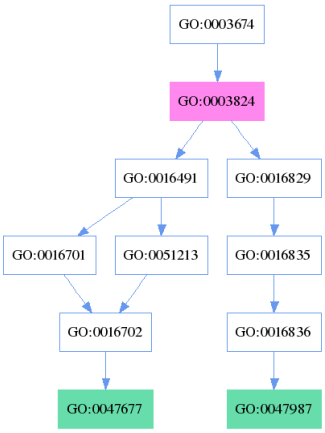

GO:0003674
GO:0005215
GO:0022857
GO:0022892
GO:0022803
GO:0022891
GO:0022804
GO:0015267
GO:0015075
GO:0022838
GO:0015562
GO:0008324
GO:0005216
GO:0046583
GO:0046873
GO:0022890
GO:0005261
GO:0046915
GO:0072509
GO:0015291
GO:0005381
GO:0005385
GO:0022883
GO:0015093
GO:0015341
GO:0015639
GO:0003674
GO:0003824
GO:0016740
GO:0016757
GO:0008194
GO:0016758
GO:0008376
GO:0008378
GO:0035250
GO:0004380
GO:0004381
GO:0003674
GO:0003824
GO:0016491
GO:0016614
GO:0016229
GO:0016616
GO:0033764
GO:0047718
GO:0047023
GO:0047042

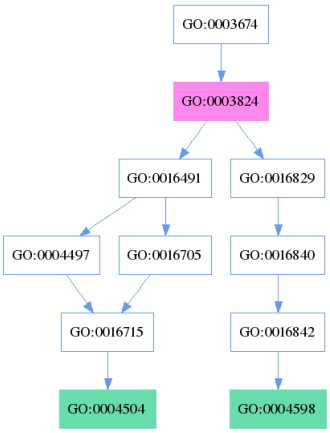

GO:0003674
GO:0003824
GO:0016740
GO:0016772
GO:0016301
GO:0016773
GO:0016779
GO:0043752
GO:0070568
GO:0008820

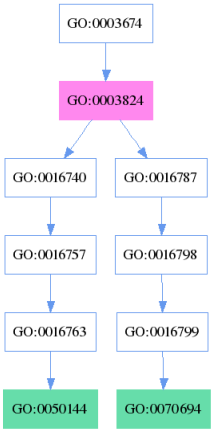

GO:0003674
GO:0003824
GO:0016491
GO:0032451
GO:0016705
GO:0035515
GO:0016706
GO:0043734

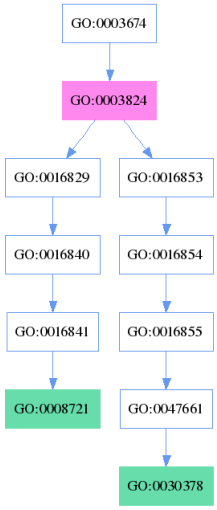

GO:0003674
GO:0005488
GO:0060089
GO:0042277
GO:0004871
GO:0001530
GO:0071723
GO:0004872
GO:0042497
GO:0008329
GO:0001875

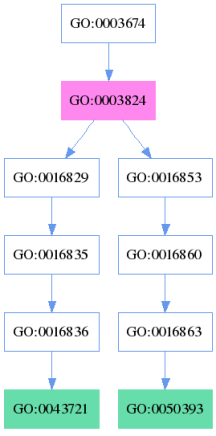

GO:0003674
GO:0003824
GO:0016740
GO:0016829
GO:0016741
GO:0016840
GO:0019239
GO:0016742
GO:0016841
GO:0030407
GO:0030412
GO:0030409

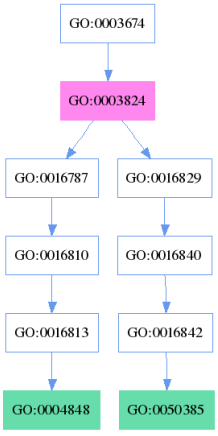

GO:0003674
GO:0003824
GO:0016491
GO:0016829
GO:0016701
GO:0051213
GO:0016835
GO:0016702
GO:0016836
GO:0047677
GO:0047987

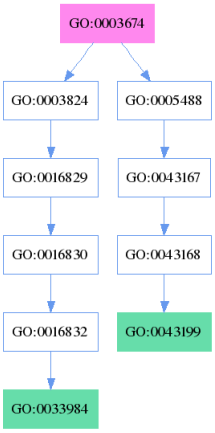

GO:0003674
GO:0003824
GO:0016491
GO:0016829
GO:0004497
GO:0016705
GO:0016840
GO:0016715
GO:0016842
GO:0004504
GO:0004598


GO:0003674
GO:0003824
GO:0016740
GO:0016787
GO:0016757
GO:0016798
GO:0016763
GO:0016799
GO:0050144
GO:0070694

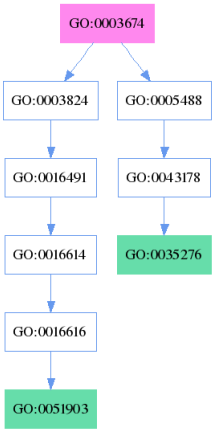

GO:0003674
GO:0003824
GO:0016829
GO:0016853
GO:0016840
GO:0016854
GO:0016841
GO:0016855
GO:0008721
GO:0047661
GO:0030378

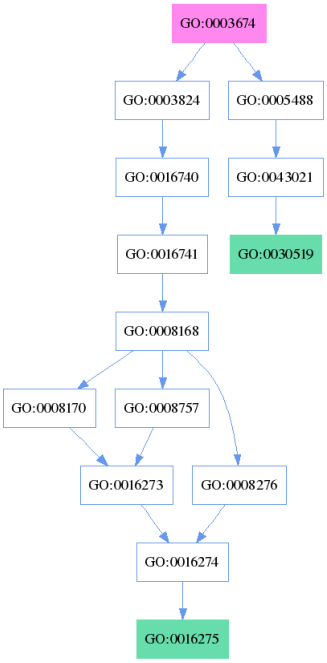

GO:0003674
GO:0003824
GO:0016829
GO:0016853
GO:0016835
GO:0016860
GO:0016836
GO:0016863
GO:0043721
GO:0050393

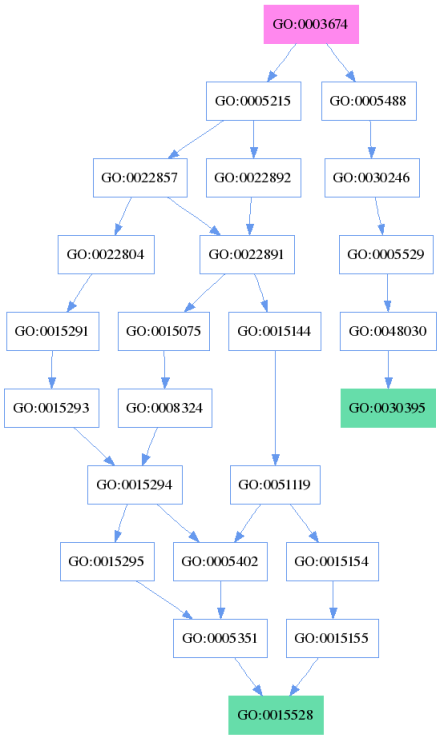

GO:0003674
GO:0003824
GO:0016787
GO:0016829
GO:0016810
GO:0016840
GO:0016813
GO:0016842
GO:0004848
GO:0050385

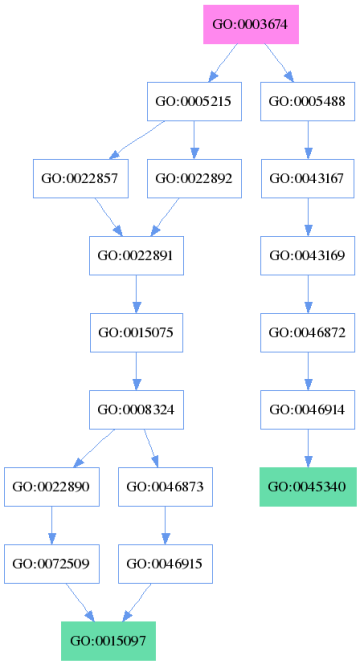

GO:0003674
GO:0003824
GO:0005488
GO:0016829
GO:0043167
GO:0016830
GO:0043168
GO:0016832
GO:0043199
GO:0033984

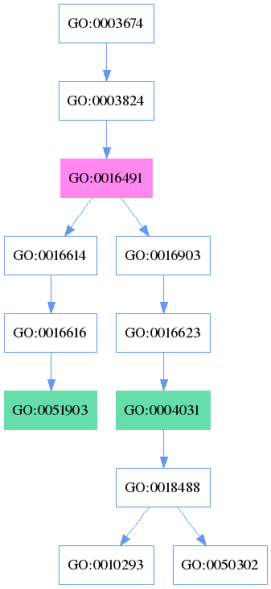

GO:0003674
GO:0003824
GO:0005488
GO:0016491
GO:0043178
GO:0016614
GO:0035276
GO:0016616
GO:0004745

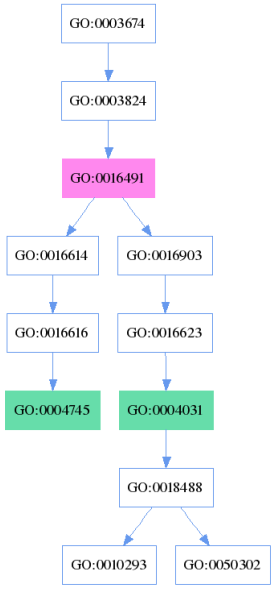

GO:0003674
GO:0003824
GO:0005488
GO:0016491
GO:0043178
GO:0016614
GO:0035276
GO:0016616
GO:0051903

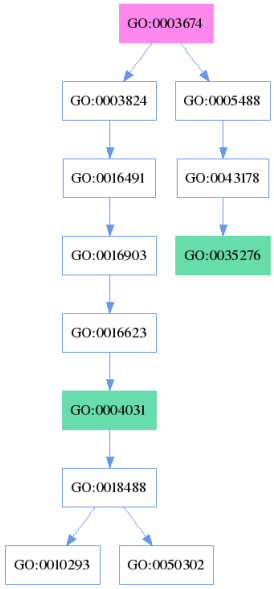

GO:0003674
GO:0003824
GO:0005488
GO:0016740
GO:0043021
GO:0016741
GO:0030519
GO:0008168
GO:0008170
GO:0008757
GO:0016273
GO:0008276
GO:0016274
GO:0016275


GO:0003674
GO:0005215
GO:0005488
GO:0022857
GO:0022892
GO:0030246
GO:0022804
GO:0022891
GO:0005529
GO:0015291
GO:0015075
GO:0015144
GO:0048030
GO:0015293
GO:0008324
GO:0030395
GO:0015294
GO:0051119
GO:0015295
GO:0005402
GO:0015154
GO:0005351
GO:0015155
GO:0015528


GO:0003674
GO:0005215
GO:0005488
GO:0022857
GO:0022892
GO:0043167
GO:0022891
GO:0043169
GO:0015075
GO:0046872
GO:0008324
GO:0046914
GO:0022890
GO:0046873
GO:0045340
GO:0072509
GO:0046915
GO:0015097




## $R_{scr}$ vs BSM

Scores derived from reference score ($R_{scr}$) only and proposed full scoring function for randomly selected 50 GO terms: GO1 and GO2 are the GO term pairs shown above. $R_{scr}$ is the score derived from reference matrices. BSM is the final score calculated by proposed method which contains two parts reference score ($R_{scr}$) and graph based score ($G_{scr}$). #GPs_OR denotes the number of GPs annotated by either GO1 or GO2. #PPIs_OR denotes the number of existing PPIs in which each of GPs must be annotated by either GO1 or GO2. #GPs_AND denotes the number of GPs annotated by both GO1 and GO2. #PPIs_AND denotes the number of existing PPIs in which each of GPs must be annotated by both GO1 and GO2. Orange and green colors denote tight and loose relations between nearest common ancestor (NCA) and given GO pairs respectively.

| GO1 | GO2 | $R_{scr}$ | BSM | #GPs_OR | #PPIs_OR | #GPs_AND | #PPIs_AND |
|---|---|---|---|---|---|---|---|
| GO:0035575 | GO:0071558 | 0.750 | 0.635 | 9 | 0 | 8 | 0 |
| GO:0004420 | GO:0042282 | 0.604 | 0.302 | 53 | 4 | 4 | 0 |
| GO:0050356 | GO:0050358 | 0.750 | 0.375 | 3 | 0 | 0 | 0 |
| GO:0004745 | GO:0051903 | 0.604 | 0.302 | 73 | 3 | 0 | 0 |
| GO:0001758 | GO:0008802 | 0.677 | 0.343 | 147 | 1 | 0 | 0 |
| GO:0004028 | GO:0008802 | 0.677 | 0.343 | 141 | 0 | 0 | 0 |
| GO:0035274 | GO:0035275 | 0.750 | 0.781 | 1 | 0 | 1 | 0 |
| GO:0047316 | GO:0047945 | 0.750 | 0.375 | 4 | 0 | 1 | 0 |
| GO:0001758 | GO:0004028 | 0.677 | 0.343 | 28 | 1 | 4 | 0 |
| GO:0004760 | GO:0008453 | 0.677 | 0.338 | 22 | 0 | 9 | 0 |
| GO:0008783 | GO:0033972 | 0.750 | 0.375 | 78 | 0 | 1 | 0 |
| GO:0004665 | GO:0008977 | 0.677 | 0.338 | 26 | 2 | 21 | 1 |
| GO:0008675 | GO:0008700 | 0.677 | 0.338 | 20 | 1 | 9 | 1 |
| GO:0010308 | GO:0010309 | 0.677 | 0.338 | 100 | 0 | 100 | 0 |
| GO:0033833 | GO:0033845 | 0.604 | 0.724 | 2 | 0 | 1 | 0 |
| GO:0004081 | GO:0008803 | 0.750 | 0.769 | 230 | 0 | 1 | 0 |
| GO:0008902 | GO:0008972 | 0.750 | 0.375 | 31 | 0 | 26 | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GO:0001758 | GO:0008886 | 0.677 | 0.343 | 28 | 1 | 0 | 0 |
| GO:0035515 | GO:0051747 | 0.750 | 0.585 | 8 | 0 | 1 | 0 |
| GO:0004028 | GO:0008886 | 0.677 | 0.343 | 22 | 0 | 0 | 0 |
| GO:0047312 | GO:0047945 | 0.750 | 0.375 | 2 | 0 | 1 | 0 |
| GO:0047312 | GO:0047316 | 0.750 | 0.375 | 5 | 0 | 1 | 0 |
| GO:0008802 | GO:0008886 | 0.677 | 0.343 | 137 | 0 | 0 | 0 |
| GO:0004557 | GO:0008456 | 0.604 | 0.302 | 89 | 2 | 2 | 1 |
| GO:0047006 | GO:0047042 | 0.713 | 0.723 | 15 | 1 | 4 | 1 |
| GO:0018575 | GO:0018576 | 0.635 | 0.318 | 15 | 0 | 5 | 0 |
| GO:0001512 | GO:0008753 | 0.750 | 0.531 | 8 | 1 | 1 | 1 |
| GO:0018800 | GO:0008704 | 0.520 | 0.436 | 4 | 0 | 3 | 0 |
| GO:0015093 | GO:0015341 | 0.604 | 0.500 | 38 | 1 | 1 | 0 |
| GO:0004380 | GO:0004381 | 0.677 | 0.573 | 5 | 0 | 4 | 0 |
| GO:0047042 | GO:0047718 | 0.598 | 0.299 | 11 | 2 | 4 | 1 |
| GO:0008820 | GO:0043752 | 0.750 | 0.610 | 6 | 0 | 5 | 0 |
| GO:0035515 | GO:0043734 | 0.635 | 0.474 | 14 | 0 | 3 | 0 |
| GO:0001875 | GO:0042497 | 0.750 | 0.441 | 22 | 5 | 1 | 0 |
| GO:0030409 | GO:0030412 | 0.677 | 0.514 | 6 | 0 | 6 | 0 |
| GO:0047677 | GO:0047987 | 0.750 | 0.550 | 8 | 0 | 1 | 0 |
| GO:0004504 | GO:0004598 | 0.677 | 0.514 | 12 | 1 | 9 | 1 |
| GO:0050144 | GO:0070694 | 0.750 | 0.550 | 21 | 1 | 3 | 0 |
| GO:0008721 | GO:0030378 | 0.750 | 0.550 | 117 | 0 | 5 | 0 |
| GO:0043721 | GO:0050393 | 0.604 | 0.477 | 1 | 0 | 1 | 0 |
| GO:0004848 | GO:0050385 | 0.750 | 0.550 | 88 | 1 | 1 | 0 |
| GO:0004031 | GO:0051903 | 0.604 | 0.458 | 55 | 2 | 0 | 0 |
| GO:0004031 | GO:0004745 | 0.604 | 0.458 | 41 | 2 | 1 | 1 |
| GO:0033984 | GO:0043199 | 0.750 | 0.545 | 4 | 0 | 0 | 0 |
| GO:0004745 | GO:0035276 | 0.604 | 0.471 | 32 | 2 | 2 | 1 |
| GO:0035276 | GO:0051903 | 0.604 | 0.471 | 47 | 2 | 0 | 0 |
| GO:0016275 | GO:0030519 | 0.640 | 0.490 | 2 | 1 | 1 | 1 |
| GO:0015528 | GO:0030395 | 0.750 | 0.545 | 10 | 0 | 1 | 0 |
| GO:0015097 | GO:0045340 | 0.677 | 0.508 | 57 | 0 | 10 | 0 |
| GO:0004031 | GO:0035276 | 0.604 | 0.471 | 15 | 1 | 1 | 1 |

**Average scores from 50 GO terms:** Tight and loose denote the geometrically close or distant relations between nearest common ancestor (NCA) and given GO pairs. Tight relation requires parent-child relationship between NCA and both GO terms. Root relation requires both GO terms are shared by root node. In other words, NCA has to be root node (i.e. GO:0003674). Loose relations are defined as all GO term pairs except 'Tight' and 'Root' relations. Each color corresponds to the colors shown above table.

| CATEGORY | $R_{scr}$ | $GO_{sim}$ | #GPs_OR | #PPIs_OR | #GPs_AND | #PPIs_AND |
|---|---|---|---|---|---|---|
| **Tight** | 0.694 | 0.435 | 56.941 | 0.706 | 10.941 | 0.118 |
| **Loose** | 0.680 | 0.469 | 29.923 | 0.769 | 2.423 | 0.231 |
| **Root** | 0.661 | 0.500 | 23.857 | 0.857 | 2.143 | 0.429 |
| **Overall** | 0.682 | 0.462 | 38.260 | 0.760 | 5.280 | 0.220 |

## [S-8] SCOP domains used in experiments

- SunID: unique SCOP ID
- Sunid_cl: SCOP level at class
- Sunid_cf: SCOP level at class family
- Sunid_sf: SCOP level at super family
- Sunid_fa: SCOP level at family
- Sunid_dm: SCOP level at domain
- Sunid_sp: SCOP level at species
- Pdbid: representing PDB ID corresponding to the sunID

| SunID | sunid_cl | sunid_cf | sunid_sf | sunid_fa | sunid_dm | sunid_sp | pdbid |
|---|---|---|---|---|---|---|---|
| 14984 | 46456 | 46457 | 46458 | 46463 | 46464 | 46465 | 3sdh |
| 15010 | 46456 | 46457 | 46458 | 46463 | 46464 | 46466 | 1b0b |
| 15018 | 46456 | 46457 | 46458 | 46463 | 46469 | 46470 | 1a6m |
| 15149 | 46456 | 46457 | 46458 | 46463 | 46469 | 46471 | 1mba |
| 15211 | 46456 | 46457 | 46458 | 46463 | 46479 | 46480 | 1ecd |
| 15212 | 46456 | 46457 | 46458 | 46463 | 46481 | 46482 | 2gdm |
| 15406 | 46456 | 46457 | 46458 | 46463 | 46486 | 46496 | 1cg5 |
| 15585 | 46456 | 46457 | 46458 | 46463 | 46500 | 46512 | 1cg5 |
| 15589 | 46456 | 46457 | 46458 | 46463 | 46500 | 46515 | 1gcv |
| 15622 | 46456 | 46457 | 46458 | 46463 | 46520 | 46521 | 1ash |
| 15623 | 46456 | 46457 | 46458 | 46463 | 46522 | 46523 | 1ith |
| 15625 | 46456 | 46457 | 46458 | 46463 | 46524 | 46525 | 1hlb |
| 15635 | 46456 | 46457 | 46458 | 46463 | 46528 | 46529 | 1cqx |
| 15702 | 46456 | 46556 | 46579 | 46580 | 46581 | 46582 | 1fxk |
| 15703 | 46456 | 46556 | 46579 | 46580 | 46583 | 46584 | 1fxk |
| 15796 | 46456 | 46625 | 46626 | 46627 | 46628 | 46629 | 1c75 |
| 15798 | 46456 | 46625 | 46626 | 46627 | 46628 | 46630 | 1ctj |
| 15801 | 46456 | 46625 | 46626 | 46627 | 46628 | 46631 | 1c53 |
| 15810 | 46456 | 46625 | 46626 | 46627 | 46636 | 46637 | 1c52 |
| 15822 | 46456 | 46625 | 46626 | 46627 | 46636 | 46639 | 1ql3 |
| 15834 | 46456 | 46625 | 46626 | 46627 | 46642 | 46643 | 1ycc |
| 15899 | 46456 | 46625 | 46626 | 46627 | 46650 | 46657 | 1cot |
| 15901 | 46456 | 46625 | 46626 | 46627 | 46658 | 46659 | 1cc5 |
| 15905 | 46456 | 46625 | 46626 | 46627 | 46660 | 46662 | 351c |
| 15909 | 46456 | 46625 | 46626 | 46627 | 46660 | 46664 | 1gks |
| 15911 | 46456 | 46625 | 46626 | 46627 | 46667 | 46668 | 1dw0 |
| 15989 | 46456 | 46688 | 46689 | 46690 | 46697 | 46698 | 1lfb |
| 16003 | 46456 | 46688 | 46689 | 46690 | 46714 | 46715 | 1bw5 |
| 16014 | 46456 | 46688 | 46689 | 46690 | 46724 | 46725 | 1vnd |
| 16017 | 46456 | 46688 | 46689 | 46690 | 46726 | 46727 | 1fjl |
| 16144 | 46456 | 46688 | 46785 | 46832 | 46835 | 46836 | 1d5v |
| 16417 | 46456 | 47094 | 47095 | 47096 | 47097 | 47098 | 1ckt |
| 16425 | 46456 | 47094 | 47095 | 47096 | 47100 | 47101 | 1qrv |
| 16433 | 46456 | 47094 | 47095 | 47096 | 47110 | 47111 | 2lef |
| 16481 | 46456 | 47112 | 47113 | 47134 | 47135 | 47136 | 1taf |
| 16482 | 46456 | 47112 | 47113 | 47134 | 47137 | 47138 | 1taf |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 16483 | 46456 | 47112 | 47113 | 47134 | 47139 | 47140 | 1bh9 |
| 16485 | 46456 | 47112 | 47113 | 47134 | 47141 | 47142 | 1bh9 |
| 16619 | 46456 | 47161 | 47216 | 47217 | 47218 | 47219 | 1avo |
| 16740 | 46456 | 47239 | 47240 | 47253 | 88792 | 88793 | 1mty |
| 16742 | 46456 | 47239 | 47240 | 47253 | 88789 | 88790 | 1mty |
| 16805 | 46456 | 47239 | 47240 | 47253 | 47261 | 47262 | 1afr |
| 16822 | 46456 | 47265 | 47266 | 47267 | 47268 | 47270 | 1bgc |
| 16826 | 46456 | 47265 | 47266 | 47267 | 47272 | 47273 | 1alu |
| 16829 | 46456 | 47265 | 47266 | 47267 | 47274 | 47275 | 1lki |
| 16831 | 46456 | 47265 | 47266 | 47267 | 47276 | 47277 | 1huw |
| 16839 | 46456 | 47265 | 47266 | 47267 | 47278 | 47279 | 1f6f |
| 16840 | 46456 | 47265 | 47266 | 47267 | 47280 | 47281 | 1cnt |
| 16844 | 46456 | 47265 | 47266 | 47267 | 47282 | 47283 | 1ax8 |
| 16845 | 46456 | 47265 | 47266 | 47267 | 47284 | 47285 | 1evs |
| 16846 | 46456 | 47265 | 47266 | 47286 | 47287 | 47288 | 1eer |
| 16849 | 46456 | 47265 | 47266 | 47286 | 47289 | 47290 | 2gmf |
| 16865 | 46456 | 47265 | 47266 | 47286 | 47293 | 47294 | 1hul |
| 16867 | 46456 | 47265 | 47266 | 47286 | 47295 | 47296 | 1hmc |
| 16869 | 46456 | 47265 | 47266 | 47286 | 47297 | 47298 | 1ete |
| 16873 | 46456 | 47265 | 47266 | 47286 | 47299 | 47300 | 1scf |
| 16884 | 46456 | 47265 | 47266 | 47286 | 47303 | 47304 | 1jli |
| 16885 | 46456 | 47265 | 47266 | 47305 | 47306 | 47307 | 2ilk |
| 16889 | 46456 | 47265 | 47266 | 47305 | 47309 | 47310 | 1au1 |
| 16900 | 46456 | 47265 | 47266 | 47305 | 47316 | 47317 | 1b5l |
| 16901 | 46456 | 47265 | 47266 | 47305 | 47318 | 47319 | 1d9c |
| 17132 | 46456 | 47458 | 47459 | 47460 | 47464 | 47465 | 1mdy |
| 17136 | 46456 | 47458 | 47459 | 47460 | 47466 | 47467 | 1an4 |
| 17138 | 46456 | 47458 | 47459 | 47460 | 47468 | 47469 | 1a0a |
| 17140 | 46456 | 47458 | 47459 | 47460 | 47470 | 47471 | 1am9 |
| 17180 | 46456 | 47472 | 47473 | 47478 | 47479 | 47485 | 1psr |
| 17185 | 46456 | 47472 | 47473 | 47478 | 47479 | 47486 | 1qls |
| 17230 | 46456 | 47472 | 47473 | 47502 | 47503 | 47504 | 1ctd |
| 17256 | 46456 | 47472 | 47473 | 47502 | 47509 | 47510 | 2scp |
| 17258 | 46456 | 47472 | 47473 | 47502 | 47509 | 47511 | 2sas |
| 17262 | 46456 | 47472 | 47473 | 47502 | 47514 | 47515 | 1c7v |
| 17299 | 46456 | 47472 | 47473 | 47502 | 47516 | 47523 | 1exr |
| 17302 | 46456 | 47472 | 47473 | 47502 | 47524 | 47525 | 1wdc |
| 17308 | 46456 | 47472 | 47473 | 47502 | 47524 | 47526 | 2mys |
| 17325 | 46456 | 47472 | 47473 | 47502 | 47530 | 47532 | 1aui |
| 17330 | 46456 | 47472 | 47473 | 47502 | 47537 | 47538 | 1jba |
| 17370 | 46456 | 47472 | 47473 | 63550 | 47552 | 47555 | 1alv |
| 17387 | 46456 | 47472 | 47565 | 47566 | 47567 | 47568 | 1c3y |
| 17388 | 46456 | 47472 | 47565 | 47566 | 47569 | 47570 | 1dqe |
| 17394 | 46456 | 47575 | 47576 | 47577 | 47578 | 47579 | 1bkr |
| 17396 | 46456 | 47575 | 47576 | 47577 | 47580 | 47581 | 1aoa |
| 17397 | 46456 | 47575 | 47576 | 47577 | 47580 | 47581 | 1aoa |
| 18146 | 46456 | 47873 | 47874 | 47875 | 47876 | 47877 | 1bo9 |
| 18392 | 46456 | 47972 | 47973 | 47974 | 47975 | 47977 | 1rss |
| 18538 | 46456 | 48096 | 48097 | 48098 | 48099 | 48100 | 1agr |
| 18542 | 46456 | 48096 | 48097 | 48098 | 48101 | 48102 | 1fqi |
| 18548 | 46456 | 48096 | 48097 | 48098 | 48105 | 48106 | 1dk8 |
| 18880 | 46456 | 48255 | 48256 | 48257 | 48258 | 48259 | 1csh |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 18968 | 46456 | 48263 | 48264 | 48265 | 48274 | 48275 | 1cpt |
| 18971 | 46456 | 48263 | 48264 | 48265 | 48278 | 48279 | 1io7 |
| 19052 | 46456 | 48316 | 48317 | 48318 | 48319 | 48320 | 1d2t |
| 19053 | 46456 | 48316 | 48317 | 48321 | 48322 | 48323 | 1qi9 |
| 19055 | 46456 | 48316 | 48317 | 48321 | 48322 | 48324 | 1qhb |
| 19061 | 46456 | 48316 | 48317 | 48325 | 48326 | 48327 | 1vns |
| 19119 | 46456 | 48370 | 48371 | 48372 | 48378 | 48379 | 1qgr |
| 19127 | 46456 | 48370 | 48371 | 48372 | 48381 | 48382 | 1qbk |
| 19155 | 53931 | 74651 | 48403 | 48404 | 48405 | 48406 | 1ycs |
| 19156 | 53931 | 74651 | 48403 | 48404 | 48407 | 48408 | 1awc |
| 19157 | 53931 | 74651 | 48403 | 48404 | 48409 | 48410 | 1bd8 |
| 19171 | 53931 | 74651 | 48403 | 48404 | 48417 | 48418 | 1ikn |
| 19174 | 53931 | 74651 | 48403 | 48404 | 48419 | 48420 | 1myo |
| 19176 | 53931 | 74651 | 48403 | 48404 | 48421 | 48422 | 1sw6 |
| 19178 | 53931 | 74651 | 48403 | 48404 | 48423 | 48424 | 1dcq |
| 19191 | 46456 | 48370 | 48439 | 48440 | 48443 | 48444 | 1dce |
| 19206 | 46456 | 48370 | 48452 | 48453 | 48454 | 48455 | 1a17 |
| 19207 | 46456 | 48370 | 48452 | 48453 | 48456 | 48457 | 1elw |
| 19209 | 46456 | 48370 | 48452 | 48453 | 48456 | 48457 | 1elr |
| 19210 | 46456 | 48370 | 48452 | 48453 | 48458 | 48459 | 1qqe |
| 19212 | 46456 | 48370 | 48452 | 48453 | 48462 | 48463 | 1fch |
| 19282 | 46456 | 48507 | 48508 | 48509 | 48515 | 48516 | 1fcy |
| 19425 | 46456 | 48556 | 48557 | 48558 | 48568 | 48569 | 1c3c |
| 19429 | 46456 | 48556 | 48557 | 48558 | 48568 | 48570 | 1dof |
| 19436 | 46456 | 48575 | 48576 | 48577 | 48578 | 48579 | 1fps |
| 19446 | 46456 | 48575 | 48576 | 48586 | 48587 | 48588 | 1di1 |
| 19450 | 46456 | 48575 | 48576 | 48586 | 48589 | 48590 | 1ps1 |
| 19630 | 46456 | 48661 | 48670 | 48671 | 48672 | 48673 | 1tbg |
| 19636 | 46456 | 48661 | 48670 | 48671 | 48672 | 48673 | 1gg2 |
| 19704 | 48724 | 48725 | 48726 | 48727 | 48728 | 48729 | 1neu |
| 19708 | 48724 | 48725 | 48726 | 48727 | 48732 | 48733 | 1qfo |
| 19740 | 48724 | 48725 | 48726 | 48727 | 48737 | 48739 | 1cid |
| 19741 | 48724 | 48725 | 48726 | 48727 | 48740 | 48741 | 1hnf |
| 19758 | 48724 | 48725 | 48726 | 48727 | 48743 | 48744 | 1ccz |
| 19762 | 48724 | 48725 | 48726 | 48727 | 48745 | 48746 | 1dr9 |
| 19763 | 48724 | 48725 | 48726 | 48727 | 48747 | 48748 | 1f97 |
| 20622 | 48724 | 48725 | 48726 | 48727 | 48933 | 48935 | 1bd2 |
| 20631 | 48724 | 48725 | 48726 | 48727 | 48933 | 48936 | 1fo0 |
| 20650 | 48724 | 48725 | 48726 | 48727 | 48933 | 48938 | 1tvd |
| 20653 | 48724 | 48725 | 48726 | 48727 | 48939 | 48941 | 1dqt |
| 21685 | 48724 | 48725 | 48726 | 49159 | 49160 | 49161 | 1vca |
| 21689 | 48724 | 48725 | 48726 | 49159 | 49162 | 49163 | 1iam |
| 21693 | 48724 | 48725 | 48726 | 49159 | 49164 | 49165 | 1zxq |
| 21694 | 48724 | 48725 | 48726 | 49159 | 49166 | 49167 | 1epf |
| 21695 | 48724 | 48725 | 48726 | 49159 | 49166 | 49167 | 1epf |
| 21706 | 48724 | 48725 | 48726 | 49159 | 49170 | 49171 | 1fhg |
| 21711 | 48724 | 48725 | 48726 | 49159 | 49172 | 49173 | 1tnn |
| 21712 | 48724 | 48725 | 48726 | 49159 | 49174 | 49175 | 1koa |
| 21713 | 48724 | 48725 | 48726 | 49159 | 49174 | 49175 | 1wiu |
| 21715 | 48724 | 48725 | 48726 | 49159 | 49174 | 49176 | 1tiu |
| 21717 | 48724 | 48725 | 48726 | 49159 | 49177 | 49178 | 1ira |
| 21718 | 48724 | 48725 | 48726 | 49159 | 49177 | 49178 | 1ira |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 21719 | 48724 | 48725 | 48726 | 49159 | 49177 | 49178 | 1ira |
| 21752 | 48724 | 48725 | 48726 | 49159 | 49182 | 49183 | 1bih |
| 21753 | 48724 | 48725 | 48726 | 49159 | 49182 | 49183 | 1bih |
| 21754 | 48724 | 48725 | 48726 | 49159 | 49182 | 49183 | 1bih |
| 21755 | 48724 | 48725 | 48726 | 49159 | 49182 | 49183 | 1bih |
| 21760 | 48724 | 48725 | 48726 | 49159 | 49184 | 49185 | 1cs6 |
| 21761 | 48724 | 48725 | 48726 | 49159 | 49184 | 49185 | 1cs6 |
| 21762 | 48724 | 48725 | 48726 | 49159 | 49184 | 49185 | 1cs6 |
| 21763 | 48724 | 48725 | 48726 | 49159 | 49184 | 49185 | 1cs6 |
| 21764 | 48724 | 48725 | 48726 | 49159 | 49186 | 49187 | 1f97 |
| 21765 | 48724 | 48725 | 48726 | 49159 | 49188 | 49189 | 1flt |
| 21778 | 48724 | 48725 | 48726 | 49159 | 49194 | 49195 | 1wwc |
| 21783 | 48724 | 48725 | 48726 | 49159 | 49196 | 49199 | 1fnl |
| 21784 | 48724 | 48725 | 48726 | 49159 | 49196 | 49199 | 1fnl |
| 21799 | 48724 | 48725 | 48726 | 49159 | 49202 | 49204 | 1nkr |
| 21800 | 48724 | 48725 | 48726 | 49159 | 49202 | 49204 | 1nkr |
| 22300 | 48724 | 48725 | 49329 | 49330 | 49341 | 49342 | 1ej8 |
| 22633 | 48724 | 49451 | 49482 | 49483 | 49484 | 49485 | 1dmh |
| 23675 | 48724 | 49763 | 49764 | 49765 | 49766 | 49767 | 1shs |
| 23928 | 48724 | 49898 | 49899 | 49900 | 49901 | 49902 | 1nls |
| 24157 | 48724 | 49898 | 49899 | 49900 | 49920 | 49921 | 1dhk |
| 24185 | 48724 | 49898 | 49899 | 49925 | 49930 | 49931 | 1dyp |
| 24198 | 48724 | 49898 | 49899 | 49932 | 100926 | 49935 | 2gal |
| 24216 | 48724 | 49898 | 49899 | 49932 | 49938 | 49939 | 1lcl |
| 24344 | 48724 | 49898 | 49899 | 49978 | 49991 | 49992 | 2nlr |
| 24656 | 48724 | 50128 | 50129 | 50130 | 50134 | 50135 | 1g31 |
| 24772 | 48724 | 50155 | 50156 | 50157 | 50160 | 50161 | 1kwa |
| 24777 | 48724 | 50155 | 50156 | 50157 | 50164 | 50165 | 1qav |
| 24779 | 48724 | 50155 | 50156 | 50157 | 50166 | 50167 | 1qau |
| 24796 | 48724 | 50181 | 50182 | 50183 | 50184 | 50185 | 1b34 |
| 24797 | 48724 | 50181 | 50182 | 50183 | 50186 | 50187 | 1b34 |
| 24798 | 48724 | 50181 | 50182 | 50183 | 50188 | 50189 | 1d3b |
| 24804 | 48724 | 50181 | 50182 | 50183 | 50190 | 50191 | 1d3b |
| 25282 | 48724 | 50198 | 50249 | 50263 | 50264 | 50265 | 3ull |
| 25284 | 48724 | 50198 | 50249 | 50263 | 50264 | 50266 | 1qvc |
| 25300 | 48724 | 50198 | 50249 | 50263 | 50267 | 50268 | 1jmc |
| 25301 | 48724 | 50198 | 50249 | 50263 | 50267 | 50268 | 1jmc |
| 25325 | 48724 | 50198 | 50249 | 50282 | 50283 | 50286 | 1c9o |
| 25327 | 48724 | 50198 | 50249 | 50282 | 50287 | 50288 | 1sro |
| 25329 | 48724 | 50198 | 50249 | 50282 | 50290 | 50291 | 1hr0 |
| 25331 | 48724 | 50198 | 50249 | 50282 | 50292 | 50293 | 1d7q |
| 25341 | 48724 | 50198 | 50249 | 50282 | 50296 | 50298 | 1bkb |
| 25342 | 48724 | 50198 | 50249 | 50282 | 50299 | 50300 | 1rl2 |
| 25560 | 48724 | 50352 | 50353 | 50362 | 50368 | 50369 | 2ila |
| 25591 | 48724 | 50352 | 50386 | 50387 | 50388 | 50389 | 1wba |
| 25599 | 48724 | 50352 | 50386 | 50387 | 50394 | 50395 | 1avw |
| 25747 | 48724 | 50474 | 50475 | 50476 | 50477 | 50478 | 1flm |
| 25753 | 48724 | 50474 | 50475 | 50476 | 50479 | 50481 | 1dnl |
| 26011 | 48724 | 50493 | 50494 | 50514 | 50515 | 50521 | 1gdn |
| 26068 | 48724 | 50493 | 50494 | 50514 | 50522 | 50524 | 1eq9 |
| 26072 | 48724 | 50493 | 50494 | 50514 | 50527 | 50528 | 1azz |
| 26074 | 48724 | 50493 | 50494 | 50514 | 50529 | 50530 | 2hlc |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 26284 | 48724 | 50493 | 50494 | 50514 | 50544 | 50545 | 1a7s |
| 26314 | 48724 | 50493 | 50494 | 50514 | 50563 | 50564 | 1bio |
| 26348 | 48724 | 50493 | 50494 | 50514 | 50579 | 50580 | 1aut |
| 26378 | 48724 | 50493 | 50494 | 50514 | 50593 | 50594 | 1fiw |
| 26819 | 48724 | 50629 | 50630 | 50646 | 50649 | 50651 | 2-Apr |
| 26828 | 48724 | 50629 | 50630 | 50646 | 50649 | 50654 | 1dpj |
| 26852 | 48724 | 50629 | 50630 | 50646 | 50658 | 50661 | 1htr |
| 26856 | 48724 | 50629 | 50630 | 50646 | 50663 | 50664 | 1mpp |
| 26878 | 48724 | 50629 | 50630 | 50646 | 50667 | 50670 | 1smr |
| 26951 | 48724 | 50728 | 50729 | 50730 | 50731 | 50732 | 1mai |
| 26952 | 48724 | 50728 | 50729 | 50730 | 50733 | 50734 | 1btn |
| 26954 | 48724 | 50728 | 50729 | 50730 | 50733 | 50735 | 1dro |
| 26957 | 48724 | 50728 | 50729 | 50730 | 50736 | 50737 | 2dyn |
| 26959 | 48724 | 50728 | 50729 | 50730 | 50738 | 50739 | 1btk |
| 26965 | 48724 | 50728 | 50729 | 50730 | 50740 | 50741 | 1pls |
| 26967 | 48724 | 50728 | 50729 | 50730 | 50742 | 50744 | 1dbh |
| 26969 | 48724 | 50728 | 50729 | 50730 | 50745 | 50746 | 1foe |
| 26974 | 48724 | 50728 | 50729 | 50730 | 50749 | 50750 | 1fao |
| 26976 | 48724 | 50728 | 50729 | 50730 | 50751 | 50752 | 1fgy |
| 26982 | 48724 | 50728 | 50729 | 50730 | 50753 | 50754 | 1fho |
| 27096 | 48724 | 50813 | 50814 | 50815 | 50821 | 50823 | 1dzk |
| 27113 | 48724 | 50813 | 50814 | 50815 | 50824 | 50825 | 1bj7 |
| 27115 | 48724 | 50813 | 50814 | 50815 | 50827 | 50828 | 1beb |
| 27126 | 48724 | 50813 | 50814 | 50815 | 50830 | 50831 | 1epb |
| 27151 | 48724 | 50813 | 50814 | 50815 | 50839 | 50840 | 1qft |
| 27173 | 48724 | 50813 | 50814 | 50847 | 50851 | 50852 | 1ifc |
| 27198 | 48724 | 50813 | 50814 | 50847 | 50858 | 50859 | 1mdc |
| 27214 | 48724 | 50813 | 50814 | 50847 | 50864 | 50865 | 1opa |
| 27223 | 48724 | 50813 | 50814 | 50847 | 50866 | 50867 | 1lfo |
| 27542 | 48724 | 50938 | 50939 | 50940 | 50941 | 50942 | 3sil |
| 27591 | 48724 | 50938 | 50939 | 50940 | 50943 | 50945 | 1nsc |
| 27611 | 48724 | 50938 | 50939 | 50940 | 50948 | 50949 | 2sli |
| 27649 | 48724 | 50964 | 50978 | 50979 | 50980 | 50981 | 1tbg |
| 27660 | 48724 | 50964 | 50978 | 50979 | 50983 | 50984 | 1erj |
| 27955 | 48724 | 51068 | 51069 | 51070 | 51071 | 51076 | 1znc |
| 27966 | 48724 | 51068 | 51069 | 51070 | 51071 | 51079 | 1kop |
| 28022 | 48724 | 51125 | 51126 | 51127 | 51128 | 51131 | 1bn8 |
| 28024 | 48724 | 51125 | 51126 | 51127 | 51128 | 51132 | 1ee6 |
| 28028 | 48724 | 51125 | 51126 | 51133 | 51134 | 51136 | 1qcx |
| 28029 | 48724 | 51125 | 51126 | 51137 | 51138 | 51139 | 1rmg |
| 28030 | 48724 | 51125 | 51126 | 51137 | 51140 | 51141 | 1bhe |
| 28090 | 48724 | 51181 | 51182 | 51187 | 51188 | 51190 | 1dgw |
| 28091 | 48724 | 51181 | 51182 | 51187 | 51188 | 51190 | 1dgw |
| 28124 | 48724 | 51181 | 51197 | 51198 | 51201 | 51202 | 1dcs |
| 28349 | 48724 | 51268 | 51283 | 51284 | 51285 | 51286 | 1euw |
| 28353 | 48724 | 51268 | 51283 | 51284 | 51285 | 51287 | 1f7d |
| 28388 | 48724 | 51305 | 51306 | 51312 | 51313 | 51314 | 1b12 |
| 28392 | 48724 | 51315 | 51316 | 51317 | 51318 | 51319 | 1hxr |
| 28515 | 51349 | 51350 | 51351 | 51352 | 51353 | 51364 | 1aw1 |
| 28539 | 51349 | 51350 | 51366 | 51375 | 51376 | 51377 | 1dbt |
| 28551 | 51349 | 51350 | 51366 | 51375 | 51376 | 51380 | 1dqw |
| 28559 | 51349 | 51350 | 51366 | 51381 | 51382 | 51383 | 1pii |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 28560 | 51349 | 51350 | 51366 | 51381 | 51382 | 51384 | 1nsj |
| 28563 | 51349 | 51350 | 51366 | 51381 | 51385 | 51386 | 1pii |
| 28564 | 51349 | 51350 | 51366 | 51381 | 51385 | 51387 | 1a53 |
| 28570 | 51349 | 51350 | 51366 | 51381 | 51388 | 51389 | 1qop |
| 28589 | 51349 | 51350 | 51366 | 51381 | 51388 | 51390 | 1geq |
| 28597 | 51349 | 51350 | 51395 | 51396 | 51397 | 51399 | 1ep3 |
| 28600 | 51349 | 51350 | 51395 | 51396 | 51397 | 51400 | 1d3g |
| 28604 | 51349 | 51350 | 51395 | 51396 | 51401 | 51402 | 1oya |
| 28614 | 51349 | 51350 | 51395 | 51396 | 51406 | 51407 | 1djq |
| 28663 | 51349 | 51350 | 51419 | 51427 | 51428 | 51429 | 1ct5 |
| 28804 | 51349 | 51350 | 51445 | 51487 | 51488 | 51489 | 1xyz |
| 28807 | 51349 | 51350 | 51445 | 51487 | 51491 | 51492 | 1edg |
| 28808 | 51349 | 51350 | 51445 | 51487 | 51493 | 51494 | 1ceo |
| 28813 | 51349 | 51350 | 51445 | 51487 | 51497 | 51498 | 1ece |
| 28817 | 51349 | 51350 | 51445 | 51487 | 51499 | 51500 | 7a3h |
| 28831 | 51349 | 51350 | 51445 | 51487 | 51502 | 51503 | 1bqc |
| 28834 | 51349 | 51350 | 51445 | 51487 | 51502 | 51504 | 1qnr |
| 28889 | 51349 | 51350 | 51445 | 51487 | 51512 | 51513 | 1bhg |
| 28946 | 51349 | 51350 | 51445 | 51521 | 51526 | 51527 | 1pbg |
| 28988 | 51349 | 51350 | 51445 | 51534 | 51537 | 51538 | 1nar |
| 28989 | 51349 | 51350 | 51445 | 51534 | 51537 | 51539 | 1cnv |
| 28990 | 51349 | 51350 | 51445 | 51534 | 51540 | 51541 | 2ebn |
| 28991 | 51349 | 51350 | 51445 | 51534 | 51540 | 51542 | 1eok |
| 28993 | 51349 | 51350 | 51445 | 51534 | 51540 | 51543 | 1edt |
| 29002 | 51349 | 51350 | 51445 | 51534 | 51544 | 51545 | 1edq |
| 29014 | 51349 | 51350 | 51556 | 51557 | 51558 | 51559 | 1a4m |
| 29079 | 51349 | 51350 | 51569 | 51570 | 51571 | 51573 | 1f74 |
| 29169 | 51349 | 51350 | 51569 | 51570 | 51588 | 51589 | 1onr |
| 29310 | 51349 | 51350 | 51621 | 51638 | 51639 | 51640 | 1dxe |
| 29800 | 51349 | 51734 | 51735 | 51751 | 51752 | 51754 | 1ek6 |
| 29805 | 51349 | 51734 | 51735 | 51751 | 51757 | 51758 | 1e6u |
| 29814 | 51349 | 51734 | 51735 | 51751 | 51761 | 51762 | 1eq2 |
| 29825 | 51349 | 51734 | 51735 | 51751 | 51765 | 51766 | 1cyd |
| 29829 | 51349 | 51734 | 51735 | 51751 | 51767 | 51768 | 1oaa |
| 29856 | 51349 | 51734 | 51735 | 51751 | 51774 | 51775 | 1fmc |
| 29862 | 51349 | 51734 | 51735 | 51751 | 51776 | 51777 | 1hdc |
| 29870 | 51349 | 51734 | 51735 | 51751 | 51778 | 51779 | 1fjh |
| 29874 | 51349 | 51734 | 51735 | 51751 | 51780 | 51781 | 1bdb |
| 29888 | 51349 | 51734 | 51735 | 51751 | 51786 | 51787 | 1geg |
| 29907 | 51349 | 51734 | 51735 | 51751 | 51791 | 51792 | 1d7o |
| 29935 | 51349 | 51734 | 51735 | 51751 | 51791 | 51794 | 1qsg |
| 29945 | 51349 | 51734 | 51735 | 51751 | 51795 | 51797 | 2ae2 |
| 30473 | 51349 | 51904 | 51905 | 51943 | 51944 | 51946 | 1ges |
| 30474 | 51349 | 51904 | 51905 | 51943 | 51944 | 51946 | 1ges |
| 30489 | 51349 | 51904 | 51905 | 51943 | 51947 | 51948 | 1fec |
| 30490 | 51349 | 51904 | 51905 | 51943 | 51947 | 51948 | 1fec |
| 30529 | 51349 | 51904 | 51905 | 51943 | 51950 | 51951 | 1trb |
| 30530 | 51349 | 51904 | 51905 | 51943 | 51950 | 51951 | 1trb |
| 30546 | 51349 | 51904 | 51905 | 51943 | 51950 | 51952 | 1vdc |
| 30549 | 51349 | 51904 | 51905 | 51943 | 51955 | 51956 | 1nhp |
| 30550 | 51349 | 51904 | 51905 | 51943 | 51955 | 51956 | 1nhp |
| 30565 | 51349 | 51904 | 51905 | 51943 | 51957 | 51958 | 1d7y |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 30567 | 51349 | 51904 | 51905 | 51943 | 51959 | 51960 | 1lvl |
| 30568 | 51349 | 51904 | 51905 | 51943 | 51959 | 51960 | 1lvl |
| 30574 | 51349 | 51904 | 51905 | 51943 | 51959 | 51962 | 3lad |
| 30578 | 51349 | 51904 | 51905 | 51943 | 51959 | 51963 | 1ebd |
| 30581 | 51349 | 51904 | 51905 | 51943 | 51959 | 51964 | 1ojt |
| 30582 | 51349 | 51904 | 51905 | 51943 | 51959 | 51964 | 1ojt |
| 30593 | 51349 | 51904 | 51905 | 51943 | 51966 | 51967 | 1fcd |
| 30594 | 51349 | 51904 | 51905 | 51943 | 51966 | 51967 | 1fcd |
| 30903 | 51349 | 52095 | 52096 | 52103 | 52104 | 52105 | 1nzy |
| 30918 | 51349 | 52095 | 52096 | 52103 | 52108 | 52109 | 1dci |
| 30921 | 51349 | 52095 | 52096 | 52103 | 52110 | 52111 | 1ef8 |
| 30999 | 51349 | 52128 | 52129 | 52130 | 52135 | 52136 | 1qtn |
| 31089 | 51349 | 52171 | 52172 | 52173 | 52178 | 52179 | 1a04 |
| 31095 | 51349 | 52171 | 52172 | 52173 | 52182 | 52183 | 1dbw |
| 31104 | 51349 | 52171 | 52172 | 52173 | 52184 | 52185 | 1qkk |
| 31105 | 51349 | 52171 | 52172 | 52173 | 52186 | 52187 | 1dz3 |
| 31120 | 51349 | 52171 | 52172 | 52173 | 52190 | 52191 | 1a2o |
| 31395 | 51349 | 52171 | 52309 | 52310 | 52311 | 52312 | 1f8y |
| 31438 | 51349 | 52171 | 52317 | 52325 | 52326 | 52327 | 1g2i |
| 31578 | 51349 | 52373 | 52374 | 52375 | 52380 | 52381 | 1gtr |
| 31591 | 51349 | 52373 | 52374 | 52375 | 52387 | 52388 | 1ile |
| 31593 | 51349 | 52373 | 52374 | 52375 | 52387 | 52389 | 1ffy |
| 31598 | 51349 | 52373 | 52374 | 52394 | 52395 | 52396 | 1coz |
| 31753 | 51349 | 52498 | 52499 | 100948 | 52501 | 52502 | 1nba |
| 31757 | 51349 | 52498 | 52499 | 100948 | 52504 | 52505 | 1yac |
| 31839 | 51349 | 52539 | 52540 | 52541 | 52542 | 52543 | 1gky |
| 31842 | 51349 | 52539 | 52540 | 52541 | 52546 | 52547 | 1dek |
| 31846 | 51349 | 52539 | 52540 | 52541 | 52548 | 52549 | 1cke |
| 31848 | 51349 | 52539 | 52540 | 52541 | 52550 | 52551 | 1qf9 |
| 31854 | 51349 | 52539 | 52540 | 52541 | 52552 | 52553 | 1e2k |
| 31886 | 51349 | 52539 | 52540 | 52541 | 52554 | 52556 | 1nks |
| 31892 | 51349 | 52539 | 52540 | 52541 | 52554 | 52557 | 2ak3 |
| 31898 | 51349 | 52539 | 52540 | 52541 | 52554 | 52559 | 3aky |
| 31915 | 51349 | 52539 | 52540 | 52541 | 52554 | 52562 | 1zin |
| 31930 | 51349 | 52539 | 52540 | 52541 | 52563 | 52565 | 4tmk |
| 31951 | 51349 | 52539 | 52540 | 52575 | 52582 | 52583 | 1nst |
| 31952 | 51349 | 52539 | 52540 | 52584 | 52585 | 52586 | 1a7j |
| 32002 | 51349 | 52539 | 52540 | 52592 | 52595 | 52596 | 1mh1 |
| 32018 | 51349 | 52539 | 52540 | 52592 | 52601 | 52602 | 3rab |
| 32140 | 51349 | 52539 | 52540 | 52592 | 52631 | 52632 | 1f60 |
| 32147 | 51349 | 52539 | 52540 | 52592 | 52635 | 52636 | 1g7s |
| 32152 | 51349 | 52539 | 52540 | 52592 | 52639 | 52640 | 1f5n |
| 32199 | 51349 | 52539 | 52540 | 52652 | 52653 | 52654 | 1byi |
| 32212 | 51349 | 52539 | 52540 | 52652 | 52655 | 52656 | 1qf5 |
| 32241 | 51349 | 52539 | 52540 | 52652 | 52659 | 52660 | 1eg7 |
| 32370 | 51349 | 52539 | 52540 | 52686 | 52687 | 52688 | 1b0u |
| 32373 | 51349 | 52539 | 52540 | 52686 | 52691 | 52692 | 1f2t |
| 32382 | 51349 | 52539 | 52540 | 52686 | 52695 | 52696 | 1qhl |
| 32423 | 51349 | 52539 | 52540 | 81269 | 52711 | 52712 | 1a5t |
| 32426 | 51349 | 52539 | 52540 | 81269 | 52715 | 52716 | 1fnn |
| 32428 | 51349 | 52539 | 52540 | 81269 | 52717 | 52718 | 1d2n |
| 32450 | 51349 | 52539 | 52540 | 81269 | 52721 | 52723 | 1g41 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 32503 | 51349 | 52742 | 52743 | 52744 | 52745 | 52750 | 1gci |
| 32606 | 51349 | 52767 | 52768 | 52773 | 52774 | 52775 | 1c3p |
| 32642 | 51349 | 52787 | 52788 | 52789 | 52790 | 52793 | 1d1q |
| 32651 | 51349 | 52798 | 52799 | 52800 | 52801 | 52802 | 1vhr |
| 32654 | 51349 | 52798 | 52799 | 52800 | 52803 | 52804 | 1mkp |
| 32697 | 51349 | 52798 | 52799 | 52800 | 52818 | 52819 | 1d5r |
| 32732 | 51349 | 52832 | 52833 | 52834 | 52835 | 52837 | 1thx |
| 32736 | 51349 | 52832 | 52833 | 52834 | 52835 | 52840 | 1f9m |
| 32739 | 51349 | 52832 | 52833 | 52834 | 52835 | 52841 | 1fb6 |
| 32760 | 51349 | 52832 | 52833 | 52834 | 52843 | 52844 | 1aba |
| 32767 | 51349 | 52832 | 52833 | 52834 | 52843 | 52845 | 1ego |
| 32771 | 51349 | 52832 | 52833 | 52834 | 52843 | 52847 | 1kte |
| 33066 | 51349 | 52832 | 52833 | 52901 | 52906 | 52908 | 1qmv |
| 33076 | 51349 | 52832 | 52833 | 52901 | 52909 | 52910 | 1prx |
| 33364 | 51349 | 53040 | 53056 | 53057 | 53058 | 53059 | 1ekj |
| 33373 | 51349 | 53040 | 53056 | 53057 | 53058 | 53060 | 1ddz |
| 33572 | 51349 | 53066 | 53098 | 53099 | 53103 | 53104 | 1eke |
| 33681 | 51349 | 53066 | 53098 | 53118 | 53119 | 53120 | 1kfs |
| 33695 | 51349 | 53066 | 53098 | 53118 | 53119 | 53121 | 1qtm |
| 33721 | 51349 | 53066 | 53098 | 53118 | 53125 | 53128 | 1tgo |
| 33763 | 51349 | 53162 | 53167 | 53168 | 53169 | 53171 | 3pnp |
| 33790 | 51349 | 53162 | 53167 | 53168 | 53174 | 53175 | 1cb0 |
| 33994 | 51349 | 53253 | 53254 | 53258 | 53263 | 53265 | 1qfx |
| 33996 | 51349 | 53253 | 53254 | 53258 | 53263 | 53266 | 1dkl |
| 34011 | 51349 | 53270 | 53271 | 53272 | 53273 | 53274 | 1nul |
| 34027 | 51349 | 53270 | 53271 | 53272 | 53275 | 53276 | 1hgx |
| 34029 | 51349 | 53270 | 53271 | 53272 | 53275 | 53277 | 1fsg |
| 34053 | 51349 | 53270 | 53271 | 53272 | 53278 | 53280 | 1ecf |
| 34067 | 51349 | 53270 | 53271 | 53272 | 53281 | 53282 | 1dqn |
| 34083 | 51349 | 53270 | 53271 | 53272 | 53286 | 53287 | 1tc1 |
| 34087 | 51349 | 53270 | 53271 | 53272 | 53288 | 53289 | 1qb7 |
| 34095 | 51349 | 53270 | 53271 | 53272 | 109612 | 53294 | 1a3c |
| 34098 | 51349 | 53270 | 53271 | 53272 | 53293 | 53295 | 1bd3 |
| 34179 | 51349 | 53334 | 53335 | 53339 | 53340 | 53341 | 1ej0 |
| 34183 | 51349 | 53334 | 53335 | 53348 | 53349 | 53350 | 1xva |
| 34244 | 51349 | 53334 | 53335 | 88786 | 53371 | 53372 | 1dct |
| 34247 | 51349 | 53334 | 53335 | 88786 | 53375 | 53376 | 1g55 |
| 34276 | 51349 | 53382 | 53383 | 53384 | 53385 | 53388 | 1ajs |
| 34385 | 51349 | 53382 | 53383 | 53384 | 53395 | 53396 | 1bw0 |
| 34395 | 51349 | 53382 | 53383 | 53402 | 53403 | 53404 | 1cl1 |
| 34399 | 51349 | 53382 | 53383 | 53402 | 53405 | 53406 | 1cs1 |
| 34403 | 51349 | 53382 | 53383 | 53402 | 53405 | 53407 | 1qgn |
| 34411 | 51349 | 53382 | 53383 | 53402 | 53408 | 53409 | 1d2f |
| 34413 | 51349 | 53382 | 53383 | 53402 | 53410 | 53411 | 1c7n |
| 34429 | 51349 | 53382 | 53383 | 53402 | 53412 | 53413 | 1eg5 |
| 34434 | 51349 | 53382 | 53383 | 53402 | 53415 | 53416 | 1elu |
| 34446 | 51349 | 53382 | 53383 | 53417 | 53420 | 53421 | 2gsa |
| 34488 | 51349 | 53382 | 53383 | 53417 | 53434 | 53435 | 1b9h |
| 34490 | 51349 | 53382 | 53383 | 53417 | 53436 | 53437 | 1bs0 |
| 34505 | 51349 | 53447 | 53448 | 53449 | 53450 | 53451 | 1qg8 |
| 34514 | 51349 | 53447 | 53448 | 68901 | 53456 | 53457 | 1eyr |
| 34524 | 51349 | 53447 | 53448 | 53464 | 53465 | 53466 | 1fxo |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 34693 | 51349 | 53473 | 53474 | 53525 | 53528 | 53529 | 1ehy |
| 34697 | 51349 | 53473 | 53474 | 53525 | 53528 | 53530 | 1qo7 |
| 34711 | 51349 | 53473 | 53474 | 53542 | 53543 | 53544 | 1tht |
| 34713 | 51349 | 53473 | 53474 | 53542 | 53545 | 53546 | 1ei9 |
| 34795 | 51349 | 53473 | 53474 | 53577 | 53578 | 53584 | 1bu8 |
| 34810 | 51349 | 53589 | 53590 | 53591 | 53592 | 53593 | 2mas |
| 34820 | 51349 | 53596 | 53597 | 53598 | 53599 | 53600 | 1ra9 |
| 34880 | 51349 | 53596 | 53597 | 53598 | 53599 | 53601 | 3dfr |
| 34885 | 51349 | 53596 | 53597 | 53598 | 53599 | 53602 | 1df7 |
| 34889 | 51349 | 53596 | 53597 | 53598 | 53599 | 53603 | 1d1g |
| 34893 | 51349 | 53596 | 53597 | 53598 | 53599 | 53604 | 1vdr |
| 34923 | 51349 | 53596 | 53597 | 53598 | 53605 | 53609 | 1aoe |
| 34928 | 51349 | 53612 | 53613 | 53614 | 53615 | 53616 | 1rkd |
| 34935 | 51349 | 53612 | 53613 | 53614 | 53617 | 53618 | 1bx4 |
| 35030 | 51349 | 53648 | 53649 | 53653 | 53654 | 53655 | 1auk |
| 35032 | 51349 | 53648 | 53649 | 53653 | 53656 | 53657 | 1fsu |
| 35061 | 51349 | 53658 | 53659 | 53660 | 53661 | 53666 | 1cnz |
| 35270 | 51349 | 53685 | 53686 | 53687 | 53688 | 53689 | 1qop |
| 35297 | 51349 | 53685 | 53686 | 53687 | 53692 | 53693 | 1tdj |
| 35335 | 51349 | 53719 | 53720 | 53721 | 53722 | 53723 | 1ad3 |
| 35369 | 51349 | 53719 | 53720 | 53721 | 53722 | 53729 | 1euh |
| 35385 | 51349 | 53719 | 53720 | 53721 | 53722 | 53730 | 1ez0 |
| 35399 | 51349 | 53731 | 53732 | 53733 | 53734 | 53736 | 1aco |
| 35575 | 51349 | 53783 | 53784 | 53785 | 53786 | 53787 | 1pfk |
| 35587 | 51349 | 53789 | 53790 | 53791 | 53792 | 53793 | 1cbf |
| 36020 | 51349 | 53926 | 53927 | 53928 | 53929 | 53930 | 1aln |
| 36021 | 51349 | 53926 | 53927 | 53928 | 53929 | 53930 | 1aln |
| 36597 | 53931 | 53954 | 53955 | 53960 | 53975 | 53977 | 1b9o |
| 37032 | 53931 | 54000 | 54001 | 54002 | 54017 | 54018 | 3gcb |
| 37035 | 53931 | 54000 | 54001 | 54002 | 54017 | 54019 | 2cb5 |
| 37065 | 53931 | 54000 | 54001 | 54002 | 54026 | 54027 | 1cs8 |
| 37085 | 53931 | 54000 | 54001 | 54002 | 54031 | 54032 | 1deu |
| 37089 | 53931 | 54000 | 54001 | 54002 | 54033 | 54034 | 1cv8 |
| 37090 | 53931 | 54000 | 54001 | 54002 | 54035 | 54036 | 1dki |
| 37121 | 53931 | 54000 | 54001 | 54047 | 54048 | 54049 | 1e2t |
| 37163 | 53931 | 54075 | 54076 | 54077 | 54078 | 54079 | 1dy5 |
| 37278 | 53931 | 54075 | 54076 | 54077 | 54084 | 54083 | 1onc |
| 37303 | 53931 | 54075 | 54076 | 54077 | 54094 | 54096 | 1agi |
| 37410 | 53931 | 54116 | 54117 | 54118 | 54132 | 54133 | 1b3a |
| 37420 | 53931 | 54116 | 54117 | 54118 | 54134 | 54135 | 1dok |
| 37428 | 53931 | 54116 | 54117 | 54118 | 54138 | 54139 | 1el0 |
| 37432 | 53931 | 54116 | 54117 | 54118 | 54142 | 54143 | 1eig |
| 37436 | 53931 | 54116 | 54117 | 54118 | 54146 | 54147 | 1f2l |
| 37441 | 53931 | 54116 | 54117 | 54118 | 54148 | 54149 | 1tvx |
| 37499 | 53931 | 54196 | 54197 | 54198 | 54201 | 54202 | 1fit |
| 37507 | 53931 | 54196 | 54197 | 54198 | 54203 | 54204 | 1kpf |
| 37642 | 53931 | 54235 | 54285 | 54286 | 54287 | 54288 | 1fm0 |
| 37676 | 53931 | 54235 | 54292 | 54293 | 54294 | 54302 | 1doi |
| 37683 | 53931 | 54235 | 54292 | 54293 | 54294 | 54309 | 1b9r |
| 37998 | 53931 | 54402 | 54403 | 54407 | 54408 | 54409 | 1eqk |
| 38006 | 53931 | 54402 | 54403 | 54407 | 54414 | 54415 | 1stf |
| 38142 | 53931 | 54446 | 54447 | 54448 | 54449 | 54450 | 1pcf |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 38431 | 53931 | 54533 | 54534 | 54535 | 54547 | 54548 | 1pin |
| 38470 | 53931 | 54592 | 54593 | 54594 | 54595 | 54596 | 1qip |
| 38484 | 53931 | 54592 | 54593 | 54594 | 54595 | 54597 | 1f9z |
| 38629 | 53931 | 54664 | 54690 | 54691 | 54692 | 54693 | 1fm0 |
| 39167 | 53931 | 54861 | 54928 | 54929 | 54936 | 54937 | 2u2f |
| 39168 | 53931 | 54861 | 54928 | 54929 | 54936 | 54937 | 1u2f |
| 39183 | 53931 | 54861 | 54928 | 54929 | 54942 | 54943 | 1fxl |
| 39184 | 53931 | 54861 | 54928 | 54929 | 54942 | 54943 | 1fxl |
| 39189 | 53931 | 54861 | 54928 | 54929 | 54946 | 54947 | 2mst |
| 39191 | 53931 | 54861 | 54928 | 54929 | 54948 | 54949 | 1cvj |
| 39192 | 53931 | 54861 | 54928 | 54929 | 54948 | 54949 | 1cvj |
| 39210 | 53931 | 54861 | 54928 | 54929 | 54952 | 54953 | 1fje |
| 39212 | 53931 | 54861 | 54928 | 54929 | 54952 | 54953 | 1fjc |
| 39299 | 53931 | 54861 | 54975 | 54976 | 54977 | 54979 | 1aps |
| 39341 | 53931 | 54861 | 55008 | 55009 | 55012 | 55013 | 2aw0 |
| 39343 | 53931 | 54861 | 55008 | 55009 | 55014 | 55015 | 1cc8 |
| 39345 | 53931 | 54861 | 55008 | 55009 | 55014 | 55016 | 1fe0 |
| 39351 | 53931 | 54861 | 55008 | 55009 | 55017 | 55018 | 1cpz |
| 39352 | 53931 | 54861 | 55008 | 55009 | 55019 | 55020 | 1qup |
| 39408 | 53931 | 54861 | 55068 | 55069 | 55070 | 55071 | 1fvg |
| 39862 | 53931 | 55330 | 55331 | 55339 | 55344 | 55345 | 1dpt |
| 40151 | 53931 | 55417 | 55418 | 55419 | 55420 | 55422 | 1ap8 |
| 40310 | 53931 | 55485 | 55486 | 55516 | 55517 | 55518 | 1ast |
| 40663 | 53931 | 55619 | 55620 | 55628 | 55629 | 55630 | 1dhn |
| 40665 | 53931 | 55619 | 55620 | 55628 | 55631 | 55632 | 1b9l |
| 40675 | 53931 | 55636 | 55637 | 55638 | 55639 | 55640 | 1puc |
| 40708 | 53931 | 55680 | 55681 | 55682 | 55683 | 55684 | 1set |
| 40714 | 53931 | 55680 | 55681 | 55682 | 55685 | 55686 | 1e1o |
| 40745 | 53931 | 55680 | 55681 | 55682 | 55692 | 55693 | 1ati |
| 40759 | 53931 | 55680 | 55681 | 55682 | 55694 | 55695 | 1qf6 |
| 40760 | 53931 | 55680 | 55681 | 55682 | 55696 | 55697 | 1eov |
| 40765 | 53931 | 55680 | 55681 | 55682 | 55696 | 55698 | 1b8a |
| 40783 | 53931 | 55680 | 55681 | 55682 | 55705 | 55706 | 12as |
| 40798 | 53931 | 55728 | 55729 | 55730 | 55733 | 55734 | 1bo4 |
| 40808 | 53931 | 55728 | 55729 | 55730 | 55740 | 55741 | 1qsm |
| 40816 | 53931 | 55728 | 55729 | 55730 | 55742 | 55743 | 1bob |
| 40817 | 53931 | 55728 | 55729 | 55730 | 55744 | 55745 | 1fy7 |
| 40818 | 53931 | 55728 | 55729 | 55730 | 55746 | 55747 | 1cjw |
| 40857 | 53931 | 55752 | 55753 | 55762 | 55763 | 55764 | 1f7s |
| 40858 | 53931 | 55752 | 55753 | 55762 | 55763 | 55765 | 1cfy |
| 40888 | 53931 | 55769 | 55770 | 55771 | 55772 | 55777 | 1ypr |
| 40891 | 53931 | 55769 | 55770 | 55771 | 55772 | 55779 | 3nul |
| 41029 | 53931 | 55830 | 55831 | 55832 | 55833 | 55836 | 1bkp |
| 41056 | 53931 | 55830 | 55831 | 55832 | 55843 | 55844 | 1b5e |
| 41062 | 53931 | 55845 | 55846 | 55847 | 55848 | 55849 | 1lba |
| 41065 | 53931 | 55855 | 55856 | 55857 | 55858 | 55859 | 1cyo |
| 41085 | 53931 | 55855 | 55856 | 55857 | 55862 | 55863 | 1cxy |
| 41104 | 53931 | 55873 | 55874 | 55879 | 55880 | 55881 | 1ei1 |
| 41107 | 53931 | 55873 | 55874 | 55879 | 55882 | 55883 | 1b63 |
| 41321 | 53931 | 55944 | 55961 | 55966 | 55967 | 55968 | 1em2 |
| 41450 | 53931 | 56023 | 56024 | 56025 | 56026 | 56027 | 1byr |
| 41634 | 53931 | 56111 | 56112 | 88854 | 56120 | 56121 | 1a06 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 41635 | 53931 | 56111 | 56112 | 88854 | 56122 | 56123 | 1phk |
| 41638 | 53931 | 56111 | 56112 | 88854 | 56124 | 56125 | 1tki |
| 41640 | 53931 | 56111 | 56112 | 88854 | 56126 | 56127 | 1kob |
| 41664 | 53931 | 56111 | 56112 | 88854 | 56139 | 56141 | 1csn |
| 41685 | 53931 | 56111 | 56112 | 88854 | 56153 | 56154 | 1qpc |
| 41766 | 53931 | 56203 | 56204 | 56205 | 56206 | 56207 | 1c4z |
| 41779 | 53931 | 56218 | 56219 | 56220 | 56221 | 56222 | 1ako |
| 41780 | 53931 | 56218 | 56219 | 56220 | 56223 | 56224 | 1hd7 |
| 41876 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41877 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41878 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41879 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41880 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41888 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41889 | 53931 | 56234 | 56235 | 56251 | 56252 | 56254 | 1ryp |
| 41932 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41933 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41934 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41935 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41936 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41937 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41938 | 53931 | 56234 | 56235 | 56251 | 56255 | 56257 | 1ryp |
| 41989 | 53931 | 56234 | 56235 | 56251 | 56258 | 56260 | 1g3k |
| 42004 | 53931 | 56234 | 56235 | 56261 | 56262 | 56263 | 1apy |
| 42012 | 53931 | 56234 | 56235 | 56261 | 56262 | 56264 | 2gac |
| 42057 | 53931 | 56280 | 56281 | 56288 | 56289 | 56290 | 1qh5 |
| 42078 | 53931 | 56299 | 56300 | 56301 | 56304 | 56306 | 1ute |
| 42088 | 53931 | 56321 | 56322 | 56323 | 56324 | 56325 | 1qdl |
| 42177 | 53931 | 56370 | 56371 | 56372 | 56373 | 56374 | 1mrj |
| 42197 | 53931 | 56370 | 56371 | 56372 | 56385 | 56386 | 1d6a |
| 42222 | 53931 | 56370 | 56371 | 56372 | 56391 | 56392 | 1hwm |
| 42324 | 53931 | 56435 | 56436 | 56437 | 56438 | 56439 | 1qdd |
| 42327 | 53931 | 56435 | 56436 | 56437 | 56442 | 56443 | 1.00E+87 |
| 42333 | 53931 | 56435 | 56436 | 56437 | 56444 | 56445 | 1egg |
| 42349 | 53931 | 56435 | 56436 | 56437 | 56450 | 56451 | 2afp |
| 42350 | 53931 | 56435 | 56436 | 56437 | 56452 | 56453 | 1qo3 |
| 42352 | 53931 | 56435 | 56436 | 56437 | 56454 | 56455 | 1dv8 |
| 42419 | 53931 | 56435 | 56436 | 56437 | 56463 | 56464 | 1byf |
| 42423 | 53931 | 56435 | 56436 | 56437 | 56465 | 56466 | 1tn3 |
| 42618 | 56572 | 56573 | 56574 | 56575 | 56576 | 56577 | 1hle |
| 42623 | 56572 | 56573 | 56574 | 56575 | 56580 | 56581 | 1as4 |
| 42955 | 56572 | 56654 | 56655 | 56656 | 56663 | 56664 | 2hhm |
| 42972 | 56572 | 56654 | 56655 | 56656 | 56667 | 56668 | 1inp |
| 43270 | 56572 | 56751 | 56752 | 56753 | 56754 | 56755 | 1daa |
| 43349 | 56572 | 56800 | 56801 | 56802 | 56803 | 56804 | 1lci |
| 43351 | 56572 | 56800 | 56801 | 56802 | 56805 | 56806 | 1amu |
| 43358 | 56572 | 56814 | 56815 | 56816 | 56817 | 56819 | 1epu |
| 43398 | 56835 | 56836 | 56854 | 56855 | 56859 | 56860 | 2bid |
| 43400 | 56835 | 56836 | 56854 | 56855 | 56862 | 56863 | 1f16 |
| 43627 | 56835 | 81443 | 81442 | 81441 | 81440 | 81439 | 1fft |
| 43651 | 56835 | 81339 | 81338 | 56895 | 56898 | 56899 | 1fx8 |
| 43988 | 56992 | 56993 | 56994 | 56995 | 57000 | 57001 | 6rlx |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 43997 | 56992 | 56993 | 56994 | 56995 | 57004 | 57005 | 1bon |
| 44102 | 56992 | 57015 | 57059 | 57072 | 57073 | 57074 | 1agg |
| 44108 | 56992 | 57015 | 57059 | 57072 | 57075 | 57076 | 1eit |
| 44109 | 56992 | 57015 | 57059 | 57072 | 57077 | 57078 | 1axh |
| 44111 | 56992 | 57015 | 57059 | 57072 | 57077 | 57078 | 1vtx |
| 44112 | 56992 | 57015 | 57059 | 57072 | 57079 | 57080 | 1dl0 |
| 44115 | 56992 | 57015 | 57059 | 57072 | 57085 | 57086 | 1i25 |
| 44116 | 56992 | 57015 | 57059 | 57072 | 57087 | 57088 | 1qk7 |
| 44119 | 56992 | 57015 | 57059 | 57072 | 57093 | 57094 | 1c6w |
| 44120 | 56992 | 57015 | 57095 | 57096 | 57097 | 57098 | 2sn3 |
| 44127 | 56992 | 57015 | 57095 | 57096 | 57097 | 57102 | 1aho |
| 44144 | 56992 | 57015 | 57095 | 57096 | 57114 | 57115 | 1bmr |
| 44154 | 56992 | 57015 | 57095 | 57116 | 57131 | 57132 | 1bah |
| 44156 | 56992 | 57015 | 57095 | 57116 | 57135 | 57136 | 1agt |
| 44161 | 56992 | 57015 | 57095 | 57116 | 57143 | 57144 | 1sis |
| 44162 | 56992 | 57015 | 57095 | 57116 | 57145 | 57146 | 1pnh |
| 44398 | 56992 | 57301 | 57302 | 57303 | 57306 | 57307 | 1tgx |
| 44403 | 56992 | 57301 | 57302 | 57303 | 57308 | 57309 | 1fas |
| 44420 | 56992 | 57301 | 57302 | 57303 | 57318 | 57319 | 2ctx |
| 44435 | 56992 | 57301 | 57302 | 57303 | 57328 | 57329 | 1f94 |
| 44549 | 56992 | 57361 | 57362 | 57363 | 57372 | 57373 | 1bik |
| 44551 | 56992 | 57361 | 57362 | 57363 | 57374 | 57375 | 1dtx |
| 44552 | 56992 | 57361 | 57362 | 57363 | 57376 | 57377 | 1bun |
| 44553 | 56992 | 57361 | 57362 | 57363 | 57378 | 57379 | 1shp |
| 44706 | 56992 | 100894 | 100895 | 57468 | 57473 | 57475 | 1tgs |
| 44715 | 56992 | 100894 | 100895 | 57468 | 57478 | 57479 | 1nub |
| 44721 | 56992 | 100894 | 100895 | 57468 | 57482 | 57483 | 1pce |
| 44722 | 56992 | 100894 | 100895 | 57468 | 57484 | 57485 | 1ldt |
| 44809 | 56992 | 57500 | 57501 | 57528 | 63395 | 63397 | 1hcn |
| 44810 | 56992 | 57500 | 57501 | 57528 | 63396 | 63398 | 1hcn |
| 44998 | 56992 | 57666 | 57667 | 57668 | 57669 | 57670 | 1a1i |
| 45025 | 56992 | 57666 | 57667 | 57668 | 57671 | 57672 | 1rmd |
| 45026 | 56992 | 57666 | 57667 | 57668 | 57673 | 57674 | 2drp |
| 45027 | 56992 | 57666 | 57667 | 57668 | 57673 | 57674 | 2drp |
| 45036 | 56992 | 57666 | 57667 | 57668 | 57677 | 57678 | 1znf |
| 45040 | 56992 | 57666 | 57667 | 57668 | 57681 | 57682 | 1ncs |
| 45041 | 56992 | 57666 | 57667 | 57668 | 57683 | 57684 | 2gli |
| 45042 | 56992 | 57666 | 57667 | 57668 | 57683 | 57684 | 2gli |
| 45044 | 56992 | 57666 | 57667 | 57668 | 57683 | 57684 | 2gli |
| 45046 | 56992 | 57666 | 57667 | 57668 | 57685 | 57686 | 1bbo |
| 45052 | 56992 | 57666 | 57667 | 57668 | 57689 | 57690 | 1bhi |
| 45053 | 56992 | 57666 | 57667 | 57668 | 57691 | 57692 | 1ubd |
| 45057 | 56992 | 57666 | 57667 | 57668 | 57693 | 57694 | 1tf3 |
| 45059 | 56992 | 57666 | 57667 | 57668 | 57693 | 57694 | 1tf3 |
| 45072 | 56992 | 57666 | 57667 | 57668 | 57695 | 57696 | 1yuj |
| 45212 | 56992 | 57769 | 57798 | 57799 | 57800 | 57801 | 1qf8 |
| 45321 | 56992 | 57849 | 57850 | 57851 | 57852 | 57853 | 1fbv |
| 45322 | 56992 | 57849 | 57850 | 57851 | 57854 | 57855 | 1rmd |
| 45323 | 56992 | 57849 | 57850 | 57851 | 57856 | 57857 | 1chc |
| 45324 | 56992 | 57849 | 57850 | 57851 | 57858 | 57859 | 1bor |
| 45325 | 56992 | 57849 | 57850 | 57851 | 57860 | 57861 | 1g25 |
| 45331 | 56992 | 57867 | 57868 | 57869 | 57870 | 57873 | 4mt2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 45337 | 56992 | 57867 | 57868 | 57869 | 57870 | 57875 | 1dmc |
| 45339 | 56992 | 57867 | 57868 | 57869 | 57870 | 57875 | 1dmf |
| 45345 | 56992 | 57867 | 57868 | 57869 | 57870 | 57877 | 1qjk |
| 45346 | 56992 | 57867 | 57868 | 57869 | 57870 | 57877 | 1qjl |
| 59161 | 46456 | 46625 | 46626 | 46627 | 63459 | 63460 | 1.00E+29 |
| 59197 | 46456 | 46688 | 46689 | 46690 | 46699 | 46700 | 1e3o |
| 59226 | 51349 | 51350 | 51445 | 51521 | 51522 | 51523 | 1e4m |
| 59262 | 51349 | 53253 | 53254 | 53255 | 53256 | 64110 | 1.00E+58 |
| 59272 | 48724 | 50813 | 50814 | 50815 | 63805 | 63806 | 1e5p |
| 59283 | 51349 | 53685 | 53686 | 53687 | 64172 | 64173 | 1e5x |
| 59296 | 51349 | 52539 | 52540 | 52566 | 52567 | 52568 | 1e6c |
| 59373 | 51349 | 51734 | 51735 | 51751 | 51769 | 63926 | 1e7w |
| 59393 | 48724 | 50198 | 50324 | 50325 | 50326 | 50327 | 1e9g |
| 59401 | 48724 | 48725 | 48726 | 48727 | 48730 | 48731 | 1eaj |
| 59414 | 51349 | 52373 | 52374 | 52397 | 52400 | 63975 | 1ej2 |
| 59454 | 48724 | 50493 | 50494 | 50514 | 63789 | 63790 | 1elv |
| 59486 | 53931 | 54533 | 54534 | 54535 | 64248 | 64249 | 1eq3 |
| 59507 | 46456 | 47239 | 47240 | 47241 | 63524 | 63525 | 1eum |
| 59642 | 46456 | 47265 | 47266 | 47267 | 63530 | 63531 | 1f45 |
| 59662 | 51349 | 64004 | 64005 | 64006 | 64007 | 64008 | 1f75 |
| 59674 | 51349 | 52373 | 52374 | 52375 | 52392 | 52393 | 1f7u |
| 59689 | 48724 | 49451 | 49472 | 49473 | 49474 | 49475 | 1f86 |
| 59694 | 48724 | 50938 | 50939 | 50940 | 50943 | 50944 | 1f8e |
| 59760 | 51349 | 53382 | 53383 | 53417 | 64128 | 64129 | 1fc4 |
| 59771 | 53931 | 54533 | 54534 | 54535 | 64250 | 64251 | 1fd9 |
| 59803 | 46456 | 46688 | 46689 | 46739 | 63467 | 63468 | 1fex |
| 59819 | 51349 | 53382 | 53383 | 53384 | 64121 | 64122 | 1fg7 |
| 59869 | 51349 | 51904 | 51905 | 51943 | 51953 | 63949 | 1fl2 |
| 59870 | 51349 | 51904 | 51905 | 51943 | 51953 | 63949 | 1fl2 |
| 59879 | 51349 | 52539 | 52540 | 52575 | 64015 | 64016 | 1fmj |
| 59922 | 51349 | 52832 | 52833 | 52834 | 64054 | 64055 | 1fo5 |
| 59927 | 51349 | 53447 | 53448 | 64136 | 64137 | 64138 | 1fo8 |
| 59976 | 51349 | 52798 | 52799 | 52800 | 64049 | 64050 | 1fpz |
| 60047 | 53931 | 56111 | 56112 | 88854 | 64406 | 64407 | 1fvr |
| 60058 | 51349 | 53747 | 53748 | 53749 | 53750 | 53751 | 1fw8 |
| 60187 | 53931 | 55810 | 55811 | 55812 | 64365 | 64366 | 1g0s |
| 60227 | 46456 | 48507 | 48508 | 48509 | 48534 | 63626 | 1g2n |
| 60228 | 51349 | 53162 | 53167 | 53168 | 53169 | 64098 | 1g2o |
| 60237 | 51349 | 52539 | 52540 | 52652 | 64019 | 64021 | 1g3q |
| 60264 | 51349 | 53040 | 53056 | 53057 | 53058 | 64084 | 1g5c |
| 60271 | 53931 | 55680 | 55681 | 55682 | 64349 | 64350 | 1g5h |
| 60288 | 51349 | 52506 | 52507 | 52508 | 63996 | 63997 | 1g5q |
| 60318 | 51349 | 52539 | 52540 | 52686 | 64025 | 64026 | 1g6h |
| 60320 | 56992 | 57361 | 57362 | 57363 | 57364 | 57365 | 1g6x |
| 60376 | 51349 | 52539 | 52540 | 81269 | 64040 | 64041 | 1g8p |
| 60400 | 48724 | 50155 | 50156 | 50157 | 63754 | 63755 | 1g9o |
| 60456 | 51349 | 53382 | 53383 | 53384 | 53392 | 64120 | 1gde |
| 60625 | 53931 | 55680 | 55681 | 55682 | 55688 | 55691 | 1h4v |
| 60643 | 51349 | 51734 | 51735 | 51751 | 63927 | 63928 | 1h5q |
| 60693 | 51349 | 51904 | 51905 | 51943 | 63947 | 63948 | 1h6v |
| 60694 | 51349 | 51904 | 51905 | 51943 | 63947 | 63948 | 1h6v |
| 60711 | 53931 | 55908 | 55909 | 64380 | 64381 | 64382 | 1h70 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 60716 | 51349 | 52832 | 52833 | 52834 | 64052 | 64053 | 1h75 |
| 60717 | 51349 | 53447 | 53448 | 68901 | 64143 | 64144 | 1h7e |
| 60726 | 51349 | 51350 | 51569 | 51594 | 51595 | 51596 | 1h7n |
| 60737 | 48724 | 50493 | 50494 | 50514 | 50531 | 50532 | 1h8d |
| 60793 | 53931 | 56435 | 56436 | 56437 | 64457 | 64458 | 1h8u |
| 60812 | 46456 | 46457 | 46458 | 46463 | 63438 | 63439 | 1h97 |
| 60887 | 46456 | 47026 | 47027 | 47028 | 47029 | 63498 | 1hbk |
| 60939 | 53931 | 55680 | 55681 | 55682 | 64347 | 64348 | 1hc7 |
| 60962 | 51349 | 52832 | 52833 | 52901 | 64066 | 64067 | 1hd2 |
| 60965 | 51349 | 51734 | 51735 | 51751 | 63931 | 63932 | 1hdo |
| 60970 | 48724 | 48725 | 48726 | 49159 | 49190 | 49191 | 1he7 |
| 61006 | 53931 | 55330 | 55331 | 55339 | 55340 | 64323 | 1hfo |
| 61042 | 46456 | 48370 | 48452 | 48453 | 48460 | 48461 | 1hh8 |
| 61127 | 53931 | 56435 | 56436 | 56437 | 64453 | 64454 | 1hq8 |
| 61154 | 51349 | 53658 | 53659 | 53660 | 53668 | 64171 | 1hqs |
| 61160 | 46456 | 47472 | 47473 | 63550 | 63551 | 63552 | 1hqv |
| 61254 | 46456 | 48096 | 48097 | 48098 | 63586 | 63587 | 1htj |
| 61366 | 46456 | 48370 | 48452 | 48453 | 48462 | 63617 | 1hxi |
| 61367 | 48724 | 48725 | 48726 | 48727 | 48933 | 63651 | 1hxm |
| 61369 | 48724 | 48725 | 48726 | 48727 | 48933 | 48938 | 1hxm |
| 61459 | 53931 | 55810 | 55811 | 64369 | 64370 | 64371 | 1hzt |
| 61515 | 53931 | 55728 | 55729 | 55730 | 64351 | 64352 | 1i12 |
| 61548 | 46456 | 47265 | 47266 | 47267 | 47272 | 63528 | 1i1r |
| 61568 | 51349 | 52539 | 52540 | 52592 | 52609 | 52611 | 1i2m |
| 61585 | 51349 | 53066 | 53098 | 53099 | 53103 | 64093 | 1i39 |
| 61605 | 56992 | 57923 | 57924 | 57925 | 57928 | 57929 | 1i3o |
| 61732 | 48724 | 50813 | 50814 | 50815 | 63807 | 63808 | 1i4u |
| 61767 | 51349 | 53447 | 53448 | 68901 | 64140 | 64141 | 1i52 |
| 61848 | 51349 | 53040 | 53056 | 53057 | 53058 | 64085 | 1i6p |
| 61889 | 53931 | 54494 | 54495 | 54496 | 54497 | 64243 | 1i7k |
| 61901 | 53931 | 56321 | 56322 | 56323 | 56324 | 64443 | 1i7q |
| 61979 | 46456 | 46625 | 46626 | 46627 | 46650 | 46655 | 1i8o |
| 62099 | 51349 | 52798 | 52799 | 52800 | 64047 | 64048 | 1i9s |
| 62119 | 46456 | 48096 | 48097 | 48098 | 63584 | 63585 | 1iap |
| 62162 | 51349 | 53382 | 53383 | 53402 | 53403 | 64125 | 1ibj |
| 62318 | 46456 | 48507 | 48508 | 48509 | 48528 | 48529 | 1ie9 |
| 62375 | 51349 | 53066 | 53098 | 53118 | 53125 | 53127 | 1ih7 |
| 62380 | 46456 | 48370 | 48452 | 48453 | 63618 | 63619 | 1ihg |
| 62382 | 48724 | 50155 | 50156 | 50157 | 63756 | 63757 | 1ihj |
| 62394 | 51349 | 52539 | 52540 | 52652 | 52668 | 52669 | 1ihu |
| 62395 | 51349 | 52539 | 52540 | 52652 | 52668 | 52669 | 1ihu |
| 62510 | 48724 | 50352 | 50353 | 50354 | 63779 | 63780 | 1ijt |
| 62591 | 56572 | 56573 | 56574 | 56575 | 64471 | 64472 | 1imv |
| 62598 | 51349 | 52539 | 52540 | 81269 | 52713 | 64037 | 1in4 |
| 62611 | 56572 | 64517 | 64518 | 64519 | 64520 | 64521 | 1io1 |
| 62650 | 51349 | 52539 | 52540 | 81269 | 64035 | 64036 | 1iqp |
| 62661 | 46456 | 47972 | 47973 | 47974 | 47975 | 63577 | 1iqv |
| 62669 | 48724 | 50629 | 50630 | 50646 | 50649 | 63795 | 1j71 |
| 62697 | 53931 | 56111 | 56112 | 64411 | 64412 | 64413 | 1j7l |
| 62700 | 46456 | 47472 | 47473 | 47502 | 47514 | 47515 | 1j7q |
| 62753 | 51349 | 56783 | 56784 | 64511 | 64512 | 64513 | 1j97 |
| 62817 | 51349 | 51734 | 51735 | 51751 | 63929 | 63930 | 1ja9 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 62819 | 53931 | 54494 | 54495 | 54496 | 54497 | 64241 | 1jat |
| 62836 | 48724 | 50198 | 50249 | 50263 | 50273 | 50274 | 1jb7 |
| 62837 | 48724 | 50198 | 50249 | 50263 | 50273 | 50274 | 1jb7 |
| 62838 | 48724 | 50198 | 50249 | 50263 | 50273 | 50274 | 1jb7 |
| 62839 | 48724 | 50198 | 50249 | 50263 | 50275 | 50276 | 1jb7 |
| 62845 | 51349 | 52171 | 52172 | 52173 | 52174 | 52175 | 1jbe |
| 62850 | 51349 | 53685 | 53686 | 53687 | 64174 | 64175 | 1jbq |
| 62888 | 48724 | 51068 | 51069 | 51070 | 51071 | 63844 | 1jd0 |
| 62936 | 46456 | 47112 | 47113 | 47134 | 63507 | 63508 | 1jfi |
| 62937 | 46456 | 47112 | 47113 | 47134 | 63509 | 63510 | 1jfi |
| 62940 | 51349 | 52832 | 52833 | 52901 | 64068 | 64069 | 1jfu |
| 63114 | 51349 | 52539 | 52540 | 52686 | 64029 | 64030 | 1jj7 |
| 63116 | 56992 | 57867 | 57868 | 57869 | 64581 | 64582 | 1jjd |
| 63155 | 53931 | 56275 | 56276 | 56277 | 56278 | 56279 | 1jl0 |
| 63177 | 48724 | 48725 | 48726 | 48727 | 63637 | 63638 | 1jma |
| 63223 | 56572 | 56654 | 56655 | 56656 | 64475 | 64476 | 1jp4 |
| 63234 | 51349 | 52539 | 52540 | 81269 | 64033 | 64034 | 1jql |
| 63279 | 48724 | 50198 | 50249 | 50282 | 63772 | 63773 | 1jt8 |
| 64709 | 48724 | 50198 | 50249 | 50282 | 68910 | 68911 | 1a62 |
| 64720 | 51349 | 53334 | 53335 | 53354 | 68927 | 68928 | 1dl5 |
| 64889 | 48724 | 50493 | 50494 | 50514 | 69284 | 69285 | 1eax |
| 64986 | 51349 | 52539 | 52540 | 52592 | 69483 | 69484 | 1f6b |
| 65078 | 48724 | 48725 | 48726 | 49159 | 49172 | 49173 | 1g1c |
| 65104 | 53931 | 56435 | 56436 | 56437 | 56456 | 56457 | 1g1t |
| 65117 | 51349 | 53270 | 53271 | 53272 | 53288 | 69541 | 1g2q |
| 65171 | 56992 | 57015 | 57059 | 57072 | 69928 | 69929 | 1g9p |
| 65288 | 48724 | 48725 | 48726 | 49159 | 69158 | 69159 | 1gl4 |
| 65305 | 48724 | 49763 | 49764 | 49765 | 49766 | 69208 | 1gme |
| 65407 | 46456 | 47768 | 47819 | 69044 | 69045 | 69046 | 1go3 |
| 65414 | 51349 | 51350 | 51445 | 51534 | 51546 | 51547 | 1goi |
| 65441 | 48724 | 51181 | 51197 | 51198 | 69346 | 69347 | 1gp6 |
| 65535 | 48724 | 48725 | 48726 | 49159 | 49168 | 49169 | 1gsm |
| 65536 | 48724 | 48725 | 48726 | 49159 | 49168 | 49169 | 1gsm |
| 65640 | 51349 | 52539 | 52540 | 52592 | 69485 | 69486 | 1h65 |
| 65643 | 46456 | 47575 | 47576 | 47577 | 69029 | 69030 | 1h67 |
| 65681 | 53931 | 64267 | 64268 | 64269 | 69718 | 69719 | 1h6h |
| 65706 | 53931 | 55873 | 55874 | 55879 | 69800 | 69801 | 1h7s |
| 65805 | 51349 | 53648 | 53649 | 53653 | 53656 | 69591 | 1hdh |
| 65830 | 48724 | 51125 | 51126 | 51137 | 51140 | 69332 | 1hg8 |
| 65835 | 48724 | 50198 | 50249 | 50282 | 69265 | 69266 | 1hh2 |
| 65845 | 48724 | 50493 | 50494 | 50514 | 50515 | 50516 | 1hj9 |
| 65915 | 53931 | 55752 | 55753 | 55762 | 69793 | 69794 | 1hqz |
| 65969 | 51349 | 52095 | 52096 | 52103 | 69440 | 69441 | 1hzd |
| 65975 | 51349 | 52820 | 52821 | 52822 | 69507 | 69508 | 1hzm |
| 66024 | 48724 | 50198 | 50324 | 50325 | 50326 | 50329 | 1i40 |
| 66041 | 51349 | 52373 | 52374 | 52375 | 52378 | 52379 | 1i6l |
| 66153 | 53931 | 53954 | 53955 | 53960 | 53961 | 69627 | 1iiz |
| 66205 | 51349 | 52832 | 52833 | 52834 | 69512 | 69513 | 1ilo |
| 66211 | 51349 | 52498 | 52499 | 100948 | 69465 | 69466 | 1im5 |
| 66220 | 51349 | 53334 | 53335 | 69557 | 69558 | 69559 | 1inl |
| 66258 | 51349 | 52373 | 52374 | 52375 | 52392 | 69454 | 1iq0 |
| 66365 | 46456 | 46457 | 46458 | 46463 | 68941 | 68942 | 1it2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 66438 | 53931 | 54116 | 54117 | 54118 | 69632 | 69633 | 1j8i |
| 66454 | 51349 | 51350 | 51726 | 51727 | 51728 | 69405 | 1j93 |
| 66482 | 48724 | 48725 | 48726 | 49159 | 69160 | 69161 | 1jbj |
| 66544 | 56992 | 57923 | 57924 | 57925 | 69974 | 69975 | 1jd5 |
| 66549 | 46456 | 48370 | 48371 | 48372 | 48373 | 48375 | 1jdh |
| 66581 | 51349 | 53162 | 53167 | 53168 | 53174 | 69536 | 1je0 |
| 66621 | 51349 | 52787 | 52788 | 52789 | 69504 | 69505 | 1jf8 |
| 66624 | 46456 | 48263 | 48264 | 48265 | 48270 | 48271 | 1jfb |
| 66638 | 46456 | 47472 | 47473 | 47502 | 69021 | 69022 | 1jfj |
| 66661 | 51349 | 53334 | 53335 | 53354 | 68927 | 69553 | 1jg1 |
| 66678 | 56992 | 57301 | 57302 | 57303 | 69949 | 69950 | 1jgk |
| 66714 | 48724 | 49784 | 49785 | 69210 | 69211 | 69212 | 1jhj |
| 66745 | 51349 | 52373 | 52374 | 52375 | 52376 | 69453 | 1jil |
| 66758 | 53931 | 55680 | 55681 | 55682 | 55701 | 55702 | 1jjc |
| 66763 | 53931 | 55680 | 55681 | 55682 | 55703 | 55704 | 1jjc |
| 66770 | 56572 | 56573 | 56574 | 56575 | 69871 | 69872 | 1jjo |
| 66795 | 53931 | 54402 | 54427 | 54431 | 69675 | 69676 | 1jkg |
| 66796 | 53931 | 54402 | 54427 | 54431 | 69677 | 69678 | 1jkg |
| 66808 | 53931 | 55810 | 55811 | 55812 | 64367 | 64368 | 1jkn |
| 66821 | 51349 | 53066 | 53098 | 53099 | 53100 | 53101 | 1jl1 |
| 66880 | 56992 | 57849 | 57850 | 57851 | 69968 | 69969 | 1jm7 |
| 66881 | 56992 | 57849 | 57850 | 57851 | 69970 | 69971 | 1jm7 |
| 67111 | 51349 | 69617 | 69618 | 69619 | 69620 | 69621 | 1jr2 |
| 67215 | 51349 | 53382 | 53383 | 69565 | 69566 | 69567 | 1js3 |
| 67377 | 51349 | 69571 | 69572 | 69573 | 69574 | 69575 | 1jw9 |
| 67384 | 53931 | 56435 | 56436 | 56437 | 88867 | 88872 | 1jwi |
| 67458 | 51349 | 53447 | 53448 | 68901 | 69569 | 69570 | 1jyk |
| 67834 | 51349 | 51350 | 51445 | 51487 | 51510 | 51511 | 1jz8 |
| 67940 | 53931 | 56321 | 56322 | 56323 | 69838 | 69839 | 1k0g |
| 67992 | 53931 | 74651 | 48403 | 48404 | 69091 | 69092 | 1k1a |
| 68253 | 51349 | 53719 | 53720 | 69602 | 69603 | 69604 | 1k75 |
| 68308 | 48724 | 50964 | 50978 | 50979 | 69313 | 69314 | 1k8k |
| 68333 | 46456 | 47472 | 47473 | 63550 | 47550 | 47551 | 1k94 |
| 68341 | 46456 | 47094 | 47095 | 47096 | 68982 | 68983 | 1k99 |
| 68356 | 56572 | 56573 | 56574 | 56575 | 64469 | 64470 | 1k9o |
| 68369 | 56572 | 56654 | 56655 | 56656 | 56669 | 69878 | 1ka1 |
| 68582 | 51349 | 52539 | 52540 | 52541 | 69474 | 69475 | 1kgd |
| 68597 | 51349 | 52171 | 52172 | 52173 | 52192 | 69445 | 1kgs |
| 68644 | 51349 | 52539 | 52540 | 52541 | 69476 | 69477 | 1kjw |
| 68790 | 56572 | 56795 | 56796 | 69892 | 69893 | 69895 | 1kq3 |
| 68810 | 56992 | 57015 | 57059 | 57072 | 69926 | 69927 | 1kqh |
| 68863 | 56992 | 57361 | 57362 | 57363 | 57366 | 57367 | 1kth |
| 70066 | 56992 | 57015 | 57059 | 57072 | 75666 | 75667 | 1cix |
| 70113 | 48724 | 50728 | 50729 | 50730 | 74993 | 74994 | 1eaz |
| 70150 | 53931 | 54116 | 54117 | 54118 | 75341 | 75342 | 1g2t |
| 70168 | 51349 | 53382 | 53383 | 53402 | 64126 | 75271 | 1gc0 |
| 70180 | 48724 | 50493 | 50494 | 50514 | 50586 | 50587 | 1gj7 |
| 70214 | 48724 | 50813 | 50814 | 50815 | 63807 | 63808 | 1gka |
| 70228 | 46456 | 48556 | 48557 | 48572 | 48573 | 48574 | 1gkm |
| 70341 | 51349 | 52767 | 52768 | 52769 | 75227 | 75228 | 1gq6 |
| 70372 | 48724 | 49784 | 49785 | 69210 | 69211 | 74892 | 1gqp |
| 70376 | 53931 | 54075 | 54076 | 54077 | 64205 | 64206 | 1gqv |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 70396 | 51349 | 52539 | 52540 | 52541 | 52563 | 69482 | 1gsi |
| 70441 | 51349 | 51350 | 51395 | 51396 | 51410 | 51411 | 1gte |
| 70597 | 51349 | 51350 | 51430 | 51431 | 75056 | 75057 | 1gve |
| 70610 | 48724 | 50493 | 50494 | 50514 | 50536 | 50538 | 1gvk |
| 70629 | 46456 | 48370 | 48371 | 74771 | 74772 | 74773 | 1gw5 |
| 70630 | 46456 | 48370 | 48371 | 74771 | 74774 | 74775 | 1gw5 |
| 70667 | 56572 | 56633 | 56634 | 56635 | 56636 | 56641 | 1gwe |
| 70722 | 48724 | 50964 | 50978 | 50979 | 75009 | 75010 | 1gxr |
| 70738 | 53931 | 54402 | 54427 | 54431 | 54432 | 75374 | 1gy7 |
| 70803 | 48724 | 49898 | 49899 | 49900 | 49904 | 74903 | 1gzc |
| 70805 | 51349 | 51350 | 51569 | 51594 | 51595 | 51597 | 1gzg |
| 70855 | 51349 | 51350 | 51445 | 51487 | 75066 | 75067 | 1h1n |
| 70857 | 51349 | 53253 | 53254 | 53255 | 69539 | 69540 | 1h2e |
| 70959 | 46456 | 47026 | 47027 | 47028 | 47029 | 47030 | 1hb6 |
| 71089 | 53931 | 54533 | 54534 | 54535 | 75388 | 75389 | 1hxv |
| 71102 | 51349 | 53334 | 53335 | 53354 | 68927 | 69554 | 1i1n |
| 71114 | 51349 | 51350 | 51366 | 51381 | 51385 | 75052 | 1i4n |
| 71122 | 53931 | 54235 | 54292 | 54293 | 75366 | 75367 | 1i7h |
| 71240 | 46456 | 47265 | 47266 | 47286 | 63532 | 63533 | 1ik0 |
| 71270 | 51349 | 51350 | 51713 | 51714 | 75096 | 75097 | 1iq8 |
| 71379 | 51349 | 52373 | 52374 | 52375 | 75159 | 75160 | 1irx |
| 71425 | 51349 | 51350 | 51445 | 51534 | 75069 | 75070 | 1itx |
| 71469 | 56992 | 100894 | 100895 | 57468 | 75678 | 75679 | 1iw4 |
| 71531 | 48724 | 51181 | 51182 | 51187 | 75033 | 75034 | 1j58 |
| 71589 | 53931 | 55680 | 55681 | 55682 | 75504 | 75505 | 1j5w |
| 71618 | 51349 | 52539 | 52540 | 52575 | 52578 | 52579 | 1j99 |
| 71631 | 51349 | 52539 | 52540 | 81269 | 75211 | 75212 | 1jbk |
| 71665 | 51349 | 52539 | 52540 | 52686 | 75207 | 75208 | 1ji0 |
| 71685 | 46456 | 47239 | 47240 | 47241 | 47250 | 74706 | 1jig |
| 71696 | 48724 | 50198 | 50249 | 50282 | 74952 | 74954 | 1jjg |
| 71710 | 53931 | 56111 | 56112 | 88854 | 75560 | 75561 | 1jks |
| 71726 | 46456 | 46457 | 46458 | 46463 | 46467 | 46468 | 1jl7 |
| 71757 | 51349 | 51350 | 51445 | 51534 | 75071 | 75072 | 1jnd |
| 71844 | 53931 | 55944 | 55961 | 55966 | 75545 | 75546 | 1jss |
| 71907 | 53931 | 54533 | 54534 | 54535 | 64250 | 75387 | 1jvw |
| 71978 | 48724 | 49784 | 49785 | 74896 | 74897 | 74898 | 1k12 |
| 72010 | 53931 | 55810 | 55811 | 55812 | 75527 | 75528 | 1k2e |
| 72015 | 53931 | 54000 | 54001 | 54002 | 75330 | 75331 | 1k3b |
| 72017 | 56835 | 56836 | 56854 | 56855 | 75637 | 75638 | 1k3k |
| 72076 | 48724 | 51125 | 51126 | 51137 | 51140 | 75025 | 1k5c |
| 72177 | 51349 | 53473 | 53474 | 53506 | 53507 | 75284 | 1k8q |
| 72181 | 46456 | 47472 | 47473 | 47478 | 47479 | 74719 | 1k8u |
| 72249 | 51349 | 52539 | 52540 | 52566 | 52567 | 75193 | 1kag |
| 72254 | 51349 | 52373 | 52374 | 52397 | 52400 | 75164 | 1kam |
| 72277 | 51349 | 51350 | 51395 | 51396 | 51408 | 51409 | 1kbi |
| 72278 | 53931 | 55855 | 55856 | 55857 | 55864 | 55865 | 1kbi |
| 72399 | 56835 | 81344 | 81343 | 81373 | 81370 | 81369 | 1kf6 |
| 72400 | 56835 | 81344 | 81343 | 81373 | 81372 | 81371 | 1kf6 |
| 72445 | 53931 | 56435 | 56436 | 56437 | 75583 | 75584 | 1kg0 |
| 72498 | 48724 | 50728 | 50729 | 50730 | 74991 | 74992 | 1ki1 |
| 72510 | 51349 | 53589 | 53590 | 53591 | 69587 | 69588 | 1kic |
| 72636 | 51349 | 52539 | 52540 | 52592 | 75204 | 75205 | 1kk1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 72663 | 51349 | 53382 | 53383 | 53417 | 53429 | 75272 | 1kl1 |
| 72678 | 51349 | 53685 | 53686 | 53687 | 64172 | 75313 | 1kl7 |
| 72722 | 56992 | 57666 | 57667 | 57668 | 57679 | 57680 | 1klr |
| 72740 | 51349 | 51350 | 51366 | 51375 | 51376 | 51379 | 1km4 |
| 72744 | 53931 | 64267 | 64268 | 64269 | 75424 | 75425 | 1kmd |
| 72757 | 51349 | 53596 | 53597 | 53598 | 53605 | 53607 | 1kmv |
| 72777 | 51349 | 52832 | 52833 | 52901 | 75237 | 75238 | 1kng |
| 72915 | 48724 | 48725 | 81296 | 81288 | 74846 | 74847 | 1ksh |
| 72972 | 53931 | 55810 | 55811 | 55812 | 64367 | 75526 | 1ktg |
| 73007 | 53931 | 55485 | 55486 | 55519 | 55520 | 75495 | 1kuf |
| 73155 | 48724 | 50198 | 50249 | 50263 | 74948 | 74949 | 1kxl |
| 73334 | 48724 | 50728 | 50729 | 50730 | 74989 | 74990 | 1kz7 |
| 73428 | 53931 | 55680 | 55681 | 55682 | 55696 | 55700 | 1l0w |
| 73452 | 48724 | 51315 | 51316 | 75041 | 75042 | 75043 | 1l1d |
| 73480 | 48724 | 50198 | 50249 | 50263 | 50267 | 50268 | 1l1o |
| 73514 | 51349 | 52539 | 52540 | 52686 | 64027 | 64028 | 1l2t |
| 73539 | 53931 | 54861 | 54928 | 54929 | 54930 | 54931 | 1l3k |
| 73540 | 53931 | 54861 | 54928 | 54929 | 54930 | 54931 | 1l3k |
| 73594 | 51349 | 53731 | 53732 | 53733 | 75319 | 75320 | 1l5j |
| 73598 | 53931 | 54235 | 54292 | 54293 | 54294 | 75365 | 1l5p |
| 73627 | 51349 | 51350 | 51569 | 51594 | 51595 | 51598 | 1l6s |
| 73632 | 51349 | 51350 | 51569 | 51570 | 75085 | 75086 | 1l6w |
| 73674 | 51349 | 52539 | 52540 | 52686 | 75209 | 75210 | 1l7v |
| 73683 | 53931 | 55417 | 55418 | 55419 | 55420 | 55421 | 1l8b |
| 73708 | 46456 | 47094 | 47095 | 47096 | 68982 | 68983 | 1l8y |
| 73813 | 56572 | 56654 | 56655 | 56656 | 56665 | 75608 | 1lbv |
| 73827 | 51349 | 53382 | 53383 | 53384 | 69563 | 69564 | 1lc5 |
| 73867 | 46456 | 46688 | 46689 | 46690 | 46693 | 46694 | 1le8 |
| 73878 | 48724 | 50813 | 50814 | 50815 | 74998 | 74999 | 1lf7 |
| 73896 | 51349 | 53270 | 53271 | 53272 | 53290 | 53291 | 1lh0 |
| 73913 | 51349 | 52373 | 52374 | 52375 | 75161 | 75162 | 1li5 |
| 73928 | 56572 | 56573 | 56574 | 56575 | 56587 | 56588 | 1lj5 |
| 73941 | 48724 | 50181 | 50182 | 50183 | 63758 | 74938 | 1ljo |
| 74005 | 51349 | 53447 | 53448 | 75273 | 75274 | 75275 | 1ll2 |
| 74038 | 53931 | 55944 | 55961 | 55966 | 75547 | 75548 | 1ln1 |
| 74144 | 53931 | 54636 | 54637 | 54638 | 54639 | 54640 | 1lo7 |
| 74192 | 46456 | 47265 | 47266 | 47305 | 47306 | 74708 | 1lqs |
| 74215 | 48724 | 51181 | 51182 | 51187 | 75030 | 75031 | 1lr5 |
| 74271 | 48724 | 50198 | 50249 | 50282 | 74952 | 74953 | 1luz |
| 74295 | 51349 | 52373 | 52374 | 52397 | 75165 | 75166 | 1lw7 |
| 74296 | 51349 | 52539 | 52540 | 52541 | 75189 | 75190 | 1lw7 |
| 74334 | 51349 | 52539 | 52540 | 52541 | 75191 | 75192 | 1ly1 |
| 74442 | 46456 | 47265 | 47266 | 47286 | 47301 | 47302 | 1m47 |
| 74457 | 53931 | 55728 | 55729 | 55730 | 75506 | 75507 | 1m4i |
| 74533 | 51349 | 51904 | 51905 | 51943 | 75126 | 75127 | 1m6i |
| 74534 | 51349 | 51904 | 51905 | 51943 | 75126 | 75127 | 1m6i |
| 74574 | 56572 | 56633 | 56634 | 56635 | 56636 | 75599 | 1m7s |
| 74579 | 53931 | 54116 | 54117 | 54118 | 54128 | 75340 | 1m8a |
| 74601 | 48724 | 50493 | 50494 | 50514 | 50536 | 74973 | 1m9u |
| 74621 | 56992 | 57015 | 57059 | 57072 | 75658 | 75659 | 1mb6 |
| 75802 | 53931 | 81607 | 81606 | 81605 | 81604 | 81603 | 1a6q |
| 76222 | 53931 | 55728 | 55729 | 55730 | 82747 | 82748 | 1ghe |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 76236 | 53931 | 54000 | 54001 | 54002 | 54022 | 54023 | 1gmy |
| 76337 | 46456 | 47094 | 47095 | 47096 | 81718 | 81719 | 1gt0 |
| 76348 | 46456 | 46625 | 46626 | 46627 | 68950 | 68951 | 1gu2 |
| 76360 | 48724 | 50493 | 50494 | 50514 | 82122 | 82123 | 1gvz |
| 76378 | 53931 | 54861 | 54975 | 54976 | 82675 | 82676 | 1gxu |
| 76410 | 51349 | 51734 | 51735 | 51751 | 82294 | 82295 | 1gz6 |
| 76572 | 48724 | 51181 | 51197 | 82194 | 82195 | 82196 | 1h2k |
| 76595 | 53931 | 54861 | 54928 | 54929 | 64274 | 64275 | 1h2v |
| 76620 | 46456 | 46625 | 46626 | 46627 | 81675 | 81676 | 1h32 |
| 76631 | 51349 | 52373 | 52374 | 52375 | 52376 | 82354 | 1h3f |
| 76643 | 51349 | 52373 | 52374 | 52375 | 82356 | 82357 | 1h3n |
| 76644 | 46456 | 47112 | 47113 | 47134 | 81720 | 81721 | 1h3o |
| 76645 | 46456 | 47112 | 47113 | 47134 | 81722 | 81723 | 1h3o |
| 76725 | 51349 | 51734 | 51735 | 51751 | 82290 | 82291 | 1hxh |
| 76732 | 51349 | 51350 | 51445 | 51487 | 51514 | 51518 | 1i1w |
| 76755 | 53931 | 54636 | 54637 | 82636 | 82637 | 82638 | 1iq6 |
| 76773 | 48724 | 49898 | 49899 | 49932 | 82034 | 82035 | 1is3 |
| 76858 | 51349 | 52373 | 52374 | 52375 | 52390 | 52391 | 1ivs |
| 76950 | 48724 | 50198 | 50249 | 50263 | 82099 | 82101 | 1iyj |
| 76951 | 48724 | 50198 | 50249 | 50263 | 82099 | 82101 | 1iyj |
| 76952 | 48724 | 50198 | 50249 | 50263 | 82099 | 82101 | 1iyj |
| 77026 | 51349 | 52373 | 52374 | 52375 | 52382 | 52383 | 1j09 |
| 77051 | 46456 | 81777 | 81778 | 81779 | 81780 | 81781 | 1j0t |
| 77075 | 53931 | 55845 | 55846 | 55847 | 82775 | 82776 | 1j3g |
| 77078 | 46456 | 47472 | 47473 | 47478 | 47479 | 81755 | 1j55 |
| 77082 | 56992 | 57015 | 57095 | 57116 | 82883 | 82884 | 1j5j |
| 77117 | 46456 | 47239 | 47240 | 47241 | 47250 | 81743 | 1ji4 |
| 77135 | 56572 | 56573 | 56574 | 56575 | 82833 | 82834 | 1jmj |
| 77216 | 51349 | 53162 | 53167 | 53168 | 82448 | 82449 | 1jys |
| 77237 | 46456 | 48255 | 48256 | 48257 | 48258 | 81862 | 1k3p |
| 77258 | 51349 | 52373 | 52374 | 52397 | 52400 | 82358 | 1k4m |
| 77270 | 46456 | 46688 | 46689 | 46690 | 46695 | 46696 | 1k61 |
| 77376 | 51349 | 51350 | 51445 | 51534 | 82249 | 82250 | 1kfw |
| 77384 | 48724 | 48725 | 48726 | 48727 | 48933 | 48935 | 1kgc |
| 77409 | 48724 | 50198 | 50249 | 50282 | 82105 | 82106 | 1khi |
| 77442 | 48724 | 48725 | 81296 | 81288 | 49241 | 49242 | 1kmt |
| 77528 | 46456 | 48370 | 48452 | 48453 | 81905 | 81907 | 1kt1 |
| 77628 | 51349 | 53719 | 53720 | 53721 | 82538 | 82539 | 1ky8 |
| 77653 | 51349 | 53270 | 53271 | 53272 | 53288 | 82455 | 1l1q |
| 77655 | 48724 | 50352 | 50353 | 50362 | 50363 | 50364 | 1l2h |
| 77768 | 51349 | 56783 | 56784 | 82388 | 82389 | 82390 | 1l6r |
| 77770 | 48724 | 48725 | 48726 | 48727 | 81942 | 81943 | 1l6z |
| 77771 | 48724 | 48725 | 48726 | 49159 | 81956 | 81957 | 1l6z |
| 77810 | 51349 | 52539 | 52540 | 81269 | 82416 | 82417 | 1l8q |
| 77873 | 46456 | 47239 | 47240 | 47241 | 47246 | 63526 | 1lb3 |
| 77926 | 46456 | 48263 | 48264 | 48265 | 81863 | 81864 | 1lfk |
| 77960 | 51349 | 53612 | 53613 | 82515 | 82516 | 82517 | 1lhp |
| 78063 | 46456 | 47239 | 47240 | 47241 | 47242 | 47243 | 1lko |
| 78171 | 51349 | 52539 | 52540 | 52652 | 52664 | 52665 | 1ls1 |
| 78184 | 51349 | 51350 | 82282 | 82283 | 82284 | 82285 | 1lt7 |
| 78232 | 51349 | 52539 | 52540 | 81269 | 82418 | 82419 | 1lv7 |
| 78265 | 51349 | 53658 | 53659 | 53660 | 82524 | 82525 | 1lwd |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 78401 | 48724 | 50036 | 50118 | 82075 | 82076 | 82077 | 1m1f |
| 78508 | 51349 | 53382 | 53383 | 53402 | 82486 | 82487 | 1m32 |
| 78598 | 46456 | 47472 | 47473 | 47502 | 81756 | 81757 | 1m45 |
| 78602 | 53931 | 55752 | 55753 | 55762 | 82752 | 82753 | 1m4j |
| 78666 | 53931 | 54533 | 54534 | 54535 | 82624 | 82625 | 1m5y |
| 78676 | 48724 | 50155 | 50156 | 50157 | 82083 | 82084 | 1m5z |
| 78701 | 51349 | 53382 | 53383 | 53384 | 64123 | 64124 | 1m6s |
| 78724 | 51349 | 52539 | 52540 | 52572 | 52573 | 52574 | 1m7g |
| 78761 | 51349 | 53382 | 53383 | 53417 | 53440 | 53441 | 1m7y |
| 78865 | 53931 | 54235 | 54236 | 54237 | 82579 | 82580 | 1m94 |
| 78907 | 51349 | 52171 | 52172 | 52173 | 82342 | 82343 | 1mb3 |
| 79008 | 56572 | 56800 | 56801 | 56802 | 82843 | 82844 | 1mdb |
| 79013 | 51349 | 53382 | 53383 | 53417 | 82488 | 82489 | 1mdo |
| 79023 | 53931 | 54000 | 54001 | 54002 | 54020 | 54021 | 1me4 |
| 79129 | 53931 | 56275 | 56276 | 56277 | 56278 | 82801 | 1mhm |
| 79150 | 46456 | 46688 | 46689 | 46690 | 81681 | 81682 | 1mij |
| 79183 | 51349 | 52095 | 52096 | 52103 | 52106 | 52107 | 1mj3 |
| 79189 | 53931 | 55855 | 55856 | 55857 | 55866 | 82777 | 1mj4 |
| 79198 | 51349 | 53334 | 53335 | 69557 | 82479 | 82480 | 1mjf |
| 79250 | 51349 | 52539 | 52540 | 52592 | 82406 | 82407 | 1mky |
| 79251 | 51349 | 52539 | 52540 | 52592 | 82406 | 82407 | 1mky |
| 79341 | 51349 | 51904 | 51905 | 51943 | 82313 | 82314 | 1mo9 |
| 79342 | 51349 | 51904 | 51905 | 51943 | 82313 | 82314 | 1mo9 |
| 79414 | 56572 | 56814 | 56815 | 56816 | 82848 | 82849 | 1mqs |
| 79469 | 51349 | 53066 | 53098 | 53099 | 53105 | 82443 | 1mu2 |
| 79617 | 53931 | 54861 | 55008 | 55009 | 82683 | 82684 | 1mwz |
| 79758 | 51349 | 52539 | 52540 | 52592 | 82404 | 82405 | 1n0u |
| 79778 | 53931 | 74651 | 48403 | 48404 | 82632 | 82633 | 1n11 |
| 79803 | 46456 | 47265 | 47266 | 47305 | 81744 | 81745 | 1n1f |
| 79813 | 46456 | 47112 | 47113 | 47134 | 81724 | 81725 | 1n1j |
| 79814 | 46456 | 47112 | 47113 | 47134 | 81726 | 81727 | 1n1j |
| 79850 | 48724 | 48725 | 48726 | 49159 | 81958 | 81959 | 1n26 |
| 79857 | 51349 | 51734 | 51735 | 51751 | 75102 | 75103 | 1n2s |
| 79966 | 51349 | 52373 | 52374 | 52375 | 52376 | 82355 | 1n3l |
| 79977 | 46456 | 48263 | 48264 | 48265 | 81865 | 81866 | 1n40 |
| 80007 | 51349 | 51350 | 51351 | 51352 | 51353 | 51360 | 1n55 |
| 80250 | 51349 | 51734 | 51735 | 51751 | 51759 | 82289 | 1n7h |
| 80304 | 51349 | 53382 | 53383 | 53402 | 82484 | 82485 | 1n8p |
| 80335 | 46456 | 48263 | 48264 | 48265 | 81869 | 81870 | 1n97 |
| 80344 | 48724 | 50181 | 50182 | 50183 | 82087 | 82088 | 1n9r |
| 80407 | 51349 | 53253 | 53254 | 53258 | 53259 | 53261 | 1nd6 |
| 80424 | 48724 | 50036 | 50118 | 82075 | 82078 | 82079 | 1ne8 |
| 80431 | 56835 | 81344 | 81343 | 81373 | 82870 | 82871 | 1nek |
| 80432 | 56835 | 81344 | 81343 | 81373 | 82872 | 82873 | 1nek |
| 80444 | 48724 | 50964 | 50978 | 50979 | 82169 | 82170 | 1nex |
| 80531 | 51349 | 52539 | 52540 | 52592 | 82410 | 82411 | 1ni3 |
| 80541 | 56992 | 57015 | 57059 | 57072 | 82879 | 82880 | 1nix |
| 80573 | 46456 | 47458 | 47459 | 47460 | 81750 | 81751 | 1nkp |
| 80574 | 46456 | 47458 | 47459 | 47460 | 47461 | 47462 | 1nkp |
| 80633 | 46456 | 47458 | 47459 | 47460 | 81752 | 81753 | 1nlw |
| 80731 | 53931 | 54861 | 54928 | 54929 | 54932 | 54933 | 1nu4 |
| 80734 | 48724 | 50352 | 50353 | 50354 | 82107 | 82108 | 1nun |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 81092 | 53931 | 56111 | 56112 | 88854 | 82791 | 82792 | 1o6l |
| 81106 | 53931 | 56111 | 56112 | 88854 | 82793 | 82794 | 1o6y |
| 81178 | 48724 | 51125 | 51126 | 51127 | 51128 | 51129 | 1o88 |
| 83054 | 48724 | 50198 | 50249 | 50282 | 88670 | 88671 | 1go3 |
| 83251 | 56992 | 57301 | 57302 | 57303 | 90152 | 90153 | 1ff4 |
| 83255 | 51349 | 51350 | 51445 | 51487 | 89469 | 89470 | 1fob |
| 83268 | 51349 | 52539 | 52540 | 52575 | 52576 | 75198 | 1g3m |
| 83336 | 46456 | 46688 | 46689 | 46739 | 46740 | 46742 | 1gv2 |
| 83338 | 46456 | 46688 | 46689 | 46739 | 46740 | 46742 | 1gvd |
| 83351 | 46456 | 48112 | 48113 | 48114 | 48125 | 48126 | 1gwu |
| 83372 | 48724 | 48725 | 48726 | 49159 | 89185 | 89186 | 1gxe |
| 83376 | 51349 | 51734 | 51735 | 51751 | 51752 | 89519 | 1gy8 |
| 83392 | 53931 | 56435 | 56436 | 56437 | 90059 | 90060 | 1gz2 |
| 83405 | 53931 | 56111 | 56112 | 88854 | 88855 | 88856 | 1gz8 |
| 83553 | 48724 | 48725 | 48726 | 48727 | 89178 | 89179 | 1hkf |
| 83581 | 51349 | 51350 | 51569 | 51570 | 51571 | 51572 | 1hl2 |
| 83663 | 51349 | 51734 | 51735 | 51751 | 51763 | 51764 | 1i24 |
| 83698 | 46456 | 48255 | 48256 | 48257 | 48258 | 89115 | 1iom |
| 83763 | 53931 | 54533 | 54534 | 54535 | 89879 | 89880 | 1ix5 |
| 83774 | 51349 | 51734 | 51735 | 51751 | 89523 | 89524 | 1iy8 |
| 83786 | 56572 | 56751 | 56752 | 56753 | 56757 | 56758 | 1iye |
| 83806 | 56992 | 57849 | 57850 | 57851 | 90220 | 90221 | 1iym |
| 83838 | 51349 | 51350 | 51621 | 51638 | 89501 | 89502 | 1izc |
| 83847 | 56572 | 90095 | 90096 | 90097 | 90098 | 90099 | 1izn |
| 83848 | 56572 | 90095 | 90096 | 90100 | 90101 | 90102 | 1izn |
| 83851 | 46456 | 48263 | 48264 | 48265 | 89116 | 89117 | 1izo |
| 83871 | 51349 | 53685 | 53686 | 53687 | 53694 | 89777 | 1j0a |
| 83980 | 51349 | 52373 | 52374 | 52375 | 52376 | 89611 | 1j1u |
| 84048 | 53931 | 56435 | 56436 | 56437 | 88861 | 88862 | 1j34 |
| 84070 | 51349 | 53596 | 53597 | 53598 | 53610 | 88963 | 1j3k |
| 84131 | 51349 | 53270 | 53271 | 53272 | 53286 | 89728 | 1j7j |
| 84138 | 51349 | 52539 | 52540 | 52592 | 82410 | 89666 | 1jal |
| 84185 | 53931 | 54533 | 54534 | 54535 | 89877 | 89878 | 1jns |
| 84205 | 51349 | 52171 | 52266 | 89594 | 89595 | 89596 | 1jrl |
| 84230 | 51349 | 51734 | 51735 | 51751 | 51772 | 51773 | 1jtv |
| 84262 | 53931 | 56435 | 56436 | 56437 | 90061 | 90062 | 1jzn |
| 84271 | 48724 | 50813 | 50814 | 50815 | 89364 | 89365 | 1jzu |
| 84401 | 51349 | 52539 | 52540 | 52541 | 52554 | 89660 | 1kht |
| 84467 | 48724 | 50813 | 50814 | 50815 | 50816 | 50817 | 1kt7 |
| 84475 | 48724 | 50813 | 50814 | 50815 | 50837 | 50838 | 1kxo |
| 84625 | 53931 | 56370 | 56371 | 56372 | 90055 | 90056 | 1lln |
| 84674 | 51349 | 51350 | 51430 | 51431 | 89461 | 89462 | 1lqa |
| 84680 | 56992 | 100894 | 100895 | 57468 | 90166 | 90167 | 1lr7 |
| 84707 | 48724 | 50493 | 50494 | 50514 | 89341 | 89342 | 1lto |
| 84743 | 56992 | 57867 | 57868 | 57869 | 57870 | 90225 | 1m0g |
| 84744 | 56992 | 57867 | 57868 | 57869 | 57870 | 90225 | 1m0j |
| 84761 | 53931 | 56370 | 56371 | 56372 | 56381 | 56382 | 1m2t |
| 84808 | 48724 | 50181 | 50182 | 50183 | 89317 | 89318 | 1m5q |
| 84979 | 53931 | 55728 | 55729 | 55730 | 90013 | 90014 | 1mk4 |
| 85092 | 56992 | 56993 | 56994 | 56995 | 56996 | 56998 | 1mso |
| 85107 | 56572 | 56573 | 56574 | 56575 | 90080 | 90081 | 1mtp |
| 85140 | 48724 | 49898 | 49899 | 49925 | 49926 | 89270 | 1mve |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 85146 | 51349 | 52506 | 52507 | 52508 | 52509 | 52510 | 1mvl |
| 85197 | 46456 | 47239 | 47240 | 47253 | 47257 | 47258 | 1mxr |
| 85241 | 46456 | 47873 | 47874 | 47875 | 89079 | 89080 | 1n00 |
| 85315 | 46456 | 48507 | 48508 | 48509 | 48530 | 48531 | 1n46 |
| 85328 | 51349 | 52171 | 52317 | 52325 | 89609 | 89610 | 1n57 |
| 85370 | 53931 | 55728 | 55729 | 55730 | 55731 | 55732 | 1n71 |
| 85374 | 51349 | 51350 | 51569 | 51570 | 69394 | 89493 | 1n7k |
| 85401 | 51349 | 52832 | 52833 | 52901 | 69516 | 69517 | 1n8j |
| 85553 | 48724 | 48725 | 48726 | 48727 | 63635 | 63636 | 1ncn |
| 85592 | 48724 | 48725 | 48726 | 48727 | 48734 | 48736 | 1nez |
| 85598 | 46456 | 47239 | 47240 | 47241 | 47244 | 89025 | 1nf4 |
| 85631 | 51349 | 52498 | 52499 | 100948 | 89643 | 89644 | 1nf9 |
| 85739 | 56572 | 56654 | 56655 | 90087 | 90088 | 90089 | 1ni9 |
| 85741 | 51349 | 52539 | 52540 | 52652 | 89671 | 89672 | 1nij |
| 85781 | 51349 | 52539 | 52540 | 81269 | 52711 | 52712 | 1njf |
| 85818 | 53931 | 54636 | 54637 | 54638 | 89900 | 89901 | 1njk |
| 85866 | 51349 | 52832 | 52833 | 52834 | 89700 | 89701 | 1nm3 |
| 85867 | 51349 | 52832 | 52833 | 52901 | 89708 | 89709 | 1nm3 |
| 85874 | 51349 | 52128 | 52129 | 52130 | 52131 | 52132 | 1nme |
| 85892 | 51349 | 52539 | 52540 | 52541 | 52563 | 64010 | 1nn5 |
| 85906 | 51349 | 56783 | 56784 | 64511 | 64512 | 89645 | 1nnl |
| 85954 | 51349 | 52539 | 52540 | 52652 | 89673 | 89674 | 1np6 |
| 86077 | 53931 | 55810 | 55811 | 55812 | 90022 | 90023 | 1nqz |
| 86078 | 48724 | 50964 | 50978 | 50979 | 89378 | 89379 | 1nr0 |
| 86079 | 48724 | 50964 | 50978 | 50979 | 89378 | 89379 | 1nr0 |
| 86125 | 51349 | 52539 | 52540 | 52592 | 89664 | 89665 | 1nrj |
| 86128 | 51349 | 56783 | 56784 | 82388 | 89646 | 89647 | 1nrw |
| 86137 | 53931 | 55728 | 55729 | 55730 | 90015 | 90016 | 1nsl |
| 86209 | 51349 | 52373 | 52374 | 52397 | 89613 | 89614 | 1nuu |
| 86211 | 56572 | 56654 | 56655 | 56656 | 56657 | 56658 | 1nuw |
| 86224 | 51349 | 53334 | 53335 | 89743 | 89744 | 89745 | 1nv8 |
| 86250 | 51349 | 51350 | 51569 | 89494 | 89495 | 89496 | 1nvm |
| 86268 | 53931 | 56111 | 56112 | 88854 | 64404 | 64405 | 1nvr |
| 86276 | 46456 | 48365 | 48366 | 48367 | 48368 | 48369 | 1nvu |
| 86294 | 51349 | 52128 | 52129 | 52130 | 69443 | 69444 | 1nw9 |
| 86295 | 53931 | 54861 | 55068 | 55069 | 55070 | 89956 | 1nwa |
| 86296 | 48724 | 89359 | 89360 | 89361 | 89362 | 89363 | 1nwb |
| 86508 | 51349 | 53719 | 53720 | 53721 | 53722 | 53726 | 1o04 |
| 86525 | 51349 | 56783 | 56784 | 75173 | 75174 | 75175 | 1o08 |
| 86535 | 53931 | 54861 | 54928 | 54929 | 54936 | 54937 | 1o0p |
| 86555 | 51349 | 51350 | 51695 | 89508 | 89509 | 89510 | 1o1z |
| 86556 | 51349 | 53719 | 53720 | 53721 | 89781 | 89782 | 1o20 |
| 86591 | 56572 | 56795 | 56796 | 69892 | 75612 | 75613 | 1o2d |
| 86620 | 51349 | 53382 | 53383 | 53384 | 53385 | 89754 | 1o4s |
| 86629 | 51349 | 88722 | 88723 | 89619 | 89620 | 89621 | 1o4w |
| 86642 | 48724 | 50198 | 50249 | 50263 | 89324 | 89325 | 1o7i |
| 86660 | 53931 | 54116 | 54117 | 54118 | 82573 | 82574 | 1o7z |
| 86681 | 51349 | 52832 | 52833 | 52901 | 52904 | 52905 | 1o8x |
| 86706 | 46456 | 47239 | 47240 | 47241 | 47250 | 89027 | 1o9r |
| 86734 | 46456 | 48112 | 48113 | 48114 | 48123 | 89092 | 1oaf |
| 86739 | 48724 | 48725 | 49329 | 49330 | 49331 | 49337 | 1oal |
| 86832 | 48724 | 51181 | 51182 | 51187 | 51188 | 89412 | 1od5 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 86833 | 48724 | 51181 | 51182 | 51187 | 51188 | 89412 | 1od5 |
| 86836 | 51349 | 52373 | 52374 | 52397 | 52398 | 89612 | 1od6 |
| 86863 | 51349 | 53162 | 53167 | 53168 | 53169 | 89721 | 1odk |
| 86875 | 48724 | 51181 | 51197 | 51198 | 51199 | 51200 | 1odm |
| 86889 | 51349 | 51350 | 51445 | 51487 | 63908 | 63909 | 1odz |
| 86924 | 48724 | 50629 | 50630 | 50646 | 50647 | 50648 | 1oew |
| 86934 | 53931 | 54402 | 54427 | 54431 | 89849 | 89850 | 1of5 |
| 86936 | 51349 | 51350 | 51395 | 51396 | 69381 | 75055 | 1ofd |
| 86990 | 48724 | 48725 | 48726 | 48727 | 48933 | 48935 | 1oga |
| 86992 | 48724 | 48725 | 48726 | 48727 | 48933 | 48937 | 1oga |
| 87001 | 53931 | 54235 | 54236 | 54237 | 54238 | 54239 | 1ogw |
| 87013 | 51349 | 52798 | 52799 | 52800 | 89698 | 89699 | 1ohe |
| 87014 | 51349 | 52798 | 52799 | 52800 | 89698 | 89699 | 1ohe |
| 87036 | 53931 | 55845 | 55846 | 55847 | 90024 | 90025 | 1oht |
| 87086 | 46456 | 48096 | 48097 | 48098 | 89090 | 89091 | 1omw |
| 87087 | 48724 | 50728 | 50729 | 50730 | 50747 | 89354 | 1omw |
| 87088 | 53931 | 56111 | 56112 | 88854 | 90034 | 90035 | 1omw |
| 87142 | 51349 | 51904 | 51905 | 51943 | 51944 | 89548 | 1onf |
| 87196 | 51349 | 51734 | 51735 | 51751 | 51769 | 89520 | 1ooe |
| 87319 | 46456 | 47472 | 47473 | 47502 | 89053 | 89055 | 1oqp |
| 87387 | 51349 | 52539 | 52540 | 52541 | 52552 | 89659 | 1osn |
| 87494 | 53931 | 54861 | 54928 | 54929 | 89940 | 89941 | 1owx |
| 87539 | 51349 | 52150 | 52151 | 89729 | 89730 | 89731 | 1oxw |
| 87554 | 53931 | 74651 | 48403 | 48404 | 82634 | 82635 | 1oy3 |
| 87648 | 51349 | 51350 | 51395 | 51396 | 89454 | 89455 | 1p0k |
| 87716 | 48724 | 50964 | 50978 | 50979 | 89376 | 89377 | 1p22 |
| 87721 | 51349 | 52171 | 52172 | 52173 | 89587 | 89588 | 1p2f |
| 87752 | 51349 | 53382 | 53383 | 53402 | 89755 | 89756 | 1p3w |
| 87775 | 53931 | 56111 | 56112 | 88854 | 69825 | 69826 | 1p4o |
| 87815 | 51349 | 52539 | 52540 | 52541 | 89657 | 89658 | 1p5z |
| 88034 | 46456 | 48507 | 48508 | 48509 | 89144 | 89145 | 1pdu |
| 88067 | 56572 | 56800 | 56801 | 56802 | 90093 | 90094 | 1pg4 |
| 88069 | 48724 | 50964 | 50978 | 50979 | 89378 | 89380 | 1pgu |
| 88070 | 48724 | 50964 | 50978 | 50979 | 89378 | 89380 | 1pgu |
| 88139 | 46456 | 48507 | 48508 | 48509 | 89147 | 89148 | 1pk5 |
| 88147 | 48724 | 48725 | 48726 | 48727 | 89174 | 89175 | 1pko |
| 88292 | 51349 | 52539 | 52540 | 52592 | 89667 | 89668 | 1pui |
| 88294 | 51349 | 52539 | 52540 | 52592 | 89669 | 89670 | 1puj |
| 88395 | 51349 | 51350 | 51569 | 51570 | 69394 | 89492 | 1ub3 |
| 88399 | 48724 | 50036 | 50118 | 82075 | 89303 | 89304 | 1ub4 |
| 88467 | 48724 | 48725 | 48726 | 49159 | 89188 | 89189 | 1uct |
| 88485 | 53931 | 55728 | 55729 | 55730 | 90011 | 90012 | 1ufh |
| 88490 | 51349 | 51350 | 51445 | 51521 | 51528 | 89476 | 1ug6 |
| 88491 | 56572 | 56573 | 56574 | 56575 | 56578 | 56579 | 1uhg |
| 90443 | 51349 | 53648 | 53649 | 102651 | 102652 | 102653 | 1ei6 |
| 90509 | 51349 | 51734 | 51735 | 51751 | 51784 | 51785 | 1gee |
| 90514 | 51349 | 51350 | 51569 | 51570 | 51586 | 51587 | 1gqn |
| 90661 | 56572 | 56751 | 56752 | 56753 | 56759 | 56760 | 1i2k |
| 90690 | 53931 | 55810 | 55811 | 55812 | 103207 | 103208 | 1iry |
| 90699 | 51349 | 53382 | 53383 | 53402 | 102599 | 102600 | 1iug |
| 90703 | 51349 | 69592 | 69593 | 69594 | 69595 | 69596 | 1iuq |
| 90717 | 53931 | 54636 | 54637 | 89902 | 102916 | 102917 | 1ixl |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 90719 | 53931 | 74651 | 48403 | 48404 | 102881 | 102883 | 1ixv |
| 90747 | 48724 | 50352 | 50353 | 50362 | 101782 | 101783 | 1j0s |
| 90788 | 51349 | 52498 | 52499 | 100948 | 102298 | 102299 | 1j2r |
| 90806 | 46456 | 47239 | 47240 | 47241 | 101130 | 101131 | 1j30 |
| 90833 | 53931 | 55769 | 103196 | 103197 | 103198 | 103199 | 1j3w |
| 90882 | 53931 | 54494 | 54495 | 54496 | 54497 | 102840 | 1jas |
| 90908 | 51349 | 51350 | 51395 | 51396 | 51397 | 51398 | 1jub |
| 90912 | 51349 | 53596 | 53597 | 53598 | 53599 | 102635 | 1juv |
| 90927 | 51349 | 51734 | 51735 | 51751 | 102148 | 102149 | 1k2w |
| 90929 | 53931 | 56234 | 56235 | 56261 | 56262 | 103315 | 1k2x |
| 90941 | 51349 | 52171 | 52172 | 52173 | 75152 | 102227 | 1k66 |
| 90943 | 51349 | 52171 | 52172 | 52173 | 75152 | 102226 | 1k68 |
| 90964 | 53931 | 54000 | 54001 | 54002 | 54005 | 54006 | 1khq |
| 91020 | 53931 | 64267 | 64268 | 64269 | 64270 | 64271 | 1kq6 |
| 91024 | 51349 | 52171 | 52172 | 52173 | 52180 | 52181 | 1krw |
| 91055 | 56992 | 57015 | 57059 | 57072 | 103531 | 103532 | 1la4 |
| 91080 | 56992 | 57015 | 57059 | 57072 | 103533 | 103534 | 1lmm |
| 91086 | 48724 | 48725 | 48726 | 48727 | 48933 | 48934 | 1lp9 |
| 91124 | 51349 | 52832 | 52833 | 52901 | 102457 | 102458 | 1lu4 |
| 91127 | 48724 | 51068 | 51069 | 51070 | 51071 | 51073 | 1lug |
| 91170 | 56992 | 57015 | 57095 | 57116 | 103542 | 103543 | 1m2s |
| 91208 | 51349 | 52128 | 52129 | 52130 | 102526 | 102527 | 1m72 |
| 91235 | 56572 | 56573 | 56574 | 56575 | 56593 | 56594 | 1m93 |
| 91242 | 48724 | 50352 | 50353 | 50362 | 101780 | 101781 | 1md6 |
| 91302 | 48724 | 48725 | 48726 | 48727 | 88543 | 88552 | 1mju |
| 91304 | 48724 | 48725 | 48726 | 48727 | 88519 | 88524 | 1mju |
| 91384 | 53931 | 56111 | 56112 | 88854 | 103292 | 103293 | 1mp8 |
| 91394 | 53931 | 55810 | 55811 | 55812 | 64365 | 103200 | 1mqe |
| 91453 | 51349 | 52373 | 52374 | 52397 | 102260 | 102261 | 1mrz |
| 91469 | 51349 | 52539 | 52540 | 52686 | 102381 | 102382 | 1mv5 |
| 91500 | 51349 | 51350 | 51430 | 51431 | 51443 | 102053 | 1mzr |
| 91636 | 46456 | 48207 | 48239 | 48246 | 48247 | 48248 | 1n4q |
| 91691 | 48724 | 50155 | 50156 | 50157 | 101717 | 101718 | 1n7e |
| 91705 | 56992 | 57015 | 57095 | 57116 | 103550 | 103551 | 1n8m |
| 91767 | 53931 | 55619 | 55620 | 55628 | 55629 | 103143 | 1nbu |
| 91815 | 53931 | 56111 | 56112 | 64411 | 103302 | 103303 | 1nd4 |
| 91817 | 53931 | 56203 | 56204 | 56205 | 103308 | 103309 | 1nd7 |
| 91849 | 51349 | 56783 | 56784 | 82388 | 102313 | 102314 | 1nf2 |
| 92005 | 53931 | 55680 | 55681 | 55682 | 103157 | 103158 | 1nnh |
| 92006 | 46456 | 47239 | 47240 | 47241 | 47242 | 101129 | 1nnq |
| 92018 | 53931 | 54861 | 54928 | 54929 | 102979 | 102980 | 1no8 |
| 92021 | 51349 | 51350 | 51445 | 51487 | 102073 | 102074 | 1nof |
| 92047 | 46456 | 48507 | 48508 | 48509 | 74794 | 74795 | 1nq7 |
| 92077 | 53931 | 54116 | 54117 | 54118 | 102730 | 102731 | 1nr4 |
| 92091 | 46456 | 48507 | 48508 | 48509 | 63624 | 63625 | 1nrl |
| 92102 | 51349 | 53253 | 53254 | 53258 | 102533 | 102534 | 1nt4 |
| 92215 | 51349 | 52832 | 52833 | 52834 | 52835 | 52839 | 1nw2 |
| 92316 | 48724 | 50198 | 50249 | 50282 | 101765 | 101766 | 1ny4 |
| 92317 | 51349 | 52171 | 52172 | 52173 | 102230 | 102231 | 1ny5 |
| 92318 | 51349 | 52539 | 52540 | 81269 | 102389 | 102390 | 1ny5 |
| 92335 | 46456 | 47472 | 47473 | 47502 | 101179 | 101180 | 1nya |
| 92377 | 51349 | 52373 | 52374 | 52375 | 102257 | 102258 | 1nzj |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 92382 | 46456 | 48370 | 48452 | 48453 | 101417 | 101418 | 1nzn |
| 92473 | 48724 | 49898 | 49899 | 49925 | 101636 | 101637 | 1o4y |
| 92484 | 51349 | 53270 | 53271 | 53272 | 102539 | 102540 | 1o57 |
| 92499 | 51349 | 51734 | 51735 | 51751 | 51788 | 102150 | 1o5i |
| 92504 | 51349 | 51350 | 51569 | 51570 | 51574 | 102090 | 1o5k |
| 92509 | 51349 | 53270 | 53271 | 53272 | 53293 | 102538 | 1o5o |
| 92560 | 51349 | 53382 | 53383 | 53417 | 102603 | 102604 | 1o69 |
| 92626 | 51349 | 53447 | 53448 | 64131 | 64132 | 64133 | 1o7q |
| 92708 | 48724 | 48725 | 48726 | 48727 | 88534 | 101504 | 1oaq |
| 92776 | 53931 | 64267 | 64268 | 64269 | 102952 | 102953 | 1ocs |
| 92782 | 51349 | 52539 | 52540 | 52584 | 102356 | 102357 | 1odf |
| 92783 | 46456 | 48263 | 48264 | 48265 | 101373 | 101374 | 1odo |
| 92826 | 46456 | 46688 | 46689 | 46739 | 100996 | 100997 | 1ofc |
| 92893 | 51349 | 102545 | 102546 | 102547 | 102548 | 102549 | 1ogd |
| 93029 | 56835 | 56836 | 56854 | 56855 | 103424 | 103425 | 1ohu |
| 93035 | 51349 | 53382 | 53383 | 53417 | 53424 | 53425 | 1ohw |
| 93164 | 51349 | 51350 | 51569 | 51570 | 102088 | 102089 | 1ojx |
| 93248 | 56835 | 103505 | 103506 | 103507 | 103508 | 103509 | 1okc |
| 93313 | 48724 | 48725 | 48726 | 49159 | 101519 | 101520 | 1oll |
| 93330 | 48724 | 49898 | 49899 | 49978 | 49991 | 101655 | 1olr |
| 93341 | 48724 | 48725 | 48726 | 49159 | 101515 | 101516 | 1olz |
| 93349 | 46456 | 47472 | 47473 | 47502 | 47533 | 47534 | 1omr |
| 93383 | 46456 | 47472 | 47565 | 47566 | 101190 | 101191 | 1ooh |
| 93447 | 46456 | 46457 | 46458 | 46463 | 100980 | 100981 | 1or4 |
| 93460 | 51349 | 51734 | 51735 | 51751 | 102138 | 102139 | 1orr |
| 93496 | 46456 | 48507 | 48508 | 48509 | 101433 | 101434 | 1osh |
| 93507 | 53931 | 74651 | 48403 | 48404 | 102879 | 102880 | 1ot8 |
| 93521 | 46456 | 47239 | 47240 | 47253 | 101136 | 101137 | 1otk |
| 93628 | 46456 | 47472 | 47565 | 47566 | 101188 | 101189 | 1ow4 |
| 93716 | 51349 | 52539 | 52540 | 52686 | 89681 | 89682 | 1oxx |
| 93775 | 46456 | 48370 | 48371 | 101410 | 101411 | 101412 | 1oyz |
| 93845 | 48724 | 50155 | 50156 | 50157 | 50168 | 74930 | 1ozi |
| 93866 | 53931 | 55728 | 55729 | 55730 | 103175 | 103176 | 1p0h |
| 93897 | 48724 | 50155 | 50156 | 50157 | 82083 | 82084 | 1p1d |
| 93899 | 53931 | 54861 | 54928 | 54929 | 102983 | 102984 | 1p1t |
| 94071 | 51349 | 52506 | 52507 | 52508 | 102301 | 102302 | 1p3y |
| 94089 | 51349 | 53648 | 53649 | 53653 | 102649 | 102650 | 1p49 |
| 94100 | 51349 | 51350 | 51395 | 51396 | 63898 | 63899 | 1p4c |
| 94142 | 51349 | 52171 | 52317 | 52325 | 89603 | 89604 | 1p5f |
| 94185 | 51349 | 53926 | 53927 | 89800 | 89801 | 89802 | 1p6o |
| 94188 | 53931 | 54861 | 55008 | 55009 | 75441 | 75442 | 1p6t |
| 94195 | 51349 | 52539 | 52540 | 52541 | 52552 | 102340 | 1p6x |
| 94216 | 56992 | 57666 | 57667 | 57668 | 103597 | 103598 | 1p7a |
| 94279 | 46456 | 46688 | 46689 | 46690 | 46691 | 46692 | 1p7i |
| 94338 | 56572 | 56633 | 56634 | 56635 | 56642 | 56643 | 1p80 |
| 94595 | 48724 | 51125 | 51126 | 51127 | 51128 | 75024 | 1pe9 |
| 94661 | 51349 | 52373 | 52374 | 52375 | 52384 | 52386 | 1pfv |
| 94758 | 51349 | 53473 | 53474 | 53542 | 102629 | 102630 | 1pja |
| 94761 | 51349 | 52095 | 52096 | 52103 | 52108 | 63954 | 1pjh |
| 94767 | 51349 | 53789 | 53790 | 53791 | 102682 | 102683 | 1pjq |
| 94794 | 51349 | 53334 | 53335 | 102566 | 102567 | 102568 | 1pjz |
| 94832 | 48724 | 51268 | 51283 | 51284 | 89425 | 89426 | 1pkh |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 94901 | 51349 | 53382 | 53383 | 69565 | 102596 | 102597 | 1pmm |
| 94919 | 51349 | 52832 | 52833 | 52901 | 52911 | 52912 | 1pn0 |
| 94930 | 53931 | 54636 | 54637 | 82636 | 102907 | 102908 | 1pn2 |
| 94931 | 53931 | 54636 | 54637 | 82636 | 102907 | 102908 | 1pn2 |
| 94963 | 46456 | 48263 | 48264 | 48265 | 101375 | 101376 | 1po5 |
| 94986 | 56835 | 56836 | 56854 | 56855 | 56856 | 103423 | 1pq1 |
| 95004 | 46456 | 48507 | 48508 | 48509 | 89149 | 89150 | 1pq9 |
| 95076 | 51349 | 51350 | 51395 | 51396 | 102042 | 102043 | 1ps9 |
| 95131 | 46456 | 46688 | 46689 | 46690 | 100994 | 100995 | 1puf |
| 95132 | 46456 | 46688 | 46689 | 46690 | 46711 | 46712 | 1puf |
| 95133 | 53931 | 55810 | 55811 | 55812 | 55813 | 55814 | 1pun |
| 95153 | 56835 | 103472 | 103473 | 103477 | 103478 | 103479 | 1pv7 |
| 95162 | 53931 | 55873 | 55874 | 55879 | 103226 | 103227 | 1pvg |
| 95198 | 56835 | 103472 | 103473 | 103474 | 103475 | 103476 | 1pw4 |
| 95211 | 48724 | 50352 | 50353 | 50354 | 101778 | 101779 | 1pwa |
| 95212 | 53931 | 56435 | 56436 | 56437 | 56461 | 56462 | 1pwb |
| 95334 | 51349 | 51350 | 51430 | 51431 | 102054 | 102055 | 1pyf |
| 95370 | 51349 | 52128 | 52129 | 52130 | 102528 | 102529 | 1pyo |
| 95390 | 51349 | 51350 | 51430 | 51431 | 102056 | 102057 | 1pz1 |
| 95470 | 51349 | 53447 | 53448 | 53452 | 53453 | 53454 | 1pzt |
| 95574 | 51349 | 53162 | 53167 | 53168 | 102500 | 102501 | 1q1g |
| 95600 | 51349 | 51904 | 51905 | 51943 | 102185 | 102186 | 1q1r |
| 95601 | 51349 | 51904 | 51905 | 51943 | 102185 | 102186 | 1q1r |
| 95608 | 48724 | 50352 | 50353 | 50354 | 101776 | 101777 | 1q1u |
| 95655 | 53931 | 55728 | 55729 | 55730 | 103169 | 103170 | 1q2y |
| 95657 | 53931 | 55810 | 55811 | 55812 | 103205 | 103206 | 1q33 |
| 95700 | 48724 | 50155 | 50156 | 50157 | 101733 | 101734 | 1q3o |
| 95771 | 53931 | 54402 | 54427 | 54431 | 89847 | 102808 | 1q40 |
| 95776 | 53931 | 54402 | 54427 | 54431 | 89849 | 102809 | 1q42 |
| 95779 | 51349 | 52539 | 52540 | 52575 | 102352 | 102353 | 1q44 |
| 95780 | 51349 | 51350 | 51395 | 51396 | 63896 | 102041 | 1q45 |
| 95828 | 53931 | 54636 | 54637 | 89902 | 102914 | 102915 | 1q4u |
| 95846 | 51349 | 52095 | 52096 | 52103 | 102208 | 102209 | 1q52 |
| 95881 | 56835 | 56836 | 56854 | 56855 | 103426 | 103427 | 1q59 |
| 95891 | 46456 | 48263 | 48264 | 48265 | 101369 | 101370 | 1q5d |
| 95901 | 46456 | 48556 | 48557 | 48558 | 101444 | 101445 | 1q5n |
| 95904 | 53931 | 56234 | 56235 | 56251 | 56255 | 103314 | 1q5q |
| 95911 | 53931 | 56234 | 56235 | 56251 | 56252 | 103313 | 1q5q |
| 95975 | 53931 | 54533 | 54534 | 54535 | 102866 | 102867 | 1q6h |
| 95986 | 51349 | 51350 | 51366 | 51375 | 75050 | 75051 | 1q6o |
| 96000 | 53931 | 54636 | 54637 | 82636 | 102905 | 102906 | 1q6w |
| 96203 | 51349 | 53589 | 53590 | 53591 | 102633 | 102634 | 1q8f |
| 96208 | 51349 | 53066 | 53098 | 53118 | 53125 | 102482 | 1q8i |
| 96214 | 53931 | 54861 | 55008 | 55009 | 55012 | 55013 | 1q8l |
| 96234 | 53931 | 56111 | 56112 | 88854 | 56148 | 56149 | 1q8y |
| 96258 | 46456 | 47112 | 47113 | 47134 | 101107 | 101108 | 1q9c |
| 96346 | 46456 | 47472 | 47473 | 47502 | 47512 | 63541 | 1qv0 |
| 96446 | 51349 | 52171 | 52317 | 52325 | 102253 | 102254 | 1qvw |
| 96467 | 51349 | 51350 | 51445 | 51487 | 102075 | 102076 | 1qw9 |
| 96471 | 48724 | 50813 | 50814 | 50815 | 101861 | 101862 | 1qwd |
| 96478 | 51349 | 53447 | 53448 | 68901 | 53456 | 102608 | 1qwj |
| 96581 | 51349 | 51734 | 51735 | 51751 | 102140 | 102141 | 1qyc |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 96611 | 48724 | 48725 | 48726 | 49159 | 49166 | 49167 | 1qz1 |
| 96621 | 51349 | 53382 | 53383 | 53402 | 102601 | 102602 | 1qz9 |
| 96643 | 48724 | 50198 | 50249 | 50263 | 101761 | 101762 | 1qzg |
| 96651 | 51349 | 52539 | 52540 | 81269 | 102393 | 102394 | 1qzm |
| 96734 | 53931 | 56111 | 56112 | 88854 | 103300 | 103301 | 1r0p |
| 96736 | 56992 | 100894 | 100895 | 57468 | 57469 | 57470 | 1r0r |
| 96754 | 51349 | 52539 | 52540 | 52686 | 102378 | 102379 | 1r0w |
| 96816 | 56992 | 57015 | 57095 | 57116 | 103544 | 103545 | 1r1g |
| 96847 | 51349 | 52832 | 52833 | 52834 | 52835 | 102430 | 1r26 |
| 96874 | 51349 | 52539 | 52540 | 52592 | 82399 | 82400 | 1r2q |
| 96895 | 51349 | 51350 | 102114 | 102115 | 102116 | 102117 | 1r30 |
| 96963 | 51349 | 51350 | 51726 | 51727 | 51728 | 51729 | 1r3s |
| 97080 | 53931 | 55728 | 55729 | 55730 | 103173 | 103174 | 1r57 |
| 97104 | 46456 | 47472 | 47565 | 47566 | 101192 | 101193 | 1r5r |
| 97134 | 51349 | 51350 | 51713 | 51714 | 51715 | 51716 | 1r5y |
| 97138 | 51349 | 52008 | 102198 | 102199 | 102200 | 102201 | 1r61 |
| 97148 | 51349 | 51734 | 51735 | 51751 | 51755 | 102133 | 1r6d |
| 97151 | 48724 | 50155 | 50156 | 50157 | 89311 | 89312 | 1r6j |
| 97161 | 51349 | 52373 | 52374 | 52375 | 52378 | 102256 | 1r6t |
| 97205 | 51349 | 52539 | 52540 | 81269 | 64038 | 64039 | 1r7r |
| 97247 | 51349 | 52539 | 52540 | 52592 | 52614 | 52615 | 1r8s |
| 97281 | 56835 | 81339 | 81338 | 56895 | 103470 | 103471 | 1rc2 |
| 97314 | 53931 | 56111 | 56112 | 88854 | 56116 | 56119 | 1rdq |
| 97455 | 48724 | 50155 | 50156 | 50157 | 101735 | 101736 | 1rgw |
| 97465 | 56835 | 81407 | 103456 | 103457 | 103458 | 103459 | 1rh5 |
| 97469 | 48724 | 48725 | 48726 | 48727 | 101508 | 101509 | 1rhf |
| 97470 | 48724 | 48725 | 48726 | 49159 | 101517 | 101518 | 1rhf |
| 97517 | 53931 | 55680 | 55681 | 55682 | 103155 | 103156 | 1riq |
| 97581 | 53931 | 54116 | 54117 | 54118 | 102732 | 102733 | 1rjt |
| 97603 | 53931 | 54861 | 54928 | 54929 | 89938 | 89939 | 1rk8 |
| 97615 | 56835 | 81407 | 103464 | 103465 | 103466 | 103467 | 1rkl |
| 97621 | 51349 | 56783 | 56784 | 82388 | 102315 | 102316 | 1rkq |
| 97631 | 51349 | 51734 | 51735 | 51751 | 102136 | 102137 | 1rkx |
| 97744 | 46456 | 48575 | 48576 | 48577 | 48578 | 101453 | 1rqj |
| 97820 | 46456 | 48575 | 48576 | 48577 | 48578 | 101454 | 1rtr |
| 98065 | 51349 | 53162 | 53167 | 53168 | 53176 | 53177 | 1rxy |
| 98132 | 51349 | 52539 | 52540 | 52584 | 102358 | 102359 | 1rz3 |
| 98236 | 48724 | 50155 | 50156 | 50157 | 89315 | 101716 | 1rzx |
| 98263 | 51349 | 53382 | 53383 | 53417 | 53438 | 53439 | 1s0a |
| 98329 | 53931 | 55873 | 55874 | 55879 | 103228 | 103229 | 1s14 |
| 98344 | 48724 | 50933 | 101887 | 101888 | 101889 | 101890 | 1s1d |
| 98377 | 51349 | 52171 | 52309 | 52310 | 102244 | 102245 | 1s2d |
| 98573 | 53931 | 54636 | 54637 | 54638 | 102901 | 102902 | 1s5u |
| 98594 | 46456 | 47472 | 47473 | 47502 | 101184 | 101185 | 1s6c |
| 98634 | 56835 | 103480 | 103481 | 103482 | 103483 | 103484 | 1s7b |
| 98777 | 46456 | 48661 | 101494 | 101495 | 101496 | 101497 | 1sa0 |
| 98790 | 53931 | 54861 | 55008 | 55009 | 55017 | 102998 | 1sb6 |
| 98814 | 56992 | 57923 | 57924 | 57925 | 69974 | 69975 | 1se0 |
| 98821 | 53931 | 54235 | 54236 | 54237 | 102782 | 102783 | 1se9 |
| 98892 | 48724 | 51268 | 51283 | 51284 | 51285 | 82213 | 1six |
| 98900 | 53931 | 55810 | 55811 | 55812 | 103209 | 103210 | 1sjy |
| 98958 | 51349 | 51734 | 51735 | 51751 | 102155 | 102156 | 1spx |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 98987 | 51349 | 52832 | 52833 | 52901 | 102455 | 102456 | 1st9 |
| 99028 | 51349 | 52832 | 52833 | 52834 | 52835 | 102431 | 1syr |
| 99066 | 53931 | 54235 | 54236 | 54237 | 102786 | 102787 | 1t0y |
| 99075 | 53931 | 55944 | 55961 | 64388 | 64389 | 64390 | 1t27 |
| 99125 | 46456 | 101477 | 101478 | 101479 | 101480 | 101481 | 1t5j |
| 99161 | 53931 | 103364 | 103365 | 103366 | 103367 | 103368 | 1uc2 |
| 99229 | 51349 | 52539 | 52540 | 52592 | 82408 | 102368 | 1udx |
| 99243 | 48724 | 50198 | 50249 | 50263 | 50264 | 101760 | 1ue1 |
| 99259 | 51349 | 64004 | 64005 | 64006 | 64007 | 64009 | 1ueh |
| 99265 | 53931 | 54235 | 54236 | 54237 | 102777 | 102778 | 1uel |
| 99270 | 48724 | 50155 | 50156 | 50157 | 101719 | 101720 | 1uep |
| 99271 | 48724 | 50155 | 50156 | 50157 | 101719 | 101720 | 1ueq |
| 99281 | 48724 | 50155 | 50156 | 50157 | 101719 | 101720 | 1uew |
| 99288 | 48724 | 50155 | 50156 | 50157 | 101721 | 101722 | 1uf1 |
| 99326 | 51349 | 52539 | 52540 | 52541 | 75187 | 102344 | 1uf9 |
| 99358 | 53931 | 54861 | 54928 | 54929 | 102985 | 102986 | 1ufw |
| 99359 | 48724 | 50155 | 50156 | 50157 | 101721 | 101722 | 1ufx |
| 99403 | 51349 | 51350 | 51445 | 51487 | 102077 | 102078 | 1uhv |
| 99429 | 51349 | 53334 | 53335 | 69557 | 69558 | 102586 | 1uir |
| 99433 | 48724 | 50155 | 50156 | 50157 | 101725 | 101726 | 1uit |
| 99448 | 51349 | 52095 | 52096 | 52103 | 52106 | 102207 | 1uiy |
| 99452 | 51349 | 52539 | 52540 | 52584 | 102360 | 102361 | 1uj2 |
| 99454 | 48724 | 50155 | 50156 | 50157 | 101727 | 101728 | 1ujd |
| 99462 | 51349 | 51350 | 51366 | 51381 | 51388 | 102038 | 1ujp |
| 99470 | 48724 | 50155 | 50156 | 50157 | 101719 | 101720 | 1ujv |
| 99549 | 48724 | 49898 | 49899 | 49932 | 101639 | 101640 | 1ule |
| 99578 | 48724 | 50155 | 50156 | 50157 | 101729 | 101730 | 1um1 |
| 99580 | 51349 | 52539 | 52540 | 81269 | 102391 | 102392 | 1um8 |
| 99607 | 46456 | 47239 | 47240 | 47241 | 47250 | 101134 | 1umn |
| 99639 | 48724 | 49898 | 49899 | 49925 | 101634 | 101635 | 1umz |
| 99686 | 51349 | 53066 | 53098 | 102492 | 102493 | 102494 | 1uoc |
| 99690 | 53931 | 74651 | 48403 | 48404 | 102881 | 102882 | 1uoh |
| 99778 | 53931 | 56370 | 56371 | 56372 | 56389 | 56390 | 1uq5 |
| 99794 | 51349 | 53755 | 53756 | 82540 | 82541 | 82542 | 1uqt |
| 99804 | 51349 | 51350 | 51445 | 51487 | 51488 | 102064 | 1ur1 |
| 99806 | 51349 | 51350 | 51430 | 51431 | 89459 | 89460 | 1ur3 |
| 99812 | 56992 | 57849 | 57850 | 57851 | 64579 | 64580 | 1ur6 |
| 99850 | 51349 | 51350 | 51430 | 51431 | 51436 | 51437 | 1us0 |
| 99881 | 48724 | 50938 | 50939 | 50940 | 63823 | 63824 | 1usr |
| 99999 | 53931 | 56111 | 56112 | 88854 | 90036 | 90037 | 1uu3 |
| 100015 | 51349 | 51350 | 51445 | 51487 | 102068 | 102069 | 1uuq |
| 100123 | 53931 | 56111 | 56112 | 88854 | 103290 | 103291 | 1uwh |
| 100220 | 51349 | 53719 | 53720 | 53721 | 102675 | 102676 | 1uzb |
| 100247 | 51349 | 53612 | 53613 | 53614 | 75295 | 102640 | 1v19 |
| 100275 | 53931 | 54235 | 54236 | 54237 | 102784 | 102785 | 1v2y |
| 100287 | 48724 | 50938 | 50939 | 50940 | 101891 | 101892 | 1v3e |
| 100296 | 46456 | 48575 | 48576 | 48577 | 101455 | 101456 | 1v4e |
| 100321 | 56835 | 81443 | 81442 | 81441 | 81433 | 81432 | 1v54 |
| 100325 | 56835 | 81407 | 81406 | 81405 | 81404 | 81403 | 1v54 |
| 100383 | 48724 | 49898 | 49899 | 49900 | 49904 | 49912 | 1v6i |
| 100424 | 51349 | 53747 | 53748 | 53749 | 53750 | 102677 | 1v6s |
| 100446 | 51349 | 53685 | 53686 | 53687 | 64172 | 102667 | 1v7c |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 100458 | 46456 | 47265 | 47266 | 47286 | 101140 | 101141 | 1v7m |
| 100495 | 53931 | 54235 | 54285 | 54286 | 102794 | 102795 | 1v8c |
| 100519 | 51349 | 88722 | 88723 | 89619 | 102270 | 102271 | 1v8p |
| 100557 | 51349 | 51350 | 51366 | 51381 | 51385 | 102037 | 1vc4 |
| 100561 | 46456 | 48556 | 48557 | 48558 | 48561 | 101443 | 1vdk |
| 100569 | 56572 | 56654 | 56655 | 56656 | 56665 | 103398 | 1vdw |
| 100637 | 51349 | 53447 | 53448 | 68901 | 64143 | 102607 | 1vh3 |
| 100640 | 48724 | 51125 | 101960 | 101961 | 101962 | 101963 | 1vh4 |
| 100642 | 53931 | 54636 | 54637 | 89902 | 102910 | 102911 | 1vh5 |
| 100661 | 51349 | 51350 | 51569 | 51570 | 102092 | 102093 | 1vhc |
| 100693 | 51349 | 51350 | 51395 | 51396 | 102044 | 102045 | 1vhn |
| 100696 | 51349 | 52171 | 52317 | 52325 | 89606 | 89607 | 1vhq |
| 100698 | 53931 | 55728 | 55729 | 55730 | 103171 | 103172 | 1vhs |
| 100700 | 51349 | 52539 | 52540 | 52541 | 75187 | 82394 | 1vht |
| 100706 | 51349 | 53162 | 53167 | 53168 | 53169 | 102499 | 1vhw |
| 100718 | 53931 | 55810 | 55811 | 55812 | 103203 | 103204 | 1vhz |
| 100736 | 51349 | 52171 | 52313 | 52314 | 52315 | 102246 | 1vi6 |
| 100750 | 51349 | 53612 | 53613 | 82515 | 102638 | 102639 | 1vi9 |
| 100754 | 51349 | 52539 | 52540 | 52566 | 52567 | 102345 | 1via |
| 100762 | 53931 | 100877 | 55120 | 75459 | 75460 | 103018 | 1vio |
| 100771 | 53931 | 55810 | 55811 | 55812 | 103201 | 103202 | 1viu |
| 100784 | 51349 | 51350 | 51395 | 51396 | 102046 | 102047 | 1viz |
| 100801 | 46456 | 109603 | 109604 | 101340 | 101341 | 101342 | 1vj7 |
| 100827 | 51349 | 53382 | 53383 | 53402 | 89757 | 102598 | 1vjo |
| 100832 | 51349 | 56783 | 56784 | 102317 | 102318 | 102319 | 1vjr |
| 100835 | 53931 | 102885 | 102886 | 102887 | 102888 | 102889 | 1vju |
| 100837 | 46456 | 47239 | 47240 | 47241 | 101132 | 101133 | 1vjx |
| 100838 | 51349 | 51350 | 51445 | 51487 | 102066 | 102067 | 1vjz |
| 100839 | 51349 | 53066 | 53098 | 53118 | 102483 | 102484 | 1vk0 |
| 100850 | 51349 | 53612 | 53613 | 53614 | 102641 | 102642 | 1vk4 |
| 103833 | 46456 | 47094 | 47095 | 47096 | 109762 | 109763 | 1j3d |
| 103874 | 48724 | 50728 | 50729 | 50730 | 110264 | 110265 | 1nty |
| 103971 | 56572 | 56795 | 56796 | 69892 | 111308 | 111309 | 1oj7 |
| 104019 | 48724 | 50493 | 50494 | 50514 | 50570 | 110238 | 1op0 |
| 104022 | 53931 | 54861 | 55008 | 55009 | 55010 | 110979 | 1osd |
| 104060 | 51349 | 51350 | 51569 | 51570 | 69394 | 69395 | 1p1x |
| 104062 | 46456 | 47575 | 47576 | 47577 | 101196 | 101197 | 1p2x |
| 104067 | 51349 | 53685 | 53686 | 53687 | 102668 | 110721 | 1p5j |
| 104074 | 51349 | 52171 | 52172 | 52173 | 52174 | 102225 | 1p6q |
| 104084 | 51349 | 52787 | 52788 | 52789 | 52790 | 110595 | 1p8a |
| 104121 | 48724 | 48725 | 48726 | 49159 | 89185 | 89186 | 1pd6 |
| 104137 | 51349 | 52171 | 52172 | 52173 | 52188 | 52189 | 1pey |
| 104333 | 56992 | 57015 | 57095 | 57116 | 111396 | 111397 | 1pvz |
| 104381 | 56992 | 57015 | 57095 | 57116 | 90137 | 90138 | 1px9 |
| 104393 | 46456 | 48507 | 48508 | 48509 | 109605 | 109987 | 1pzl |
| 104397 | 48724 | 48725 | 49329 | 49330 | 49331 | 110065 | 1pzs |
| 104475 | 53931 | 54533 | 54534 | 54535 | 82619 | 82620 | 1q1c |
| 104476 | 46456 | 46457 | 46458 | 46463 | 100978 | 109625 | 1q1f |
| 104525 | 51349 | 52539 | 52540 | 52541 | 52548 | 110525 | 1q3t |
| 104526 | 48724 | 50493 | 50494 | 50514 | 110241 | 110242 | 1q3x |
| 104537 | 53931 | 56111 | 56112 | 88854 | 69823 | 69824 | 1q5k |
| 104555 | 53931 | 55752 | 55753 | 55762 | 55763 | 111104 | 1q8g |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 104620 | 51349 | 53253 | 53254 | 53258 | 53263 | 110657 | 1qwo |
| 104812 | 51349 | 53926 | 53927 | 53928 | 75327 | 110764 | 1r5t |
| 104820 | 51349 | 52539 | 52540 | 81269 | 82421 | 82422 | 1r6b |
| 104839 | 51349 | 51350 | 51445 | 51487 | 51488 | 69386 | 1r85 |
| 104851 | 48724 | 50352 | 50386 | 50387 | 110213 | 110214 | 1r8n |
| 104852 | 48724 | 50352 | 50386 | 50387 | 110215 | 110216 | 1r8o |
| 105024 | 53931 | 54402 | 54403 | 54407 | 110819 | 110820 | 1roa |
| 105027 | 46456 | 110018 | 110019 | 110020 | 110021 | 110022 | 1rp4 |
| 105058 | 51349 | 51350 | 51569 | 89494 | 110366 | 110367 | 1rqb |
| 105064 | 51349 | 52373 | 52374 | 52375 | 52384 | 110490 | 1rqg |
| 105081 | 56572 | 56795 | 56796 | 69892 | 111310 | 111311 | 1rrm |
| 105095 | 46456 | 47575 | 47576 | 47577 | 47580 | 109828 | 1rt8 |
| 105140 | 51349 | 53755 | 53756 | 110734 | 110735 | 110736 | 1rzu |
| 105250 | 53931 | 55728 | 55729 | 55730 | 111098 | 111099 | 1s3z |
| 105310 | 46456 | 47472 | 47473 | 47502 | 109821 | 109822 | 1s6i |
| 105317 | 53931 | 74651 | 48403 | 48404 | 110891 | 110892 | 1s70 |
| 105374 | 51349 | 52171 | 52172 | 52173 | 110464 | 110465 | 1s8n |
| 105377 | 48724 | 89359 | 89360 | 89361 | 101849 | 101850 | 1s98 |
| 105379 | 48724 | 49451 | 49482 | 49483 | 110096 | 110097 | 1s9a |
| 105410 | 51349 | 51734 | 51735 | 51751 | 110407 | 110408 | 1sb8 |
| 105412 | 51349 | 100949 | 100950 | 110520 | 110523 | 110524 | 1sbq |
| 105416 | 51349 | 52128 | 52129 | 52130 | 52133 | 52134 | 1sc3 |
| 105448 | 48724 | 50198 | 50249 | 50263 | 50264 | 110194 | 1se8 |
| 105455 | 51349 | 53596 | 53597 | 53598 | 53610 | 102636 | 1sej |
| 105465 | 51349 | 52832 | 52833 | 52834 | 110604 | 110605 | 1sen |
| 105482 | 51349 | 53382 | 53383 | 53417 | 53424 | 110686 | 1sff |
| 105535 | 51349 | 51350 | 51621 | 51638 | 110375 | 110376 | 1sgj |
| 105541 | 51349 | 52539 | 52540 | 52686 | 110560 | 110561 | 1sgw |
| 105543 | 46456 | 47575 | 47576 | 47577 | 89056 | 89057 | 1sh5 |
| 105549 | 53931 | 54636 | 54637 | 89902 | 110900 | 110901 | 1sh8 |
| 105764 | 48724 | 48725 | 48726 | 48727 | 101506 | 101507 | 1smo |
| 105772 | 48724 | 50198 | 50249 | 50282 | 110198 | 110199 | 1smx |
| 105778 | 56572 | 64483 | 64484 | 64485 | 64488 | 75609 | 1smy |
| 105779 | 56572 | 64483 | 64484 | 64490 | 64493 | 75610 | 1smy |
| 105833 | 51349 | 51734 | 51735 | 51751 | 110413 | 110414 | 1sny |
| 105858 | 51349 | 100949 | 100950 | 110520 | 110521 | 110522 | 1sou |
| 105881 | 48724 | 48725 | 48726 | 48727 | 110042 | 110043 | 1sq2 |
| 105882 | 51349 | 52539 | 52540 | 52584 | 52587 | 52588 | 1sq5 |
| 105889 | 48724 | 50964 | 50978 | 50979 | 110287 | 110288 | 1sq9 |
| 105961 | 51349 | 51350 | 51569 | 89494 | 110364 | 110365 | 1sr9 |
| 105966 | 56992 | 57666 | 57667 | 57668 | 111434 | 111435 | 1srk |
| 106048 | 51349 | 52539 | 52540 | 52592 | 89667 | 110534 | 1svi |
| 106051 | 51349 | 52539 | 52540 | 52592 | 52623 | 52625 | 1svs |
| 106054 | 51349 | 53382 | 53383 | 53384 | 64123 | 110676 | 1svv |
| 106082 | 51349 | 52539 | 52540 | 81269 | 110566 | 110567 | 1sxj |
| 106090 | 51349 | 52539 | 52540 | 81269 | 110574 | 110575 | 1sxj |
| 106101 | 53931 | 55845 | 55846 | 55847 | 111139 | 111140 | 1sxr |
| 106126 | 46456 | 47239 | 47240 | 47253 | 47257 | 109786 | 1syy |
| 106164 | 51349 | 52095 | 52096 | 52103 | 82331 | 82332 | 1szo |
| 106208 | 51349 | 52539 | 52540 | 52541 | 110527 | 110528 | 1t0h |
| 106238 | 53931 | 54533 | 54534 | 54535 | 75388 | 110869 | 1t11 |
| 106354 | 51349 | 53382 | 53383 | 53402 | 110684 | 110685 | 1t3i |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 106357 | 51349 | 52820 | 52821 | 52822 | 110598 | 110599 | 1t3k |
| 106397 | 53931 | 56111 | 56112 | 88854 | 103296 | 103297 | 1t46 |
| 106420 | 53931 | 56111 | 56112 | 88854 | 111196 | 111197 | 1t4h |
| 106459 | 51349 | 100949 | 100950 | 110513 | 110518 | 110519 | 1t5o |
| 106543 | 51349 | 52767 | 52768 | 52773 | 110587 | 110588 | 1t64 |
| 106564 | 48724 | 50629 | 50630 | 50646 | 110247 | 110248 | 1t6e |
| 106635 | 46456 | 48507 | 48508 | 48509 | 63621 | 109986 | 1t7r |
| 106639 | 53931 | 54636 | 54637 | 89902 | 110898 | 110899 | 1t82 |
| 106679 | 51349 | 53162 | 53167 | 53168 | 110645 | 110646 | 1t8s |
| 106685 | 51349 | 52539 | 52540 | 52575 | 110532 | 110533 | 1t8t |
| 106729 | 51349 | 56783 | 56784 | 110509 | 110510 | 110511 | 1ta0 |
| 106731 | 51349 | 51350 | 51445 | 51534 | 89478 | 89479 | 1ta3 |
| 106785 | 53931 | 56435 | 56436 | 56437 | 111260 | 111261 | 1tdq |
| 106792 | 51349 | 56783 | 56784 | 75173 | 110504 | 110505 | 1te2 |
| 106794 | 46456 | 48370 | 48371 | 101410 | 109963 | 109964 | 1te4 |
| 106818 | 51349 | 52539 | 52540 | 52575 | 110530 | 110531 | 1tex |
| 106854 | 53931 | 54000 | 54001 | 110773 | 110774 | 110775 | 1tff |
| 107010 | 53931 | 55728 | 55729 | 55730 | 111100 | 111101 | 1tiq |
| 107014 | 46456 | 47472 | 47473 | 47502 | 109823 | 109824 | 1tiz |
| 107054 | 46456 | 48556 | 48557 | 48558 | 48564 | 48566 | 1tjv |
| 107161 | 51349 | 51350 | 51569 | 51570 | 110358 | 110359 | 1to3 |
| 107201 | 51349 | 52539 | 52540 | 52592 | 110546 | 110547 | 1tq4 |
| 107228 | 51349 | 51350 | 51366 | 51372 | 51373 | 110343 | 1tqj |
| 107238 | 46456 | 48263 | 48264 | 48265 | 109958 | 109959 | 1tqn |
| 107309 | 51349 | 52832 | 52833 | 52834 | 110606 | 110607 | 1ttz |
| 107320 | 46456 | 46457 | 46458 | 46463 | 109628 | 109629 | 1tu9 |
| 107323 | 51349 | 52539 | 52540 | 81269 | 110576 | 110577 | 1tue |
| 107423 | 48724 | 50728 | 50729 | 50730 | 110270 | 110271 | 1txd |
| 107424 | 48724 | 51315 | 51316 | 63873 | 63874 | 110332 | 1txj |
| 107476 | 51349 | 53447 | 53448 | 68901 | 110689 | 110690 | 1tzf |
| 107530 | 46456 | 47112 | 47113 | 47114 | 47115 | 47116 | 1tzy |
| 107531 | 46456 | 47112 | 47113 | 47114 | 47119 | 47120 | 1tzy |
| 107532 | 46456 | 47112 | 47113 | 47114 | 47122 | 47123 | 1tzy |
| 107543 | 51349 | 53382 | 53383 | 53384 | 110679 | 110680 | 1u08 |
| 107546 | 51349 | 52539 | 52540 | 81269 | 110564 | 110565 | 1u0j |
| 107551 | 48724 | 50198 | 50249 | 50282 | 110200 | 110201 | 1u0l |
| 107566 | 51349 | 52171 | 52172 | 52173 | 52174 | 52177 | 1u0s |
| 107567 | 56572 | 111330 | 111331 | 111332 | 111333 | 111335 | 1u0t |
| 107656 | 53931 | 56111 | 56112 | 88854 | 111200 | 111201 | 1u46 |
| 107714 | 53931 | 54533 | 54534 | 54535 | 110870 | 110871 | 1u79 |
| 107723 | 56835 | 111351 | 111352 | 111353 | 111354 | 111355 | 1u7g |
| 107783 | 48724 | 50198 | 50249 | 50282 | 110196 | 110197 | 1ueb |
| 107784 | 48724 | 50198 | 50249 | 50282 | 110196 | 110197 | 1ueb |
| 107823 | 46456 | 46688 | 46689 | 46739 | 109649 | 109650 | 1ug2 |
| 107828 | 48724 | 48725 | 48726 | 49159 | 49206 | 49207 | 1ugn |
| 107853 | 46456 | 46688 | 46689 | 46690 | 109647 | 109648 | 1uhs |
| 107900 | 46456 | 47575 | 47576 | 47577 | 109830 | 109831 | 1ujo |
| 107968 | 48724 | 50728 | 50729 | 50730 | 89355 | 89356 | 1unq |
| 107998 | 53931 | 54861 | 54975 | 54976 | 110974 | 110975 | 1urr |
| 108073 | 51349 | 51350 | 51445 | 51521 | 51531 | 51532 | 1uws |
| 108275 | 56572 | 56800 | 56801 | 56802 | 111314 | 111315 | 1v25 |
| 108281 | 51349 | 53382 | 53383 | 53384 | 110677 | 110678 | 1v2d |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 108379 | 48724 | 50728 | 50729 | 50730 | 110256 | 110257 | 1v5m |
| 108381 | 53931 | 54235 | 54236 | 54237 | 110796 | 110797 | 1v5o |
| 108382 | 48724 | 50728 | 50729 | 50730 | 110258 | 110259 | 1v5p |
| 108383 | 48724 | 50155 | 50156 | 50157 | 82083 | 110177 | 1v5q |
| 108385 | 53931 | 54235 | 54236 | 54237 | 110798 | 110799 | 1v5t |
| 108386 | 48724 | 50728 | 50729 | 50730 | 110260 | 110261 | 1v5u |
| 108389 | 48724 | 50728 | 50729 | 50730 | 110262 | 110263 | 1v61 |
| 108390 | 48724 | 50155 | 50156 | 50157 | 110180 | 110181 | 1v62 |
| 108392 | 46456 | 47094 | 47095 | 47096 | 68982 | 109764 | 1v64 |
| 108394 | 48724 | 50155 | 50156 | 50157 | 110182 | 110183 | 1v6b |
| 108407 | 53931 | 56435 | 56436 | 56437 | 88867 | 103347 | 1v7p |
| 108415 | 51349 | 53447 | 53448 | 53467 | 110687 | 110688 | 1v82 |
| 108421 | 53931 | 54235 | 54236 | 54237 | 110800 | 110801 | 1v86 |
| 108422 | 56992 | 57849 | 57850 | 57851 | 111457 | 111458 | 1v87 |
| 108423 | 48724 | 50728 | 50729 | 50730 | 110266 | 110267 | 1v88 |
| 108424 | 48724 | 50728 | 50729 | 50730 | 110268 | 110269 | 1v89 |
| 108448 | 53931 | 100877 | 55120 | 75459 | 103019 | 103020 | 1v9f |
| 108449 | 53931 | 100877 | 55120 | 75459 | 111010 | 111011 | 1v9k |
| 108451 | 48724 | 50890 | 50891 | 50892 | 50901 | 50902 | 1v9t |
| 108466 | 48724 | 50155 | 50156 | 50157 | 110184 | 110185 | 1vae |
| 108474 | 48724 | 50155 | 50156 | 50157 | 110186 | 110187 | 1vb7 |
| 108475 | 51349 | 53334 | 53335 | 53354 | 68927 | 110668 | 1vbf |
| 108509 | 51349 | 52539 | 52540 | 52652 | 110551 | 110553 | 1vco |
| 108554 | 53931 | 55769 | 103196 | 103197 | 111123 | 111124 | 1vet |
| 108631 | 53931 | 54235 | 54285 | 54286 | 54287 | 110806 | 1vjk |
| 108633 | 53931 | 56111 | 56112 | 88854 | 56144 | 56145 | 1vjy |
| 108643 | 53931 | 110856 | 110857 | 110858 | 110859 | 110860 | 1vkb |
| 108669 | 53931 | 55752 | 55753 | 55762 | 111110 | 111111 | 1vkk |
| 108704 | 51349 | 51734 | 51735 | 51751 | 110409 | 110410 | 1vl0 |
| 108720 | 51349 | 53334 | 53335 | 110671 | 110672 | 110673 | 1vl5 |
| 108732 | 48724 | 50474 | 50475 | 50476 | 110234 | 110235 | 1vl7 |
| 108733 | 51349 | 51734 | 51735 | 51751 | 110411 | 110412 | 1vl8 |
| 108816 | 46456 | 47239 | 47240 | 47241 | 63524 | 109783 | 1vlg |
| 108832 | 56572 | 56795 | 56796 | 69892 | 111312 | 111313 | 1vlj |
| 108836 | 51349 | 53334 | 53335 | 110671 | 110674 | 110675 | 1vlm |
| 108864 | 51349 | 53719 | 53720 | 53721 | 89781 | 110731 | 1vlu |
| 108884 | 51349 | 53612 | 53613 | 53614 | 53615 | 110706 | 1vm7 |
| 108897 | 53931 | 56218 | 56219 | 56220 | 111211 | 111212 | 1vyb |
| 108902 | 48724 | 50813 | 50814 | 50847 | 110276 | 110277 | 1vyf |
| 108906 | 51349 | 51350 | 51395 | 51396 | 63900 | 63901 | 1vyr |
| 109027 | 51349 | 51350 | 51351 | 51352 | 51353 | 110342 | 1w0m |
| 109040 | 48724 | 50938 | 50939 | 50940 | 50950 | 50951 | 1w0p |
| 109125 | 53931 | 56103 | 56104 | 111188 | 111189 | 111190 | 1w2f |
| 109130 | 51349 | 100949 | 100950 | 110513 | 110516 | 110517 | 1w2w |
| 109149 | 46456 | 48370 | 48452 | 48453 | 109973 | 109974 | 1w3b |
| 109152 | 51349 | 51350 | 51569 | 51570 | 110360 | 110361 | 1w3i |
| 109171 | 51349 | 51904 | 51905 | 51943 | 110440 | 110441 | 1w4x |
| 109172 | 51349 | 51904 | 51905 | 51943 | 110440 | 110441 | 1w4x |
| 109216 | 46456 | 47239 | 47240 | 47253 | 47257 | 47260 | 1w68 |
| 109232 | 56992 | 57849 | 57850 | 57851 | 111459 | 111460 | 1wd2 |
| 109253 | 51349 | 52095 | 52096 | 52103 | 110450 | 110451 | 1wdk |
| 109286 | 53931 | 56218 | 56219 | 56220 | 111215 | 111216 | 1wdu |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 109330 | 48724 | 50128 | 50129 | 50130 | 50131 | 110172 | 1we3 |
| 109393 | 51349 | 53926 | 53927 | 89800 | 110765 | 110766 | 1wkq |
| 109395 | 48724 | 50629 | 50630 | 50646 | 50649 | 110246 | 1wkr |
| 109406 | 56992 | 57015 | 57095 | 57116 | 111398 | 111399 | 1wm7 |
| 109409 | 51349 | 52742 | 52743 | 52744 | 110578 | 110579 | 1wmd |
| 109414 | 51349 | 52539 | 52540 | 52592 | 110537 | 110539 | 1wms |
| 109422 | 53931 | 56435 | 56436 | 56437 | 111262 | 111263 | 1wmz |
| 109446 | 51349 | 52767 | 52768 | 52769 | 110585 | 110586 | 1woh |
| 109494 | 48724 | 50198 | 50249 | 50282 | 50296 | 110195 | 1x6o |
| 109527 | 51349 | 52498 | 52499 | 100948 | 110502 | 110503 | 1x9g |
| 111575 | 48724 | 50198 | 50249 | 50282 | 74950 | 74951 | 1kl9 |
| 111639 | 51349 | 53473 | 53474 | 117714 | 117715 | 117716 | 1pv1 |
| 111643 | 48724 | 51125 | 51126 | 51127 | 117306 | 117307 | 1pxz |
| 111694 | 53931 | 55485 | 55486 | 55519 | 118046 | 118047 | 1r55 |
| 111710 | 51349 | 52742 | 52743 | 52744 | 117569 | 117570 | 1r6v |
| 111772 | 51349 | 51350 | 51366 | 51381 | 51388 | 117361 | 1rd5 |
| 111789 | 46456 | 48263 | 48264 | 48265 | 48266 | 48267 | 1re9 |
| 111796 | 48724 | 50474 | 50475 | 50476 | 117232 | 117233 | 1rfe |
| 111801 | 48724 | 50352 | 50353 | 50354 | 50357 | 50359 | 1rg8 |
| 111835 | 56992 | 57867 | 57868 | 57869 | 57870 | 57876 | 1rju |
| 111854 | 51349 | 52539 | 52540 | 52541 | 52554 | 117527 | 1rkb |
| 111859 | 53931 | 56370 | 56371 | 56372 | 103335 | 103336 | 1rl0 |
| 111860 | 51349 | 56783 | 56784 | 82388 | 117503 | 117504 | 1rlm |
| 111924 | 48724 | 50493 | 50494 | 50514 | 50561 | 50562 | 1rrk |
| 111959 | 51349 | 52798 | 52799 | 52800 | 102418 | 117578 | 1rxd |
| 112008 | 46456 | 48263 | 48264 | 48265 | 116991 | 116992 | 1s1f |
| 112040 | 51349 | 53066 | 53098 | 53118 | 53125 | 117647 | 1s5j |
| 112044 | 46456 | 46688 | 46689 | 46690 | 116780 | 116781 | 1s7e |
| 112049 | 53931 | 56111 | 56112 | 88854 | 118137 | 118138 | 1s9j |
| 112068 | 51349 | 52506 | 52507 | 52508 | 117500 | 117501 | 1sbz |
| 112077 | 51349 | 51350 | 51569 | 51570 | 51586 | 117378 | 1sfl |
| 112091 | 48724 | 50493 | 50494 | 50514 | 110239 | 110240 | 1si5 |
| 112106 | 48724 | 50938 | 50939 | 50940 | 117274 | 117275 | 1so7 |
| 112126 | 51349 | 52539 | 52540 | 81269 | 89688 | 89689 | 1svm |
| 112213 | 56992 | 57015 | 57095 | 57096 | 57097 | 118240 | 1t0z |
| 112259 | 46456 | 48556 | 48557 | 48572 | 117028 | 117029 | 1t6j |
| 112358 | 48724 | 50198 | 50249 | 50282 | 110200 | 117197 | 1t9h |
| 112359 | 51349 | 52539 | 52540 | 52592 | 110542 | 117534 | 1t9h |
| 112364 | 48724 | 50474 | 50475 | 50476 | 50479 | 117226 | 1t9m |
| 112429 | 51349 | 52832 | 52833 | 52834 | 52835 | 117590 | 1ti3 |
| 112440 | 46456 | 48370 | 48452 | 48453 | 117007 | 117008 | 1tjc |
| 112461 | 46456 | 47239 | 47240 | 47241 | 47250 | 101135 | 1tjo |
| 112614 | 51349 | 51350 | 51366 | 51372 | 51373 | 117359 | 1tqx |
| 112683 | 48724 | 49784 | 49785 | 69210 | 117104 | 117105 | 1tvg |
| 112690 | 48724 | 50813 | 50814 | 50847 | 89368 | 89369 | 1tw4 |
| 112725 | 56572 | 64483 | 64484 | 64490 | 64486 | 64487 | 1twf |
| 112726 | 56572 | 64483 | 64484 | 64485 | 64491 | 64492 | 1twf |
| 112731 | 46456 | 63561 | 63562 | 55294 | 55295 | 64318 | 1twf |
| 112736 | 53931 | 55247 | 55257 | 64311 | 64312 | 64313 | 1twf |
| 112779 | 53931 | 102885 | 102886 | 102887 | 110896 | 110897 | 1txn |
| 112781 | 53931 | 81607 | 81606 | 81605 | 118149 | 118150 | 1txo |
| 112850 | 51349 | 53612 | 53613 | 53614 | 117720 | 117721 | 1tyy |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 112852 | 51349 | 53685 | 53686 | 53687 | 53694 | 110722 | 1tyz |
| 112953 | 56835 | 81322 | 81321 | 81320 | 56876 | 56877 | 1u19 |
| 113037 | 51349 | 51350 | 51621 | 51638 | 110375 | 117388 | 1u5h |
| 113050 | 53931 | 56111 | 56112 | 88854 | 118127 | 118128 | 1u5r |
| 113069 | 53931 | 55728 | 55729 | 55730 | 118070 | 118071 | 1u6m |
| 113093 | 56572 | 118202 | 118203 | 118204 | 118205 | 118206 | 1u7l |
| 113233 | 46456 | 47071 | 47072 | 117033 | 117034 | 117035 | 1u97 |
| 113234 | 51349 | 52171 | 52317 | 52325 | 117481 | 117482 | 1u9c |
| 113283 | 53931 | 54861 | 54975 | 54976 | 54977 | 117977 | 1ulr |
| 113392 | 48724 | 50728 | 50729 | 50730 | 117244 | 117245 | 1upq |
| 113394 | 48724 | 49898 | 49899 | 49925 | 117134 | 117135 | 1ups |
| 113402 | 51349 | 51350 | 51445 | 51487 | 89469 | 117367 | 1ur4 |
| 113414 | 46456 | 46457 | 46458 | 46463 | 109626 | 109627 | 1urv |
| 113422 | 53931 | 55680 | 55681 | 55682 | 118058 | 118059 | 1usy |
| 113490 | 48724 | 50198 | 50249 | 50263 | 117191 | 117192 | 1v1q |
| 113496 | 53931 | 110856 | 110857 | 110858 | 117851 | 117852 | 1v30 |
| 113497 | 51349 | 53253 | 53254 | 53255 | 117669 | 117670 | 1v37 |
| 113544 | 51349 | 52742 | 52743 | 52744 | 117571 | 117572 | 1v6c |
| 113547 | 53931 | 54235 | 54236 | 54237 | 102786 | 117798 | 1v6e |
| 113548 | 56992 | 57301 | 57302 | 57303 | 57341 | 57342 | 1v6p |
| 113561 | 56992 | 57015 | 57059 | 57072 | 118233 | 118234 | 1v7f |
| 113581 | 53931 | 55810 | 55811 | 55812 | 64365 | 118083 | 1v8y |
| 113604 | 51349 | 100949 | 100950 | 110513 | 117517 | 117518 | 1vb5 |
| 113637 | 53931 | 54861 | 54913 | 54914 | 54915 | 102972 | 1vfj |
| 113663 | 53931 | 55728 | 55729 | 55730 | 118072 | 118073 | 1vkc |
| 113676 | 46456 | 101385 | 101386 | 116993 | 116994 | 116995 | 1vmg |
| 113935 | 51349 | 53382 | 53383 | 53384 | 117691 | 117692 | 1vp4 |
| 113948 | 51349 | 53447 | 53448 | 68901 | 64140 | 117699 | 1vpa |
| 113966 | 51349 | 52539 | 52540 | 52686 | 117548 | 117549 | 1vpl |
| 113967 | 53931 | 54636 | 54637 | 54638 | 117883 | 117884 | 1vpm |
| 114004 | 51349 | 53926 | 53927 | 89800 | 117756 | 117757 | 1vq2 |
| 114060 | 51349 | 53658 | 53659 | 53660 | 53661 | 117725 | 1w0d |
| 114064 | 51349 | 53066 | 53098 | 53118 | 117648 | 117649 | 1w0h |
| 114076 | 48724 | 50728 | 50729 | 50730 | 117246 | 117247 | 1w1h |
| 114080 | 51349 | 52539 | 52540 | 52686 | 52693 | 117544 | 1w1w |
| 114088 | 51349 | 53382 | 53383 | 53417 | 53426 | 117697 | 1w23 |
| 114090 | 51349 | 52171 | 52172 | 52173 | 117472 | 117473 | 1w25 |
| 114091 | 51349 | 52171 | 52172 | 52173 | 117472 | 117473 | 1w25 |
| 114096 | 46456 | 48556 | 48557 | 48572 | 117028 | 117030 | 1w27 |
| 114120 | 51349 | 51350 | 51445 | 51487 | 51514 | 51517 | 1w32 |
| 114197 | 51349 | 53447 | 53448 | 68901 | 117700 | 117701 | 1w55 |
| 114252 | 51349 | 52539 | 52540 | 81269 | 117556 | 117557 | 1w5s |
| 114275 | 48724 | 49898 | 49899 | 49932 | 100925 | 101638 | 1w6n |
| 114288 | 51349 | 51734 | 51735 | 51751 | 117421 | 117422 | 1w6u |
| 114317 | 46456 | 47873 | 47874 | 47875 | 116947 | 116948 | 1w7b |
| 114368 | 51349 | 88722 | 88723 | 89619 | 117492 | 117493 | 1w8i |
| 114376 | 48724 | 50938 | 50939 | 50940 | 50946 | 50947 | 1w8o |
| 114408 | 48724 | 50474 | 50475 | 50476 | 117230 | 117231 | 1w9a |
| 114412 | 51349 | 51350 | 51445 | 51534 | 51548 | 117368 | 1w9p |
| 114432 | 46456 | 48370 | 48371 | 48372 | 48383 | 48384 | 1wa5 |
| 114433 | 46456 | 48370 | 48371 | 48372 | 116998 | 116999 | 1wa5 |
| 114439 | 51349 | 52539 | 52540 | 52592 | 117537 | 117538 | 1wb1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 114524 | 51349 | 53270 | 53271 | 53272 | 117673 | 117674 | 1wd5 |
| 114526 | 51349 | 69617 | 69618 | 69619 | 117738 | 117739 | 1wd7 |
| 114534 | 51349 | 53789 | 53790 | 53791 | 102684 | 117735 | 1wde |
| 114541 | 53931 | 74651 | 48403 | 48404 | 117879 | 117880 | 1wdy |
| 114546 | 53931 | 54235 | 54236 | 54237 | 117799 | 117801 | 1we6 |
| 114561 | 56992 | 57849 | 57850 | 57851 | 118294 | 118295 | 1weo |
| 114568 | 53931 | 54861 | 54928 | 54929 | 117951 | 117952 | 1wex |
| 114569 | 53931 | 54861 | 54928 | 54929 | 117953 | 117954 | 1wey |
| 114571 | 53931 | 54861 | 54928 | 54929 | 117957 | 117958 | 1wf0 |
| 114573 | 53931 | 54861 | 54928 | 54929 | 117961 | 117962 | 1wf2 |
| 114574 | 51349 | 52539 | 52540 | 52592 | 52637 | 117533 | 1wf3 |
| 114589 | 48724 | 50198 | 50249 | 50282 | 117198 | 117199 | 1wfq |
| 114599 | 53931 | 54861 | 54928 | 54929 | 117963 | 117964 | 1wg1 |
| 114602 | 53931 | 54861 | 54928 | 54929 | 117955 | 117956 | 1wg5 |
| 114603 | 48724 | 50155 | 50156 | 50157 | 117178 | 117179 | 1wg6 |
| 114604 | 48724 | 50728 | 50729 | 50730 | 117248 | 117249 | 1wg7 |
| 114610 | 53931 | 54235 | 54236 | 54237 | 117802 | 117803 | 1wgd |
| 114611 | 46456 | 47094 | 47095 | 47096 | 68982 | 109764 | 1wgf |
| 114613 | 53931 | 54235 | 54236 | 54237 | 117806 | 117807 | 1wgh |
| 114620 | 48724 | 50728 | 50729 | 50730 | 117250 | 117251 | 1wgq |
| 114626 | 46456 | 46688 | 46689 | 46739 | 116782 | 116783 | 1wgx |
| 114629 | 48724 | 50155 | 50156 | 50157 | 101723 | 101724 | 1wh1 |
| 114631 | 53931 | 54235 | 54236 | 54237 | 117808 | 117809 | 1wh3 |
| 114635 | 46456 | 46688 | 46689 | 46690 | 116774 | 116775 | 1wh7 |
| 114638 | 48724 | 50155 | 50156 | 50157 | 101731 | 101732 | 1wha |
| 114652 | 53931 | 54861 | 54928 | 54929 | 117967 | 117968 | 1whv |
| 114653 | 53931 | 54861 | 54928 | 54929 | 117969 | 117970 | 1whw |
| 114654 | 53931 | 54861 | 54928 | 54929 | 117969 | 117970 | 1whx |
| 114655 | 53931 | 54861 | 54928 | 54929 | 117971 | 117972 | 1why |
| 114658 | 48724 | 50728 | 50729 | 50730 | 117252 | 117253 | 1wi1 |
| 114661 | 48724 | 50198 | 50249 | 50282 | 117200 | 117201 | 1wi5 |
| 114662 | 53931 | 54861 | 54928 | 54929 | 117973 | 117974 | 1wi8 |
| 114664 | 53931 | 54235 | 54236 | 54237 | 117810 | 117811 | 1wia |
| 114669 | 48724 | 50155 | 50156 | 50157 | 117184 | 117185 | 1wif |
| 114675 | 51349 | 52832 | 52833 | 52834 | 117591 | 117592 | 1wik |
| 114677 | 56992 | 57849 | 57850 | 57851 | 118296 | 118297 | 1wim |
| 114699 | 48724 | 50198 | 50249 | 50263 | 117193 | 117194 | 1wjj |
| 114700 | 51349 | 52832 | 52833 | 52834 | 117593 | 117594 | 1wjk |
| 114703 | 53931 | 54235 | 54236 | 54237 | 117812 | 117813 | 1wjn |
| 114704 | 46456 | 47575 | 47576 | 47577 | 47580 | 47581 | 1wjo |
| 114705 | 56992 | 57666 | 57667 | 57668 | 118267 | 118268 | 1wjp |
| 114706 | 56992 | 57666 | 57667 | 57668 | 118267 | 118268 | 1wjp |
| 114707 | 56992 | 57666 | 57667 | 57668 | 118267 | 118268 | 1wjp |
| 114712 | 53931 | 54235 | 54236 | 54237 | 117814 | 117815 | 1wju |
| 114719 | 53931 | 56435 | 56436 | 56437 | 118166 | 118167 | 1wk1 |
| 114724 | 51349 | 100949 | 100950 | 110520 | 117519 | 117520 | 1wkc |
| 114734 | 51349 | 53066 | 53098 | 53118 | 117652 | 117653 | 1wlj |
| 114736 | 53931 | 54235 | 54236 | 54237 | 117816 | 117817 | 1wm3 |
| 114751 | 51349 | 53719 | 53720 | 53721 | 117733 | 117734 | 1wnd |
| 114830 | 51349 | 53658 | 53659 | 53660 | 53661 | 117724 | 1wpw |
| 114836 | 51349 | 56783 | 56784 | 82388 | 82389 | 117502 | 1wr8 |
| 114838 | 53931 | 54235 | 54292 | 54293 | 54294 | 54301 | 1wri |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 114915 | 51349 | 56783 | 56784 | 102317 | 117507 | 117508 | 1wvi |
| 114966 | 48724 | 50813 | 50814 | 50815 | 50845 | 50846 | 1x8q |
| 114983 | 46456 | 46457 | 46458 | 46463 | 116752 | 116753 | 1x9f |
| 114984 | 46456 | 46457 | 46458 | 46463 | 116754 | 116755 | 1x9f |
| 114985 | 46456 | 46457 | 46458 | 46463 | 116756 | 116757 | 1x9f |
| 114999 | 51349 | 53066 | 53098 | 53118 | 53123 | 53124 | 1x9m |
| 115041 | 48724 | 48725 | 48726 | 49159 | 117041 | 117042 | 1xau |
| 115077 | 53931 | 56111 | 56112 | 88854 | 118131 | 118132 | 1xbb |
| 115145 | 53931 | 54000 | 54001 | 54050 | 54051 | 54052 | 1xd3 |
| 115160 | 51349 | 51904 | 51905 | 51943 | 51959 | 117441 | 1xdi |
| 115219 | 53931 | 55728 | 55729 | 55730 | 118066 | 118067 | 1xeb |
| 115229 | 48724 | 48725 | 48726 | 48727 | 117038 | 117039 | 1xed |
| 115235 | 51349 | 52539 | 52540 | 52686 | 52693 | 117545 | 1xew |
| 115264 | 51349 | 52767 | 52768 | 52769 | 117573 | 117574 | 1xfk |
| 115277 | 51349 | 51734 | 51735 | 51751 | 117411 | 117412 | 1xg5 |
| 115286 | 51349 | 51734 | 51735 | 51751 | 69407 | 69408 | 1xgk |
| 115291 | 51349 | 53447 | 53448 | 117702 | 117703 | 117704 | 1xhb |
| 115292 | 51349 | 51904 | 51905 | 51943 | 117442 | 117443 | 1xhc |
| 115293 | 51349 | 51904 | 51905 | 51943 | 117442 | 117443 | 1xhc |
| 115304 | 53931 | 110856 | 110857 | 110858 | 117853 | 117854 | 1xhs |
| 115355 | 51349 | 53382 | 53383 | 53384 | 117693 | 117694 | 1xi9 |
| 115367 | 48724 | 48725 | 48726 | 49159 | 69162 | 110051 | 1xiw |
| 115368 | 48724 | 48725 | 48726 | 49159 | 117043 | 117044 | 1xiw |
| 115387 | 51349 | 52539 | 52540 | 52652 | 89673 | 117539 | 1xjc |
| 115396 | 48724 | 50198 | 50249 | 50263 | 101761 | 117190 | 1xjv |
| 115397 | 48724 | 50198 | 50249 | 50263 | 101761 | 117190 | 1xjv |
| 115408 | 48724 | 50813 | 50814 | 50815 | 117265 | 117266 | 1xki |
| 115409 | 53931 | 56111 | 56112 | 88854 | 82795 | 82796 | 1xkk |
| 115561 | 53931 | 55728 | 55729 | 55730 | 118064 | 118065 | 1xmt |
| 115580 | 46456 | 48370 | 48452 | 48453 | 117009 | 117010 | 1xnf |
| 115680 | 46456 | 47472 | 47473 | 47502 | 47541 | 47542 | 1xo5 |
| 115768 | 53931 | 55247 | 55257 | 64311 | 118020 | 118021 | 1xpp |
| 115879 | 51349 | 52798 | 52799 | 52800 | 117579 | 117580 | 1xri |
| 115892 | 48724 | 51268 | 51283 | 51284 | 117335 | 117336 | 1xs1 |
| 116017 | 53931 | 64267 | 64268 | 64269 | 117932 | 117933 | 1xte |
| 116031 | 51349 | 53334 | 53335 | 117685 | 117686 | 117687 | 1xtp |
| 116052 | 51349 | 51734 | 51735 | 51751 | 117423 | 117424 | 1xu9 |
| 116088 | 51349 | 56783 | 56784 | 82388 | 117505 | 117506 | 1xvi |
| 116091 | 46456 | 48507 | 48508 | 48509 | 117014 | 117016 | 1xvp |
| 116094 | 51349 | 52832 | 52833 | 52901 | 102452 | 117600 | 1xvq |
| 116133 | 53931 | 56111 | 56112 | 88854 | 118133 | 118134 | 1xws |
| 116143 | 46456 | 109603 | 109604 | 101340 | 116971 | 116972 | 1xx7 |
| 116203 | 51349 | 53334 | 53335 | 110671 | 117683 | 117684 | 1xxl |
| 116322 | 46456 | 47768 | 47819 | 69044 | 69045 | 116939 | 1y14 |
| 116323 | 48724 | 50198 | 50249 | 50282 | 88670 | 117195 | 1y14 |
| 116341 | 51349 | 52787 | 52788 | 52789 | 69504 | 117577 | 1y1l |
| 116374 | 46456 | 47472 | 47473 | 63550 | 116903 | 116904 | 1y1x |
| 116394 | 53931 | 54196 | 54197 | 54198 | 117789 | 117790 | 1y23 |
| 116400 | 56992 | 57015 | 57059 | 57072 | 118235 | 118236 | 1y29 |
| 116494 | 51349 | 52539 | 52540 | 52541 | 117528 | 117529 | 1y63 |
| 116563 | 51349 | 53334 | 53335 | 117688 | 117689 | 117690 | 1y8c |
| 116599 | 51349 | 51734 | 51735 | 51751 | 117426 | 117427 | 1yb1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 116609 | 51349 | 53162 | 53167 | 53168 | 110645 | 117658 | 1ybf |
| 116646 | 53931 | 55153 | 55154 | 118007 | 118008 | 118009 | 1yem |
| 118730 | 46456 | 47472 | 47473 | 140538 | 140539 | 140540 | 1pul |
| 118773 | 51349 | 51350 | 51445 | 51487 | 51502 | 141774 | 1rh9 |
| 118846 | 51349 | 56783 | 56784 | 82388 | 142163 | 142164 | 1s2o |
| 118878 | 53931 | 55728 | 55729 | 55730 | 143678 | 143679 | 1s7k |
| 118890 | 53931 | 54636 | 54637 | 82636 | 102907 | 143168 | 1s9c |
| 118930 | 51349 | 51734 | 51735 | 51751 | 51782 | 51783 | 1sby |
| 118957 | 51349 | 52095 | 52096 | 52103 | 52108 | 141999 | 1sg4 |
| 119053 | 51349 | 53334 | 53335 | 53336 | 142573 | 142574 | 1sus |
| 119136 | 48724 | 50493 | 50494 | 50514 | 50548 | 50549 | 1t32 |
| 119141 | 53931 | 55752 | 55753 | 55762 | 111105 | 111107 | 1t3y |
| 119171 | 53931 | 56435 | 56436 | 56437 | 143957 | 143958 | 1t8d |
| 119205 | 51349 | 51350 | 51395 | 51396 | 141761 | 141762 | 1tb3 |
| 119264 | 48724 | 50181 | 50182 | 50183 | 63758 | 141290 | 1th7 |
| 119305 | 48724 | 50155 | 50156 | 50157 | 50162 | 50163 | 1tp5 |
| 119340 | 53931 | 54235 | 54236 | 54237 | 142954 | 142955 | 1ttn |
| 119347 | 51349 | 51350 | 51395 | 51396 | 51397 | 141759 | 1tv5 |
| 119360 | 51349 | 51350 | 51445 | 51487 | 51499 | 141773 | 1tvn |
| 119463 | 53931 | 55810 | 55811 | 55812 | 143764 | 143765 | 1u20 |
| 119536 | 48724 | 50728 | 50729 | 50730 | 141403 | 141404 | 1u5d |
| 119540 | 48724 | 50728 | 50729 | 50730 | 141417 | 141418 | 1u5e |
| 119572 | 53931 | 54861 | 54928 | 54929 | 143322 | 143323 | 1u6f |
| 119811 | 51349 | 51734 | 51735 | 51751 | 51788 | 141875 | 1uzm |
| 119838 | 51349 | 51904 | 51905 | 51943 | 51959 | 63950 | 1v59 |
| 119839 | 51349 | 51904 | 51905 | 51943 | 51959 | 63950 | 1v59 |
| 119859 | 51349 | 53685 | 53686 | 53687 | 142749 | 142750 | 1v71 |
| 119860 | 51349 | 53382 | 53383 | 53384 | 142659 | 142660 | 1v72 |
| 119877 | 51349 | 88722 | 88723 | 89619 | 142112 | 142113 | 1v96 |
| 119903 | 48724 | 50155 | 50156 | 50157 | 141275 | 141276 | 1va8 |
| 119942 | 51349 | 53685 | 53686 | 53687 | 64172 | 142748 | 1vb3 |
| 119970 | 53931 | 55810 | 55811 | 55812 | 143758 | 143759 | 1vcd |
| 119974 | 51349 | 51350 | 51395 | 51396 | 89454 | 141760 | 1vcf |
| 119980 | 51349 | 53270 | 53271 | 53272 | 142558 | 142559 | 1vch |
| 119988 | 51349 | 51350 | 51569 | 51570 | 69394 | 141830 | 1vcv |
| 119991 | 51349 | 51350 | 51695 | 89508 | 141864 | 141865 | 1vd6 |
| 119994 | 51349 | 53270 | 53271 | 53272 | 142560 | 142561 | 1vdm |
| 120007 | 51349 | 53685 | 53686 | 53687 | 53690 | 142743 | 1ve1 |
| 120008 | 51349 | 53789 | 53790 | 53791 | 117736 | 142780 | 1ve2 |
| 120010 | 51349 | 53334 | 53335 | 117688 | 142599 | 142600 | 1ve3 |
| 120013 | 51349 | 53685 | 53686 | 53687 | 53692 | 142747 | 1ve5 |
| 120018 | 51349 | 53382 | 53383 | 53417 | 142674 | 142675 | 1vef |
| 120029 | 51349 | 51350 | 51445 | 51534 | 63910 | 63911 | 1vf8 |
| 120033 | 51349 | 51350 | 51445 | 51521 | 51528 | 141785 | 1vff |
| 120362 | 48724 | 50198 | 50249 | 50282 | 50299 | 50301 | 1vqo |
| 120373 | 51349 | 52079 | 52080 | 52081 | 52082 | 52083 | 1vqo |
| 120375 | 51349 | 53066 | 53137 | 53138 | 53139 | 53140 | 1vqo |
| 120376 | 51349 | 52079 | 52080 | 52081 | 52084 | 52085 | 1vqo |
| 120378 | 48724 | 50036 | 50104 | 50105 | 50108 | 50109 | 1vqo |
| 120381 | 48724 | 50036 | 50104 | 50105 | 50106 | 50107 | 1vqo |
| 120385 | 53931 | 54574 | 54575 | 54576 | 54577 | 54578 | 1vqo |
| 120417 | 51349 | 51350 | 51366 | 51375 | 51376 | 141746 | 1vqt |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 120427 | 53931 | 55680 | 55681 | 143642 | 143645 | 143646 | 1vqz |
| 120431 | 48724 | 51181 | 51182 | 82191 | 82192 | 141601 | 1vr3 |
| 120548 | 48724 | 51268 | 51283 | 51284 | 51285 | 141654 | 1vyq |
| 120639 | 53931 | 54000 | 54001 | 54047 | 54048 | 142855 | 1w4t |
| 120643 | 53931 | 54000 | 54001 | 54047 | 54048 | 75335 | 1w5r |
| 120652 | 53931 | 55680 | 55681 | 143642 | 143643 | 143644 | 1w66 |
| 120679 | 51349 | 53447 | 53448 | 68901 | 64140 | 142690 | 1w77 |
| 120779 | 48724 | 50155 | 50156 | 50157 | 89311 | 89312 | 1w9e |
| 120801 | 48724 | 51181 | 51197 | 51198 | 141621 | 141622 | 1w9y |
| 120802 | 51349 | 51350 | 51569 | 51570 | 51584 | 110357 | 1wa3 |
| 120809 | 46456 | 47239 | 140453 | 140454 | 140455 | 140456 | 1wa8 |
| 120810 | 46456 | 47239 | 140453 | 140454 | 140457 | 140458 | 1wa8 |
| 120834 | 51349 | 52539 | 52540 | 52686 | 52697 | 52699 | 1wb9 |
| 120845 | 51349 | 51350 | 51569 | 51570 | 51584 | 51585 | 1wbh |
| 120888 | 51349 | 51350 | 51445 | 51521 | 141786 | 141787 | 1wcg |
| 120965 | 53931 | 54861 | 54928 | 54929 | 143349 | 143350 | 1wel |
| 120967 | 48724 | 50155 | 50156 | 50157 | 141277 | 141278 | 1wf8 |
| 120968 | 53931 | 54235 | 54236 | 54237 | 142934 | 142935 | 1wf9 |
| 120973 | 48724 | 50155 | 50156 | 50157 | 141271 | 141272 | 1wi4 |
| 120989 | 51349 | 53685 | 53686 | 53687 | 53690 | 142742 | 1wkv |
| 121008 | 46456 | 47472 | 47473 | 140538 | 140541 | 140542 | 1wlm |
| 121015 | 53931 | 54636 | 54637 | 89902 | 89903 | 102909 | 1wlu |
| 121042 | 51349 | 51734 | 51735 | 51751 | 63935 | 141878 | 1wma |
| 121132 | 51349 | 52832 | 52833 | 52901 | 102459 | 142371 | 1wp0 |
| 121276 | 53931 | 55680 | 55681 | 55682 | 55688 | 143637 | 1wu7 |
| 121335 | 46456 | 46625 | 46626 | 46627 | 46669 | 46670 | 1wve |
| 121358 | 53931 | 54861 | 54928 | 54929 | 143330 | 143331 | 1wwh |
| 121366 | 51349 | 53926 | 53927 | 89800 | 142840 | 142843 | 1wwr |
| 121380 | 53931 | 55728 | 55729 | 55730 | 143648 | 143649 | 1wwz |
| 121382 | 51349 | 51350 | 51569 | 51570 | 75085 | 141836 | 1wx0 |
| 121393 | 53931 | 54235 | 54236 | 54237 | 142950 | 142951 | 1wx8 |
| 121394 | 53931 | 54235 | 54236 | 54237 | 142952 | 142953 | 1wx9 |
| 121406 | 51349 | 52539 | 52540 | 52592 | 142251 | 142252 | 1wxq |
| 121409 | 53931 | 54235 | 54236 | 54237 | 142932 | 142933 | 1wxv |
| 121471 | 51349 | 53789 | 53790 | 53791 | 142783 | 142784 | 1wyz |
| 121483 | 51349 | 52095 | 52096 | 52103 | 142000 | 142001 | 1wz8 |
| 121491 | 51349 | 56783 | 56784 | 82388 | 117505 | 142156 | 1wzc |
| 121525 | 51349 | 53334 | 53335 | 117688 | 142601 | 142602 | 1wzn |
| 121529 | 53931 | 54494 | 54495 | 54496 | 54497 | 143052 | 1wzv |
| 121579 | 48724 | 50728 | 50729 | 50730 | 141405 | 141406 | 1x1f |
| 121592 | 53931 | 54235 | 54236 | 54237 | 142942 | 142943 | 1x1m |
| 121595 | 51349 | 52539 | 52540 | 52592 | 142283 | 142284 | 1x1r |
| 121597 | 51349 | 51734 | 51735 | 51751 | 141886 | 141887 | 1x1t |
| 121645 | 46456 | 46688 | 46689 | 46690 | 140147 | 140148 | 1x2m |
| 121646 | 46456 | 46688 | 46689 | 46690 | 140153 | 140154 | 1x2n |
| 121660 | 56992 | 57666 | 57667 | 57668 | 144133 | 144134 | 1x3c |
| 121676 | 46456 | 46688 | 46689 | 46690 | 140145 | 140146 | 1x41 |
| 121678 | 48724 | 48725 | 48726 | 49159 | 141013 | 141014 | 1x44 |
| 121679 | 48724 | 50155 | 50156 | 50157 | 141282 | 141283 | 1x45 |
| 121683 | 53931 | 54861 | 54928 | 54929 | 143346 | 143347 | 1x4a |
| 121686 | 53931 | 54861 | 54928 | 54929 | 143306 | 143307 | 1x4d |
| 121687 | 53931 | 54861 | 54928 | 54929 | 143320 | 143321 | 1x4e |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 121688 | 53931 | 54861 | 54928 | 54929 | 143306 | 143307 | 1x4f |
| 121689 | 53931 | 54861 | 54928 | 54929 | 143298 | 143299 | 1x4g |
| 121690 | 53931 | 54861 | 54928 | 54929 | 143358 | 143359 | 1x4h |
| 121705 | 46456 | 46688 | 46689 | 46690 | 140157 | 140158 | 1x58 |
| 121714 | 48724 | 50155 | 50156 | 50157 | 110182 | 141263 | 1x5n |
| 121715 | 53931 | 54861 | 54928 | 54929 | 143356 | 143357 | 1x5o |
| 121721 | 53931 | 54861 | 54928 | 54929 | 143318 | 143319 | 1x5u |
| 121722 | 56992 | 57666 | 57667 | 57668 | 144149 | 144150 | 1x5w |
| 121723 | 56992 | 57666 | 57667 | 57668 | 144149 | 144150 | 1x5w |
| 121742 | 56992 | 57666 | 57667 | 57668 | 144137 | 144138 | 1x6f |
| 121743 | 56992 | 57666 | 57667 | 57668 | 144153 | 144154 | 1x6h |
| 121768 | 48724 | 50813 | 50814 | 50815 | 50835 | 50836 | 1x71 |
| 121840 | 53931 | 55297 | 55315 | 55316 | 55319 | 111033 | 1xbi |
| 121850 | 46456 | 46688 | 46689 | 46739 | 140171 | 140172 | 1xc5 |
| 122007 | 48724 | 50474 | 50475 | 50476 | 141356 | 141357 | 1xhn |
| 122022 | 51349 | 52832 | 52833 | 52901 | 52909 | 142367 | 1xiy |
| 122055 | 46456 | 47472 | 47473 | 47478 | 47479 | 47487 | 1xk4 |
| 122057 | 46456 | 47472 | 47473 | 47478 | 47479 | 69020 | 1xk4 |
| 122081 | 53931 | 54000 | 54001 | 54002 | 142848 | 142849 | 1xkg |
| 122094 | 51349 | 51350 | 51569 | 51570 | 51574 | 141832 | 1xky |
| 122119 | 53931 | 54235 | 54292 | 54293 | 54294 | 54307 | 1xlq |
| 122143 | 56835 | 81443 | 81442 | 81441 | 81438 | 81437 | 1xme |
| 122211 | 53931 | 56435 | 56436 | 56437 | 69857 | 69858 | 1xph |
| 122297 | 51349 | 52539 | 52540 | 52592 | 142275 | 142276 | 1xtq |
| 122300 | 51349 | 53270 | 53271 | 53272 | 53293 | 142555 | 1xtt |
| 122387 | 51349 | 52832 | 52833 | 52901 | 142381 | 142382 | 1xvw |
| 122433 | 51349 | 51350 | 51569 | 51570 | 51574 | 141831 | 1xxx |
| 122474 | 51349 | 52832 | 52833 | 52901 | 102459 | 102460 | 1xzo |
| 122476 | 51349 | 53270 | 53271 | 53272 | 142556 | 142557 | 1y0b |
| 122550 | 51349 | 51734 | 51735 | 51751 | 117413 | 117414 | 1y1p |
| 122608 | 53931 | 56280 | 56281 | 143916 | 143917 | 143918 | 1y44 |
| 122674 | 51349 | 53648 | 53649 | 53650 | 53651 | 53652 | 1y6v |
| 122700 | 48724 | 50155 | 50156 | 50157 | 141282 | 141283 | 1y7n |
| 122716 | 53931 | 55728 | 55729 | 55730 | 143674 | 143675 | 1y7r |
| 122769 | 53931 | 54494 | 54495 | 54496 | 54497 | 143058 | 1y8x |
| 122779 | 53931 | 55728 | 55729 | 55730 | 143696 | 143697 | 1y9k |
| 122783 | 53931 | 55728 | 55729 | 55730 | 143650 | 143651 | 1y9w |
| 122785 | 46456 | 48370 | 48452 | 48453 | 140843 | 140844 | 1ya0 |
| 122837 | 53931 | 56234 | 56235 | 56251 | 56255 | 56256 | 1yar |
| 122844 | 53931 | 56234 | 56235 | 56251 | 56252 | 56253 | 1yar |
| 122865 | 53931 | 55845 | 55846 | 55847 | 143780 | 143781 | 1yb0 |
| 122980 | 51349 | 51734 | 51735 | 51751 | 141907 | 141908 | 1yde |
| 123006 | 51349 | 51350 | 51695 | 89508 | 110382 | 110383 | 1ydy |
| 123054 | 53931 | 54494 | 54495 | 54496 | 54497 | 143049 | 1yf9 |
| 123214 | 53931 | 56111 | 56112 | 88854 | 56146 | 56147 | 1yhw |
| 123381 | 51349 | 52539 | 52540 | 52541 | 142213 | 142214 | 1yj5 |
| 123493 | 53931 | 55728 | 55729 | 55730 | 143668 | 143669 | 1yk3 |
| 123519 | 46456 | 140717 | 140718 | 140719 | 140720 | 140721 | 1ykh |
| 123651 | 53931 | 54636 | 54637 | 54638 | 102903 | 102904 | 1yli |
| 123699 | 51349 | 52820 | 52821 | 52822 | 52825 | 52826 | 1ymk |
| 123718 | 46456 | 109603 | 109604 | 101340 | 140763 | 140764 | 1ynb |
| 123769 | 51349 | 51734 | 51735 | 51751 | 141913 | 141914 | 1yo6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 123778 | 53931 | 54636 | 54637 | 89902 | 143182 | 143183 | 1yoc |
| 123785 | 51349 | 69571 | 69572 | 89763 | 89766 | 89767 | 1yov |
| 123786 | 51349 | 69571 | 69572 | 89763 | 89764 | 89765 | 1yov |
| 123799 | 51349 | 53447 | 53448 | 53464 | 142688 | 142689 | 1yp2 |
| 123833 | 53931 | 56435 | 56436 | 56437 | 143955 | 143956 | 1ypq |
| 123876 | 53931 | 54235 | 54236 | 54237 | 142946 | 142947 | 1yqb |
| 123915 | 51349 | 52539 | 52540 | 52652 | 142297 | 142298 | 1yrb |
| 123919 | 53931 | 55728 | 55729 | 55730 | 143686 | 143687 | 1yre |
| 123941 | 53931 | 54494 | 54495 | 54496 | 54497 | 143053 | 1yrv |
| 123994 | 51349 | 53066 | 53098 | 53118 | 142495 | 142496 | 1yt3 |
| 124066 | 46456 | 47472 | 47473 | 47478 | 47479 | 140535 | 1yut |
| 124077 | 46456 | 47239 | 47240 | 47241 | 140440 | 140441 | 1yuz |
| 124099 | 46456 | 47161 | 47226 | 47230 | 140413 | 140414 | 1yvi |
| 124106 | 53931 | 55728 | 55729 | 55730 | 143694 | 143695 | 1yvo |
| 124131 | 51349 | 53162 | 53187 | 142526 | 142527 | 142531 | 1yw4 |
| 124132 | 51349 | 53162 | 53187 | 142526 | 142527 | 142530 | 1yw6 |
| 124168 | 53931 | 55728 | 55729 | 55730 | 143654 | 143655 | 1yx0 |
| 124192 | 51349 | 51734 | 51735 | 51751 | 141884 | 141885 | 1yxm |
| 124264 | 48724 | 51315 | 51316 | 63873 | 63874 | 117337 | 1yz1 |
| 124275 | 51349 | 52171 | 52266 | 89594 | 142056 | 142057 | 1yzf |
| 124277 | 51349 | 53334 | 53335 | 142647 | 142648 | 142649 | 1yzh |
| 124301 | 51349 | 52539 | 52540 | 52592 | 142279 | 142280 | 1z06 |
| 124328 | 56572 | 111330 | 111331 | 111332 | 111333 | 111334 | 1z0s |
| 124369 | 48724 | 50813 | 50814 | 50815 | 141462 | 141463 | 1z24 |
| 124374 | 51349 | 52539 | 52540 | 52592 | 142231 | 142232 | 1z2a |
| 124387 | 53931 | 54235 | 54236 | 54237 | 142936 | 142937 | 1z2m |
| 124388 | 53931 | 54235 | 54236 | 54237 | 142936 | 142937 | 1z2m |
| 124389 | 53931 | 54494 | 54495 | 54496 | 54497 | 143060 | 1z2u |
| 124421 | 51349 | 51350 | 51395 | 51396 | 141763 | 141764 | 1z41 |
| 124431 | 53931 | 55728 | 55729 | 55730 | 143658 | 143659 | 1z4e |
| 124444 | 53931 | 55728 | 55729 | 55730 | 55737 | 143647 | 1z4r |
| 124450 | 53931 | 54636 | 54637 | 54638 | 143152 | 143153 | 1z54 |
| 124498 | 51349 | 52832 | 52833 | 52901 | 75237 | 142370 | 1z5y |
| 124530 | 51349 | 52832 | 52833 | 52834 | 142353 | 142354 | 1z6n |
| 124531 | 46456 | 47239 | 47240 | 47241 | 47246 | 140436 | 1z6o |
| 124543 | 46456 | 47239 | 47240 | 47241 | 47246 | 140437 | 1z6o |
| 124624 | 51349 | 53270 | 53271 | 53272 | 53283 | 53284 | 1z7g |
| 124631 | 56992 | 100894 | 100895 | 57468 | 57469 | 57470 | 1z7k |
| 124632 | 53931 | 55680 | 55681 | 55682 | 118058 | 143639 | 1z7m |
| 124722 | 46456 | 48263 | 48264 | 48265 | 48272 | 48273 | 1z8o |
| 124774 | 46456 | 47239 | 47240 | 47253 | 140442 | 140443 | 1za0 |
| 124850 | 51349 | 53066 | 53098 | 53099 | 142490 | 142491 | 1zbf |
| 124881 | 48724 | 50728 | 50729 | 50730 | 141399 | 141400 | 1zc3 |
| 124903 | 51349 | 51350 | 51695 | 89508 | 141866 | 141867 | 1zcc |
| 124931 | 51349 | 53473 | 53474 | 53525 | 53526 | 102626 | 1zd3 |
| 124938 | 51349 | 52539 | 52540 | 52592 | 52614 | 142219 | 1zd9 |
| 124943 | 53931 | 54494 | 54495 | 54496 | 54497 | 143048 | 1zdn |
| 124980 | 51349 | 53648 | 53649 | 53650 | 53651 | 64165 | 1zed |
| 124981 | 46456 | 140958 | 140959 | 140963 | 140964 | 140965 | 1zee |
| 124984 | 51349 | 51734 | 51735 | 51751 | 141882 | 141883 | 1zem |
| 124994 | 51349 | 52171 | 52172 | 52173 | 52192 | 52193 | 1zes |
| 125065 | 51349 | 52171 | 52172 | 52173 | 142033 | 142034 | 1zgz |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 125074 | 53931 | 54861 | 54928 | 54929 | 89940 | 89941 | 1zh5 |
| 125112 | 53931 | 143999 | 144000 | 144001 | 144002 | 144003 | 1zhx |
| 125176 | 51349 | 51734 | 51735 | 51751 | 89521 | 89522 | 1zk4 |
| 125182 | 48724 | 50890 | 50891 | 50892 | 141505 | 141506 | 1zkc |
| 125198 | 53931 | 54235 | 54236 | 54237 | 117799 | 142931 | 1zkh |
| 125199 | 53931 | 54636 | 54637 | 89902 | 143175 | 143176 | 1zki |
| 125354 | 51349 | 52832 | 52833 | 52834 | 142355 | 142356 | 1zma |
| 125363 | 51349 | 51734 | 51735 | 51751 | 102153 | 102154 | 1zmt |
| 125380 | 48724 | 50813 | 50814 | 50815 | 50832 | 50833 | 1znd |
| 125416 | 53931 | 55728 | 55729 | 143714 | 143715 | 143716 | 1zo0 |
| 125437 | 51349 | 53382 | 53383 | 53417 | 53418 | 53419 | 1zod |
| 125555 | 48724 | 51181 | 51182 | 82191 | 82192 | 82193 | 1zrr |
| 125573 | 46456 | 47239 | 47240 | 47241 | 47250 | 140433 | 1zs3 |
| 125600 | 51349 | 56783 | 56784 | 142186 | 142189 | 142190 | 1zs9 |
| 125662 | 46456 | 48370 | 48452 | 48453 | 140839 | 140840 | 1zu2 |
| 125679 | 51349 | 52539 | 52540 | 52592 | 142269 | 142270 | 1zun |
| 125680 | 53931 | 54494 | 54495 | 54496 | 143068 | 143069 | 1zuo |
| 125757 | 51349 | 53334 | 53335 | 69550 | 69551 | 142580 | 1zx0 |
| 125812 | 56835 | 56836 | 56854 | 56855 | 90103 | 90104 | 1zy3 |
| 125938 | 46456 | 46688 | 46785 | 46832 | 140265 | 140266 | 2a07 |
| 125944 | 48724 | 50813 | 50814 | 50847 | 141467 | 141468 | 2a0a |
| 125960 | 51349 | 52767 | 52768 | 52769 | 52770 | 142347 | 2a0m |
| 125997 | 56572 | 56751 | 56752 | 56753 | 56757 | 64508 | 2a1h |
| 126033 | 48724 | 50474 | 50475 | 50476 | 50479 | 141347 | 2a2j |
| 126072 | 51349 | 51734 | 51735 | 51751 | 141911 | 141912 | 2a35 |
| 126091 | 51349 | 51350 | 51556 | 51557 | 141800 | 141801 | 2a3l |
| 126139 | 53931 | 56218 | 56219 | 56220 | 56225 | 56226 | 2a40 |
| 126150 | 51349 | 51350 | 51569 | 51570 | 51576 | 141833 | 2a4a |
| 126163 | 51349 | 52832 | 52833 | 52901 | 142387 | 142388 | 2a4v |
| 126293 | 53931 | 143119 | 143120 | 143121 | 143122 | 143123 | 2a6q |
| 126341 | 51349 | 52095 | 52096 | 52103 | 142004 | 142005 | 2a7k |
| 126389 | 51349 | 52373 | 52374 | 63976 | 63977 | 89615 | 2a84 |
| 126408 | 51349 | 53926 | 53927 | 89800 | 142835 | 142836 | 2a8n |
| 126528 | 51349 | 53612 | 53613 | 53614 | 142712 | 142713 | 2abq |
| 126534 | 51349 | 53612 | 53613 | 53614 | 53617 | 53619 | 2abs |
| 126545 | 53931 | 55153 | 55154 | 118007 | 143487 | 143488 | 2aca |
| 126581 | 53931 | 54861 | 54928 | 54929 | 54950 | 54951 | 2adb |
| 126582 | 53931 | 54861 | 54928 | 54929 | 54950 | 54951 | 2adc |
| 126583 | 53931 | 54861 | 54928 | 54929 | 54950 | 54951 | 2adc |
| 126599 | 48724 | 50728 | 50729 | 50730 | 141411 | 141412 | 2adz |
| 126600 | 53931 | 55728 | 55729 | 55730 | 143664 | 143665 | 2ae6 |
| 126617 | 51349 | 52767 | 52768 | 52769 | 52770 | 142346 | 2aeb |
| 126620 | 51349 | 53270 | 53271 | 53272 | 53290 | 142553 | 2aee |
| 126674 | 51349 | 53612 | 53613 | 53614 | 75295 | 75296 | 2afb |
| 126683 | 51349 | 52539 | 52540 | 52652 | 52661 | 52662 | 2afh |
| 126707 | 51349 | 53162 | 53187 | 142532 | 142533 | 142534 | 2afw |
| 126719 | 51349 | 51734 | 51735 | 51751 | 141894 | 141895 | 2ag5 |
| 126740 | 48724 | 50938 | 50939 | 50940 | 82164 | 89371 | 2ah2 |
| 126742 | 51349 | 56783 | 56784 | 75173 | 142145 | 142146 | 2ah5 |
| 126771 | 48724 | 50198 | 50249 | 50282 | 74950 | 141310 | 2aho |
| 126841 | 53931 | 55728 | 55729 | 55730 | 143690 | 143691 | 2aj6 |
| 126892 | 51349 | 53612 | 53613 | 53614 | 82518 | 82519 | 2ajr |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 126914 | 51349 | 52539 | 52540 | 52592 | 69488 | 142228 | 2aka |
| 126976 | 53931 | 54636 | 54637 | 54638 | 143162 | 143163 | 2ali |
| 127024 | 51349 | 51350 | 51556 | 51557 | 51558 | 141799 | 2amx |
| 127135 | 56992 | 57769 | 144210 | 144211 | 144214 | 144216 | 2apo |
| 127168 | 56992 | 57769 | 144210 | 144211 | 144212 | 144213 | 2aqa |
| 127225 | 48724 | 50474 | 50475 | 50476 | 141350 | 141351 | 2arz |
| 127244 | 48724 | 50198 | 50249 | 50282 | 69265 | 69267 | 2asb |
| 127253 | 48724 | 50474 | 50475 | 50476 | 141362 | 141363 | 2asf |
| 127274 | 51349 | 52046 | 52047 | 52055 | 52056 | 52057 | 2ast |
| 127305 | 53931 | 55728 | 55729 | 55730 | 143680 | 143681 | 2atr |
| 127377 | 51349 | 53334 | 53335 | 53336 | 142571 | 142572 | 2avd |
| 127380 | 51349 | 53334 | 53335 | 110671 | 142595 | 142596 | 2avn |
| 127432 | 53931 | 54494 | 54495 | 54496 | 54497 | 143056 | 2awf |
| 127497 | 51349 | 52539 | 52540 | 52541 | 142211 | 142212 | 2axp |
| 127571 | 53931 | 143112 | 143113 | 143114 | 143115 | 143116 | 2ayu |
| 127574 | 51349 | 52171 | 52172 | 52173 | 142023 | 142024 | 2ayx |
| 127615 | 53931 | 55810 | 55811 | 55812 | 143756 | 143757 | 2azw |
| 127624 | 53931 | 55810 | 55811 | 55812 | 143762 | 143763 | 2b06 |
| 127632 | 51349 | 52008 | 102198 | 102199 | 141972 | 141973 | 2b0a |
| 127650 | 53931 | 55810 | 55811 | 55812 | 143760 | 143761 | 2b0v |
| 127759 | 51349 | 56783 | 56784 | 82388 | 142165 | 142166 | 2b30 |
| 127786 | 48724 | 50198 | 50249 | 50263 | 50267 | 50268 | 2b3g |
| 127792 | 53931 | 54636 | 54637 | 82636 | 143171 | 143172 | 2b3n |
| 127796 | 51349 | 53334 | 53335 | 89743 | 89744 | 110666 | 2b3t |
| 127804 | 51349 | 53731 | 53732 | 53733 | 142758 | 142759 | 2b3y |
| 127817 | 51349 | 52171 | 52172 | 52173 | 142029 | 142030 | 2b4a |
| 127892 | 53931 | 55728 | 55729 | 55730 | 143652 | 143653 | 2b5g |
| 127905 | 56572 | 56573 | 56574 | 56575 | 56584 | 56586 | 2b5t |
| 127914 | 51349 | 52832 | 52833 | 52901 | 142373 | 142374 | 2b5x |
| 127974 | 51349 | 51734 | 51735 | 51751 | 141880 | 141881 | 2b69 |
| 128042 | 51349 | 52832 | 52833 | 52901 | 102459 | 142372 | 2b7k |
| 128093 | 51349 | 82543 | 82544 | 82545 | 82546 | 82547 | 2b8n |
| 128103 | 53931 | 54861 | 54919 | 54920 | 54921 | 143286 | 2b8q |
| 128117 | 46456 | 47265 | 47266 | 47286 | 47291 | 47292 | 2b8u |
| 128257 | 51349 | 53789 | 53790 | 53791 | 142785 | 142786 | 2bb3 |
| 128302 | 51349 | 53162 | 53187 | 142526 | 142527 | 142529 | 2bco |
| 128321 | 51349 | 51734 | 51735 | 51751 | 141892 | 141893 | 2bd0 |
| 128456 | 51349 | 53755 | 53756 | 110734 | 110735 | 142768 | 2bfw |
| 128480 | 51349 | 51734 | 51735 | 51751 | 141898 | 141899 | 2bgk |
| 128505 | 53931 | 55845 | 55846 | 55847 | 143782 | 143783 | 2bgx |
| 128565 | 53931 | 54636 | 54637 | 82636 | 143169 | 143170 | 2bi0 |
| 128566 | 53931 | 54636 | 54637 | 82636 | 143169 | 143170 | 2bi0 |
| 128633 | 48724 | 50629 | 50630 | 50646 | 50673 | 50674 | 2bju |
| 128666 | 51349 | 51734 | 51735 | 51751 | 141904 | 141905 | 2bka |
| 128714 | 51349 | 53382 | 53383 | 53402 | 89757 | 142664 | 2bkw |
| 128733 | 51349 | 51734 | 51735 | 51751 | 117419 | 117420 | 2bll |
| 128797 | 51349 | 52539 | 52540 | 52592 | 142249 | 142250 | 2bmj |
| 128817 | 51349 | 52832 | 52833 | 52901 | 69516 | 142368 | 2bmx |
| 128869 | 51349 | 53447 | 53448 | 142691 | 142692 | 142693 | 2bo4 |
| 128951 | 46456 | 48370 | 48371 | 48372 | 48378 | 140814 | 2bpt |
| 129090 | 56992 | 144121 | 144122 | 144123 | 144124 | 144125 | 2bsk |
| 129091 | 56992 | 144121 | 144122 | 144123 | 144126 | 144127 | 2bsk |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 129129 | 48724 | 51268 | 51283 | 51284 | 141655 | 141656 | 2bsy |
| 129130 | 48724 | 51268 | 51283 | 51284 | 141655 | 141656 | 2bsy |
| 129152 | 53931 | 54235 | 54292 | 54293 | 54310 | 54311 | 2bt6 |
| 129227 | 48724 | 49451 | 49482 | 49483 | 49486 | 49488 | 2bur |
| 129228 | 48724 | 49451 | 49482 | 49483 | 49489 | 49491 | 2bur |
| 129238 | 51349 | 52539 | 52540 | 52592 | 52633 | 52634 | 2bv3 |
| 129324 | 48724 | 49898 | 49899 | 49978 | 49991 | 89275 | 2bw8 |
| 129363 | 53931 | 54235 | 54236 | 54237 | 142940 | 142941 | 2bwf |
| 129368 | 51349 | 53382 | 53383 | 53417 | 142676 | 142677 | 2bwn |
| 129486 | 46456 | 47112 | 47113 | 47134 | 140400 | 140401 | 2byk |
| 129487 | 46456 | 47112 | 47113 | 47134 | 140398 | 140399 | 2byk |
| 129490 | 51349 | 53382 | 53383 | 53417 | 53422 | 53423 | 2byl |
| 129533 | 53931 | 54861 | 54928 | 54929 | 143302 | 143303 | 2bz2 |
| 129552 | 48724 | 50493 | 50494 | 50514 | 50550 | 50551 | 2bz6 |
| 129559 | 51349 | 53334 | 53335 | 102566 | 102567 | 142579 | 2bzg |
| 129577 | 51349 | 51734 | 51735 | 51751 | 51788 | 141877 | 2c07 |
| 129594 | 51349 | 51350 | 51445 | 51487 | 141781 | 141782 | 2c0h |
| 129644 | 53931 | 142876 | 142877 | 142878 | 142879 | 142880 | 2c1w |
| 129670 | 53931 | 54636 | 54637 | 82636 | 143173 | 143174 | 2c2i |
| 129672 | 46456 | 48370 | 48452 | 48453 | 140837 | 140838 | 2c2l |
| 129711 | 46456 | 47768 | 47819 | 69044 | 69045 | 140642 | 2c35 |
| 129837 | 51349 | 56783 | 56784 | 102317 | 142169 | 142170 | 2c4n |
| 129894 | 51349 | 51734 | 51735 | 51751 | 141890 | 141891 | 2c5a |
| 130039 | 51349 | 52539 | 52540 | 52592 | 52626 | 52629 | 2c78 |
| 130080 | 51349 | 53334 | 53335 | 88786 | 53367 | 53368 | 2c7p |
| 130123 | 46456 | 46625 | 46626 | 46627 | 109645 | 109646 | 2c8s |
| 130129 | 48724 | 48725 | 48726 | 49159 | 141015 | 141016 | 2c9a |
| 130143 | 48724 | 48725 | 49329 | 49330 | 49331 | 49333 | 2c9v |
| 130145 | 53931 | 54235 | 54236 | 54237 | 54246 | 54247 | 2c9w |
| 130540 | 46456 | 46688 | 46689 | 46739 | 140167 | 140168 | 2cjj |
| 130570 | 51349 | 53334 | 53335 | 53336 | 53337 | 53338 | 2cl5 |
| 130672 | 48724 | 50728 | 50729 | 50730 | 141419 | 141420 | 2coa |
| 130674 | 48724 | 50728 | 50729 | 50730 | 141415 | 141416 | 2coc |
| 130675 | 48724 | 50728 | 50729 | 50730 | 141395 | 141396 | 2cod |
| 130676 | 48724 | 50728 | 50729 | 50730 | 141407 | 141408 | 2cof |
| 130697 | 53931 | 54861 | 54928 | 54929 | 143326 | 143327 | 2cpd |
| 130698 | 53931 | 54861 | 54928 | 54929 | 143340 | 143341 | 2cpe |
| 130699 | 53931 | 54861 | 54928 | 54929 | 117969 | 117970 | 2cpf |
| 130700 | 53931 | 54861 | 54928 | 54929 | 117969 | 117970 | 2cph |
| 130701 | 53931 | 54861 | 54928 | 54929 | 143294 | 143295 | 2cpi |
| 130702 | 53931 | 54861 | 54928 | 54929 | 143354 | 143355 | 2cpj |
| 130708 | 53931 | 54861 | 54928 | 54929 | 143296 | 143297 | 2cpx |
| 130709 | 53931 | 54861 | 54928 | 54929 | 143349 | 143350 | 2cpy |
| 130710 | 53931 | 54861 | 54928 | 54929 | 143336 | 143337 | 2cpz |
| 130711 | 53931 | 54861 | 54928 | 54929 | 143304 | 143305 | 2cq0 |
| 130712 | 53931 | 54861 | 54928 | 54929 | 143332 | 143333 | 2cq1 |
| 130713 | 53931 | 54861 | 54928 | 54929 | 143334 | 143335 | 2cq2 |
| 130714 | 53931 | 54861 | 54928 | 54929 | 143314 | 143315 | 2cq3 |
| 130715 | 53931 | 54861 | 54928 | 54929 | 143352 | 143353 | 2cq4 |
| 130717 | 53931 | 54861 | 54928 | 54929 | 143310 | 143311 | 2cqb |
| 130718 | 53931 | 54861 | 54928 | 54929 | 143312 | 143313 | 2cqc |
| 130719 | 53931 | 54861 | 54928 | 54929 | 143360 | 143361 | 2cqd |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 130722 | 53931 | 54861 | 54928 | 54929 | 117957 | 117958 | 2cqg |
| 130723 | 53931 | 54861 | 54928 | 54929 | 143344 | 143345 | 2cqh |
| 130724 | 53931 | 54861 | 54928 | 54929 | 143298 | 143299 | 2cqi |
| 130727 | 53931 | 54861 | 54928 | 54929 | 143349 | 143351 | 2cqp |
| 130728 | 46456 | 46688 | 46689 | 46739 | 140169 | 140170 | 2cqq |
| 130729 | 46456 | 46688 | 46689 | 46739 | 140169 | 140170 | 2cqr |
| 130730 | 48724 | 48725 | 48726 | 49159 | 49170 | 141012 | 2cqv |
| 130739 | 46456 | 46688 | 46689 | 46739 | 140163 | 140164 | 2crg |
| 130747 | 48724 | 50155 | 50156 | 50157 | 141286 | 141287 | 2cs5 |
| 130769 | 56992 | 57666 | 57667 | 57668 | 144145 | 144146 | 2csh |
| 130770 | 56992 | 57666 | 57667 | 57668 | 144145 | 144146 | 2csh |
| 130771 | 48724 | 50155 | 50156 | 50157 | 141265 | 141266 | 2csj |
| 130780 | 48724 | 50155 | 50156 | 50157 | 141279 | 141281 | 2css |
| 130788 | 56992 | 57666 | 57667 | 57668 | 144153 | 144154 | 2ct1 |
| 130789 | 56992 | 57666 | 57667 | 57668 | 144153 | 144154 | 2ct1 |
| 130792 | 56992 | 57666 | 57667 | 57668 | 144135 | 144136 | 2ctd |
| 130793 | 56992 | 57666 | 57667 | 57668 | 144135 | 144136 | 2ctd |
| 130800 | 51349 | 53382 | 53383 | 53402 | 142665 | 142666 | 2ctz |
| 130805 | 46456 | 46688 | 46689 | 46739 | 140173 | 140174 | 2cu7 |
| 130809 | 46456 | 46688 | 46689 | 46690 | 140159 | 140160 | 2cuf |
| 130859 | 51349 | 52832 | 52833 | 52901 | 142389 | 142390 | 2cvb |
| 130973 | 51349 | 52832 | 52833 | 52901 | 142377 | 142378 | 2cx4 |
| 131013 | 51349 | 52539 | 52540 | 52592 | 142257 | 142258 | 2cxx |
| 131016 | 53931 | 55728 | 55729 | 55730 | 118068 | 118069 | 2cy2 |
| 131021 | 53931 | 54636 | 54637 | 54638 | 143150 | 143151 | 2cye |
| 131025 | 51349 | 51350 | 51445 | 51487 | 51507 | 141777 | 2cyg |
| 131035 | 53931 | 54861 | 54913 | 54914 | 143281 | 143282 | 2cz4 |
| 131093 | 46456 | 140958 | 140959 | 140963 | 140966 | 140967 | 2d0t |
| 131151 | 51349 | 51734 | 51735 | 51751 | 141896 | 141897 | 2d1y |
| 131269 | 51349 | 52373 | 52374 | 52375 | 52384 | 52385 | 2d5b |
| 131357 | 48724 | 48725 | 48726 | 49159 | 141013 | 141014 | 2dav |
| 131374 | 51349 | 53612 | 53613 | 53614 | 142714 | 142715 | 2dcn |
| 131430 | 51349 | 53789 | 53790 | 53791 | 102684 | 142779 | 2dek |
| 131470 | 48724 | 49898 | 49899 | 49978 | 49979 | 49985 | 2dfb |
| 131515 | 53931 | 54861 | 54928 | 54929 | 143324 | 143325 | 2dgx |
| 131532 | 53931 | 54861 | 54928 | 54929 | 143328 | 143329 | 2dis |
| 131533 | 53931 | 54861 | 54928 | 54929 | 143290 | 143291 | 2dit |
| 131558 | 56992 | 57666 | 57667 | 57668 | 144139 | 144140 | 2dlk |
| 131559 | 56992 | 57666 | 57667 | 57668 | 144139 | 144140 | 2dlk |
| 131563 | 56992 | 57666 | 57667 | 57668 | 144147 | 144148 | 2dlq |
| 131565 | 56992 | 57666 | 57667 | 57668 | 144147 | 144148 | 2dlq |
| 131573 | 56992 | 57666 | 57667 | 57668 | 144149 | 144150 | 2dmd |
| 131793 | 51349 | 51350 | 51556 | 141819 | 141822 | 141823 | 2dvt |
| 131980 | 48724 | 50198 | 50242 | 50243 | 50246 | 50248 | 2e2d |
| 132022 | 46456 | 48112 | 48113 | 48114 | 88935 | 48118 | 2.00E+39 |
| 132104 | 46456 | 46688 | 46689 | 46690 | 116778 | 116779 | 2ecc |
| 132304 | 51349 | 52539 | 52540 | 52592 | 142267 | 142268 | 2erx |
| 132351 | 48724 | 51181 | 51182 | 51187 | 63853 | 63854 | 2et1 |
| 132360 | 51349 | 53066 | 53098 | 53099 | 53103 | 142489 | 2etj |
| 132387 | 53931 | 55728 | 55729 | 55730 | 143692 | 143693 | 2eui |
| 132407 | 46456 | 48112 | 48113 | 48114 | 48119 | 48120 | 2eut |
| 132447 | 51349 | 51734 | 51735 | 51751 | 141902 | 141903 | 2ew8 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 132516 | 51349 | 53334 | 53335 | 117685 | 142597 | 142598 | 2ex4 |
| 132649 | 51349 | 53612 | 53613 | 53614 | 142710 | 142711 | 2f02 |
| 132661 | 48724 | 50155 | 50156 | 50157 | 89313 | 89314 | 2f0a |
| 132673 | 53931 | 54636 | 54637 | 89902 | 143179 | 143181 | 2f0x |
| 132782 | 53931 | 56218 | 56219 | 56220 | 111213 | 143901 | 2f1n |
| 132843 | 53931 | 55845 | 55846 | 55847 | 143784 | 143785 | 2f2l |
| 132844 | 53931 | 55845 | 55846 | 55847 | 143784 | 143785 | 2f2l |
| 132904 | 53931 | 54636 | 54637 | 89902 | 143184 | 143185 | 2f41 |
| 132915 | 51349 | 53783 | 53784 | 53785 | 75321 | 75322 | 2f48 |
| 132958 | 53931 | 54494 | 54495 | 54496 | 54497 | 143050 | 2f4w |
| 133015 | 48724 | 50155 | 50156 | 50157 | 117180 | 141264 | 2f5y |
| 133021 | 51349 | 52171 | 52309 | 52310 | 52311 | 142064 | 2f62 |
| 133047 | 51349 | 51350 | 51556 | 141819 | 141828 | 141829 | 2f6k |
| 133055 | 51349 | 52095 | 52096 | 52103 | 142006 | 142007 | 2f6q |
| 133061 | 51349 | 51350 | 51395 | 51396 | 141765 | 141766 | 2f6u |
| 133119 | 51349 | 52832 | 52833 | 52901 | 52902 | 142365 | 2f8a |
| 133128 | 51349 | 53382 | 53383 | 53384 | 64121 | 102594 | 2f8j |
| 133153 | 51349 | 53066 | 53098 | 53118 | 142497 | 142498 | 2f96 |
| 133159 | 53931 | 54861 | 54928 | 54929 | 143316 | 143317 | 2f9d |
| 133161 | 51349 | 53755 | 53756 | 110734 | 142769 | 142770 | 2f9f |
| 133223 | 53931 | 54235 | 54236 | 54237 | 142948 | 142949 | 2faz |
| 133249 | 53931 | 55153 | 55154 | 118007 | 143489 | 143490 | 2fbl |
| 133254 | 46456 | 48370 | 48452 | 48453 | 140841 | 140842 | 2fbn |
| 133272 | 53931 | 55728 | 55729 | 55730 | 143684 | 143685 | 2fck |
| 133299 | 48724 | 48725 | 48726 | 49159 | 49179 | 49180 | 2fdb |
| 133311 | 48724 | 51181 | 51197 | 141628 | 141629 | 141630 | 2fdi |
| 133319 | 51349 | 56783 | 56784 | 75173 | 142141 | 142142 | 2fdr |
| 133320 | 51349 | 51350 | 51366 | 51375 | 141749 | 141752 | 2fds |
| 133322 | 51349 | 88722 | 88723 | 89619 | 142114 | 142115 | 2fe1 |
| 133323 | 48724 | 50155 | 50156 | 50157 | 101713 | 101714 | 2fe5 |
| 133325 | 53931 | 55728 | 55729 | 55730 | 143676 | 143677 | 2fe7 |
| 133354 | 51349 | 52171 | 52317 | 52325 | 142072 | 142073 | 2fex |
| 133385 | 51349 | 51350 | 51556 | 141819 | 141820 | 141821 | 2ffi |
| 133426 | 48724 | 50474 | 50475 | 50476 | 141364 | 141365 | 2fg9 |
| 133478 | 51349 | 52539 | 52540 | 52592 | 89664 | 142224 | 2fh5 |
| 133497 | 48724 | 50474 | 50475 | 50476 | 141358 | 141359 | 2fhq |
| 133504 | 53931 | 54116 | 54117 | 54118 | 54128 | 54131 | 2fht |
| 133514 | 53931 | 55728 | 55729 | 55730 | 143662 | 143663 | 2fia |
| 133539 | 53931 | 55728 | 55729 | 55730 | 143708 | 143709 | 2fiw |
| 133559 | 46456 | 47239 | 47240 | 47241 | 47250 | 140435 | 2fjc |
| 133611 | 48724 | 50728 | 50729 | 50730 | 141413 | 141414 | 2fjl |
| 133648 | 51349 | 53334 | 53335 | 69560 | 142591 | 142592 | 2fk8 |
| 133650 | 53931 | 55810 | 55811 | 64369 | 143770 | 143771 | 2fkb |
| 133688 | 46456 | 47239 | 47240 | 47241 | 47244 | 140431 | 2fkz |
| 133706 | 53931 | 55728 | 55729 | 55730 | 143682 | 143683 | 2fl4 |
| 133799 | 53931 | 56321 | 56322 | 56323 | 143943 | 143944 | 2fn0 |
| 133810 | 51349 | 52539 | 52540 | 81269 | 142326 | 142327 | 2fna |
| 133831 | 51349 | 53382 | 53383 | 53417 | 142672 | 142673 | 2fnu |
| 133868 | 53931 | 54494 | 54495 | 54496 | 143070 | 143071 | 2fo3 |
| 133961 | 51349 | 51734 | 51735 | 51751 | 141909 | 141910 | 2fr1 |
| 133962 | 51349 | 51734 | 51735 | 51751 | 141909 | 141910 | 2fr1 |
| 133965 | 51349 | 53926 | 53927 | 53928 | 75327 | 142830 | 2fr5 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 134003 | 53931 | 54636 | 54637 | 89902 | 89903 | 89904 | 2fs2 |
| 134005 | 48724 | 50813 | 50814 | 50847 | 50861 | 50862 | 2fs6 |
| 134036 | 53931 | 55728 | 55729 | 55730 | 143704 | 143705 | 2fsr |
| 134119 | 51349 | 56783 | 56784 | 82388 | 142161 | 142162 | 2fue |
| 134124 | 51349 | 52832 | 52833 | 142405 | 142406 | 142407 | 2fug |
| 134129 | 56572 | 56761 | 56762 | 144028 | 144029 | 144030 | 2fug |
| 134130 | 53931 | 143242 | 143243 | 143244 | 143245 | 143246 | 2fug |
| 134131 | 56572 | 56769 | 56770 | 144031 | 144032 | 144033 | 2fug |
| 134175 | 53931 | 54636 | 54637 | 54638 | 143146 | 143147 | 2fuj |
| 134185 | 48724 | 50474 | 50475 | 50476 | 141366 | 141367 | 2fur |
| 134197 | 51349 | 53612 | 53613 | 53614 | 53615 | 142709 | 2fv7 |
| 134226 | 53931 | 55810 | 55811 | 55812 | 143766 | 143767 | 2fvv |
| 134231 | 51349 | 52095 | 52096 | 52103 | 142002 | 142003 | 2fw2 |
| 134242 | 51349 | 52832 | 52833 | 52834 | 102435 | 102436 | 2fwh |
| 134245 | 48724 | 50181 | 50182 | 50183 | 141291 | 141292 | 2fwk |
| 134344 | 51349 | 52832 | 52833 | 52901 | 142385 | 142386 | 2fy6 |
| 134391 | 51349 | 53334 | 53335 | 53351 | 142581 | 142582 | 2fyt |
| 134437 | 46456 | 47239 | 47240 | 47241 | 140438 | 140439 | 2fzf |
| 134448 | 51349 | 53596 | 53597 | 53598 | 53605 | 53608 | 2fzi |
| 134499 | 51349 | 52539 | 52540 | 52652 | 142299 | 142300 | 2g0t |
| 134557 | 53931 | 55728 | 55729 | 55730 | 143656 | 143657 | 2g3a |
| 134654 | 53931 | 56321 | 56322 | 56323 | 143945 | 143946 | 2g5f |
| 134674 | 48724 | 48725 | 48726 | 48727 | 89176 | 89177 | 2g5r |
| 134741 | 51349 | 56783 | 56784 | 142186 | 142187 | 142188 | 2g80 |
| 134766 | 51349 | 53926 | 53927 | 89800 | 142833 | 142834 | 2g84 |
| 134779 | 53931 | 55830 | 55831 | 55832 | 55833 | 55834 | 2g8o |
| 134889 | 53931 | 55728 | 55729 | 55730 | 143666 | 143667 | 2gan |
| 134903 | 51349 | 53382 | 53383 | 53384 | 142661 | 142662 | 2gb3 |
| 135015 | 53931 | 55330 | 55331 | 55339 | 55340 | 55343 | 2gdg |
| 135043 | 51349 | 51734 | 51735 | 51751 | 141888 | 141889 | 2gdz |
| 135044 | 53931 | 55728 | 55729 | 55730 | 143660 | 143661 | 2ge3 |
| 135077 | 53931 | 54636 | 54637 | 54638 | 143156 | 143157 | 2gf6 |
| 135096 | 51349 | 56783 | 56784 | 75173 | 142137 | 142138 | 2gfh |
| 135109 | 53931 | 56111 | 56112 | 88854 | 56129 | 56130 | 2gfs |
| 135155 | 53931 | 54861 | 55008 | 55009 | 64279 | 64280 | 2ggp |
| 135175 | 56992 | 57666 | 57667 | 57668 | 144141 | 144142 | 2ghf |
| 135176 | 56992 | 57666 | 57667 | 57668 | 144141 | 144142 | 2ghf |
| 135184 | 53931 | 54861 | 54928 | 54929 | 143338 | 143339 | 2ghp |
| 135185 | 53931 | 54861 | 54928 | 54929 | 143338 | 143339 | 2ghp |
| 135186 | 53931 | 54861 | 54928 | 54929 | 143338 | 143339 | 2ghp |
| 135254 | 48724 | 48725 | 48726 | 48727 | 141008 | 141009 | 2giy |
| 135270 | 51349 | 52539 | 52540 | 52592 | 102366 | 102367 | 2gj8 |
| 135287 | 51349 | 52539 | 52540 | 52592 | 142295 | 142296 | 2gjs |
| 135415 | 51349 | 52539 | 52540 | 81269 | 64031 | 142324 | 2gno |
| 135434 | 51349 | 56783 | 56784 | 75173 | 142139 | 142140 | 2go7 |
| 135630 | 46456 | 101385 | 101386 | 116993 | 140790 | 140791 | 2gta |
| 135737 | 51349 | 53066 | 53098 | 53118 | 82444 | 82445 | 2gui |
| 135759 | 51349 | 51904 | 51905 | 51943 | 117444 | 117445 | 2gv8 |
| 135760 | 51349 | 51904 | 51905 | 51943 | 117444 | 117445 | 2gv8 |
| 135776 | 53931 | 54636 | 54637 | 54638 | 143148 | 143149 | 2gvh |
| 135777 | 53931 | 54636 | 54637 | 54638 | 143148 | 143149 | 2gvh |
| 135805 | 51349 | 51350 | 51556 | 141819 | 141826 | 141827 | 2gwg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 135850 | 51349 | 52079 | 52080 | 52081 | 52082 | 141994 | 2gyc |
| 135853 | 48724 | 50036 | 50104 | 50105 | 50106 | 141241 | 2gyc |
| 135944 | 51349 | 88722 | 88723 | 89619 | 142110 | 142111 | 2h1c |
| 136046 | 48724 | 50155 | 50156 | 50157 | 82085 | 82086 | 2h3l |
| 136185 | 46456 | 48370 | 48439 | 48440 | 48441 | 69093 | 2h6f |
| 136186 | 46456 | 48207 | 48239 | 48246 | 48247 | 69090 | 2h6f |
| 136219 | 51349 | 51734 | 51735 | 51751 | 51791 | 51793 | 2h7m |
| 136231 | 56992 | 57301 | 57302 | 57303 | 144103 | 144104 | 2h7z |
| 136246 | 56992 | 57301 | 57302 | 57303 | 111404 | 111405 | 2h8u |
| 136314 | 51349 | 53066 | 53098 | 53118 | 142493 | 142494 | 2hbk |
| 136319 | 53931 | 54636 | 54637 | 89902 | 143177 | 143178 | 2hbo |
| 136320 | 51349 | 51350 | 51556 | 141819 | 141824 | 141825 | 2hbv |
| 136333 | 51349 | 56783 | 56784 | 75173 | 142147 | 142148 | 2hcf |
| 136345 | 51349 | 56783 | 56784 | 75173 | 142135 | 142136 | 2hdo |
| 136354 | 46456 | 109603 | 109604 | 101340 | 140765 | 140766 | 2hek |
| 136514 | 51349 | 53066 | 53098 | 53118 | 53119 | 53122 | 2hhv |
| 136555 | 53931 | 55873 | 55874 | 55879 | 82778 | 82779 | 2hkj |
| 136570 | 53931 | 54636 | 54637 | 54638 | 143144 | 143145 | 2hlj |
| 136633 | 48724 | 50813 | 50814 | 50847 | 50856 | 110275 | 2hnx |
| 136644 | 51349 | 53382 | 53383 | 53384 | 82482 | 82483 | 2hox |
| 136657 | 48724 | 50474 | 50475 | 50476 | 141360 | 141361 | 2hq7 |
| 136659 | 48724 | 50474 | 50475 | 50476 | 141354 | 141355 | 2hq9 |
| 136726 | 51349 | 56783 | 56784 | 75173 | 142143 | 142144 | 2hsz |
| 136744 | 48724 | 50474 | 50475 | 50476 | 141348 | 141349 | 2hti |
| 136776 | 46456 | 47112 | 47113 | 47114 | 47125 | 47127 | 2hue |
| 136837 | 53931 | 54636 | 54637 | 54638 | 143160 | 143161 | 2hx5 |
| 136860 | 51349 | 53926 | 53927 | 89800 | 142837 | 142839 | 2hxv |
| 136936 | 48724 | 50474 | 50475 | 50476 | 141352 | 141353 | 2i02 |
| 137095 | 53931 | 55728 | 55729 | 55730 | 143700 | 143701 | 2i6c |
| 137195 | 46456 | 109603 | 109604 | 140769 | 140770 | 140771 | 2ibn |
| 137266 | 51349 | 52742 | 52743 | 52744 | 89692 | 89693 | 2id4 |
| 137335 | 48724 | 48725 | 48726 | 49159 | 49190 | 49191 | 2ifg |
| 137343 | 51349 | 52832 | 52833 | 52834 | 52835 | 52842 | 2ifq |
| 137449 | 46456 | 48263 | 48264 | 48265 | 48268 | 48269 | 2ij2 |
| 137521 | 53931 | 54494 | 54495 | 143072 | 143073 | 143074 | 2in1 |
| 137524 | 46456 | 47239 | 47240 | 47253 | 109788 | 109789 | 2inc |
| 137525 | 46456 | 47239 | 47240 | 47253 | 109790 | 109791 | 2inc |
| 137709 | 48724 | 51181 | 51197 | 141628 | 141631 | 141632 | 2iuw |
| 137738 | 51349 | 53755 | 53756 | 110734 | 142771 | 142772 | 2iw1 |
| 137739 | 46456 | 46688 | 46689 | 46739 | 140165 | 140166 | 2iw5 |
| 137812 | 51349 | 52539 | 52540 | 52566 | 52567 | 75194 | 2iyv |
| 137920 | 48724 | 50198 | 50242 | 50243 | 50244 | 50245 | 2j0t |
| 138009 | 53931 | 56111 | 56112 | 88854 | 90038 | 90039 | 2j4z |
| 138130 | 46456 | 48507 | 48508 | 48509 | 48521 | 48523 | 2j7y |
| 138131 | 53931 | 54116 | 54117 | 54118 | 54150 | 54151 | 2j7z |
| 138240 | 46456 | 48370 | 48371 | 140819 | 140820 | 140821 | 2jak |
| 138243 | 53931 | 56111 | 56112 | 88854 | 143893 | 143894 | 2jav |
| 138356 | 56835 | 56836 | 56854 | 56855 | 118212 | 118213 | 2jm6 |
| 138392 | 48724 | 49898 | 49899 | 49932 | 49940 | 49941 | 2nn8 |
| 138562 | 51349 | 51350 | 51445 | 51487 | 89473 | 89474 | 2nt0 |
| 138608 | 53931 | 54636 | 54637 | 54638 | 143154 | 143155 | 2nuj |
| 138714 | 46456 | 140958 | 140959 | 140960 | 140961 | 140962 | 2nw8 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 138721 | 56835 | 118214 | 118215 | 118216 | 118217 | 118218 | 2nww |
| 138745 | 48724 | 48725 | 48726 | 48727 | 48737 | 48738 | 2nxy |
| 138862 | 51349 | 53334 | 53335 | 69557 | 69558 | 142590 | 2o07 |
| 138884 | 51349 | 51734 | 51735 | 51751 | 63933 | 102152 | 2o23 |
| 138920 | 46456 | 109603 | 109604 | 101340 | 140767 | 140768 | 2o6i |
| 139030 | 53931 | 143119 | 143120 | 143121 | 143124 | 143125 | 2odk |
| 139092 | 46456 | 101385 | 101386 | 116993 | 140792 | 140793 | 2oie |
| 139443 | 53931 | 56111 | 56112 | 88854 | 82789 | 82790 | 2oza |
| 139497 | 46456 | 48507 | 48508 | 48509 | 69109 | 69110 | 2p54 |
| 139643 | 46456 | 109603 | 109604 | 101340 | 116969 | 116970 | 2paq |
| 139645 | 51349 | 51350 | 51445 | 51487 | 51495 | 51496 | 2pb1 |
| 139762 | 51349 | 52742 | 52743 | 52744 | 52762 | 52763 | 2pwa |
| 139825 | 51349 | 51734 | 51735 | 51751 | 117417 | 117418 | 2q46 |
| 139941 | 53931 | 54861 | 54999 | 55000 | 55001 | 55002 | 2uub |
| 139942 | 51349 | 53066 | 53137 | 53138 | 53141 | 53142 | 2uub |
| 139943 | 48724 | 50198 | 50249 | 50282 | 50302 | 50303 | 2uub |
| 144534 | 48724 | 50964 | 50978 | 50979 | 159257 | 159258 | 1vyh |
| 144580 | 56992 | 57500 | 57501 | 57528 | 64562 | 64563 | 1xwd |
| 144607 | 51349 | 69571 | 69572 | 89763 | 89764 | 89765 | 1y8x |
| 144615 | 46456 | 47161 | 47216 | 47217 | 158392 | 158393 | 1yar |
| 144649 | 48724 | 48725 | 48726 | 48727 | 158865 | 158866 | 1yjd |
| 144710 | 48724 | 48725 | 48726 | 48727 | 48933 | 63651 | 1ypz |
| 144788 | 51349 | 52539 | 52540 | 81269 | 159577 | 159578 | 2a5y |
| 144808 | 48724 | 48725 | 48726 | 48727 | 48933 | 48935 | 2ak4 |
| 144837 | 48724 | 48725 | 48726 | 48727 | 48734 | 158864 | 2atp |
| 144996 | 48724 | 48725 | 48726 | 48727 | 48933 | 48935 | 2bnq |
| 145078 | 48724 | 50728 | 50729 | 50730 | 159209 | 159210 | 2dfk |
| 145109 | 48724 | 48725 | 48726 | 49159 | 158879 | 158880 | 2dyp |
| 145162 | 48724 | 50352 | 50353 | 50354 | 159145 | 159146 | 2fdb |
| 145176 | 53931 | 74651 | 48403 | 48404 | 160152 | 160153 | 2fo1 |
| 145232 | 51349 | 52171 | 52313 | 52314 | 52315 | 159491 | 2gy9 |
| 145237 | 48724 | 50198 | 50249 | 50282 | 50304 | 159088 | 2gy9 |
| 145271 | 51349 | 53066 | 53137 | 53138 | 53139 | 159642 | 2gyc |
| 145525 | 48724 | 48725 | 48726 | 48727 | 48933 | 48937 | 2ij0 |
| 145572 | 51349 | 52079 | 52080 | 52081 | 52082 | 159454 | 2j01 |
| 145656 | 48724 | 50728 | 50729 | 50730 | 159213 | 159214 | 2j59 |
| 145694 | 53931 | 54861 | 54913 | 54914 | 54917 | 160309 | 2ns1 |
| 145736 | 48724 | 50964 | 50978 | 50979 | 159259 | 159260 | 2ovr |
| 145768 | 53931 | 54494 | 54495 | 54496 | 54497 | 54503 | 2uyz |
| 145841 | 53931 | 55680 | 55681 | 143642 | 160607 | 160608 | 1x2g |
| 145891 | 53931 | 110848 | 110849 | 160095 | 160096 | 160097 | 1xw3 |
| 145918 | 56835 | 81339 | 81338 | 56895 | 103468 | 118226 | 1ymg |
| 146035 | 53931 | 56111 | 56112 | 64411 | 160818 | 160819 | 1zyl |
| 146036 | 51349 | 100949 | 100950 | 110513 | 159543 | 159544 | 2a0u |
| 146040 | 56835 | 161069 | 161070 | 161071 | 161072 | 161073 | 2a65 |
| 146047 | 51349 | 53755 | 53756 | 159785 | 159788 | 159789 | 2acv |
| 146061 | 53931 | 55297 | 55315 | 55316 | 160504 | 160505 | 2ale |
| 146085 | 53931 | 54861 | 54928 | 54929 | 54932 | 160313 | 2b0g |
| 146096 | 48724 | 50198 | 50249 | 50282 | 159100 | 159103 | 2ba0 |
| 146161 | 53931 | 55297 | 55315 | 55316 | 55317 | 89986 | 2bo1 |
| 146163 | 56572 | 111330 | 111331 | 160984 | 160985 | 160987 | 2bon |
| 146219 | 51349 | 53447 | 53448 | 159726 | 159731 | 159732 | 2bvl |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 146225 | 51349 | 53755 | 53756 | 159785 | 159786 | 159787 | 2c1x |
| 146379 | 53931 | 55680 | 55681 | 143642 | 160604 | 160606 | 2c8m |
| 146398 | 46456 | 48263 | 48264 | 48265 | 48276 | 48277 | 2cib |
| 146411 | 46456 | 46688 | 46689 | 46739 | 158244 | 158245 | 2ckx |
| 146423 | 48724 | 48725 | 48726 | 48727 | 158867 | 158868 | 2cry |
| 146459 | 53931 | 55728 | 55729 | 55730 | 160640 | 160641 | 2d4p |
| 146466 | 51349 | 53066 | 53098 | 102492 | 159633 | 159634 | 2d5r |
| 146608 | 51349 | 52539 | 52540 | 52592 | 52633 | 142226 | 2dy1 |
| 146654 | 46456 | 48507 | 48508 | 48509 | 81916 | 81917 | 2e2r |
| 146687 | 53931 | 143112 | 143113 | 143114 | 160140 | 160141 | 2.00E+50 |
| 146706 | 53931 | 56280 | 56281 | 143916 | 143917 | 143920 | 2e7y |
| 146789 | 46456 | 46688 | 46689 | 46690 | 158242 | 158243 | 2ecb |
| 146896 | 48724 | 50728 | 50729 | 50730 | 159211 | 159212 | 2elb |
| 146971 | 56992 | 57666 | 57667 | 57668 | 161152 | 161153 | 2epp |
| 146973 | 56992 | 57666 | 57667 | 57668 | 161152 | 161153 | 2epr |
| 146974 | 56992 | 57666 | 57667 | 57668 | 161152 | 161153 | 2eps |
| 146989 | 56992 | 100894 | 100895 | 57468 | 57476 | 161138 | 2erw |
| 147177 | 51349 | 53162 | 53187 | 142526 | 159649 | 159650 | 2gu2 |
| 147196 | 46456 | 109603 | 109604 | 101340 | 158725 | 158726 | 2gz4 |
| 147239 | 56835 | 161083 | 161084 | 161085 | 161088 | 161089 | 2h8a |
| 147437 | 53931 | 56299 | 56300 | 160875 | 160876 | 160877 | 2hy1 |
| 147460 | 53931 | 55769 | 103196 | 103197 | 118074 | 118075 | 2hz5 |
| 147512 | 48724 | 50728 | 50729 | 50730 | 50740 | 50741 | 2i5f |
| 147657 | 51349 | 53066 | 53098 | 53118 | 89719 | 142492 | 2igi |
| 147818 | 48724 | 50198 | 50249 | 50282 | 159089 | 159090 | 2ix0 |
| 147819 | 48724 | 50198 | 50249 | 50282 | 159089 | 159090 | 2ix0 |
| 147820 | 48724 | 50198 | 50249 | 50282 | 159089 | 159090 | 2ix0 |
| 147903 | 56992 | 57666 | 57667 | 57668 | 57693 | 57694 | 2j7j |
| 147905 | 56992 | 57666 | 57667 | 57668 | 57693 | 57694 | 2j7j |
| 147950 | 48724 | 50198 | 50249 | 50282 | 159091 | 159092 | 2ja9 |
| 147962 | 51349 | 53473 | 53474 | 159762 | 159763 | 159764 | 2jbw |
| 147977 | 53931 | 55728 | 55729 | 55730 | 143670 | 143671 | 2jdc |
| 147995 | 48724 | 50198 | 50249 | 50282 | 159100 | 159102 | 2je6 |
| 148002 | 51349 | 51350 | 51445 | 51487 | 159385 | 159386 | 2je8 |
| 148032 | 53931 | 56111 | 56112 | 88854 | 160812 | 160813 | 2jfl |
| 148140 | 53931 | 55153 | 55154 | 118007 | 160428 | 160429 | 2jmu |
| 148163 | 48724 | 50036 | 50104 | 159031 | 159032 | 159033 | 2joy |
| 148164 | 48724 | 50813 | 50814 | 50815 | 159228 | 159229 | 2joz |
| 148310 | 48724 | 50198 | 50249 | 50282 | 159096 | 159097 | 2nn6 |
| 148313 | 48724 | 50198 | 50249 | 50282 | 159098 | 159099 | 2nn6 |
| 148392 | 48724 | 48725 | 48726 | 48727 | 48933 | 48937 | 2nts |
| 148532 | 51349 | 53755 | 53756 | 159794 | 159795 | 159796 | 2nzw |
| 148541 | 46456 | 48370 | 48445 | 48446 | 48449 | 48450 | 2o02 |
| 148591 | 51349 | 53334 | 53335 | 69560 | 159680 | 159681 | 2o57 |
| 148596 | 53931 | 55810 | 55811 | 64369 | 103217 | 160692 | 2o5f |
| 148616 | 53931 | 54636 | 54637 | 54638 | 143158 | 143159 | 2o5u |
| 148679 | 46456 | 48370 | 48445 | 48446 | 158768 | 158769 | 2o8p |
| 148695 | 53931 | 54636 | 54637 | 54638 | 160180 | 160181 | 2oaf |
| 148714 | 51349 | 102545 | 102546 | 102547 | 159677 | 159678 | 2ob5 |
| 148780 | 53931 | 54196 | 54197 | 54198 | 159880 | 159881 | 2oik |
| 148794 | 53931 | 54636 | 54637 | 54638 | 160182 | 160183 | 2oiw |
| 148903 | 51349 | 52539 | 52540 | 52686 | 159569 | 159570 | 2onk |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 148997 | 46456 | 47265 | 47266 | 47286 | 158425 | 158426 | 2oqp |
| 149025 | 53931 | 54000 | 54001 | 54002 | 142846 | 142847 | 2oul |
| 149034 | 53931 | 54636 | 54637 | 89902 | 160184 | 160185 | 2ov9 |
| 149100 | 48724 | 48725 | 48726 | 49159 | 49162 | 49163 | 2oz4 |
| 149101 | 48724 | 48725 | 48726 | 49159 | 49162 | 49163 | 2oz4 |
| 149135 | 53931 | 55680 | 55681 | 143642 | 160604 | 160605 | 2p0l |
| 149171 | 48724 | 89359 | 89360 | 89361 | 159226 | 159227 | 2p2e |
| 149189 | 48724 | 50493 | 50494 | 50514 | 50574 | 50575 | 2p3u |
| 149207 | 48724 | 48725 | 48726 | 48727 | 88563 | 88564 | 2p49 |
| 149248 | 53931 | 55680 | 55681 | 143642 | 160609 | 160610 | 2p5i |
| 149285 | 51349 | 53334 | 53335 | 110671 | 159682 | 159683 | 2p7i |
| 149354 | 53931 | 55769 | 55770 | 55771 | 55772 | 55774 | 2pbd |
| 149493 | 48724 | 50198 | 50249 | 50263 | 50269 | 50270 | 2pi2 |
| 149497 | 48724 | 50198 | 50249 | 50263 | 50271 | 50272 | 2pi2 |
| 149569 | 46456 | 109603 | 109604 | 101340 | 158729 | 158730 | 2pjq |
| 149669 | 51349 | 52539 | 52540 | 52686 | 89683 | 89684 | 2pmk |
| 149781 | 53931 | 54533 | 54534 | 54535 | 54536 | 54537 | 2ppn |
| 149782 | 53931 | 56111 | 56112 | 64411 | 160816 | 160817 | 2ppq |
| 149785 | 51349 | 53755 | 53756 | 159785 | 159790 | 159791 | 2pq6 |
| 149786 | 46456 | 109603 | 109604 | 101340 | 158719 | 158720 | 2pq7 |
| 149793 | 46456 | 48370 | 48452 | 48453 | 101417 | 140836 | 2pqr |
| 149824 | 53931 | 55944 | 55961 | 55966 | 160730 | 160731 | 2pso |
| 149865 | 53931 | 56111 | 56112 | 64411 | 160814 | 160815 | 2pul |
| 149878 | 53931 | 54533 | 54534 | 54535 | 82624 | 82625 | 2pv2 |
| 149928 | 51349 | 53066 | 53098 | 53118 | 117650 | 117651 | 2py5 |
| 150034 | 53931 | 56370 | 56371 | 56372 | 160917 | 160918 | 2q3n |
| 150097 | 56835 | 161083 | 161084 | 161085 | 161086 | 161087 | 2q7r |
| 150105 | 51349 | 53382 | 53383 | 53384 | 53385 | 53390 | 2q7w |
| 150112 | 46456 | 48575 | 48576 | 48577 | 158813 | 158814 | 2q80 |
| 150320 | 51349 | 51350 | 51569 | 51570 | 51576 | 51582 | 2qap |
| 150676 | 53931 | 56280 | 56281 | 56288 | 56289 | 160851 | 2qed |
| 150733 | 56992 | 57923 | 57924 | 57925 | 57930 | 57931 | 2qfa |
| 150738 | 46456 | 46457 | 46458 | 46463 | 46530 | 46531 | 2qfk |
| 150751 | 53931 | 56299 | 56300 | 56301 | 56302 | 56303 | 2qfr |
| 150782 | 46456 | 109603 | 109604 | 101340 | 158723 | 158724 | 2qgs |
| 150808 | 53931 | 54861 | 55008 | 55009 | 55017 | 69736 | 2qif |
| 150815 | 53931 | 54402 | 54427 | 54431 | 159959 | 159960 | 2qiy |
| 150835 | 46456 | 47575 | 47576 | 47577 | 101194 | 101195 | 2qjz |
| 150874 | 51349 | 52539 | 52540 | 52652 | 159564 | 159565 | 2qm8 |
| 151197 | 48724 | 50629 | 50630 | 50646 | 50671 | 50672 | 2qp8 |
| 151215 | 53931 | 55944 | 55961 | 160742 | 160745 | 160746 | 2qpv |
| 151386 | 56572 | 111330 | 111331 | 160984 | 160988 | 160989 | 2qv7 |
| 151441 | 51349 | 52171 | 52218 | 52235 | 52240 | 52241 | 2qwx |
| 151455 | 51349 | 53066 | 53098 | 53118 | 53132 | 159630 | 2qxf |
| 151470 | 51349 | 52539 | 52540 | 52652 | 52666 | 52667 | 2qy9 |
| 151529 | 46456 | 47161 | 47226 | 47230 | 47231 | 47232 | 2r25 |
| 151530 | 51349 | 52171 | 52172 | 52173 | 102228 | 102229 | 2r25 |
| 151572 | 46456 | 48507 | 48508 | 48509 | 101436 | 101437 | 2r40 |
| 151588 | 51349 | 53382 | 53383 | 53384 | 110681 | 142654 | 2r5e |
| 151610 | 56835 | 161097 | 161098 | 161099 | 161106 | 161107 | 2r6g |
| 151611 | 56835 | 161097 | 161098 | 161099 | 161104 | 161105 | 2r6g |
| 151629 | 48724 | 50198 | 50249 | 50282 | 159094 | 159095 | 2r7d |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 151850 | 51349 | 56783 | 56784 | 82388 | 142157 | 142158 | 2rbk |
| 152048 | 51349 | 51734 | 51735 | 51751 | 51788 | 117408 | 2rhc |
| 152184 | 56835 | 161083 | 161084 | 161085 | 161090 | 161091 | 2uui |
| 152457 | 51349 | 51350 | 51445 | 51487 | 141783 | 141784 | 2v3g |
| 152460 | 51349 | 75216 | 75217 | 159513 | 159516 | 159517 | 2v3k |
| 152871 | 53931 | 53954 | 53955 | 53960 | 53961 | 53962 | 2vb1 |
| 152917 | 51349 | 53755 | 53756 | 159785 | 159792 | 159793 | 2vch |
| 153217 | 51349 | 53447 | 53448 | 159726 | 159729 | 159730 | 2vk9 |
| 153344 | 48724 | 50198 | 50249 | 50282 | 159104 | 159105 | 2vnu |
| 153345 | 48724 | 50198 | 50249 | 50282 | 159104 | 159105 | 2vnu |
| 153346 | 48724 | 50198 | 50249 | 50282 | 159104 | 159105 | 2vnu |
| 153399 | 51349 | 52539 | 52540 | 52541 | 69478 | 69479 | 2vp4 |
| 153421 | 48724 | 50474 | 50475 | 50476 | 117228 | 117229 | 2vpa |
| 153731 | 51349 | 51350 | 51445 | 51487 | 159383 | 159384 | 2vzs |
| 153794 | 46456 | 47239 | 47240 | 47241 | 47250 | 109784 | 2yw6 |
| 153994 | 51349 | 53926 | 53927 | 53928 | 142831 | 142832 | 2z3g |
| 154002 | 46456 | 47265 | 47266 | 47286 | 158422 | 158423 | 2z3q |
| 154184 | 48724 | 50890 | 50891 | 50892 | 141511 | 141512 | 2z6w |
| 154234 | 51349 | 52171 | 52218 | 52235 | 110470 | 110471 | 2z98 |
| 154343 | 51349 | 52832 | 52833 | 52901 | 117601 | 117602 | 2zct |
| 154405 | 53931 | 54235 | 54236 | 54237 | 89831 | 89833 | 2zeq |
| 154406 | 46456 | 47472 | 47473 | 47502 | 101182 | 101183 | 2zfd |
| 154589 | 48724 | 50728 | 50729 | 50730 | 159207 | 159208 | 2zkm |
| 154835 | 46456 | 109603 | 109604 | 101340 | 158721 | 158722 | 3b57 |
| 154838 | 48724 | 48725 | 48726 | 49159 | 158877 | 158878 | 3b5h |
| 154839 | 48724 | 48725 | 48726 | 49159 | 158877 | 158878 | 3b5h |
| 154901 | 51349 | 53066 | 53098 | 53118 | 159631 | 159632 | 3b6o |
| 155055 | 53931 | 54861 | 54919 | 54920 | 54921 | 54922 | 3bbb |
| 155061 | 51349 | 75216 | 75217 | 159513 | 159514 | 159515 | 3bbd |
| 155161 | 53931 | 56435 | 56436 | 56437 | 56440 | 56441 | 3bdw |
| 155185 | 53931 | 54861 | 54928 | 54929 | 143346 | 143347 | 3beg |
| 155376 | 56835 | 56836 | 56854 | 56855 | 144081 | 144082 | 3bl2 |
| 155393 | 53931 | 56111 | 56112 | 88854 | 160810 | 160811 | 3blh |
| 155446 | 51349 | 51350 | 82282 | 82283 | 102108 | 102109 | 3bof |
| 155462 | 48724 | 48725 | 48726 | 48727 | 101510 | 101511 | 3bp5 |
| 155490 | 46456 | 46688 | 46785 | 46832 | 46833 | 46834 | 3bpy |
| 155497 | 53931 | 56111 | 56112 | 88854 | 56142 | 160809 | 3bqc |
| 155704 | 48724 | 50352 | 50386 | 50387 | 50396 | 50397 | 3bx1 |
| 155780 | 48724 | 50198 | 50249 | 50282 | 159106 | 159107 | 3bzk |
| 155784 | 53931 | 56321 | 56322 | 56323 | 143941 | 143942 | 3bzn |
| 155847 | 51349 | 52767 | 52768 | 52773 | 159582 | 159583 | 3c10 |
| 156848 | 53931 | 55944 | 55961 | 160742 | 160743 | 160744 | 3cnw |
| 156903 | 53931 | 55769 | 103196 | 103197 | 111120 | 111121 | 3cpt |
| 156985 | 53931 | 56370 | 56371 | 56372 | 160915 | 160916 | 3ctk |
| 157035 | 56572 | 56800 | 56801 | 56802 | 111316 | 160961 | 3cw9 |
| 157149 | 46456 | 48263 | 48264 | 48265 | 158749 | 158750 | 3czh |
| 157161 | 53931 | 56299 | 56300 | 160875 | 160878 | 160879 | 3d03 |
| 157208 | 46456 | 46457 | 46458 | 46463 | 46486 | 46497 | 3d1k |
| 157209 | 46456 | 46457 | 46458 | 46463 | 46500 | 46513 | 3d1k |
| 157265 | 56835 | 161097 | 161098 | 161099 | 161100 | 161101 | 3d31 |
| 157439 | 48724 | 48725 | 48726 | 49159 | 63665 | 63666 | 3d85 |
| 157463 | 56992 | 57923 | 57924 | 57925 | 57926 | 57927 | 3d9t |

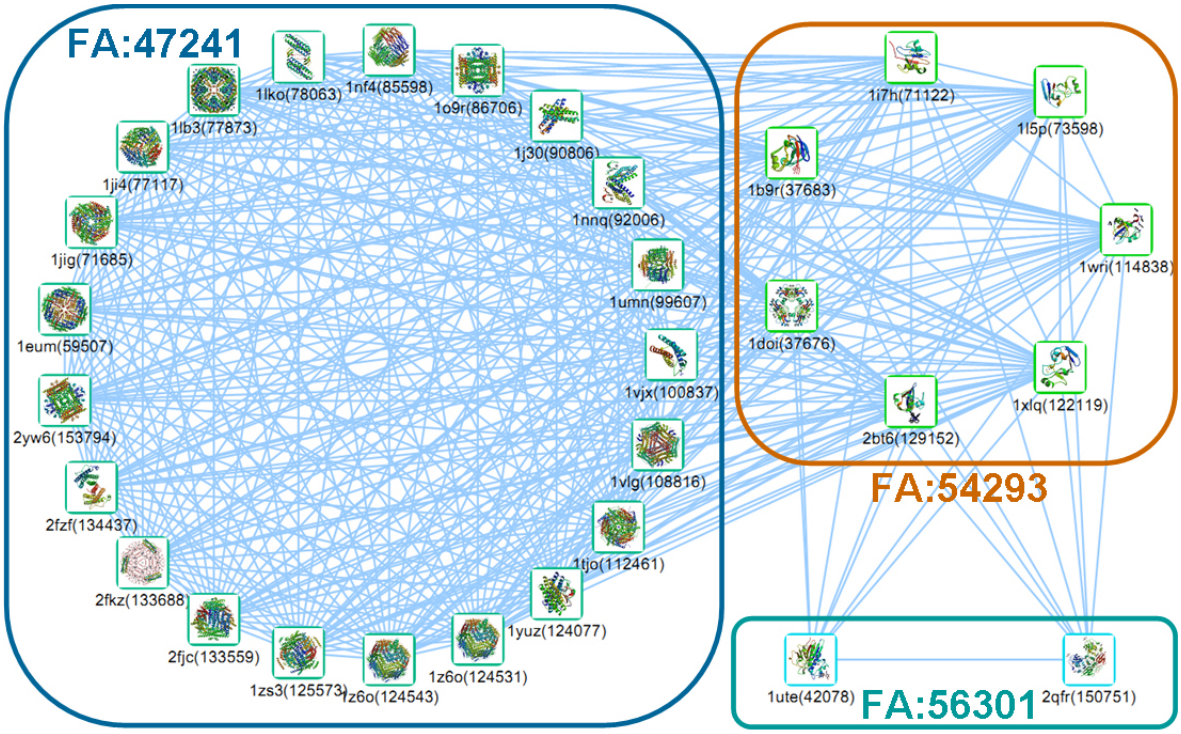
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 157466 | 48724 | 48725 | 48726 | 49159 | 49179 | 49181 | 3dar |
| 157727 | 56835 | 161097 | 161098 | 161099 | 161108 | 161109 | 3dhw |
| 157729 | 51349 | 52539 | 52540 | 52686 | 159571 | 159572 | 3dhw |
| 157741 | 46456 | 47239 | 47240 | 47253 | 47257 | 69006 | 3dhz |
| 157802 | 46456 | 47265 | 47266 | 47305 | 89036 | 89037 | 3dlq |
| 157824 | 56992 | 57849 | 57850 | 57851 | 75693 | 75694 | 3dpl |
| 157841 | 46456 | 48207 | 48239 | 48246 | 48249 | 48250 | 3dss |
| 158080 | 51349 | 51350 | 51430 | 51431 | 51434 | 51435 | 3eau |
| 158144 | 46456 | 48370 | 48445 | 48446 | 158770 | 158771 | 3efz |
| 158154 | 48724 | 51268 | 51283 | 51284 | 51285 | 102010 | 3ehw |

# [S-9] 25 species used for creating phylogenetic profiles

| Taxonomy | Scientific Name |
|---|---|
| 9606 | Homo sapiens |
| 10090 | Mus musculus |
| 10116 | Rattus norvegicus |
| 3702 | Arabidopsis thaliana |
| 9913 | Bos taurus |
| 83333 | Escherichia coli (strain K12) |
| 559292 | Saccharomyces cerevisiae (strain ATCC 204508 / S288c) |
| 7227 | Drosophila melanogaster |
| 9031 | Gallus gallus |
| 1423 | Bacillus subtilis |
| 39947 | Oryza sativa subsp. japonica |
| 9823 | Sus scrofa |
| 7955 | Danio rerio |
| 6239 | Caenorhabditis elegans |
| 284812 | Schizosaccharomyces pombe (strain ATCC 38366 / 972) |
| 44689 | Dictyostelium discoideum |
| 83334 | Escherichia coli O157:H7 |
| 1773 | Mycobacterium tuberculosis |
| 90371 | Salmonella typhimurium |
| 71421 | Haemophilus influenzae (strain ATCC 51907 / DSM 11121 / KW20 / Rd) |
| 287 | Pseudomonas aeruginosa |
| 9615 | Canis familiaris |
| 70601 | Pyrococcus horikoshii (strain ATCC 700860 / DSM 12428 / JCM 9974 / NBRC 100139 / OT-3) |
| 90370 | Salmonella typhi |
| 243232 | Methanocaldococcus jannaschii (strain ATCC 43067 / DSM 2661 / JAL-1 / JCM 10045 / NBRC 100440) |

## [S-10] Analysis on domain clusters

From BSM score with 70% threshold, 1UTE and 2QFR are associated with FA:56301 which is Purple acid phosphatase-like. 1B9R, 1I7H, 1L5P, 1WRI, 1XLQ, 2BT6, and 1DOI are associated with FA:54293 which is 2Fe-2S ferredoxin-related. All others are associated with FA:47241 which is Ferritin




| PDB1 | PDB2 | Sunid1 | Sunid2 | GPsim | dmTM | dmDali | dmSEQ | dmPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1B9R | 1EUM | 37683 | 59507 | 0.72 | 0.215 | 0 | 0.286 | 0 | 0.215 | 0 | 0.286 |
| 1B9R | 1JIG | 37683 | 71685 | 0.72 | 0.223 | 0 | 0 | 0 | 0.223 | 0 | 0 |
| 1B9R | 1JI4 | 37683 | 77117 | 0.72 | 0.214 | 0 | 0 | 0 | 0.214 | 0 | 0 |
| 1B9R | 1LB3 | 37683 | 77873 | 0.72 | 0.203 | 0 | 0.235 | 0 | 0.203 | 0 | 0.235 |
| 1B9R | 1LKO | 37683 | 78063 | 0.72 | 0.247 | 0 | 0.75 | 0 | 0.182 | 0 | 0.4 |
| 1B9R | 1NF4 | 37683 | 85598 | 0.72 | 0.189 | 0 | 0.276 | 0 | 0.189 | 0 | 0.276 |
| 1B9R | 1O9R | 37683 | 86706 | 0.72 | 0.236 | 0 | 0 | 0 | 0.236 | 0 | 0 |
| 1B9R | 1J30 | 37683 | 90806 | 0.72 | 0.216 | 0 | 0.364 | 0 | 0.216 | 0 | 0.222 |
| 1B9R | 1NNQ | 37683 | 92006 | 0.72 | 0.202 | 0 | 0.461 | 0 | 0.201 | 0 | 0.461 |
| 1B9R | 1UMN | 37683 | 99607 | 0.72 | 0.197 | 0 | 0 | 0 | 0.197 | 0 | 0 |
| 1B9R | 1VJX | 37683 | 100837 | 0.72 | 0.217 | 0 | 0 | 0 | 0.217 | 0 | 0 |
| 1B9R | 1VLG | 37683 | 108816 | 0.72 | 0.218 | 0 | 0 | 0 | 0.218 | 0 | 0 |
| 1B9R | 1TJO | 37683 | 112461 | 0.72 | 0.181 | 0 | 0 | 0 | 0.181 | 0 | 0 |
| 1B9R | 1YUZ | 37683 | 124077 | 0.72 | 0.224 | 0 | 0 | 0 | 0.229 | 0 | 0 |
| 1B9R | 1Z6O | 37683 | 124531 | 0.72 | 0.183 | 0 | 0 | 0 | 0.175 | 0 | 0.353 |
| 1B9R | 1Z6O | 37683 | 124543 | 0.72 | 0.196 | 0 | 0.353 | 0 | 0.175 | 0 | 0.353 |
| 1B9R | 1ZS3 | 37683 | 125573 | 0.72 | 0.195 | 0 | 0.333 | 0 | 0.195 | 0 | 0.333 |
| 1B9R | 2FJC | 37683 | 133559 | 0.72 | 0.243 | 0 | 0 | 0 | 0.243 | 0 | 0.227 |

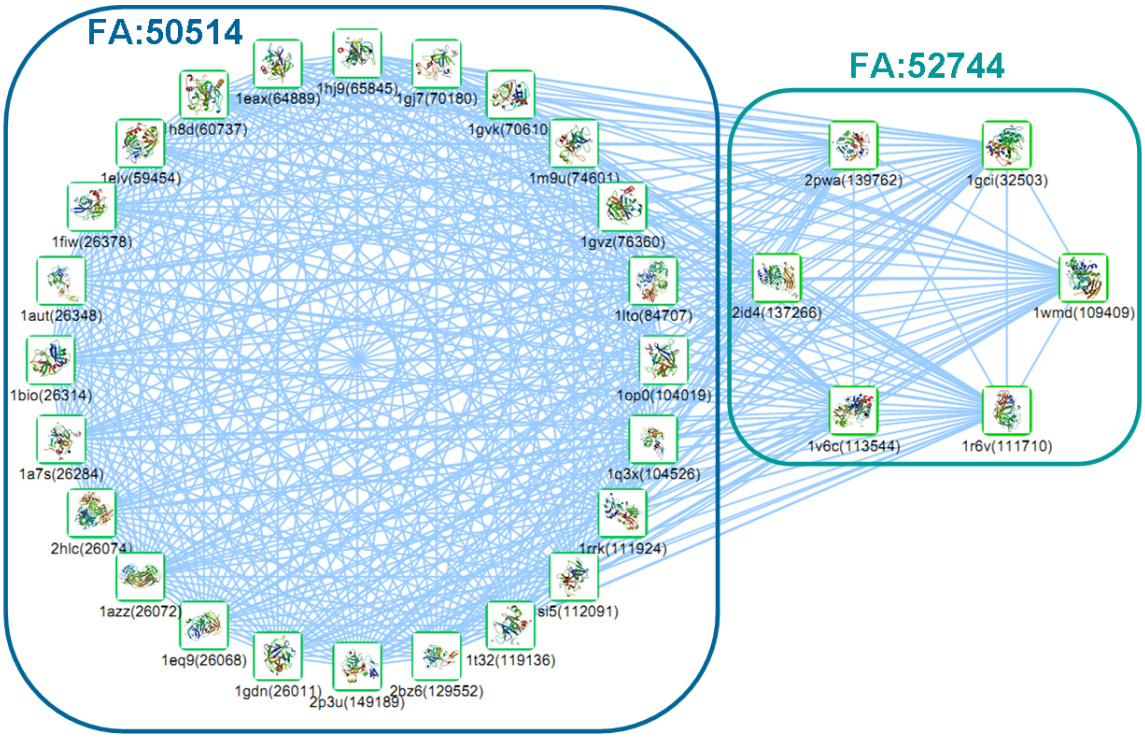
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1B9R | 2FKZ | 37683 | 133688 | 0.72 | 0.217 | 0 | 0 | 0 | 0.213 | 0 | 0.179 |
| 1B9R | 2FZF | 37683 | 134437 | 0.72 | 0.264 | 0 | 0 | 0 | 0.264 | 0 | 0 |
| 1B9R | 2YW6 | 37683 | 153794 | 0.72 | 0.204 | 0 | 0 | 0 | 0.211 | 0 | 0 |
| 1DOI | 1EUM | 37676 | 59507 | 0.72 | 0.19 | 0 | 0 | 0 | 0.19 | 0 | 0 |
| 1DOI | 1JIG | 37676 | 71685 | 0.72 | 0.234 | 0 | 0 | 0 | 0.234 | 0 | 0 |
| 1DOI | 1JI4 | 37676 | 77117 | 0.72 | 0.224 | 0 | 0.267 | 0 | 0.224 | 0 | 0.267 |
| 1DOI | 1LB3 | 37676 | 77873 | 0.72 | 0.211 | 0 | 0.444 | 0 | 0.211 | 0 | 0.444 |
| 1DOI | 1LKO | 37676 | 78063 | 0.72 | 0.239 | 0 | 0 | 0 | 0.196 | 0 | 0 |
| 1DOI | 1NF4 | 37676 | 85598 | 0.72 | 0.223 | 0 | 0.615 | 0 | 0.223 | 0 | 0.615 |
| 1DOI | 1O9R | 37676 | 86706 | 0.72 | 0.221 | 0 | 0.3 | 0 | 0.221 | 0 | 0.3 |
| 1DOI | 1J30 | 37676 | 90806 | 0.72 | 0.24 | 0 | 0 | 0 | 0.24 | 0 | 0 |
| 1DOI | 1NNQ | 37676 | 92006 | 0.72 | 0.262 | 0 | 0 | 0 | 0.239 | 0 | 0 |
| 1DOI | 1UMN | 37676 | 99607 | 0.72 | 0.225 | 0 | 0.233 | 0 | 0.225 | 0 | 0.233 |
| 1DOI | 1VJX | 37676 | 100837 | 0.72 | 0.232 | 0 | 0 | 0 | 0.232 | 0 | 0 |
| 1DOI | 1VLG | 37676 | 108816 | 0.72 | 0.217 | 0 | 0.667 | 0 | 0.217 | 0 | 0.667 |
| 1DOI | 1TJO | 37676 | 112461 | 0.72 | 0.206 | 0 | 0.5 | 0 | 0.206 | 0 | 0.5 |
| 1DOI | 1YUZ | 37676 | 124077 | 0.72 | 0.246 | 0 | 0.467 | 0 | 0.235 | 0 | 0.467 |
| 1DOI | 1Z6O | 37676 | 124531 | 0.72 | 0.199 | 0 | 0.25 | 0 | 0.185 | 0 | 0.25 |
| 1DOI | 1Z6O | 37676 | 124543 | 0.72 | 0.19 | 0 | 0.24 | 0 | 0.185 | 0 | 0.25 |
| 1DOI | 1ZS3 | 37676 | 125573 | 0.72 | 0.209 | 0 | 0.385 | 0 | 0.209 | 0 | 0.455 |
| 1DOI | 2FJC | 37676 | 133559 | 0.72 | 0.233 | 0 | 0 | 0 | 0.233 | 0 | 0 |
| 1DOI | 2FKZ | 37676 | 133688 | 0.72 | 0.205 | 0 | 0 | 0 | 0.204 | 0 | 0 |
| 1DOI | 2FZF | 37676 | 134437 | 0.72 | 0.237 | 0 | 0.235 | 0 | 0.237 | 0 | 0.5 |
| 1DOI | 2YW6 | 37676 | 153794 | 0.72 | 0.247 | 0 | 0.667 | 0 | 0.234 | 0 | 0.571 |
| 1I7H | 1EUM | 71122 | 59507 | 0.72 | 0.286 | 0 | 0.5 | 0.14 | 0.286 | 0 | 0.5 |
| 1I7H | 1JIG | 71122 | 71685 | 0.72 | 0.221 | 0 | 0 | 0.037 | 0.221 | 0 | 0 |
| 1I7H | 1JI4 | 71122 | 77117 | 0.72 | 0.213 | 0 | 0 | 0.037 | 0.213 | 0 | 0 |
| 1I7H | 1LB3 | 71122 | 77873 | 0.72 | 0.209 | 0 | 0 | 0.143 | 0.209 | 0 | 0 |
| 1I7H | 1LKO | 71122 | 78063 | 0.72 | 0.219 | 0 | 0 | 0 | 0.163 | 0 | 0.245 |
| 1I7H | 1NF4 | 71122 | 85598 | 0.72 | 0.174 | 0 | 0 | 0.017 | 0.174 | 0 | 0 |
| 1I7H | 1O9R | 71122 | 86706 | 0.72 | 0.215 | 0 | 0.229 | 0.037 | 0.215 | 0 | 0.229 |
| 1I7H | 1J30 | 71122 | 90806 | 0.72 | 0.203 | 0 | 0 | 0 | 0.203 | 0 | 0 |
| 1I7H | 1NNQ | 71122 | 92006 | 0.72 | 0.193 | 0 | 0 | 0 | 0.187 | 0 | 0 |
| 1I7H | 1UMN | 71122 | 99607 | 0.72 | 0.185 | 0 | 0.714 | 0.017 | 0.185 | 0 | 0.714 |
| 1I7H | 1VJX | 71122 | 100837 | 0.72 | 0.215 | 0 | 0 | 0 | 0.215 | 0 | 0 |
| 1I7H | 1VLG | 71122 | 108816 | 0.72 | 0.21 | 0 | 0.429 | 0.017 | 0.21 | 0 | 0.429 |
| 1I7H | 1TJO | 71122 | 112461 | 0.72 | 0.189 | 0 | 0.461 | 0 | 0.189 | 0 | 0.461 |
| 1I7H | 1YUZ | 71122 | 124077 | 0.72 | 0.252 | 0 | 1 | 0 | 0.175 | 0 | 0 |
| 1I7H | 1Z6O | 71122 | 124531 | 0.72 | 0.193 | 0 | 0 | 0 | 0.172 | 0 | 0 |
| 1I7H | 1Z6O | 71122 | 124543 | 0.72 | 0.195 | 0 | 0 | 0.094 | 0.172 | 0 | 0 |
| 1I7H | 1ZS3 | 71122 | 125573 | 0.72 | 0.183 | 0 | 0.833 | 0 | 0.183 | 0 | 0.833 |
| 1I7H | 2FJC | 71122 | 133559 | 0.72 | 0.193 | 0 | 0 | 0.017 | 0.193 | 0 | 0 |
| 1I7H | 2FKZ | 71122 | 133688 | 0.72 | 0.226 | 0 | 0 | 0.053 | 0.228 | 0 | 0 |
| 1I7H | 2FZF | 71122 | 134437 | 0.72 | 0.214 | 0 | 0 | 0 | 0.214 | 0 | 0.225 |
| 1I7H | 2YW6 | 71122 | 153794 | 0.72 | 0.235 | 0 | 0.333 | 0.037 | 0.212 | 0 | 0.333 |
| 1L5P | 1EUM | 73598 | 59507 | 0.72 | 0.302 | 0 | 0.364 | 0 | 0.302 | 0 | 0.364 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1L5P | 1JIG | 73598 | 71685 | 0.72 | 0.248 | 0 | 0 | 0 | 0.248 | 0 | 0 |
| 1L5P | 1JI4 | 73598 | 77117 | 0.72 | 0.228 | 0 | 0 | 0 | 0.228 | 0 | 0 |
| 1L5P | 1LB3 | 73598 | 77873 | 0.72 | 0.199 | 0 | 0 | 0 | 0.199 | 0 | 0 |
| 1L5P | 1LKO | 73598 | 78063 | 0.72 | 0.211 | 0 | 0 | 0 | 0.192 | 0 | 0 |
| 1L5P | 1NF4 | 73598 | 85598 | 0.72 | 0.191 | 0 | 0 | 0 | 0.191 | 0 | 0 |
| 1L5P | 1O9R | 73598 | 86706 | 0.72 | 0.174 | 0 | 0.32 | 0 | 0.174 | 0 | 0.32 |
| 1L5P | 1J30 | 73598 | 90806 | 0.72 | 0.209 | 0 | 0 | 0 | 0.209 | 0 | 0 |
| 1L5P | 1NNQ | 73598 | 92006 | 0.72 | 0.195 | 0 | 0.368 | 0 | 0.146 | 0 | 0.429 |
| 1L5P | 1UMN | 73598 | 99607 | 0.72 | 0.242 | 0 | 0 | 0 | 0.242 | 0 | 0 |
| 1L5P | 1VJX | 73598 | 100837 | 0.72 | 0.215 | 0 | 0 | 0 | 0.215 | 0 | 0 |
| 1L5P | 1VLG | 73598 | 108816 | 0.72 | 0.209 | 0 | 0 | 0 | 0.209 | 0 | 0 |
| 1L5P | 1TJO | 73598 | 112461 | 0.72 | 0.189 | 0 | 0.296 | 0 | 0.189 | 0 | 0.296 |
| 1L5P | 1YUZ | 73598 | 124077 | 0.72 | 0.228 | 0 | 0 | 0 | 0.166 | 0 | 0 |
| 1L5P | 1Z6O | 73598 | 124531 | 0.72 | 0.166 | 0 | 0.391 | 0 | 0.159 | 0 | 0.391 |
| 1L5P | 1Z6O | 73598 | 124543 | 0.72 | 0.178 | 0 | 0.25 | 0 | 0.159 | 0 | 0.391 |
| 1L5P | 1ZS3 | 73598 | 125573 | 0.72 | 0.201 | 0 | 0.5 | 0 | 0.201 | 0 | 0.5 |
| 1L5P | 2FJC | 73598 | 133559 | 0.72 | 0.225 | 0 | 0.625 | 0 | 0.225 | 0 | 0.625 |
| 1L5P | 2FKZ | 73598 | 133688 | 0.72 | 0.201 | 0 | 0.222 | 0 | 0.212 | 0 | 0.222 |
| 1L5P | 2FZF | 73598 | 134437 | 0.72 | 0.208 | 0 | 0 | 0 | 0.208 | 0 | 0 |
| 1L5P | 2YW6 | 73598 | 153794 | 0.72 | 0.191 | 0 | 0 | 0 | 0.221 | 0 | 0.6 |
| 1UTE | 1DOI | 42078 | 37676 | 0.717 | 0.194 | 0 | 0 | 0 | 0.194 | 0 | 0 |
| 1UTE | 1B9R | 42078 | 37683 | 0.717 | 0.183 | 0 | 0 | 0 | 0.183 | 0 | 0 |
| 1UTE | 1I7H | 42078 | 71122 | 0.717 | 0.34 | 0 | 0 | 0.094 | 0.34 | 0 | 0.217 |
| 1UTE | 1L5P | 42078 | 73598 | 0.717 | 0.338 | 0 | 0.444 | 0 | 0.338 | 0 | 0.444 |
| 1UTE | 1WRI | 42078 | 114838 | 0.717 | 0.336 | 0 | 0.389 | 0.024 | 0.336 | 0 | 0.389 |
| 1UTE | 1XLQ | 42078 | 122119 | 0.717 | 0.435 | 0 | 0 | 0.027 | 0.435 | 0 | 0 |
| 1UTE | 2BT6 | 42078 | 129152 | 0.717 | 0.398 | 0 | 0.444 | 0.091 | 0.398 | 0 | 0.444 |
| 1WRI | 1EUM | 114838 | 59507 | 0.72 | 0.268 | 0 | 0.357 | 0.025 | 0.268 | 0 | 0.357 |
| 1WRI | 1JIG | 114838 | 71685 | 0.72 | 0.261 | 0 | 0 | 0.016 | 0.261 | 0 | 0 |
| 1WRI | 1JI4 | 114838 | 77117 | 0.72 | 0.248 | 0 | 0.266 | 0.016 | 0.248 | 0 | 0.266 |
| 1WRI | 1LB3 | 114838 | 77873 | 0.72 | 0.296 | 0 | 0.353 | 0.008 | 0.296 | 0 | 0.353 |
| 1WRI | 1LKO | 114838 | 78063 | 0.72 | 0.27 | 0 | 0.25 | 0 | 0.269 | 0 | 0.455 |
| 1WRI | 1NF4 | 114838 | 85598 | 0.72 | 0.269 | 0 | 0.667 | 0.008 | 0.269 | 0 | 0.667 |
| 1WRI | 1O9R | 114838 | 86706 | 0.72 | 0.255 | 0 | 0 | 0.016 | 0.255 | 0 | 0 |
| 1WRI | 1J30 | 114838 | 90806 | 0.72 | 0.233 | 0 | 0 | 0 | 0.233 | 0 | 0 |
| 1WRI | 1NNQ | 114838 | 92006 | 0.72 | 0.196 | 0 | 0.258 | 0 | 0.232 | 0 | 0.256 |
| 1WRI | 1UMN | 114838 | 99607 | 0.72 | 0.246 | 0 | 0.286 | 0.008 | 0.246 | 0 | 0.286 |
| 1WRI | 1VJX | 114838 | 100837 | 0.72 | 0.242 | 0 | 0 | 0 | 0.242 | 0 | 0 |
| 1WRI | 1VLG | 114838 | 108816 | 0.72 | 0.27 | 0 | 0 | 0.066 | 0.27 | 0 | 0 |
| 1WRI | 1TJO | 114838 | 112461 | 0.72 | 0.274 | 0 | 0 | 0 | 0.274 | 0 | 0 |
| 1WRI | 1YUZ | 114838 | 124077 | 0.72 | 0.22 | 0 | 0.265 | 0 | 0.193 | 0 | 0.265 |
| 1WRI | 1Z6O | 114838 | 124531 | 0.72 | 0.162 | 0 | 0 | 0 | 0.145 | 0 | 0.417 |
| 1WRI | 1Z6O | 114838 | 124543 | 0.72 | 0.176 | 0 | 0.417 | 0.044 | 0.145 | 0 | 0.417 |
| 1WRI | 1ZS3 | 114838 | 125573 | 0.72 | 0.18 | 0 | 0 | 0 | 0.18 | 0 | 0 |
| 1WRI | 2FJC | 114838 | 133559 | 0.72 | 0.192 | 0 | 0 | 0.008 | 0.192 | 0 | 0 |
| 1WRI | 2FKZ | 114838 | 133688 | 0.72 | 0.198 | 0 | 0 | 0.025 | 0.2 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1WRI | 2FZF | 114838 | 134437 | 0.72 | 0.212 | 0 | 0 | 0 | 0.212 | 0 | 0 |
| 1WRI | 2YW6 | 114838 | 153794 | 0.72 | 0.214 | 0 | 0.556 | 0.016 | 0.199 | 0 | 0.556 |
| 1XLQ | 1EUM | 122119 | 59507 | 0.72 | 0.267 | 0 | 0.3 | 0.068 | 0.267 | 0 | 0.3 |
| 1XLQ | 1JIG | 122119 | 71685 | 0.72 | 0.27 | 0 | 0 | 0.044 | 0.27 | 0 | 0 |
| 1XLQ | 1JI4 | 122119 | 77117 | 0.72 | 0.256 | 0 | 0 | 0.044 | 0.256 | 0 | 0 |
| 1XLQ | 1LB3 | 122119 | 77873 | 0.72 | 0.276 | 0 | 0 | 0.064 | 0.276 | 0 | 0 |
| 1XLQ | 1LKO | 122119 | 78063 | 0.72 | 0.265 | 0 | 0 | 0 | 0.285 | 0 | 0.75 |
| 1XLQ | 1NF4 | 122119 | 85598 | 0.72 | 0.271 | 0 | 0 | 0.021 | 0.271 | 0 | 0 |
| 1XLQ | 1O9R | 122119 | 86706 | 0.72 | 0.247 | 0 | 0 | 0.044 | 0.247 | 0 | 0 |
| 1XLQ | 1J30 | 122119 | 90806 | 0.72 | 0.273 | 0 | 0 | 0 | 0.273 | 0 | 0 |
| 1XLQ | 1NNQ | 122119 | 92006 | 0.72 | 0.191 | 0 | 0 | 0 | 0.285 | 0 | 0 |
| 1XLQ | 1UMN | 122119 | 99607 | 0.72 | 0.281 | 0 | 0 | 0.021 | 0.281 | 0 | 0 |
| 1XLQ | 1VJX | 122119 | 100837 | 0.72 | 0.291 | 0 | 0 | 0 | 0.291 | 0 | 0 |
| 1XLQ | 1VLG | 122119 | 108816 | 0.72 | 0.315 | 0 | 0 | 0.064 | 0.315 | 0 | 0 |
| 1XLQ | 1TJO | 122119 | 112461 | 0.72 | 0.266 | 0 | 0 | 0 | 0.266 | 0 | 0 |
| 1XLQ | 1YUZ | 122119 | 124077 | 0.72 | 0.255 | 0 | 0.385 | 0 | 0.192 | 0 | 0.225 |
| 1XLQ | 1Z6O | 122119 | 124531 | 0.72 | 0.171 | 0 | 0 | 0 | 0.163 | 0 | 0 |
| 1XLQ | 1Z6O | 122119 | 124543 | 0.72 | 0.211 | 0 | 0 | 0.144 | 0.163 | 0 | 0 |
| 1XLQ | 1ZS3 | 122119 | 125573 | 0.72 | 0.214 | 0 | 0 | 0 | 0.214 | 0 | 0 |
| 1XLQ | 2FJC | 122119 | 133559 | 0.72 | 0.226 | 0 | 0.333 | 0.021 | 0.226 | 0 | 0.333 |
| 1XLQ | 2FKZ | 122119 | 133688 | 0.72 | 0.224 | 0 | 0 | 0.068 | 0.235 | 0 | 0 |
| 1XLQ | 2FZF | 122119 | 134437 | 0.72 | 0.241 | 0 | 0 | 0 | 0.241 | 0 | 0 |
| 1XLQ | 2YW6 | 122119 | 153794 | 0.72 | 0.237 | 0 | 0 | 0.044 | 0.224 | 0 | 0 |
| 2BT6 | 1EUM | 129152 | 59507 | 0.72 | 0.278 | 0 | 0.75 | 0.169 | 0.278 | 0 | 0.75 |
| 2BT6 | 1JIG | 129152 | 71685 | 0.72 | 0.257 | 0 | 0.5 | 0.108 | 0.257 | 0 | 0.5 |
| 2BT6 | 1JI4 | 129152 | 77117 | 0.72 | 0.31 | 0 | 0 | 0.108 | 0.31 | 0 | 0.625 |
| 2BT6 | 1LB3 | 129152 | 77873 | 0.72 | 0.274 | 0 | 0.191 | 0.214 | 0.274 | 0 | 0.191 |
| 2BT6 | 1LKO | 129152 | 78063 | 0.72 | 0.279 | 0 | 0 | 0 | 0.31 | 0 | 0 |
| 2BT6 | 1NF4 | 129152 | 85598 | 0.72 | 0.293 | 0 | 0 | 0.052 | 0.293 | 0 | 0 |
| 2BT6 | 1O9R | 129152 | 86706 | 0.72 | 0.297 | 0 | 0.5 | 0.108 | 0.297 | 0 | 0.5 |
| 2BT6 | 1J30 | 129152 | 90806 | 0.72 | 0.22 | 0 | 0 | 0 | 0.22 | 0 | 0 |
| 2BT6 | 1NNQ | 129152 | 92006 | 0.72 | 0.256 | 0 | 0 | 0 | 0.286 | 0 | 0.353 |
| 2BT6 | 1UMN | 129152 | 99607 | 0.72 | 0.289 | 0 | 0.262 | 0.052 | 0.289 | 0 | 0.262 |
| 2BT6 | 1VJX | 129152 | 100837 | 0.72 | 0.304 | 0 | 0 | 0 | 0.304 | 0 | 0 |
| 2BT6 | 1VLG | 129152 | 108816 | 0.72 | 0.337 | 0 | 0.246 | 0.073 | 0.337 | 0 | 0.246 |
| 2BT6 | 1TJO | 129152 | 112461 | 0.72 | 0.25 | 0 | 0.375 | 0 | 0.25 | 0 | 0.461 |
| 2BT6 | 1YUZ | 129152 | 124077 | 0.72 | 0.263 | 0 | 0 | 0 | 0.321 | 0 | 0.286 |
| 2BT6 | 1Z6O | 129152 | 124531 | 0.72 | 0.304 | 0 | 0.5 | 0 | 0.299 | 0 | 0.5 |
| 2BT6 | 1Z6O | 129152 | 124543 | 0.72 | 0.306 | 0 | 0.357 | 0.311 | 0.299 | 0 | 0.5 |
| 2BT6 | 1ZS3 | 129152 | 125573 | 0.72 | 0.251 | 0 | 0 | 0 | 0.251 | 0 | 0 |
| 2BT6 | 2FJC | 129152 | 133559 | 0.72 | 0.23 | 0 | 0 | 0.052 | 0.23 | 0 | 0 |
| 2BT6 | 2FKZ | 129152 | 133688 | 0.72 | 0.186 | 0 | 0 | 0.169 | 0.195 | 0 | 0 |
| 2BT6 | 2FZF | 129152 | 134437 | 0.72 | 0.232 | 0 | 0.235 | 0 | 0.232 | 0 | 0.235 |
| 2BT6 | 2YW6 | 129152 | 153794 | 0.72 | 0.244 | 0 | 0.4 | 0.108 | 0.196 | 0 | 0.244 |
| 2QFR | 1DOI | 150751 | 37676 | 0.717 | 0.208 | 0 | 0.625 | 0 | 0.149 | 0 | 0.625 |
| 2QFR | 1B9R | 150751 | 37683 | 0.717 | 0.18 | 0 | 0 | 0 | 0.127 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2QFR | 1I7H | 150751 | 71122 | 0.717 | 0.176 | 0 | 0 | 0.017 | 0.139 | 0 | 0 |
| 2QFR | 1L5P | 150751 | 73598 | 0.717 | 0.165 | 0 | 0 | 0 | 0.137 | 0 | 0 |
| 2QFR | 1WRI | 150751 | 114838 | 0.717 | 0.147 | 0 | 0 | 0.186 | 0.117 | 0 | 0.444 |
| 2QFR | 1XLQ | 150751 | 122119 | 0.717 | 0.186 | 0 | 0.312 | 0.021 | 0.143 | 0 | 0.4 |
| 2QFR | 2BT6 | 150751 | 129152 | 0.717 | 0.179 | 0 | 0 | 0.052 | 0.146 | 0 | 0 |

From BSM score with 70% threshold, 2PWA, 1GCI, 1WMD, 1R6V, 1V6C, and 2ID4 are associated with FA:52744 which is  Subtilases family and all others are associated with FA:50514 which is Eukaryotic proteases family

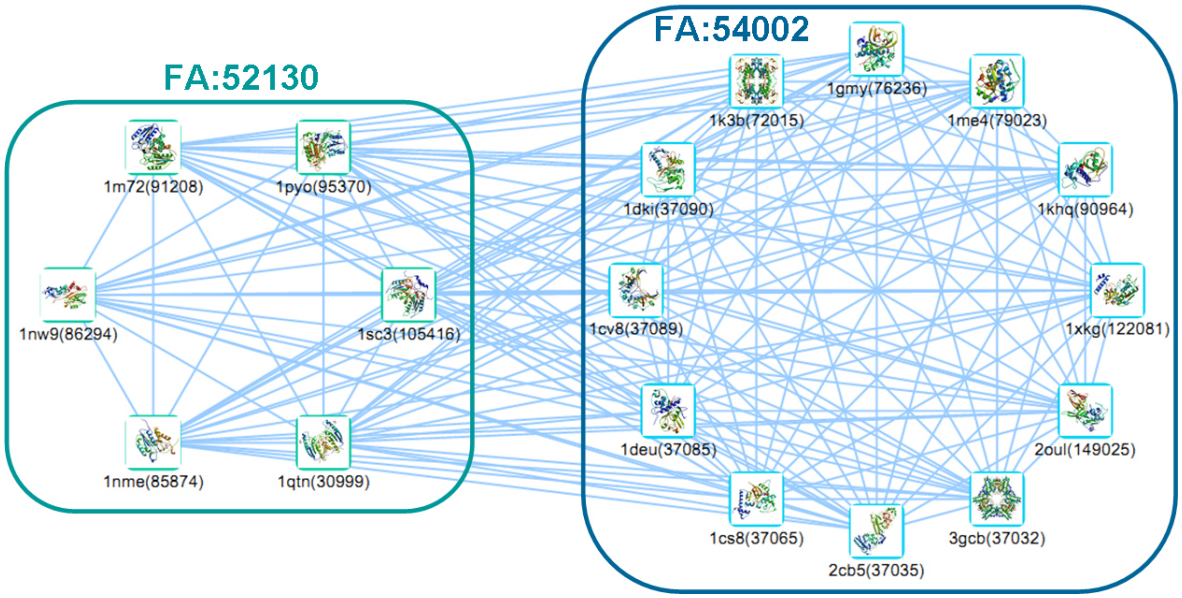



| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1A7S | 1WMD | 26284 | 109409 | 0.72 | 0.248 | 0 | 0.232 | 0.055 | 0.189 | 0 | 0.333 |
| 1A7S | 1R6V | 26284 | 111710 | 0.72 | 0.141 | 0.015 | 0.294 | 0.026 | 0.141 | 0.015 | 0.294 |
| 1A7S | 1V6C | 26284 | 113544 | 0.72 | 0.178 | 0 | 0.375 | 0 | 0.178 | 0 | 0.375 |
| 1A7S | 2ID4 | 26284 | 137266 | 0.72 | 0.228 | 0 | 0.29 | 0.091 | 0.166 | 0 | 0.29 |
| 1A7S | 2PWA | 26284 | 139762 | 0.72 | 0.265 | 0 | 0.857 | 0.055 | 0.265 | 0 | 0.857 |
| 1A7S | 1GCI | 26284 | 32503 | 0.72 | 0.261 | 0 | 0.3 | 0.026 | 0.261 | 0 | 0.3 |
| 1AUT | 1WMD | 26348 | 109409 | 0.72 | 0.266 | 0 | 0 | 0.044 | 0.202 | 0 | 0.256 |
| 1AUT | 1R6V | 26348 | 111710 | 0.72 | 0.147 | 0 | 0.308 | 0.021 | 0.147 | 0 | 0.308 |
| 1AUT | 1V6C | 26348 | 113544 | 0.72 | 0.199 | 0 | 0.857 | 0 | 0.199 | 0 | 0.857 |
| 1AUT | 2ID4 | 26348 | 137266 | 0.72 | 0.249 | 0 | 0.571 | 0.163 | 0.192 | 0 | 0.278 |
| 1AUT | 2PWA | 26348 | 139762 | 0.72 | 0.276 | 0 | 0.625 | 0.044 | 0.276 | 0 | 0.625 |
| 1AUT | 1GCI | 26348 | 32503 | 0.72 | 0.279 | 0 | 0.296 | 0.021 | 0.279 | 0 | 0.296 |
| 1AZZ | 1WMD | 26072 | 109409 | 0.72 | 0.249 | 0 | 0.344 | 0.066 | 0.201 | 0 | 0.344 |
| 1AZZ | 1R6V | 26072 | 111710 | 0.72 | 0.147 | 0.021 | 0.4 | 0.032 | 0.147 | 0.021 | 0.4 |
| 1AZZ | 2PWA | 26072 | 139762 | 0.72 | 0.275 | 0 | 0 | 0.066 | 0.275 | 0 | 0 |
| 1AZZ | 1V6C | 26072 | 113544 | 0.72 | 0.188 | 0 | 0.255 | 0 | 0.188 | 0 | 0.255 |
| 1AZZ | 1GCI | 26072 | 32503 | 0.72 | 0.259 | 0 | 0.6 | 0.032 | 0.259 | 0 | 0.6 |
| 1AZZ | 2ID4 | 26072 | 137266 | 0.72 | 0.208 | 0 | 0.4 | 0.162 | 0.182 | 0 | 0.417 |
| 1BIO | 1WMD | 26314 | 109409 | 0.72 | 0.255 | 0 | 0.312 | 0.044 | 0.196 | 0 | 0.312 |
| 1BIO | 1R6V | 26314 | 111710 | 0.72 | 0.136 | 0 | 0.222 | 0.021 | 0.136 | 0 | 0.222 |
| 1BIO | 1V6C | 26314 | 113544 | 0.72 | 0.195 | 0 | 0 | 0 | 0.195 | 0 | 0 |
| 1BIO | 2ID4 | 26314 | 137266 | 0.72 | 0.239 | 0 | 0 | 0.163 | 0.173 | 0 | 0 |
| 1BIO | 2PWA | 26314 | 139762 | 0.72 | 0.272 | 0 | 0.222 | 0.044 | 0.272 | 0 | 0.222 |
| 1BIO | 1GCI | 26314 | 32503 | 0.72 | 0.281 | 0 | 0 | 0.021 | 0.281 | 0 | 0 |
| 1EAX | 1WMD | 64889 | 109409 | 0.72 | 0.27 | 0 | 0 | 0.034 | 0.222 | 0 | 0 |
| 1EAX | 1R6V | 64889 | 111710 | 0.72 | 0.143 | 0.013 | 0 | 0.017 | 0.143 | 0.013 | 0 |
| 1EAX | 1V6C | 64889 | 113544 | 0.72 | 0.229 | 0 | 0.5 | 0 | 0.229 | 0 | 0.375 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1EAX | 2ID4 | 64889 | 137266 | 0.72 | 0.249 | 0 | 0 | 0.322 | 0.193 | 0 | 0.461 |
| 1EAX | 2PWA | 64889 | 139762 | 0.72 | 0.269 | 0 | 0.308 | 0.034 | 0.269 | 0 | 0.308 |
| 1ELV | 1WMD | 59454 | 109409 | 0.72 | 0.239 | 0 | 0.387 | 0.055 | 0.204 | 0 | 0.387 |
| 1ELV | 1R6V | 59454 | 111710 | 0.72 | 0.15 | 0.001 | 0.286 | 0.026 | 0.18 | 0.015 | 0.5 |
| 1ELV | 1V6C | 59454 | 113544 | 0.72 | 0.194 | 0 | 0 | 0 | 0.226 | 0 | 0.194 |
| 1ELV | 2ID4 | 59454 | 137266 | 0.72 | 0.23 | 0 | 1 | 0.091 | 0.2 | 0 | 0.184 |
| 1ELV | 2PWA | 59454 | 139762 | 0.72 | 0.282 | 0 | 0.4 | 0.055 | 0.28 | 0 | 0.5 |
| 1EQ9 | 1WMD | 26068 | 109409 | 0.72 | 0.236 | 0 | 0.312 | 0 | 0.202 | 0 | 0.333 |
| 1EQ9 | 1R6V | 26068 | 111710 | 0.72 | 0.143 | 0.004 | 0.385 | 0 | 0.143 | 0.004 | 0.385 |
| 1EQ9 | 1V6C | 26068 | 113544 | 0.72 | 0.197 | 0 | 0.286 | 0 | 0.197 | 0 | 0.286 |
| 1EQ9 | 2ID4 | 26068 | 137266 | 0.72 | 0.228 | 0 | 0.389 | 0 | 0.168 | 0 | 0.389 |
| 1EQ9 | 2PWA | 26068 | 139762 | 0.72 | 0.276 | 0 | 0.45 | 0 | 0.276 | 0 | 0.45 |
| 1EQ9 | 1GCI | 26068 | 32503 | 0.72 | 0.265 | 0 | 0.429 | 0 | 0.265 | 0 | 0.429 |
| 1FIW | 1WMD | 26378 | 109409 | 0.72 | 0.262 | 0 | 0.333 | 0.044 | 0.21 | 0 | 0.231 |
| 1FIW | 1R6V | 26378 | 111710 | 0.72 | 0.137 | 0.02 | 0.438 | 0.021 | 0.148 | 0.02 | 0.438 |
| 1FIW | 1V6C | 26378 | 113544 | 0.72 | 0.223 | 0 | 0.385 | 0 | 0.208 | 0 | 0.385 |
| 1FIW | 2ID4 | 26378 | 137266 | 0.72 | 0.249 | 0 | 0.32 | 0.163 | 0.188 | 0 | 0.32 |
| 1FIW | 2PWA | 26378 | 139762 | 0.72 | 0.289 | 0 | 0.269 | 0.044 | 0.283 | 0 | 0.269 |
| 1FIW | 1GCI | 26378 | 32503 | 0.72 | 0.288 | 0 | 0 | 0.021 | 0.283 | 0 | 0 |
| 1GCI | 1OP0 | 32503 | 104019 | 0.72 | 0.285 | 0 | 0.529 | 0.026 | 0.285 | 0 | 0.529 |
| 1GCI | 1Q3X | 32503 | 104526 | 0.72 | 0.306 | 0 | 0.4 | 0.021 | 0.25 | 0 | 0.375 |
| 1GCI | 1RRK | 32503 | 111924 | 0.72 | 0.271 | 0 | 0.3 | 0.017 | 0.23 | 0.049 | 0.455 |
| 1GCI | 1SI5 | 32503 | 112091 | 0.72 | 0.307 | 0 | 0.353 | 0.021 | 0.307 | 0 | 0.353 |
| 1GCI | 1T32 | 32503 | 119136 | 0.72 | 0.29 | 0 | 0 | 0.008 | 0.29 | 0 | 0 |
| 1GCI | 2BZ6 | 32503 | 129552 | 0.72 | 0.296 | 0 | 0 | 0.026 | 0.296 | 0 | 0 |
| 1GCI | 2P3U | 32503 | 149189 | 0.72 | 0.274 | 0 | 0 | 0.026 | 0.268 | 0 | 0 |
| 1GCI | 1GVK | 32503 | 70610 | 0.72 | 0.301 | 0 | 0.4 | 0.026 | 1 | 0 | 0.4 |
| 1GCI | 1GVZ | 32503 | 76360 | 0.72 | 0.293 | 0 | 0 | 0.017 | 0.293 | 0 | 0 |
| 1GCI | 1LTO | 32503 | 84707 | 0.72 | 0.3 | 0 | 0.429 | 0.008 | 0.3 | 0 | 0.429 |
| 1GCI | 1H8D | 32503 | 60737 | 0.72 | 0.279 | 0 | 0 | 0.026 | 0.298 | 0 | 0 |
| 1GCI | 1EAX | 32503 | 64889 | 0.72 | 0.305 | 0 | 0.312 | 0.017 | 0.305 | 0 | 0.312 |
| 1GCI | 1ELV | 32503 | 59454 | 0.72 | 0.275 | 0 | 0.625 | 0.026 | 0.251 | 0 | 0.318 |
| 1GCI | 1M9U | 32503 | 74601 | 0.72 | 0.291 | 0 | 0.333 | 0.021 | 0.291 | 0 | 0.333 |
| 1GCI | 1GJ7 | 32503 | 70180 | 0.72 | 0.292 | 0 | 0.286 | 0.026 | 0.351 | 0 | 0.286 |
| 1GCI | 1HJ9 | 32503 | 65845 | 0.72 | 0.284 | 0 | 0.522 | 0.026 | 0.284 | 0 | 0.522 |
| 1GDN | 1WMD | 26011 | 109409 | 0.72 | 0.233 | 0 | 0.35 | 0 | 0.199 | 0 | 0.35 |
| 1GDN | 1R6V | 26011 | 111710 | 0.72 | 0.143 | 0.006 | 0.333 | 0 | 0.143 | 0.006 | 0.333 |
| 1GDN | 1V6C | 26011 | 113544 | 0.72 | 0.199 | 0 | 0.36 | 0 | 0.199 | 0 | 0.36 |
| 1GDN | 2ID4 | 26011 | 137266 | 0.72 | 0.23 | 0 | 0 | 0 | 0.179 | 0 | 0.385 |
| 1GDN | 2PWA | 26011 | 139762 | 0.72 | 0.274 | 0 | 0.4 | 0 | 0.274 | 0 | 0.4 |
| 1GDN | 1GCI | 26011 | 32503 | 0.72 | 0.268 | 0 | 0.345 | 0 | 0.268 | 0 | 0.345 |
| 1GJ7 | 1WMD | 70180 | 109409 | 0.72 | 0.253 | 0 | 0.205 | 0.055 | 0.019 | 0 | 0.205 |
| 1GJ7 | 1R6V | 70180 | 111710 | 0.72 | 0.144 | 0 | 0.257 | 0.026 | 0.013 | 0 | 0.375 |
| 1GJ7 | 1V6C | 70180 | 113544 | 0.72 | 0.215 | 0 | 0.556 | 0 | 0.011 | 0 | 0.556 |
| 1GJ7 | 2ID4 | 70180 | 137266 | 0.72 | 0.246 | 0 | 0.8 | 0.091 | 0.019 | 0 | 0.556 |
| 1GJ7 | 2PWA | 70180 | 139762 | 0.72 | 0.283 | 0 | 0.294 | 0.055 | 0.028 | 0 | 0.294 |
| 1GVK | 1WMD | 70610 | 109409 | 0.72 | 0.269 | 0 | 0.188 | 0.055 | 0.007 | 0 | 0.333 |
| 1GVK | 1R6V | 70610 | 111710 | 0.72 | 0.147 | 0.018 | 0.371 | 0.026 | 0.004 | 0.018 | 0.371 |
| 1GVK | 1V6C | 70610 | 113544 | 0.72 | 0.2 | 0 | 0.4 | 0 | 0.007 | 0 | 0.4 |
| 1GVK | 2ID4 | 70610 | 137266 | 0.72 | 0.244 | 0 | 0.286 | 0.091 | 0.006 | 0 | 0.333 |
| 1GVK | 2PWA | 70610 | 139762 | 0.72 | 0.28 | 0 | 0.273 | 0.055 | 0.011 | 0 | 0.273 |
| 1GVZ | 1WMD | 76360 | 109409 | 0.72 | 0.237 | 0 | 0 | 0.034 | 0.203 | 0 | 0.571 |
| 1GVZ | 1R6V | 76360 | 111710 | 0.72 | 0.149 | 0.008 | 0.571 | 0.017 | 0.149 | 0.008 | 0.571 |
| 1GVZ | 1V6C | 76360 | 113544 | 0.72 | 0.195 | 0 | 0.4 | 0 | 0.195 | 0 | 0.4 |
| 1GVZ | 2ID4 | 76360 | 137266 | 0.72 | 0.237 | 0 | 0.667 | 0.092 | 0.176 | 0 | 0.461 |
| 1GVZ | 2PWA | 76360 | 139762 | 0.72 | 0.258 | 0 | 0.296 | 0.034 | 0.258 | 0 | 0.296 |
| 1H8D | 1WMD | 60737 | 109409 | 0.72 | 0.259 | 0 | 0.412 | 0.055 | 0.223 | 0 | 0.444 |
| 1H8D | 1R6V | 60737 | 111710 | 0.72 | 0.151 | 0.013 | 0.243 | 0.026 | 0.148 | 0.013 | 0.243 |
| 1H8D | 1V6C | 60737 | 113544 | 0.72 | 0.209 | 0 | 0.257 | 0 | 0.205 | 0 | 0.257 |
| 1H8D | 2ID4 | 60737 | 137266 | 0.72 | 0.274 | 0 | 0.273 | 0.091 | 0.192 | 0 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1H8D | 2PWA | 60737 | 139762 | 0.72 | 0.282 | 0 | 0.455 | 0.055 | 0.286 | 0 | 0.375 |
| 1HJ9 | 1WMD | 65845 | 109409 | 0.72 | 0.245 | 0 | 0.467 | 0.055 | 0.181 | 0 | 0.258 |
| 1HJ9 | 1R6V | 65845 | 111710 | 0.72 | 0.134 | 0.019 | 0.35 | 0.026 | 0.134 | 0.019 | 0.35 |
| 1HJ9 | 1V6C | 65845 | 113544 | 0.72 | 0.187 | 0 | 0.295 | 0 | 0.187 | 0 | 0.295 |
| 1HJ9 | 2ID4 | 65845 | 137266 | 0.72 | 0.228 | 0 | 0.344 | 0.091 | 0.176 | 0 | 0.344 |
| 1HJ9 | 2PWA | 65845 | 139762 | 0.72 | 0.269 | 0 | 0.333 | 0.055 | 0.269 | 0 | 0.333 |
| 1LTO | 1WMD | 84707 | 109409 | 0.72 | 0.24 | 0 | 0.444 | 0.016 | 0.194 | 0 | 0.245 |
| 1LTO | 1R6V | 84707 | 111710 | 0.72 | 0.149 | 0.014 | 0.352 | 0.008 | 0.149 | 0.014 | 0.352 |
| 1LTO | 1V6C | 84707 | 113544 | 0.72 | 0.212 | 0 | 0.545 | 0 | 0.212 | 0 | 0.545 |
| 1LTO | 2ID4 | 84707 | 137266 | 0.72 | 0.242 | 0 | 0.333 | 0.144 | 0.184 | 0 | 0.333 |
| 1LTO | 2PWA | 84707 | 139762 | 0.72 | 0.281 | 0 | 0.4 | 0.016 | 0.281 | 0 | 0.4 |
| 1M9U | 1WMD | 74601 | 109409 | 0.72 | 0.251 | 0 | 1 | 0.044 | 0.193 | 0 | 0.333 |
| 1M9U | 1R6V | 74601 | 111710 | 0.72 | 0.143 | 0 | 0.282 | 0.021 | 0.143 | 0 | 0.282 |
| 1M9U | 1V6C | 74601 | 113544 | 0.72 | 0.226 | 0 | 0.262 | 0 | 0.226 | 0 | 0.262 |
| 1M9U | 2ID4 | 74601 | 137266 | 0.72 | 0.241 | 0 | 0.302 | 0.163 | 0.186 | 0 | 0.302 |
| 1M9U | 2PWA | 74601 | 139762 | 0.72 | 0.283 | 0 | 0.8 | 0.044 | 0.283 | 0 | 0.8 |
| 1OP0 | 1WMD | 104019 | 109409 | 0.72 | 0.274 | 0 | 0.333 | 0.055 | 0.215 | 0 | 0.333 |
| 1OP0 | 1V6C | 104019 | 113544 | 0.72 | 0.188 | 0 | 0.556 | 0 | 0.188 | 0 | 0.556 |
| 1OP0 | 2ID4 | 104019 | 137266 | 0.72 | 0.209 | 0 | 0.316 | 0.091 | 0.176 | 0 | 0.316 |
| 1OP0 | 1R6V | 104019 | 111710 | 0.72 | 0.145 | 0 | 0.5 | 0.026 | 0.145 | 0 | 0.5 |
| 1OP0 | 2PWA | 104019 | 139762 | 0.72 | 0.277 | 0 | 0.444 | 0.055 | 0.277 | 0 | 0.444 |
| 1Q3X | 2ID4 | 104526 | 137266 | 0.72 | 0.239 | 0 | 0.289 | 0.04 | 0.236 | 0 | 0.289 |
| 1Q3X | 1WMD | 104526 | 109409 | 0.72 | 0.254 | 0 | 0.333 | 0.044 | 0.208 | 0 | 0.333 |
| 1Q3X | 1V6C | 104526 | 113544 | 0.72 | 0.224 | 0 | 0.241 | 0 | 0.225 | 0 | 0.241 |
| 1Q3X | 1R6V | 104526 | 111710 | 0.72 | 0.147 | 0.005 | 0.5 | 0.021 | 0.151 | 0.003 | 0.5 |
| 1Q3X | 2PWA | 104526 | 139762 | 0.72 | 0.277 | 0 | 0.222 | 0.044 | 0.279 | 0 | 0.625 |
| 1R6V | 2BZ6 | 111710 | 129552 | 0.72 | 0.289 | 0.003 | 0.545 | 0.026 | 0.289 | 0.003 | 0.545 |
| 1R6V | 1T32 | 111710 | 119136 | 0.72 | 0.314 | 0.008 | 0.4 | 0.008 | 0.314 | 0.008 | 0.4 |
| 1R6V | 2P3U | 111710 | 149189 | 0.72 | 0.306 | 0 | 0.267 | 0.026 | 0.267 | 0 | 0.267 |
| 1R6V | 1RRK | 111710 | 111924 | 0.72 | 0.31 | 0 | 0.444 | 0.017 | 0.294 | 0.045 | 0.538 |
| 1R6V | 1SI5 | 111710 | 112091 | 0.72 | 0.308 | 0.009 | 0.278 | 0.021 | 0.308 | 0.009 | 0.278 |
| 1RRK | 2PWA | 111924 | 139762 | 0.72 | 0.266 | 0 | 0 | 0.034 | 0.382 | 0.05 | 0 |
| 1RRK | 1V6C | 111924 | 113544 | 0.72 | 0.23 | 0 | 0.4 | 0 | 0.279 | 0.05 | 0.4 |
| 1RRK | 2ID4 | 111924 | 137266 | 0.72 | 0.239 | 0 | 0.8 | 0.092 | 0.251 | 0.043 | 0.667 |
| 1SI5 | 2PWA | 112091 | 139762 | 0.72 | 0.274 | 0 | 0 | 0.044 | 0.274 | 0 | 0 |
| 1SI5 | 2ID4 | 112091 | 137266 | 0.72 | 0.231 | 0 | 0.242 | 0.163 | 0.152 | 0 | 0.242 |
| 1SI5 | 1V6C | 112091 | 113544 | 0.72 | 0.189 | 0 | 0.318 | 0 | 0.189 | 0 | 0.318 |
| 1T32 | 2PWA | 119136 | 139762 | 0.72 | 0.273 | 0 | 0.389 | 0.016 | 0.273 | 0 | 0.389 |
| 1T32 | 2ID4 | 119136 | 137266 | 0.72 | 0.227 | 0 | 0.833 | 0.144 | 0.166 | 0 | 0.389 |
| 1V6C | 2BZ6 | 113544 | 129552 | 0.72 | 0.31 | 0 | 0 | 0 | 0.31 | 0 | 0.178 |
| 1V6C | 2P3U | 113544 | 149189 | 0.72 | 0.306 | 0 | 0.556 | 0 | 0.318 | 0 | 0.312 |
| 1V6C | 1T32 | 113544 | 119136 | 0.72 | 0.322 | 0 | 0.333 | 0 | 0.322 | 0 | 0.333 |
| 1WMD | 1SI5 | 109409 | 112091 | 0.72 | 0.3 | 0 | 0 | 0.044 | 0.316 | 0 | 0 |
| 1WMD | 2BZ6 | 109409 | 129552 | 0.72 | 0.271 | 0 | 0 | 0.055 | 0.298 | 0 | 0.389 |
| 1WMD | 2P3U | 109409 | 149189 | 0.72 | 0.305 | 0 | 0.533 | 0.055 | 0.309 | 0 | 0.5 |
| 1WMD | 1RRK | 109409 | 111924 | 0.72 | 0.286 | 0 | 0 | 0.034 | 0.237 | 0.033 | 0 |
| 1WMD | 1T32 | 109409 | 119136 | 0.72 | 0.287 | 0 | 0.26 | 0.016 | 0.303 | 0 | 0.26 |
| 2BZ6 | 2PWA | 129552 | 139762 | 0.72 | 0.283 | 0 | 0.353 | 0.055 | 0.283 | 0 | 0.353 |
| 2BZ6 | 2ID4 | 129552 | 137266 | 0.72 | 0.231 | 0 | 0.227 | 0.091 | 0.186 | 0 | 0.227 |
| 2HLC | 1GCI | 26074 | 32503 | 0.72 | 0.274 | 0 | 0.538 | 0.004 | 0.274 | 0 | 0.538 |
| 2HLC | 2ID4 | 26074 | 137266 | 0.72 | 0.237 | 0 | 0.438 | 0.069 | 0.182 | 0.004 | 0.438 |
| 2HLC | 2PWA | 26074 | 139762 | 0.72 | 0.275 | 0 | 0 | 0.008 | 0.275 | 0 | 0 |
| 2HLC | 1WMD | 26074 | 109409 | 0.72 | 0.241 | 0 | 0.282 | 0.008 | 0.192 | 0 | 0.282 |
| 2HLC | 1R6V | 26074 | 111710 | 0.72 | 0.145 | 0.019 | 0.288 | 0.004 | 0.145 | 0.019 | 0.288 |
| 2HLC | 1V6C | 26074 | 113544 | 0.72 | 0.209 | 0 | 0 | 0 | 0.209 | 0 | 0 |
| 2ID4 | 2P3U | 137266 | 149189 | 0.72 | 0.316 | 0 | 0.333 | 0.091 | 0.369 | 0 | 0.667 |
| 2PWA | 2P3U | 139762 | 149189 | 0.72 | 0.308 | 0 | 0 | 0.055 | 0.299 | 0 | 0 |

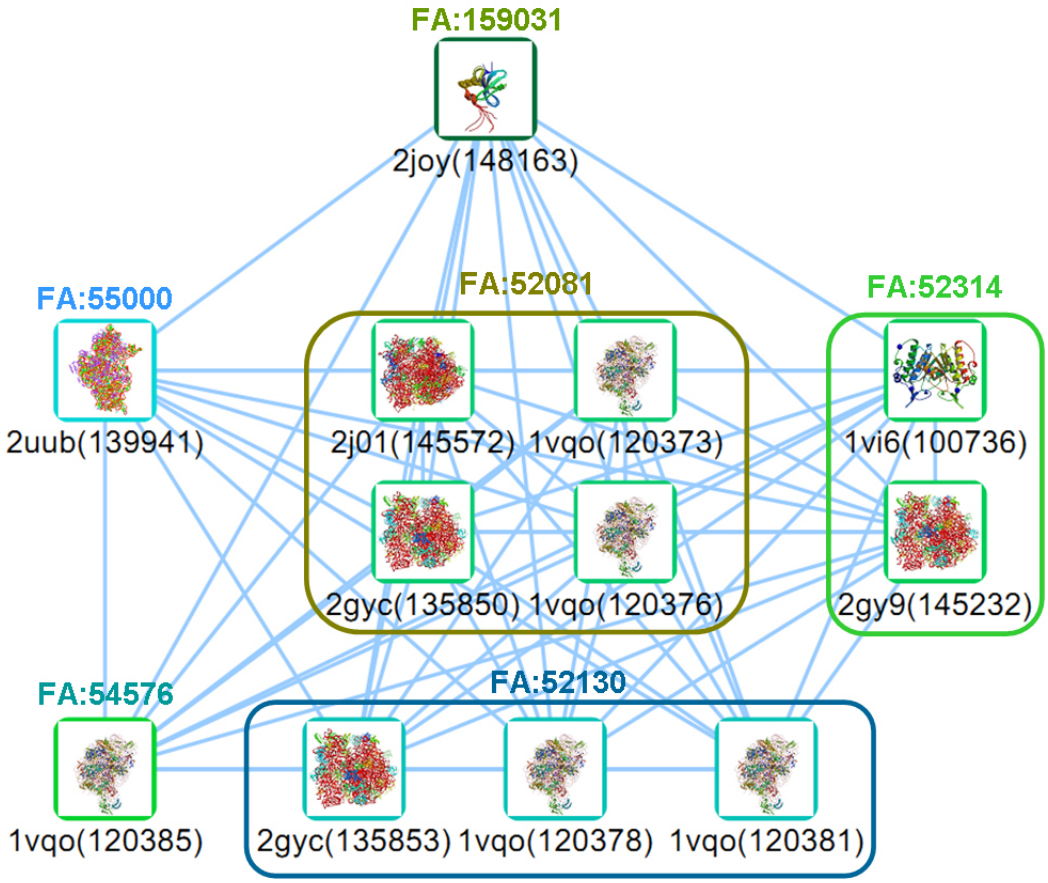
From BSM score with 70% threshold, 1M72, 1PYO, 1SC3, 1QTN, 1NME, and 1NW9 are associated with FA:52130 which is  Caspase catalytic domain family and all others are associated with FA:54002 which is Papain-like family




| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1CS8 | 1NME | 37065 | 85874 | 0.731 | 0.261 | 0 | 0.375 | 0.296 | 0.311 | 0 | 0.375 |
| 1CS8 | 1SC3 | 37065 | 105416 | 0.731 | 0.299 | 0 | 0.28 | 0.052 | 0.296 | 0 | 0.28 |
| 1CS8 | 1NW9 | 37065 | 86294 | 0.731 | 0.31 | 0 | 0.294 | 0.396 | 0.307 | 0 | 0.294 |
| 1CS8 | 1M72 | 37065 | 91208 | 0.731 | 0.266 | 0 | 0.25 | 0.296 | 0.266 | 0 | 0.25 |
| 1CS8 | 1PYO | 37065 | 95370 | 0.731 | 0.295 | 0 | 0.296 | 0.214 | 0.263 | 0 | 0.296 |
| 1CV8 | 1NME | 37089 | 85874 | 0.731 | 0.277 | 0 | 0.625 | 0 | 0.274 | 0 | 0.5 |
| 1CV8 | 1SC3 | 37089 | 105416 | 0.731 | 0.264 | 0.001 | 0.438 | 0 | 0.273 | 0.001 | 0.438 |
| 1CV8 | 1PYO | 37089 | 95370 | 0.731 | 0.279 | 0 | 0.243 | 0 | 0.291 | 0 | 0.243 |
| 1CV8 | 1M72 | 37089 | 91208 | 0.731 | 0.266 | 0.001 | 0.375 | 0 | 0.266 | 0.001 | 0.375 |
| 1CV8 | 1NW9 | 37089 | 86294 | 0.731 | 0.273 | 0.001 | 0.333 | 0 | 0.351 | 0.001 | 0.333 |
| 1DEU | 1NME | 37085 | 85874 | 0.731 | 0.273 | 0 | 0.133 | 0.322 | 0.308 | 0 | 0.231 |
| 1DEU | 1SC3 | 37085 | 105416 | 0.731 | 0.291 | 0 | 0.278 | 0.124 | 0.273 | 0 | 0.267 |
| 1DEU | 1M72 | 37085 | 91208 | 0.731 | 0.275 | 0 | 0.444 | 0.322 | 0.275 | 0 | 0.444 |
| 1DEU | 1NW9 | 37085 | 86294 | 0.731 | 0.315 | 0 | 0 | 0.446 | 0.372 | 0 | 0 |
| 1DEU | 1PYO | 37085 | 95370 | 0.731 | 0.28 | 0 | 0.179 | 0.377 | 0.275 | 0 | 0.179 |
| 1DKI | 1NME | 37090 | 85874 | 0.731 | 0.294 | 0 | 0.22 | 0 | 0.285 | 0 | 0.22 |
| 1DKI | 1SC3 | 37090 | 105416 | 0.731 | 0.293 | 0 | 0.345 | 0 | 0.302 | 0 | 0.429 |
| 1DKI | 1PYO | 37090 | 95370 | 0.731 | 0.267 | 0 | 0 | 0 | 0.298 | 0 | 0 |
| 1DKI | 1M72 | 37090 | 91208 | 0.731 | 0.276 | 0 | 0.259 | 0 | 0.276 | 0 | 0.259 |
| 1DKI | 1NW9 | 37090 | 86294 | 0.731 | 0.299 | 0 | 0 | 0 | 0.402 | 0 | 0.545 |
| 1GMY | 1NME | 76236 | 85874 | 0.731 | 0.257 | 0 | 0.8 | 0.505 | 0.285 | 0 | 0.8 |
| 1GMY | 1SC3 | 76236 | 105416 | 0.731 | 0.29 | 0 | 0.286 | 0.069 | 0.277 | 0 | 0.389 |
| 1GMY | 1NW9 | 76236 | 86294 | 0.731 | 0.317 | 0 | 0.294 | 0.557 | 0.35 | 0 | 0.75 |
| 1GMY | 1M72 | 76236 | 91208 | 0.731 | 0.249 | 0 | 0 | 0.505 | 0.249 | 0 | 0 |
| 1GMY | 1PYO | 76236 | 95370 | 0.731 | 0.283 | 0 | 0.375 | 0.717 | 0.255 | 0 | 0.4 |
| 1K3B | 1NME | 72015 | 85874 | 0.731 | 0.245 | 0 | 0 | 0.557 | 0.242 | 0 | 0.364 |
| 1K3B | 1SC3 | 72015 | 105416 | 0.731 | 0.281 | 0 | 0.271 | 0.091 | 0.2 | 0 | 0.265 |
| 1K3B | 1M72 | 72015 | 91208 | 0.731 | 0.264 | 0 | 0.4 | 0.557 | 0.185 | 0 | 0.333 |
| 1K3B | 1NW9 | 72015 | 86294 | 0.731 | 0.272 | 0 | 0.571 | 0.426 | 0.316 | 0 | 0.333 |
| 1K3B | 1PYO | 72015 | 95370 | 0.731 | 0.272 | 0 | 0.282 | 0.674 | 0.218 | 0 | 0.222 |
| 1KHQ | 1SC3 | 90964 | 105416 | 0.731 | 0.243 | 0 | 0.389 | 0.06 | 0.251 | 0 | 0.389 |
| 1KHQ | 1PYO | 90964 | 95370 | 0.731 | 0.28 | 0 | 0 | 0.112 | 0.285 | 0 | 0.24 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1KHQ | 1M72 | 90964 | 91208 | 0.731 | 0.253 | 0 | 0.5 | 0.181 | 0.253 | 0 | 0.5 |
| 1M72 | 1XKG | 91208 | 122081 | 0.731 | 0.259 | 0 | 0 | 0 | 0.209 | 0 | 0 |
| 1M72 | 2OUL | 91208 | 149025 | 0.731 | 0.299 | 0 | 0.267 | 0 | 0.299 | 0 | 0.267 |
| 1ME4 | 1NME | 79023 | 85874 | 0.731 | 0.278 | 0 | 0.5 | 0.037 | 0.275 | 0 | 0.5 |
| 1ME4 | 1SC3 | 79023 | 105416 | 0.731 | 0.249 | 0 | 0 | 0.149 | 0.25 | 0 | 0 |
| 1ME4 | 1PYO | 79023 | 95370 | 0.731 | 0.273 | 0 | 0.5 | 0 | 0.274 | 0 | 0.5 |
| 1ME4 | 1NW9 | 79023 | 86294 | 0.731 | 0.289 | 0 | 0.429 | 0.001 | 0.379 | 0 | 0.256 |
| 1ME4 | 1M72 | 79023 | 91208 | 0.731 | 0.27 | 0 | 0 | 0.037 | 0.27 | 0 | 0 |
| 1NME | 1KHQ | 85874 | 90964 | 0.731 | 0.292 | 0 | 0.571 | 0.181 | 0.212 | 0 | 0.571 |
| 1NME | 1XKG | 85874 | 122081 | 0.731 | 0.227 | 0.001 | 0.385 | 0 | 0.152 | 0.001 | 0.385 |
| 1NME | 2OUL | 85874 | 149025 | 0.731 | 0.267 | 0 | 0.263 | 0 | 0.203 | 0 | 0.263 |
| 1NW9 | 1XKG | 86294 | 122081 | 0.731 | 0.254 | 0 | 0.195 | 0 | 0.145 | 0 | 0.31 |
| 1NW9 | 2OUL | 86294 | 149025 | 0.731 | 0.306 | 0.006 | 0.206 | 0.012 | 0.167 | 0 | 0.226 |
| 1NW9 | 1KHQ | 86294 | 90964 | 0.731 | 0.277 | 0.01 | 0 | 0.194 | 0.199 | 0 | 0.438 |
| 1PYO | 1XKG | 95370 | 122081 | 0.731 | 0.266 | 0.003 | 0.417 | 0 | 0.177 | 0.003 | 0.385 |
| 1PYO | 2OUL | 95370 | 149025 | 0.731 | 0.289 | 0 | 0 | 0.025 | 0.201 | 0 | 0.267 |
| 1QTN | 3GCB | 30999 | 37032 | 0.731 | 0.21 | 0 | 0.625 | 0.163 | 0.137 | 0 | 0.625 |
| 1QTN | 1XKG | 30999 | 122081 | 0.731 | 0.24 | 0.001 | 0.281 | 0 | 0.193 | 0.001 | 0.455 |
| 1QTN | 2OUL | 30999 | 149025 | 0.731 | 0.287 | 0 | 0.316 | 0 | 0.213 | 0 | 0.333 |
| 1QTN | 1DEU | 30999 | 37085 | 0.731 | 0.276 | 0 | 0.5 | 0.322 | 0.202 | 0 | 0.391 |
| 1QTN | 1CS8 | 30999 | 37065 | 0.731 | 0.224 | 0 | 0.196 | 0.296 | 0.16 | 0 | 0.196 |
| 1QTN | 1CV8 | 30999 | 37089 | 0.731 | 0.326 | 0 | 0.348 | 0 | 0.276 | 0 | 0.273 |
| 1QTN | 1DKI | 30999 | 37090 | 0.731 | 0.228 | 0 | 0.267 | 0 | 0.18 | 0 | 0.267 |
| 1QTN | 1GMY | 30999 | 76236 | 0.731 | 0.245 | 0 | 0.417 | 0.505 | 0.185 | 0 | 0.417 |
| 1QTN | 1K3B | 30999 | 72015 | 0.731 | 0.257 | 0 | 0 | 0.557 | 0.303 | 0 | 0.25 |
| 1QTN | 1ME4 | 30999 | 79023 | 0.731 | 0.304 | 0 | 0.273 | 0.037 | 0.213 | 0 | 0.194 |
| 1QTN | 1KHQ | 30999 | 90964 | 0.731 | 0.259 | 0 | 0 | 0.181 | 0.218 | 0 | 0 |
| 1QTN | 2CB5 | 30999 | 37035 | 0.731 | 0.194 | 0 | 0.461 | 0.163 | 0.139 | 0 | 0.444 |
| 1SC3 | 1XKG | 105416 | 122081 | 0.731 | 0.251 | 0.003 | 0.333 | 0 | 0.179 | 0.001 | 0.333 |
| 1SC3 | 2OUL | 105416 | 149025 | 0.731 | 0.287 | 0.003 | 0 | 0.021 | 0.224 | 0.003 | 0.216 |
| 2CB5 | 1NME | 37035 | 85874 | 0.731 | 0.284 | 0 | 0.28 | 0.163 | 0.323 | 0 | 0.28 |
| 2CB5 | 1SC3 | 37035 | 105416 | 0.731 | 0.319 | 0 | 0.467 | 0.106 | 0.355 | 0 | 0.263 |
| 2CB5 | 1PYO | 37035 | 95370 | 0.731 | 0.286 | 0 | 0.28 | 0.212 | 0.346 | 0 | 0.28 |
| 2CB5 | 1M72 | 37035 | 91208 | 0.731 | 0.285 | 0 | 0.417 | 0.163 | 0.285 | 0 | 0.417 |
| 2CB5 | 1NW9 | 37035 | 86294 | 0.731 | 0.311 | 0 | 0.5 | 0.276 | 0.422 | 0 | 0.246 |
| 3GCB | 1NME | 37032 | 85874 | 0.731 | 0.267 | 0 | 0.167 | 0.163 | 0.329 | 0 | 0.167 |
| 3GCB | 1M72 | 37032 | 91208 | 0.731 | 0.279 | 0 | 0.5 | 0.163 | 0.279 | 0 | 0.5 |
| 3GCB | 1PYO | 37032 | 95370 | 0.731 | 0.275 | 0 | 0.268 | 0.212 | 0.336 | 0 | 0.37 |
| 3GCB | 1NW9 | 37032 | 86294 | 0.731 | 0.323 | 0 | 0.4 | 0.276 | 0.338 | 0 | 0.278 |
| 3GCB | 1SC3 | 37032 | 105416 | 0.731 | 0.319 | 0 | 0 | 0.106 | 0.336 | 0 | 0 |

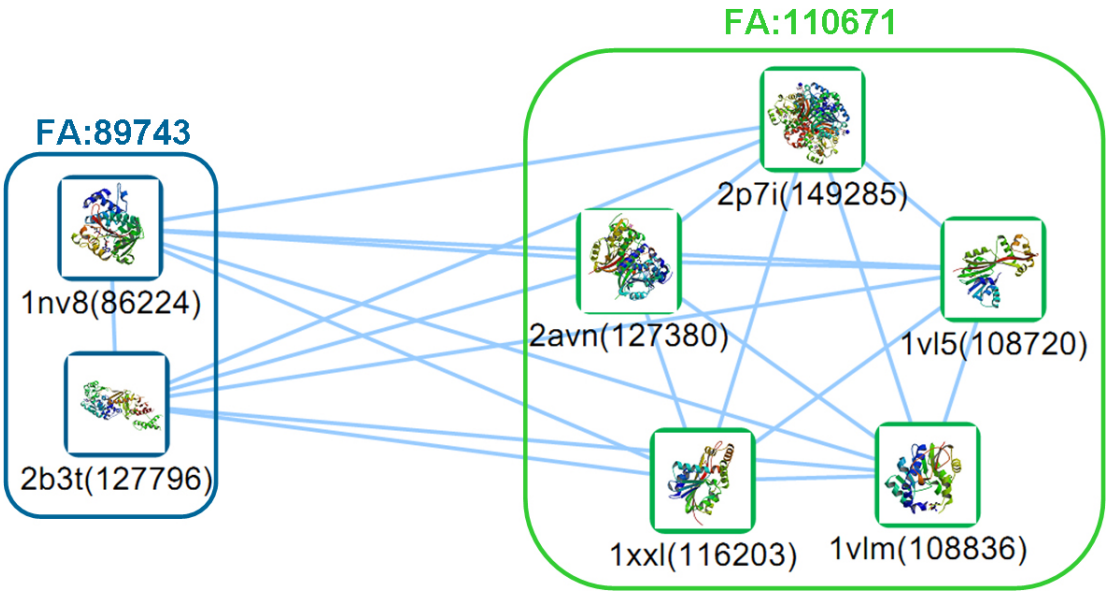
From BSM score with 70% threshold, 2JOY is associated with FA:159031 which is Ribosomal protein L14e. 2UUB is associated with FA:55000 which is Ribosomal protein S10. 1VQO (suid:120385) is associated with FA:54576 which is Ribosomal protein L31e. 1VI6 and 2GY9 are associated with FA52314 which is Ribosomal protein S2. 2GYC and two 1VQOs consisting of sunid 120378 and 120381 are associated with FA:50105 which is Ribosomal proteins L24p and L21e. 2J01, 1VQO (sunid: 120373, 120376), and 2GYC are associated with FA:52081 which is Ribosomal proteins L15p and L18e.




| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1VI6 | 2JOY | 100736 | 148163 | 0.703 | 0.356 | 0 | 0 | 0 | 0.356 | 0 | 0 |
| 1VI6 | 1VQO | 100736 | 120381 | 0.703 | 0.275 | 0 | 0 | 0.026 | 0.205 | 0.028 | 0.75 |
| 1VI6 | 1VQO | 100736 | 120373 | 0.703 | 0.272 | 0 | 0.312 | 0 | 0.205 | 0.028 | 0.75 |
| 1VI6 | 1VQO | 100736 | 120378 | 0.703 | 0.283 | 0 | 0 | 0 | 0.205 | 0.028 | 0.75 |
| 1VI6 | 1VQO | 100736 | 120376 | 0.703 | 0.37 | 0.005 | 0 | 0 | 0.205 | 0.028 | 0.75 |
| 1VI6 | 2J01 | 100736 | 145572 | 0.703 | 0.256 | 0 | 0 | 0.001 | 0.296 | 0.011 | 0.4 |
| 1VI6 | 2GYC | 100736 | 135853 | 0.703 | 0.266 | 0 | 0 | 0.162 | 0.308 | 0.03 | 0.467 |
| 1VI6 | 2UUB | 100736 | 139941 | 0.703 | 0.273 | 0 | 0 | 0.097 | 0.707 | 0.286 | 0.25 |
| 1VI6 | 1VQO | 100736 | 120385 | 0.703 | 0.35 | 0 | 0.381 | 0 | 0.205 | 0.028 | 0.75 |
| 1VI6 | 2GYC | 100736 | 135850 | 0.703 | 0.193 | 0 | 0 | 0.162 | 0.308 | 0.03 | 0.467 |
| 1VQO | 2JOY | 120381 | 148163 | 0.703 | 0.385 | 0.024 | 0.5 | 0 | 0.405 | 0 | 0.417 |
| 1VQO | 2JOY | 120376 | 148163 | 0.703 | 0.277 | 0 | 0 | 0 | 0.405 | 0 | 0.417 |
| 1VQO | 2JOY | 120385 | 148163 | 0.703 | 0.301 | 0 | 0.227 | 0 | 0.405 | 0 | 0.417 |
| 1VQO | 2J01 | 120381 | 145572 | 0.703 | 0.157 | 0 | 0.455 | 0.149 | 0.286 | 0 | 0.444 |
| 1VQO | 1VQO | 120381 | 120385 | 0.703 | 0.316 | 0 | 0.409 | 0 | 1 | 0.125 | 1 |
| 1VQO | 1VQO | 120376 | 120381 | 0.703 | 0.266 | 0 | 0 | 0 | 1 | 0.125 | 1 |

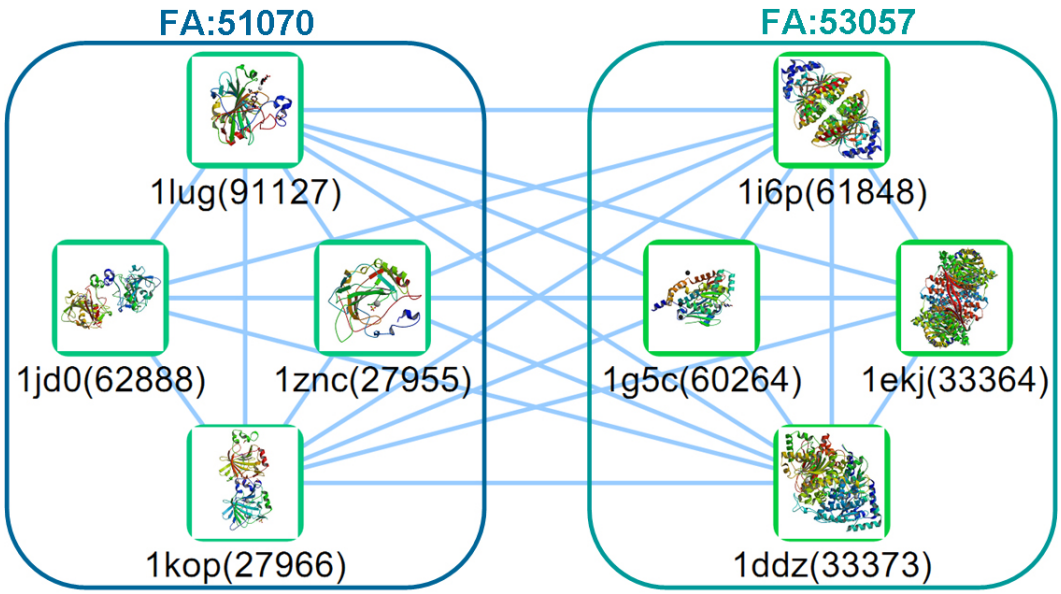
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1VQO | 2GY9 | 120381 | 145232 | 0.703 | 0.191 | 0 | 0 | 0.012 | 0.248 | 0 | 0.333 |
| 1VQO | 2JOY | 120378 | 148163 | 0.703 | 0.392 | 0.033 | 0.242 | 0 | 0.405 | 0 | 0.417 |
| 1VQO | 2UUB | 120381 | 139941 | 0.703 | 0.237 | 0 | 0.667 | 0.106 | 0.248 | 0 | 0.353 |
| 1VQO | 2JOY | 120373 | 148163 | 0.703 | 0.269 | 0 | 0.226 | 0 | 0.405 | 0 | 0.417 |
| 1VQO | 1VQO | 120373 | 120381 | 0.703 | 0.215 | 0 | 0.303 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2GYC | 120381 | 135850 | 0.703 | 0.19 | 0 | 0 | 0.017 | 0.255 | 0 | 0.385 |
| 1VQO | 1VQO | 120373 | 120378 | 0.703 | 0.216 | 0 | 0.283 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2UUB | 120373 | 139941 | 0.703 | 0.227 | 0 | 0 | 0 | 0.248 | 0 | 0.353 |
| 1VQO | 2GY9 | 120373 | 145232 | 0.703 | 0.185 | 0 | 0 | 0 | 0.248 | 0 | 0.333 |
| 1VQO | 2GYC | 120373 | 135853 | 0.703 | 0.198 | 0 | 0 | 0 | 0.255 | 0 | 0.385 |
| 1VQO | 1VQO | 120373 | 120385 | 0.703 | 0.34 | 0 | 0 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2UUB | 120376 | 139941 | 0.703 | 0.288 | 0 | 0 | 0 | 0.248 | 0 | 0.353 |
| 1VQO | 1VQO | 120376 | 120378 | 0.703 | 0.314 | 0 | 0.263 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2GYC | 120376 | 135853 | 0.703 | 0.227 | 0 | 0 | 0 | 0.255 | 0 | 0.385 |
| 1VQO | 1VQO | 120376 | 120385 | 0.703 | 0.354 | 0 | 0 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2GY9 | 120376 | 145232 | 0.703 | 0.216 | 0.006 | 0 | 0 | 0.248 | 0 | 0.333 |
| 1VQO | 2UUB | 120378 | 139941 | 0.703 | 0.222 | 0 | 0 | 0 | 0.248 | 0 | 0.353 |
| 1VQO | 2J01 | 120378 | 145572 | 0.703 | 0.16 | 0 | 0.309 | 0 | 0.286 | 0 | 0.444 |
| 1VQO | 1VQO | 120378 | 120385 | 0.703 | 0.275 | 0 | 0.25 | 0 | 1 | 0.125 | 1 |
| 1VQO | 2GYC | 120378 | 135850 | 0.703 | 0.187 | 0 | 0 | 0 | 0.255 | 0 | 0.385 |
| 1VQO | 2GY9 | 120378 | 145232 | 0.703 | 0.16 | 0 | 0 | 0 | 0.248 | 0 | 0.333 |
| 1VQO | 2GYC | 120385 | 135850 | 0.703 | 0.185 | 0 | 0 | 0 | 0.255 | 0 | 0.385 |
| 1VQO | 2GYC | 120385 | 135853 | 0.703 | 0.237 | 0 | 0 | 0 | 0.255 | 0 | 0.385 |
| 1VQO | 2GY9 | 120385 | 145232 | 0.703 | 0.14 | 0 | 0 | 0 | 0.248 | 0 | 0.333 |
| 1VQO | 2UUB | 120385 | 139941 | 0.703 | 0.296 | 0 | 0.75 | 0 | 0.248 | 0 | 0.353 |
| 1VQO | 2J01 | 120385 | 145572 | 0.703 | 0.185 | 0 | 0 | 0 | 0.286 | 0 | 0.444 |
| 2GY9 | 2JOY | 145232 | 148163 | 0.703 | 0.346 | 0 | 0.263 | 0 | 0.323 | 0 | 0.4 |
| 2GY9 | 2J01 | 145232 | 145572 | 0.703 | 0.229 | 0 | 0 | 0.028 | 0.297 | 0 | 0.35 |
| 2GYC | 2JOY | 135850 | 148163 | 0.703 | 0.245 | 0 | 0 | 0 | 0.295 | 0 | 0.545 |
| 2GYC | 2JOY | 135853 | 148163 | 0.703 | 0.383 | 0.024 | 0.444 | 0 | 0.295 | 0 | 0.545 |
| 2GYC | 2GYC | 135850 | 135853 | 0.703 | 0.21 | 0 | 0.444 | 0.722 | 1 | 0.147 | 1 |
| 2GYC | 2GY9 | 135850 | 145232 | 0.703 | 0.152 | 0 | 0 | 0.448 | 0.296 | 0 | 0.364 |
| 2GYC | 2UUB | 135850 | 139941 | 0.703 | 0.211 | 0 | 0.375 | 0.214 | 0.318 | 0 | 0.467 |
| 2GYC | 2J01 | 135853 | 145572 | 0.703 | 0.156 | 0 | 0 | 0.013 | 0.276 | 0.054 | 0.25 |
| 2GYC | 2GY9 | 135853 | 145232 | 0.703 | 0.167 | 0 | 0.323 | 0.448 | 0.296 | 0 | 0.364 |
| 2GYC | 2UUB | 135853 | 139941 | 0.703 | 0.204 | 0 | 0.667 | 0.214 | 0.318 | 0 | 0.467 |
| 2J01 | 2JOY | 145572 | 148163 | 0.703 | 0.256 | 0 | 0 | 0 | 0.225 | 0 | 0.5 |
| 2UUB | 2JOY | 139941 | 148163 | 0.703 | 0.257 | 0 | 0 | 0 | 0.351 | 0 | 0.3 |
| 2UUB | 2GY9 | 139941 | 145232 | 0.703 | 0.148 | 0 | 0 | 0.367 | 0.916 | 0.382 | 0.444 |
| 2UUB | 2J01 | 139941 | 145572 | 0.703 | 0.165 | 0 | 0.219 | 0.367 | 0.283 | 0.014 | 0.346 |

From BSM score with 70% threshold, 1NV8 and 2B3T are associated with FA:89743 which is N5-glutamine methyltransferase, HemK family and all others are associated with FA:110671 which is UbiE/COQ5 methyltransferase family



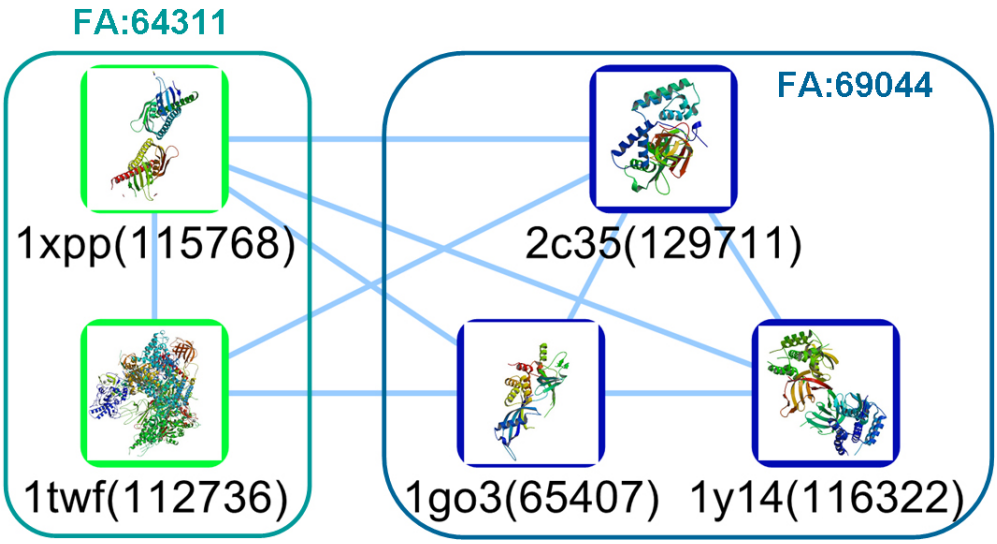

| PDB1 | PDB2 | Sunid1 | Sunid2 | GPsim | dmTM | dmDali | dmSEQ | dmPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1NV8 | 2AVN | 86224 | 127380 | 0.723 | 0.529 | 0.132 | 0.267 | 0 | 0.527 | 0.132 | 0.267 |
| 1NV8 | 1VLM | 86224 | 108836 | 0.723 | 0.481 | 0.126 | 0.316 | 0.008 | 0.481 | 0.126 | 0.316 |
| 1NV8 | 1VL5 | 86224 | 108720 | 0.723 | 0.547 | 0.171 | 0.269 | 0.09 | 0.547 | 0.171 | 0.269 |
| 1NV8 | 1XXL | 86224 | 116203 | 0.723 | 0.543 | 0.177 | 0.302 | 0.09 | 0.543 | 0.177 | 0.302 |
| 1NV8 | 2P7I | 86224 | 149285 | 0.723 | 0.565 | 0.133 | 0.37 | 0 | 0.565 | 0.132 | 0.37 |
| 2B3T | 1VL5 | 127796 | 108720 | 0.723 | 0.489 | 0.204 | 0.229 | 0.057 | 0.489 | 0.206 | 0.23 |
| 2B3T | 1VLM | 127796 | 108836 | 0.723 | 0.469 | 0.157 | 0.209 | 0.012 | 0.469 | 0.158 | 0.209 |
| 2B3T | 1XXL | 127796 | 116203 | 0.723 | 0.489 | 0.208 | 0.25 | 0.057 | 0.489 | 0.208 | 0.25 |
| 2B3T | 2AVN | 127796 | 127380 | 0.723 | 0.489 | 0.15 | 0.344 | 0 | 0.489 | 0.15 | 0.312 |
| 2B3T | 2P7I | 127796 | 149285 | 0.723 | 0.601 | 0.166 | 0.4 | 0 | 0.604 | 0.166 | 0.4 |

From BSM score with 70% threshold, 1JD0, 1KOP, 1LUG, and 1ZNC are associated with FA:51070 which is Carbonic anhydrase family and all others are associated with FA:53057 which is beta-carbonic anhydrase, cab family



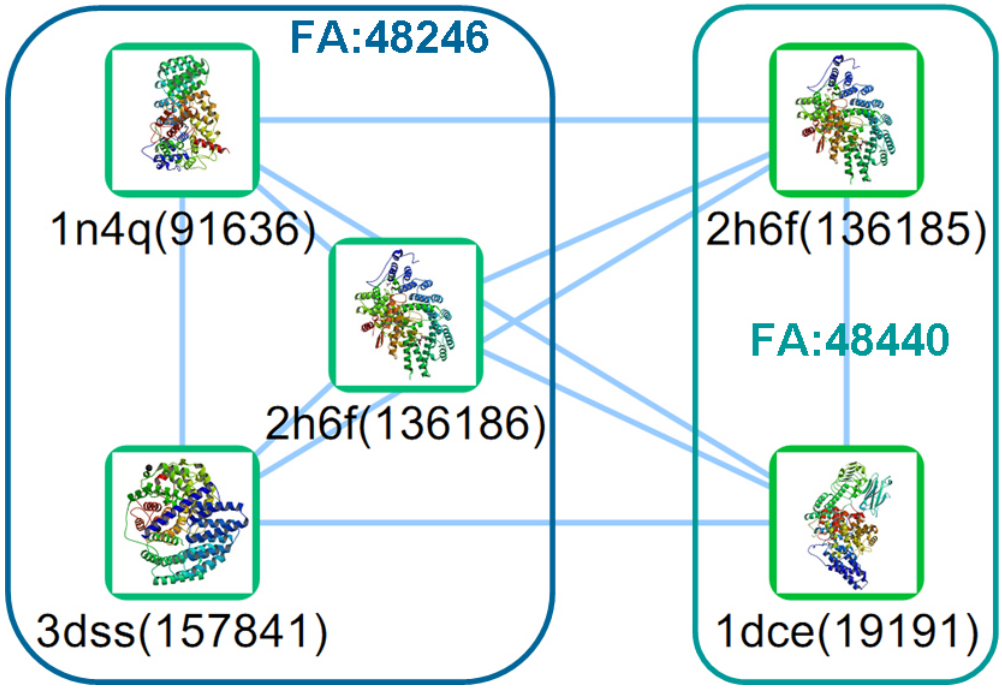

| PDB1 | PDB2 | Sunid1 | Sunid2 | GPsim | dmTM | dmDali | dmSEQ | dmPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1JD0 | 1EKJ | 62888 | 33364 | 0.739 | 0.211 | 0 | 0.231 | 0.123 | 0.211 | 0 | 0.231 |
| 1JD0 | 1DDZ | 62888 | 33373 | 0.739 | 0.221 | 0 | 0.4 | 0.223 | 0.327 | 0 | 0.667 |
| 1JD0 | 1G5C | 62888 | 60264 | 0.739 | 0.203 | 0 | 0 | 0.079 | 0.203 | 0 | 0 |
| 1JD0 | 1I6P | 62888 | 61848 | 0.739 | 0.244 | 0 | 0.333 | 0.223 | 0.244 | 0 | 0.333 |
| 1KOP | 1G5C | 27966 | 60264 | 0.739 | 0.254 | 0 | 0.357 | 0.066 | 0.254 | 0 | 0.357 |
| 1KOP | 1EKJ | 27966 | 33364 | 0.739 | 0.238 | 0 | 0 | 0.103 | 0.238 | 0 | 0 |
| 1KOP | 1I6P | 27966 | 61848 | 0.739 | 0.216 | 0 | 0.571 | 0.186 | 0.216 | 0 | 0.571 |
| 1KOP | 1DDZ | 27966 | 33373 | 0.739 | 0.199 | 0 | 0 | 0.186 | 0.167 | 0 | 0 |
| 1LUG | 1EKJ | 91127 | 33364 | 0.739 | 0.236 | 0 | 0.833 | 0.123 | 0.236 | 0 | 0.833 |
| 1LUG | 1DDZ | 91127 | 33373 | 0.739 | 0.201 | 0 | 0.6 | 0.223 | 0.323 | 0 | 0.364 |
| 1LUG | 1G5C | 91127 | 60264 | 0.739 | 0.221 | 0 | 0.667 | 0.079 | 0.221 | 0 | 0.667 |
| 1LUG | 1I6P | 91127 | 61848 | 0.739 | 0.247 | 0 | 0.296 | 0.223 | 0.247 | 0 | 0.296 |
| 1ZNC | 1G5C | 27955 | 60264 | 0.739 | 0.284 | 0 | 0 | 0.034 | 0.284 | 0 | 0 |
| 1ZNC | 1DDZ | 27955 | 33373 | 0.739 | 0.239 | 0 | 0.385 | 0.094 | 0.196 | 0 | 0 |
| 1ZNC | 1I6P | 27955 | 61848 | 0.739 | 0.279 | 0 | 0.538 | 0.094 | 0.279 | 0 | 0.538 |
| 1ZNC | 1EKJ | 27955 | 33364 | 0.739 | 0.277 | 0 | 0.417 | 0.053 | 0.277 | 0 | 0.417 |

From BSM score with 70% threshold, 1XPP and 1TWF are associated with FA:64311 which is RBP11/RpoL and all others are associated with FA:69044 which is RNA polymerase II subunit RBP4 (RpoF)



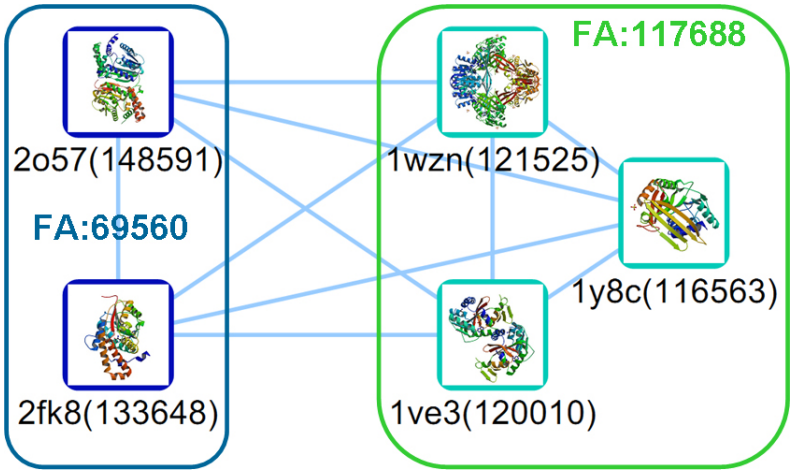

| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1TWF | 1GO3 | 112736 | 65407 | 0.719 | 0.269 | 0 | 0 | 0 | 0.071 | 0.028 | 0.333 |
| 1TWF | 2C35 | 112736 | 129711 | 0.719 | 0.263 | 0 | 0 | 0.429 | 0.365 | 0.024 | 0.357 |
| 1TWF | 1Y14 | 112736 | 116322 | 0.719 | 0.263 | 0 | 0 | 0.069 | 0.374 | 0.029 | 0.221 |
| 1XPP | 1GO3 | 115768 | 65407 | 0.719 | 0.293 | 0 | 0 | 0 | 0.28 | 0.033 | 0.364 |
| 1XPP | 2C35 | 115768 | 129711 | 0.719 | 0.26 | 0 | 0 | 0 | 0.26 | 0.018 | 0 |
| 1XPP | 1Y14 | 115768 | 116322 | 0.719 | 0.25 | 0 | 0.5 | 0 | 0.25 | 0.023 | 0.5 |

From BSM score with 70% threshold, 2H6F and 1DCE are associated with FA:48440 which is Protein prenylyltransferase and all others are associated with FA:48246 which is Protein prenyltransferases




| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1DCE | 3DSS | 19191 | 157841 | 0.729 | 0.347 | 0.021 | 0.375 | 0.541 | 0.94 | 0.345 | 1 |
| 1DCE | 1N4Q | 19191 | 91636 | 0.729 | 0.344 | 0.023 | 0.308 | 0.541 | 0.743 | 0.219 | 0.269 |
| 1DCE | 2H6F | 19191 | 136186 | 0.729 | 0.301 | 0.038 | 0.412 | 0.541 | 0.741 | 0.221 | 0.269 |
| 1N4Q | 2H6F | 91636 | 136185 | 0.729 | 0.346 | 0.029 | 0.379 | 0.322 | 0.979 | 0.458 | 0.928 |
| 2H6F | 3DSS | 136185 | 157841 | 0.729 | 0.311 | 0.021 | 0.237 | 0.322 | 0.767 | 0.279 | 0.269 |
| 2H6F | 2H6F | 136185 | 136186 | 0.729 | 0.295 | 0.04 | 0.556 | 0.322 | 1 | 0.507 | 1 |

From BSM score with 70% threshold, 2O57 and 2FK8 are associated with FA:69560 which is Mycolic acid cyclopropane synthase and all others are associated with FA:117688 which is CAC2371-like family and also known to be similar overall fold to the Glycine N-methyltransferase (scop_fa 53348) and mRNA cap (Guanine N-7) methyltransferase (scop_fa 102560) families.




| PDB1 | PDB2 | DM1 | DM2 | GPsim | domainTM | domainDali | domainSEQ | domainPHY | gpTM | gpDali | gpSEQ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2FK8 | 1VE3 | 133648 | 120010 | 0.703 | 0.573 | 0.205 | 0.198 | 0 | 0.572 | 0.204 | 0.198 |
| 2FK8 | 1WZN | 133648 | 121525 | 0.703 | 0.592 | 0.206 | 0.393 | 0.008 | 0.59 | 0.205 | 0.393 |
| 2FK8 | 1Y8C | 133648 | 116563 | 0.703 | 0.594 | 0.192 | 0.268 | 0.008 | 0.592 | 0.191 | 0.308 |
| 2O57 | 1VE3 | 148591 | 120010 | 0.703 | 0.586 | 0.239 | 0.294 | 0 | 0.586 | 0.245 | 0.294 |
| 2O57 | 1WZN | 148591 | 121525 | 0.703 | 0.589 | 0.223 | 0.333 | 0 | 0.589 | 0.224 | 0.333 |
| 2O57 | 1Y8C | 148591 | 116563 | 0.703 | 0.585 | 0.216 | 0.379 | 0 | 0.585 | 0.224 | 0.379 |

# [S-11] OMIM dataset

Disease and gene product information used in OMIM-based experiments

| OMIM | DISEASE | UniProt ID | Name of Protein |
|---|---|---|---|
| 104300 | ALZHEIMER DISEASE | P78380 | Oxidized low-density lipoprotein receptor 1 |
| 104300 | ALZHEIMER DISEASE | P49810 | Presenilin-2 |
| 104300 | ALZHEIMER DISEASE | P05067 | Amyloid beta A4 protein |
| 105200 | AMYLOIDOSIS | P02647 | Apolipoprotein A-I |
| 105201 | AMYLOIDOSIS | P02671 | Fibrinogen alpha chain |
| 105202 | AMYLOIDOSIS | P61626 | Lysozyme C |
| 105830 | ANGELMAN SYNDROME | P51608 | Methyl-CpG-binding protein 2 |
| 105830 | ANGELMAN SYNDROME | Q05086 | Ubiquitin-protein ligase E3A |
| 105830 | ANGELMAN SYNDROME | O60312 | Probable phospholipid-transporting ATPase VA |
| 109100 | AUTOIMMUNE DISEASE | P16410 | Cytotoxic T-lymphocyte protein 4 |
| 109100 | AUTOIMMUNE DISEASE | Q15116 | Programmed cell death protein 1 |
| 109100 | AUTOIMMUNE DISEASE | O60602 | Toll-like receptor 5 |
| 109100 | AUTOIMMUNE DISEASE | O43918 | Autoimmune regulator |
| 109800 | BLADDER CANCER | P22607 | Fibroblast growth factor receptor 3 |
| 109800 | BLADDER CANCER | P06400 | Retinoblastoma-associated protein |
| 109800 | BLADDER CANCER | P01112 | GTPase HRas |
| 114480 | BREAST CANCER | P38398 | Breast cancer type 1 susceptibility protein |
| 114480 | BREAST CANCER | P51587 | Breast cancer type 2 susceptibility protein |
| 114480 | BREAST CANCER | Q9BX63 | Fanconi anemia group J protein |
| 114480 | BREAST CANCER | Q96BI1 | Solute carrier family 22 member 18 |
| 114480 | BREAST CANCER | P42336 | Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha isoform |
| 114480 | BREAST CANCER | Q06609 | DNA repair protein RAD51 homolog 1 |
| 114480 | BREAST CANCER | O60934 | Nibrin |
| 114550 | HEPATOCELLULAR CARCINOMA | Q16667 | Cyclin-dependent kinase inhibitor 3 |
| 114550 | HEPATOCELLULAR CARCINOMA | P08581 | Hepatocyte growth factor receptor |
| 114550 | HEPATOCELLULAR CARCINOMA | Q9ULD2 | Mitochondrial tumor suppressor 1 |
| 114550 | HEPATOCELLULAR CARCINOMA | O15169 | Axin-1 |
| 115150 | CFC SYNDROME | P15056 | B-Raf proto-oncogene serine/threonine-protein kinase |
| 115150 | CFC SYNDROME | Q02750 | Dual specificity mitogen-activated protein kinase kinase 1 |
| 115150 | CFC SYNDROME | P01116 | GTPase KRas |
| 115150 | CFC SYNDROME | P36507 | Dual specificity mitogen-activated protein kinase kinase 2 |
| 115200 | DILATED CARDIOMYOPATHY 1A | O75112 | LIM domain-binding protein 3 |
| 115200 | DILATED CARDIOMYOPATHY 1A | P02545 | Lamin-A/C |
| 115200 | DILATED CARDIOMYOPATHY 1A | P12883 | Myosin-7 |
| 115200 | DILATED CARDIOMYOPATHY 1A | P68032 | Actin, alpha cardiac muscle 1 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | Q9HC96 | Calpain-10 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P14672 | Solute carrier family 2, facilitated glucose transporter member 4 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P35680 | Hepatocyte nuclear factor 1-beta |

| | | | |
|---|---|---|---|
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P06213 | Insulin receptor |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P35568 | Insulin receptor substrate 1 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | Q9UQF2 | C-jun-amino-terminal kinase-interacting protein 1 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | Q13562 | Neurogenic differentiation factor 1 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P52945 | Pancreas/duodenum homeobox protein 1 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | Q16821 | Protein phosphatase 1 regulatory subunit 3A |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | O15357 | Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 2 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | Q8IWU4 | Zinc transporter 8 |
| 125853 | NONINSULIN-DEPENDENT DIABETES MELLITUS | P22413 | Ectonucleotide pyrophosphatase/phosphodiesterase family member 1 |
| 130600 | ELLIPTOCYTOSIS 2 | P02730 | Band 3 anion transport protein |
| 130600 | ELLIPTOCYTOSIS 2 | P11277 | Spectrin beta chain, erythrocyte |
| 130600 | ELLIPTOCYTOSIS 2 | P02549 | Spectrin alpha chain, erythrocyte |
| 133239 | ESCC | Q9Y238 | Deleted in lung and esophageal cancer protein 1 |
| 133239 | ESCC | P37173 | TGF-beta receptor type-2 |
| 133239 | ESCC | Q9NZC7 | WW domain-containing oxidoreductase |
| 133239 | ESCC | P04637 | Cellular tumor antigen p53 |
| 142623 | HIRSCHSPRUNG DISEASE | P42892 | Endothelin-converting enzyme 1 |
| 142623 | HIRSCHSPRUNG DISEASE | P14138 | Endothelin-3 |
| 142623 | HIRSCHSPRUNG DISEASE | P24530 | Endothelin B receptor |
| 142623 | HIRSCHSPRUNG DISEASE | P32004 | Neural cell adhesion molecule L1 |
| 142623 | HIRSCHSPRUNG DISEASE | Q99748 | Neurturin |
| 142623 | HIRSCHSPRUNG DISEASE | P07949 | Proto-oncogene tyrosine-protein kinase receptor ret |
| 142623 | HIRSCHSPRUNG DISEASE | P39905 | Glial cell line-derived neurotrophic factor |
| 144700 | RENAL CELL CARCINOMA 1 | Q96SL1 | Disrupted in renal carcinoma protein 2 |
| 144700 | RENAL CELL CARCINOMA 1 | P49789 | Bis(5'-adenosyl)-triphosphatase |
| 144700 | RENAL CELL CARCINOMA 1 | Q96EW2 | HSPB1-associated protein 1 |
| 144700 | RENAL CELL CARCINOMA 1 | O15527 | N-glycosylase/DNA lyase |
| 144700 | RENAL CELL CARCINOMA 1 | P40337 | Von Hippel-Lindau disease tumor suppressor |
| 144700 | RENAL CELL CARCINOMA 1 | Q8WU17 | RING finger protein 139 |
| 145900 | DEJERINE-SOTTAS SYNDROME | P08034 | Gap junction beta-1 protein |
| 145900 | DEJERINE-SOTTAS SYNDROME | P11161 | Early growth response protein 2 |
| 145900 | DEJERINE-SOTTAS SYNDROME | Q01453 | Peripheral myelin protein 22 |
| 145900 | DEJERINE-SOTTAS SYNDROME | Q9BXM0 | Periaxin |
| 145900 | DEJERINE-SOTTAS SYNDROME | P25189 | Myelin P0 protein |
| 146110 | HYPOGONADOTROPIC HYPOGONADISM | P11362 | Basic fibroblast growth factor receptor 1 |
| 146110 | HYPOGONADOTROPIC HYPOGONADISM | P30968 | Gonadotropin-releasing hormone receptor |
| 146110 | HYPOGONADOTROPIC HYPOGONADISM | Q6X4W1 | Nasal embryonic luteinizing hormone-releasing hormone factor |
| 146110 | HYPOGONADOTROPIC HYPOGONADISM | Q969F8 | KiSS-1 receptor |
| 149730 | LACRIMOAURICULODENTODIGITAL SYNDROME | P21802 | Fibroblast growth factor receptor 2 |

| | | | |
|---|---|---|---|
| 149730 | LACRIMOAURICULODENTODIGITAL SYNDROME | P22607 | Fibroblast growth factor receptor 3 |
| 149730 | LACRIMOAURICULODENTODIGITAL SYNDROME | O15520 | Fibroblast growth factor 10 |
| 155255 | MEDULLOBLASTOMA | P25054 | Adenomatous polyposis coli protein |
| 155255 | MEDULLOBLASTOMA | Q9Y6C5 | Protein patched homolog 2 |
| 155255 | MEDULLOBLASTOMA | Q9UMX1 | Suppressor of fused homolog |
| 155255 | MEDULLOBLASTOMA | P35222 | Catenin beta-1 |
| 158000 | MONILETHRIX | Q14533 | Keratin, type II cuticular Hb1 |
| 158000 | MONILETHRIX | P78385 | Keratin, type II cuticular Hb3 |
| 158000 | MONILETHRIX | O43790 | Keratin, type II cuticular Hb6 |
| 158810 | BETHLEM MYOPATHY | P12109 | Collagen alpha-1(VI) chain |
| 158810 | BETHLEM MYOPATHY | P12111 | Collagen alpha-3(VI) chain |
| 158810 | BETHLEM MYOPATHY | P12110 | Collagen alpha-2(VI) chain |
| 168600 | PARKINSON DISEASE | P04062 | Glucosylceramidase |
| 168600 | PARKINSON DISEASE | O43464 | Serine protease HTRA2, mitochondrial |
| 168600 | PARKINSON DISEASE | Q5S007 | Leucine-rich repeat serine/threonine-protein kinase 2 |
| 168600 | PARKINSON DISEASE | Q99497 | Protein DJ-1 |
| 168600 | PARKINSON DISEASE | Q9BXM7 | Serine/threonine-protein kinase PINK1, mitochondrial |
| 168600 | PARKINSON DISEASE | O60260 | E3 ubiquitin-protein ligase parkin |
| 168600 | PARKINSON DISEASE | Q9Y6H5 | Synphilin-1 |
| 168600 | PARKINSON DISEASE | P37840 | Alpha-synuclein |
| 170400 | HYPOKALEMIC PERIODIC PARALYSIS | Q9Y6H6 | Potassium voltage-gated channel subfamily E member 3 |
| 170400 | HYPOKALEMIC PERIODIC PARALYSIS | P35499 | Sodium channel protein type 4 subunit alpha |
| 170400 | HYPOKALEMIC PERIODIC PARALYSIS | Q13698 | Voltage-dependent L-type calcium channel subunit alpha-1S |
| 171300 | PHEOCHROMOCYTOMA | P21912 | Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial |
| 171300 | PHEOCHROMOCYTOMA | O14521 | Succinate dehydrogenase [ubiquinone] cytochrome b small subunit, mitochondrial |
| 171300 | PHEOCHROMOCYTOMA | P07949 | Proto-oncogene tyrosine-protein kinase receptor ret |
| 171300 | PHEOCHROMOCYTOMA | P40337 | Von Hippel-Lindau disease tumor suppressor |
| 176807 | PROSTATE CANCER | O96017 | Serine/threonine-protein kinase Chk2 |
| 176807 | PROSTATE CANCER | P29323 | Ephrin type-B receptor 2 |
| 176807 | PROSTATE CANCER | P60484 | Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN |
| 176807 | PROSTATE CANCER | Q05823 | 2-5A-dependent ribonuclease |
| 176807 | PROSTATE CANCER | Q9BQ52 | Zinc phosphodiesterase ELAC protein 2 |
| 176807 | PROSTATE CANCER | P50539 | MAX-interacting protein 1 |
| 180300 | RHEUMATOID ARTHRITIS | Q9UBC1 | NF-kappa-B inhibitor-like protein 1 |
| 180300 | RHEUMATOID ARTHRITIS | Q9UM07 | Protein-arginine deiminase type-4 |
| 180300 | RHEUMATOID ARTHRITIS | Q8TDQ0 | Hepatitis A virus cellular receptor 2 |
| 180300 | RHEUMATOID ARTHRITIS | P11021 | 78 kDa glucose-regulated protein |
| 187500 | TETRALOGY OF FALLOT | Q8WW38 | Zinc finger protein ZFPM2 |
| 187500 | TETRALOGY OF FALLOT | P78504 | Protein jagged-1 |
| 187500 | TETRALOGY OF FALLOT | P52952 | Homeobox protein Nkx-2.5 |
| 188050 | THROMBOPHILIA VENOUS THROMBOEMBOLISM | P01008 | Antithrombin-III |

| | | | |
|---|---|---|---|
| 188050 | THROMBOPHILIA VENOUS THROMBOEMBOLISM | P05546 | Heparin cofactor 2 |
| 188050 | THROMBOPHILIA VENOUS THROMBOEMBOLISM | P05121 | Plasminogen activator inhibitor 1 |
| 188050 | THROMBOPHILIA VENOUS THROMBOEMBOLISM | P00747 | Plasminogen |
| 188050 | THROMBOPHILIA VENOUS THROMBOEMBOLISM | P04070 | Vitamin K-dependent protein C |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q16204 | Coiled-coil domain-containing protein 6 |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q8TBA6 | Golgin subfamily A member 5 |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q13772 | Nuclear receptor coactivator 4 |
| 188550 | PAPILLARY CARCINOMA OF THYROID | P04629 | High affinity nerve growth factor receptor |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q15154 | Pericentriolar material 1 protein |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q8IUD2 | ELKS/RAB6-interacting/CAST family member 1 |
| 188550 | PAPILLARY CARCINOMA OF THYROID | P07949 | Proto-oncogene tyrosine-protein kinase receptor ret |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q92734 | Protein TFG |
| 188550 | PAPILLARY CARCINOMA OF THYROID | O15164 | Transcription intermediary factor 1-alpha |
| 188550 | PAPILLARY CARCINOMA OF THYROID | P06753 | Tropomyosin alpha-3 chain |
| 188550 | PAPILLARY CARCINOMA OF THYROID | P12270 | Nucleoprotein TPR |
| 188550 | PAPILLARY CARCINOMA OF THYROID | Q9UPN9 | E3 ubiquitin-protein ligase TRIM33 |
| 188550 | PAPILLARY CARCINOMA OF THYROID | P14373 | Zinc finger protein RFP |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P56539 | Caveolin-3 |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P10916 | Myosin regulatory light chain 2, ventricular/cardiac muscle isoform |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P13533 | Myosin-6 |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P12883 | Myosin-7 |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P08590 | Myosin light chain 3 |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | Q9H1R3 | Myosin light chain kinase 2, skeletal/cardiac muscle |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | Q9UM54 | Myosin-VI |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | Q14896 | Myosin-binding protein C, cardiac-type |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | O15273 | Telethonin |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P19429 | Troponin I, cardiac muscle |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P45379 | Troponin T, cardiac muscle |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P09493 | Tropomyosin alpha-1 chain |
| 192600 | FAMILIAL HYPERTROPHIC CARDIOMYOPATHY | P68032 | Actin, alpha cardiac muscle 1 |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9UIG0 | Bromodomain adjacent to zinc finger domain protein 1B |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9BQE9 | B-cell CLL/lymphoma 7 protein family member B |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9UDT6 | CAP-Gly domain-containing linker protein 2 |

| | | | |
|---|---|---|---|
| 194050 | WILLIAMS-BEUREN SYNDROME | P15502 | Elastin |
| 194050 | WILLIAMS-BEUREN SYNDROME | O75344 | FK506-binding protein 6 |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9UHL9 | General transcription factor II-I repeat domain-containing protein 1 |
| 194050 | WILLIAMS-BEUREN SYNDROME | P78347 | General transcription factor II-I |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q15056 | Eukaryotic translation initiation factor 4H |
| 194050 | WILLIAMS-BEUREN SYNDROME | P53667 | LIM domain kinase 1 |
| 194050 | WILLIAMS-BEUREN SYNDROME | P35250 | Replication factor C subunit 2 |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9Y4P3 | Transducin beta-like protein 2 |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9NP71 | Williams-Beuren syndrome chromosomal region 14 protein |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q96I51 | Williams-Beuren syndrome chromosomal region 16 protein |
| 194050 | WILLIAMS-BEUREN SYNDROME | O43709 | Uncharacterized methyltransferase WBSCR22 |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9H6D5 | Williams-Beuren syndrome chromosomal region 23 protein |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q8N6F8 | Williams-Beuren syndrome chromosomal region 27 protein |
| 194050 | WILLIAMS-BEUREN SYNDROME | Q9GZY6 | Linker for activation of T-cells family member 2 |
| 202370 | ADRENOLEUKODYSTROPHY NEONATAL | Q92968 | Peroxisomal membrane protein PEX13 |
| 202370 | ADRENOLEUKODYSTROPHY NEONATAL | O43933 | Peroxisome biogenesis factor 1 |
| 202370 | ADRENOLEUKODYSTROPHY NEONATAL | Q7Z412 | Peroxisome assembly protein 26 |
| 202370 | ADRENOLEUKODYSTROPHY NEONATAL | P50542 | Peroxisomal targeting signal 1 receptor |
| 202370 | ADRENOLEUKODYSTROPHY | O60683 | Peroxisome assembly protein 10 |
| 203300 | HERMANSKY-PUDLAK SYNDROME | O00203 | AP-3 complex subunit beta-1 |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q6QNY0 | Biogenesis of lysosome-related organelles complex 1 subunit 3 |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q96EV8 | Dysbindin |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q969F9 | Hermansky-Pudlak syndrome 3 protein |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q9NQG7 | Hermansky-Pudlak syndrome 4 protein |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q9UPZ3 | Hermansky-Pudlak syndrome 5 protein |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q86YV9 | Hermansky-Pudlak syndrome 6 protein |
| 203300 | HERMANSKY-PUDLAK SYNDROME | Q92902 | Hermansky-Pudlak syndrome 1 protein |
| 203450 | ALEXANDER DISEASE | P14136 | Glial fibrillary acidic protein |
| 203450 | ALEXANDER DISEASE | P49821 | NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial |
| 203450 | ALEXANDER DISEASE | P02511 | Alpha-crystallin B chain |
| 209880 | CONGENITAL ONDINE CURSE | P23560 | Brain-derived neurotrophic factor |
| 209880 | CONGENITAL ONDINE CURSE | P14138 | Endothelin-3 |
| 209880 | CONGENITAL ONDINE CURSE | Q99453 | Paired mesoderm homeobox protein 2B |
| 209880 | CONGENITAL ONDINE CURSE | P07949 | Proto-oncogene tyrosine-protein kinase receptor ret |
| 209880 | CONGENITAL ONDINE CURSE | P39905 | Glial cell line-derived neurotrophic factor |
| 209900 | BARDET-BIEDL SYNDROME | Q9H0F7 | ADP-ribosylation factor-like protein 6 |
| 209900 | BARDET-BIEDL SYNDROME | Q8TAM1 | Bardet-Biedl syndrome 10 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q6ZW61 | Bardet-Biedl syndrome 12 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q8NFJ9 | Bardet-Biedl syndrome 1 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q9BXC9 | Bardet-Biedl syndrome 2 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q8N3I7 | Bardet-Biedl syndrome 5 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q8IWZ6 | Bardet-Biedl syndrome 7 protein |
| 209900 | BARDET-BIEDL SYNDROME | Q9NPJ1 | McKusick-Kaufman/Bardet-Biedl syndromes putative chaperonin |

| | | | |
|---|---|---|---|
| 209900 | BARDET-BIEDL SYNDROME | Q3SYG4 | Protein PTHB1 |
| 209900 | BARDET-BIEDL SYNDROME | Q8TAM2 | Tetratricopeptide repeat protein 8 |
| 209900 | BARDET-BIEDL SYNDROME | Q96RK4 | Bardet-Biedl syndrome 4 protein |
| 209950 | FAMILIAL ATYPICAL MYCOBACTERIOSIS | P42701 | Interleukin-12 receptor beta-1 chain |
| 209950 | FAMILIAL ATYPICAL MYCOBACTERIOSIS | P29460 | Interleukin-12 subunit beta |
| 209950 | FAMILIAL ATYPICAL MYCOBACTERIOSIS | P15260 | Interferon-gamma receptor alpha chain |
| 209950 | FAMILIAL ATYPICAL MYCOBACTERIOSIS | P38484 | Interferon-gamma receptor beta chain |
| 209950 | FAMILIAL ATYPICAL MYCOBACTERIOSIS | P42224 | Signal transducer and activator of transcription 1-alpha/beta |
| 211980 | ADENOCARCINOMA OF LUNG | P15056 | B-Raf proto-oncogene serine/threonine-protein kinase |
| 211980 | ADENOCARCINOMA OF LUNG | Q9Y238 | Deleted in lung and esophageal cancer protein 1 |
| 211980 | ADENOCARCINOMA OF LUNG | P00533 | Epidermal growth factor receptor |
| 211980 | ADENOCARCINOMA OF LUNG | Q96BI1 | Solute carrier family 22 member 18 |
| 211980 | ADENOCARCINOMA OF LUNG | P04637 | Cellular tumor antigen p53 |
| 214100 | ZELLWEGER SYNDROME | O60683 | Peroxisome assembly protein 10 |
| 214100 | ZELLWEGER SYNDROME | O75381 | Peroxisomal membrane protein PEX14 |
| 214100 | ZELLWEGER SYNDROME | Q9Y5Y5 | Peroxisomal membrane protein PEX16 |
| 214100 | ZELLWEGER SYNDROME | P40855 | Peroxisomal biogenesis factor 19 |
| 214100 | ZELLWEGER SYNDROME | Q7Z412 | Peroxisome assembly protein 26 |
| 214100 | ZELLWEGER SYNDROME | P28328 | Peroxisome assembly factor 1 |
| 214100 | ZELLWEGER SYNDROME | P56589 | Peroxisomal biogenesis factor 3 |
| 214100 | ZELLWEGER SYNDROME | P50542 | Peroxisomal targeting signal 1 receptor |
| 214100 | ZELLWEGER SYNDROME | Q13608 | Peroxisome assembly factor 2 |
| 214100 | ZELLWEGER SYNDROME | O00623 | Peroxisome assembly protein 12 |
| 219080 | ACTH-INDEPENDENT MACRONODULAR ADRENAL HYPERPLASIA | P84996 | Protein ALEX |
| 219080 | ACTH-INDEPENDENT MACRONODULAR ADRENAL HYPERPLASIA | Q5JWF2 | Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas |
| 219080 | ACTH-INDEPENDENT MACRONODULAR ADRENAL HYPERPLASIA | O95467 | Neuroendocrine secretory protein 55 |
| 219080 | ACTH-INDEPENDENT MACRONODULAR ADRENAL HYPERPLASIA | P63092 | Guanine nucleotide-binding protein G(s) subunit alpha isoforms short |
| 219100 | AUTOSOMAL RECESSIVE CUTIS LAXA | O95967 | EGF-containing fibulin-like extracellular matrix protein 2 |
| 219100 | AUTOSOMAL RECESSIVE CUTIS LAXA | P28300 | Protein-lysine 6-oxidase |
| 219100 | AUTOSOMAL RECESSIVE CUTIS LAXA | Q9UBX5 | Fibulin-5 |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | Q12887 | Protoheme IX farnesyltransferase, mitochondrial |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | Q7KZN9 | Cytochrome c oxidase assembly protein COX15 homolog |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | P00395 | Cytochrome c oxidase subunit 1 |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | P00414 | Cytochrome c oxidase subunit 3 |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | O75880 | Protein SCO1 homolog, mitochondrial |

| | | | |
|---|---|---|---|
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | O43819 | Protein SCO2 homolog, mitochondrial |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | Q15526 | Surfeit locus protein 1 |
| 220110 | MITOCHONDRIAL COMPLEX IV DEFICIENCY | P00403 | Cytochrome c oxidase subunit 2 |
| 226650 | PROGRESSIVE EPIDERMOLYSIS BULLOSA JUNCTIONALIS | Q9UMD9 | Collagen alpha-1(XVII) chain |
| 226650 | PROGRESSIVE EPIDERMOLYSIS BULLOSA JUNCTIONALIS | Q13751 | Laminin subunit beta-3 |
| 226650 | PROGRESSIVE EPIDERMOLYSIS BULLOSA JUNCTIONALIS | P16144 | Integrin beta-4 |
| 226700 | EPIDERMOLYSIS BULLOSA LETALIS | Q13751 | Laminin subunit beta-3 |
| 226700 | EPIDERMOLYSIS BULLOSA LETALIS | Q13753 | Laminin subunit gamma-2 |
| 226700 | EPIDERMOLYSIS BULLOSA LETALIS | Q16787 | Laminin subunit alpha-3 |
| 227650 | FANCONI ANEMIA | P51587 | Breast cancer type 2 susceptibility protein |
| 227650 | FANCONI ANEMIA | Q9BXW9 | Fanconi anemia group D2 protein |
| 227650 | FANCONI ANEMIA | O15360 | Fanconi anemia group A protein |
| 227650 | FANCONI ANEMIA | Q8NB91 | Fanconi anemia group B protein |
| 227650 | FANCONI ANEMIA | Q9HB96 | Fanconi anemia group E protein |
| 227650 | FANCONI ANEMIA | Q9NPI8 | Fanconi anemia group F protein |
| 227650 | FANCONI ANEMIA | O15287 | Fanconi anemia group G protein |
| 227650 | FANCONI ANEMIA | Q9NVI1 | Fanconi anemia group I protein |
| 227650 | FANCONI ANEMIA | Q9BX63 | Fanconi anemia group J protein |
| 227650 | FANCONI ANEMIA | Q9NW38 | E3 ubiquitin-protein ligase FANCL |
| 227650 | FANCONI ANEMIA | Q8IYD8 | Fanconi anemia group M protein |
| 227650 | FANCONI ANEMIA | Q86YC2 | Partner and localizer of BRCA2 |
| 227650 | FANCONI ANEMIA | Q00597 | Fanconi anemia group C protein |
| 231200 | GIANT PLATELET SYNDROME | P07359 | Platelet glycoprotein Ib alpha chain |
| 231200 | GIANT PLATELET SYNDROME | P14770 | Platelet glycoprotein IX |
| 231200 | GIANT PLATELET SYNDROME | P13224 | Platelet glycoprotein Ib beta chain |
| 235400 | ATYPICAL HEMOLYTIC UREMIC SYNDROME | P08603 | Complement factor H |
| 235400 | ATYPICAL HEMOLYTIC UREMIC SYNDROME | P05156 | Complement factor I |
| 235400 | ATYPICAL HEMOLYTIC UREMIC SYNDROME | P15529 | Membrane cofactor protein |
| 235400 | ATYPICAL HEMOLYTIC UREMIC SYNDROME | Q76LX8 | A disintegrin and metalloproteinase with thrombospondin motifs 13 |
| 236670 | WALKER-WARBURG SYNDROME | O75072 | Fukutin |
| 236670 | WALKER-WARBURG SYNDROME | Q9H9S5 | Fukutin-related protein |
| 236670 | WALKER-WARBURG SYNDROME | Q8WZA1 | Protein O-linked-mannose beta-1,2-N-acetylglucosaminyltransferase 1 |
| 236670 | WALKER-WARBURG SYNDROME | Q9UKY4 | Protein O-mannosyl-transferase 2 |
| 236670 | WALKER-WARBURG SYNDROME | Q9Y6A1 | Protein O-mannosyl-transferase 1 |
| 236750 | IDIOPATHIC HYDROPS FETALIS | P08236 | Beta-glucuronidase |
| 236750 | IDIOPATHIC HYDROPS FETALIS | P04062 | Glucosylceramidase |
| 236750 | IDIOPATHIC HYDROPS FETALIS | Q04446 | 1,4-alpha-glucan-branching enzyme |
| 236750 | IDIOPATHIC HYDROPS FETALIS | P10746 | Uroporphyrinogen-III synthase |
| 236750 | IDIOPATHIC HYDROPS FETALIS | Q9NRA2 | Sialin |
| 236750 | IDIOPATHIC HYDROPS FETALIS | P69905 | Hemoglobin subunit alpha |
| 242100 | CONGENITAL NONBULLOUS | Q9BYJ1 | Epidermis-type lipoxygenase 3 |

| | | | |
|---|---|---|---|
| | ICHTHYOSIFORM ERYTHRODERMA | | |
| 242100 | CONGENITAL NONBULLOUS ICHTHYOSIFORM ERYTHRODERMA | O75342 | Arachidonate 12-lipoxygenase, 12R type |
| 242100 | CONGENITAL NONBULLOUS ICHTHYOSIFORM ERYTHRODERMA | P22735 | Protein-glutamine gamma-glutamyltransferase K |
| 248600 | MAPLE SYRUP URINE DISEASE | P09622 | Dihydrolipoyl dehydrogenase, mitochondrial |
| 248600 | MAPLE SYRUP URINE DISEASE | P11182 | Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial |
| 248600 | MAPLE SYRUP URINE DISEASE | P12694 | 2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial |
| 248600 | MAPLE SYRUP URINE DISEASE | P21953 | 2-oxoisovalerate dehydrogenase subunit beta, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | Q8N183 | Mimitin, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | O15239 | NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 1 |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | P28331 | NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | O43181 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | O75251 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | P49821 | NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | P03897 | NADH-ubiquinone oxidoreductase chain 3 |
| 252010 | MITOCHONDRIAL COMPLEX I DEFICIENCY | O75306 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial |
| 252150 | MOLYBDENUM COFACTOR DEFICIENCY | Q9NQX3 | Gephyrin |
| 252150 | MOLYBDENUM COFACTOR DEFICIENCY | Q9NZB8 | Molybdenum cofactor biosynthesis protein 1 B |
| 252150 | MOLYBDENUM COFACTOR DEFICIENCY | O96007 | Molybdenum cofactor synthesis protein 2 large subunit |
| 252150 | MOLYBDENUM COFACTOR DEFICIENCY | O96033 | Molybdenum cofactor synthesis protein 2 small subunit |
| 252150 | MOLYBDENUM COFACTOR DEFICIENCY | O14940 | Molybdenum cofactor biosynthesis protein 1 A |
| 254200 | MYASTHENIA GRAVIS | P02708 | Acetylcholine receptor subunit alpha |
| 254200 | MYASTHENIA GRAVIS | P11230 | Acetylcholine receptor subunit beta |
| 254200 | MYASTHENIA GRAVIS | Q04844 | Acetylcholine receptor subunit epsilon |
| 254200 | MYASTHENIA GRAVIS | P28329 | Choline O-acetyltransferase |
| 254200 | MYASTHENIA GRAVIS | Q07001 | Acetylcholine receptor subunit delta |
| 254500 | PRIMARY MYELOMA, MULTIPLE AMYLOIDOSIS | P24385 | G1/S-specific cyclin-D1 |
| 254500 | PRIMARY MYELOMA, MULTIPLE AMYLOIDOSIS | P01857 | Ig gamma-1 chain C region |
| 254500 | PRIMARY MYELOMA, MULTIPLE AMYLOIDOSIS | Q15306 | Interferon regulatory factor 4 |
| 254500 | PRIMARY MYELOMA, MULTIPLE AMYLOIDOSIS | P22607 | Fibroblast growth factor receptor 3 |
| 256000 | LEIGH SYNDROME | P00846 | ATP synthase subunit a |
| 256000 | LEIGH SYNDROME | Q12887 | Protoheme IX farnesyltransferase, mitochondrial |
| 256000 | LEIGH SYNDROME | Q7KZN9 | Cytochrome c oxidase assembly protein COX15 homolog |
| 256000 | LEIGH SYNDROME | P31040 | Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial |
| 256000 | LEIGH SYNDROME | O00217 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial |
| 256000 | LEIGH SYNDROME | P49821 | NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial |
| 256000 | LEIGH SYNDROME | P03897 | NADH-ubiquinone oxidoreductase chain 3 |

| | | | |
|---|---|---|---|
| 256000 | LEIGH SYNDROME | Q15526 | Surfeit locus protein 1 |
| 256000 | LEIGH SYNDROME | O75251 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial |
| 259700 | OSTEOPETROSIS, AUTOSOMAL RECESSIVE 1 | P51798 | Chloride channel protein 7 |
| 259700 | OSTEOPETROSIS, AUTOSOMAL RECESSIVE 1 | Q86WC4 | Osteopetrosis-associated transmembrane protein 1 |
| 259700 | OSTEOPETROSIS, AUTOSOMAL RECESSIVE 1 | Q13488 | V-type proton ATPase 116 kDa subunit a isoform 3 |
| 262600 | PITUITARY DWARFISM III | Q9UBX0 | Homeobox expressed in ES cells 1 |
| 262600 | PITUITARY DWARFISM III | O75360 | Homeobox protein prophet of Pit-1 |
| 262600 | PITUITARY DWARFISM III | Q9UBR4 | LIM/homeobox protein Lhx3 |
| 265120 | SURFACTANT METABOLISM DYSFUNCTION, PULMONARY, 1 | P07988 | Pulmonary surfactant-associated protein B |
| 265120 | SURFACTANT METABOLISM DYSFUNCTION, PULMONARY, 1 | P11686 | Pulmonary surfactant-associated protein C |
| 265120 | SURFACTANT METABOLISM DYSFUNCTION, PULMONARY, 1 | P32927 | Cytokine receptor common subunit beta |
| 266600 | INFLAMMATORY BOWEL DISEASE 1 | P22301 | Interleukin-10 |
| 266600 | INFLAMMATORY BOWEL DISEASE 1 | Q5VWK5 | Interleukin-23 receptor |
| 266600 | INFLAMMATORY BOWEL DISEASE 1 | Q9H015 | Solute carrier family 22 member 4 |
| 266600 | INFLAMMATORY BOWEL DISEASE 1 | O76082 | Solute carrier family 22 member 5 |
| 266600 | INFLAMMATORY BOWEL DISEASE 1 | Q9HC29 | Nucleotide-binding oligomerization domain-containing protein 2 |
| 267430 | RENAL TUBULAR DYSGENESIS | P12821 | Angiotensin-converting enzyme |
| 267430 | RENAL TUBULAR DYSGENESIS | P00797 | Renin |
| 267430 | RENAL TUBULAR DYSGENESIS | P30556 | Type-1 angiotensin II receptor |
| 268000 | RETINITIS PIGMENTOSA | P29973 | cGMP-gated cation channel alpha-1 |
| 268000 | RETINITIS PIGMENTOSA | P82279 | Crumbs homolog 1 |
| 268000 | RETINITIS PIGMENTOSA | Q5IJ48 | Crumbs homolog 2 |
| 268000 | RETINITIS PIGMENTOSA | O43186 | Cone-rod homeobox protein |
| 268000 | RETINITIS PIGMENTOSA | Q12866 | Proto-oncogene tyrosine-protein kinase MER |
| 268000 | RETINITIS PIGMENTOSA | P08100 | Rhodopsin |
| 268000 | RETINITIS PIGMENTOSA | P16499 | Rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit alpha |
| 268000 | RETINITIS PIGMENTOSA | P35913 | Rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit beta |
| 268000 | RETINITIS PIGMENTOSA | P23942 | Peripherin-2 |
| 268000 | RETINITIS PIGMENTOSA | P47804 | RPE-retinal G protein-coupled receptor |
| 268000 | RETINITIS PIGMENTOSA | P12271 | Retinaldehyde-binding protein 1 |
| 268000 | RETINITIS PIGMENTOSA | Q03395 | Rod outer segment membrane protein 1 |
| 268000 | RETINITIS PIGMENTOSA | Q8TA86 | Retinitis pigmentosa 9 protein |
| 268000 | RETINITIS PIGMENTOSA | O75445 | Usherin |
| 268220 | RHABDOMYOSARCOMA 2 | P23760 | Paired box protein Pax-3 |
| 268220 | RHABDOMYOSARCOMA 2 | P23759 | Paired box protein Pax-7 |
| 268220 | RHABDOMYOSARCOMA 2 | Q12778 | Forkhead box protein O1 |
| 272200 | MULTIPLE SULFATASE DEFICIENCY | P15289 | Arylsulfatase A |
| 272200 | MULTIPLE SULFATASE DEFICIENCY | Q8NBK3 | Sulfatase-modifying factor 1 |
| 272200 | MULTIPLE SULFATASE DEFICIENCY | P15848 | Arylsulfatase B |
| 275355 | SQUAMOUS CELL CARCINOMA | Q9UK53 | Inhibitor of growth protein 1 |
| 275355 | SQUAMOUS CELL CARCINOMA | Q9NXR8 | Inhibitor of growth protein 3 |
| 275355 | SQUAMOUS CELL CARCINOMA | P04637 | Cellular tumor antigen p53 |

| | | | |
|---|---|---|---|
| 275355 | SQUAMOUS CELL CARCINOMA | O14763 | Tumor necrosis factor receptor superfamily member 10B |
| 275355 | SQUAMOUS CELL CARCINOMA | P60484 | Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN |
| 276300 | TURCOT SYNDROME | P25054 | Adenomatous polyposis coli protein |
| 276300 | TURCOT SYNDROME | P54278 | Mismatch repair endonuclease PMS2 |
| 276300 | TURCOT SYNDROME | P40692 | DNA mismatch repair protein Mlh1 |
| 276900 | USHER SYNDROME, TYPE I | Q9H251 | Cadherin-23 |
| 276900 | USHER SYNDROME, TYPE I | Q13402 | Myosin-VIIa |
| 276900 | USHER SYNDROME, TYPE I | Q96QU1 | Protocadherin-15 |
| 276900 | USHER SYNDROME, TYPE I | Q495M9 | Usher syndrome type-1G protein |
| 276900 | USHER SYNDROME, TYPE I | Q9Y6N9 | Harmonin |
| 277580 | WAARDENBURG-SHAH SYNDROME | P24530 | Endothelin B receptor |
| 277580 | WAARDENBURG-SHAH SYNDROME | P56693 | Transcription factor SOX-10 |
| 277580 | WAARDENBURG-SHAH SYNDROME | P14138 | Endothelin-3 |
| 535000 | LEBER OPTIC ATROPHY | P00846 | ATP synthase subunit a |
| 535000 | LEBER OPTIC ATROPHY | P00395 | Cytochrome c oxidase subunit 1 |
| 535000 | LEBER OPTIC ATROPHY | P00414 | Cytochrome c oxidase subunit 3 |
| 535000 | LEBER OPTIC ATROPHY | P00156 | Cytochrome b |
| 535000 | LEBER OPTIC ATROPHY | P03886 | NADH-ubiquinone oxidoreductase chain 1 |
| 535000 | LEBER OPTIC ATROPHY | P03901 | NADH-ubiquinone oxidoreductase chain 4L |
| 535000 | LEBER OPTIC ATROPHY | P03905 | NADH-ubiquinone oxidoreductase chain 4 |
| 535000 | LEBER OPTIC ATROPHY | P03915 | NADH-ubiquinone oxidoreductase chain 5 |
| 535000 | LEBER OPTIC ATROPHY | P03923 | NADH-ubiquinone oxidoreductase chain 6 |
| 535000 | LEBER OPTIC ATROPHY | P03891 | NADH-ubiquinone oxidoreductase chain 2 |
| 540000 | MELAS SYNDROME | P03905 | NADH-ubiquinone oxidoreductase chain 4 |
| 540000 | MELAS SYNDROME | P03923 | NADH-ubiquinone oxidoreductase chain 6 |
| 540000 | MELAS SYNDROME | P03886 | NADH-ubiquinone oxidoreductase chain 1 |
| 580000 | STREPTOMYCIN OTOTOXICITY | Q969Y2 | tRNA modification GTPase GTPBP3, mitochondrial |
| 580000 | STREPTOMYCIN OTOTOXICITY | Q8WVM0 | Dimethyladenosine transferase 1, mitochondrial |
| 580000 | STREPTOMYCIN OTOTOXICITY | O75648 | Mitochondrial tRNA-specific 2-thiouridylase 1 |
| 601367 | ISCHEMIC STROKE | P12821 | Angiotensin-converting enzyme |
| 601367 | ISCHEMIC STROKE | P20292 | Arachidonate 5-lipoxygenase-activating protein |
| 601367 | ISCHEMIC STROKE | P12259 | Coagulation factor V |
| 601367 | ISCHEMIC STROKE | P05112 | Interleukin-4 |
| 601367 | ISCHEMIC STROKE | P24723 | Protein kinase C eta type |
| 601367 | ISCHEMIC STROKE | P42898 | Methylenetetrahydrofolate reductase |
| 601367 | ISCHEMIC STROKE | P00734 | Prothrombin |
| 601367 | ISCHEMIC STROKE | P16109 | P-selectin |
| 601462 | MYASTHENIC SYNDROME, CONGENITAL, SLOW-CHANNEL | P11230 | Acetylcholine receptor subunit beta |
| 601462 | MYASTHENIC SYNDROME, CONGENITAL, SLOW-CHANNEL | Q07001 | Acetylcholine receptor subunit delta |
| 601462 | MYASTHENIC SYNDROME, CONGENITAL, SLOW-CHANNEL | Q04844 | Acetylcholine receptor subunit epsilon |
| 601462 | MYASTHENIC SYNDROME, CONGENITAL, SLOW-CHANNEL | P02708 | Acetylcholine receptor subunit alpha |
| 601495 | AGAMMAGLOBULINEMIA, NON-BRUTON TYPE, AUTOSOMAL | P15814 | Immunoglobulin lambda-like polypeptide 1 |

| | | | |
|---|---|---|---|
| | RECESSIVE | | |
| 601495 | AGAMMAGLOBULINEMIA, NON-BRUTON TYPE, AUTOSOMAL RECESSIVE | Q8IWT6 | Leucine-rich repeat-containing protein 8A |
| 601495 | AGAMMAGLOBULINEMIA, NON-BRUTON TYPE, AUTOSOMAL RECESSIVE | P11912 | B-cell antigen receptor complex-associated protein alpha-chain |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | O60683 | Peroxisome assembly protein 10 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | O00623 | Peroxisome assembly protein 12 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | Q9Y5Y5 | Peroxisomal membrane protein PEX16 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | P40855 | Peroxisomal biogenesis factor 19 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | O43933 | Peroxisome biogenesis factor 1 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | Q7Z412 | Peroxisome assembly protein 26 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | P28328 | Peroxisome assembly factor 1 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | P56589 | Peroxisomal biogenesis factor 3 |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | O00628 | Peroxisomal targeting signal 2 receptor |
| 601539 | PEROXISOME BIOGENESIS DISORDERS | Q13608 | Peroxisome assembly factor 2 |
| 601634 | NEURAL TUBE DEFECTS, FOLATE-SENSITIVE | P11586 | C-1-tetrahydrofolate synthase, cytoplasmic |
| 601634 | NEURAL TUBE DEFECTS, FOLATE-SENSITIVE | Q99707 | Methionine synthase |
| 601634 | NEURAL TUBE DEFECTS, FOLATE-SENSITIVE | P42898 | Methylenetetrahydrofolate reductase |
| 601634 | NEURAL TUBE DEFECTS, FOLATE-SENSITIVE | Q9UBK8 | Methionine synthase reductase, mitochondrial |
| 601665 | OBESITY LEANNESS | O00253 | Agouti-related protein |
| 601665 | OBESITY LEANNESS | P41159 | Leptin |
| 601665 | OBESITY LEANNESS | P32245 | Melanocortin receptor 4 |
| 601665 | OBESITY LEANNESS | P37231 | Peroxisome proliferator-activated receptor gamma |
| 601665 | OBESITY LEANNESS | Q15466 | Nuclear receptor subfamily 0 group B member 2 |
| 601665 | OBESITY LEANNESS | P55916 | Mitochondrial uncoupling protein 3 |
| 601665 | OBESITY LEANNESS | P01189 | Corticotropin-lipotropin |
| 601675 | TRICHOTHIODYSTROPHY, PHOTOSENSITIVE | P19447 | TFIIH basal transcription factor complex helicase XPB subunit |
| 601675 | TRICHOTHIODYSTROPHY, PHOTOSENSITIVE | Q6ZYL4 | General transcription factor IIH subunit 5 |
| 601675 | TRICHOTHIODYSTROPHY, PHOTOSENSITIVE | P18074 | TFIIH basal transcription factor complex helicase subunit |
| 603174 | HOMOCYSTEINEMIA | Q99707 | Methionine synthase |
| 603174 | HOMOCYSTEINEMIA | Q9Y4U1 | Methylmalonic aciduria and homocystinuria type C protein |
| 603174 | HOMOCYSTEINEMIA | P42898 | Methylenetetrahydrofolate reductase |
| 603174 | HOMOCYSTEINEMIA | P35520 | Cystathionine beta-synthase |
| 603233 | PSEUDOHYPOPARATHYROIDISM, TYPE IB | P84996 | Protein ALEX |
| 603233 | PSEUDOHYPOPARATHYROIDISM, TYPE IB | Q5JWF2 | Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas |
| 603233 | PSEUDOHYPOPARATHYROIDISM, TYPE IB | P63092 | Guanine nucleotide-binding protein G(s) subunit alpha isoforms short |

| | | | |
|---|---|---|---|
| 603233 | PSEUDOHYPOPARATHYROIDISM, TYPE IB | O95467 | Neuroendocrine secretory protein 55 |
| 603233 | PSEUDOHYPOPARATHYROIDISM, TYPE IB | O14662 | Syntaxin-16 |
| 603554 | OMENN SYNDROME | Q96SD1 | Protein artemis |
| 603554 | OMENN SYNDROME | P15918 | V(D)J recombination-activating protein 1 |
| 603554 | OMENN SYNDROME | P55895 | V(D)J recombination-activating protein 2 |
| 603896 | OVARIOLEUKODYSTROPHY | Q14232 | Translation initiation factor eIF-2B subunit alpha |
| 603896 | OVARIOLEUKODYSTROPHY | P49770 | Translation initiation factor eIF-2B subunit beta |
| 603896 | OVARIOLEUKODYSTROPHY | Q9UI10 | Translation initiation factor eIF-2B subunit delta |
| 603896 | OVARIOLEUKODYSTROPHY | Q13144 | Translation initiation factor eIF-2B subunit epsilon |
| 603896 | OVARIOLEUKODYSTROPHY | Q9NR50 | Translation initiation factor eIF-2B subunit gamma |
| 604219 | CATARACT, AUTOSOMAL DOMINo MatchT | Q13515 | Phakinin |
| 604219 | CATARACT, AUTOSOMAL DOMINo MatchT | P43320 | Beta-crystallin B2 |
| 604219 | CATARACT, AUTOSOMAL DOMINo MatchT | P07315 | Gamma-crystallin C |
| 604219 | CATARACT, AUTOSOMAL DOMINo MatchT | P07320 | Gamma-crystallin D |
| 604219 | CATARACT, AUTOSOMAL DOMINo MatchT | P02489 | Alpha-crystallin A chain |
| 604229 | PETERS ANOMALY | Q16678 | Cytochrome P450 1B1 |
| 604229 | PETERS ANOMALY | Q12948 | Forkhead box protein C1 |
| 604229 | PETERS ANOMALY | P26367 | Paired box protein Pax-6 |
| 604229 | PETERS ANOMALY | Q99697 | Pituitary homeobox 2 |
| 604229 | PETERS ANOMALY | P61812 | Transforming growth factor beta-2 |
| 604229 | PETERS ANOMALY | Q9UKV0 | Histone deacetylase 9 |
| 604233 | GENERALIZED EPILEPSY WITH FEBRILE SEIZURES PLUS | P35498 | Sodium channel protein type 1 subunit alpha |
| 604233 | GENERALIZED EPILEPSY WITH FEBRILE SEIZURES PLUS | Q07699 | Sodium channel subunit beta-1 |
| 604233 | GENERALIZED EPILEPSY WITH FEBRILE SEIZURES PLUS | Q99250 | Sodium channel protein type 2 subunit alpha |
| 604233 | GENERALIZED EPILEPSY WITH FEBRILE SEIZURES PLUS | P18507 | Gamma-aminobutyric acid receptor subunit gamma-2 |
| 604271 | SHORT STATURE, IDIOPATHIC, AUTOSOMAL | P10912 | Growth hormone receptor |
| 604271 | SHORT STATURE, IDIOPATHIC, AUTOSOMAL | P01241 | Somatotropin |
| 604271 | SHORT STATURE, IDIOPATHIC, AUTOSOMAL | Q92847 | Growth hormone secretagogue receptor type 1 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5F3 | Protocadherin beta-1 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E7 | Protocadherin beta-2 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E6 | Protocadherin beta-3 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E5 | Protocadherin beta-4 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E4 | Protocadherin beta-5 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E3 | Protocadherin beta-6 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E2 | Protocadherin beta-7 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9UN66 | Protocadherin beta-8 |

| | | | |
|---|---|---|---|
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E1 | Protocadherin beta-9 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9UN67 | Protocadherin beta-10 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5F1 | Protocadherin beta-12 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5F0 | Protocadherin beta-13 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E9 | Protocadherin beta-14 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5E8 | Protocadherin beta-15 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9NRJ7 | Protocadherin beta-16 |
| 604967 | PROTOCADHERIN-BETA GENE CLUSTER | Q9Y5F2 | Protocadherin beta-11 |
| 605074 | RENAL CELL CARCINOMA, PAPILLARY | Q9BZE9 | Tether containing UBX domain for GLUT4 |
| 605074 | RENAL CELL CARCINOMA, PAPILLARY | Q92733 | Proline-rich protein PRCC |
| 605074 | RENAL CELL CARCINOMA, PAPILLARY | P19532 | Transcription factor E3 |
| 605074 | RENAL CELL CARCINOMA, PAPILLARY | P08581 | Hepatocyte growth factor receptor |
| 605899 | GLYCINE ENCEPHALOPATHY | P23378 | Glycine dehydrogenase [decarboxylating], mitochondrial |
| 605899 | GLYCINE ENCEPHALOPATHY | P48728 | Aminomethyltransferase, mitochondrial |
| 605899 | GLYCINE ENCEPHALOPATHY | P23434 | Glycine cleavage system H protein, mitochondrial |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | P19835 | Bile salt-activated lipase |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | P20823 | Hepatocyte nuclear factor 1-alpha |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | P35680 | Hepatocyte nuclear factor 1-beta |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | P41235 | Hepatocyte nuclear factor 4-alpha |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | P35557 | Glucokinase |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | Q13562 | Neurogenic differentiation factor 1 |
| 606391 | MATURITY-ONSET DIABETES OF THE YOUNG | O14901 | Krueppel-like factor 11 |
| 606904 | JUVENILE MYOCLONIC EPILEPSY | O00305 | Voltage-dependent L-type calcium channel subunit beta-4 |
| 606904 | JUVENILE MYOCLONIC EPILEPSY | Q5JVL4 | EF-hand domain-containing protein 1 |
| 606904 | JUVENILE MYOCLONIC EPILEPSY | P14867 | Gamma-aminobutyric acid receptor subunit alpha-1 |
| 607748 | FAMILIAL HYPERCHOLANEMIA | P07099 | Epoxide hydrolase 1 |
| 607748 | FAMILIAL HYPERCHOLANEMIA | Q9UDY2 | Tight junction protein ZO-2 |
| 607748 | FAMILIAL HYPERCHOLANEMIA | Q14032 | Bile acid-CoA:amino acid N-acyltransferase |
| 607785 | JUVENILE MYELOMONOCYTIC LEUKEMIA | Q06124 | Tyrosine-protein phosphatase non-receptor type 11 |
| 607785 | JUVENILE MYELOMONOCYTIC LEUKEMIA | P01116 | GTPase KRas |
| 607785 | JUVENILE MYELOMONOCYTIC LEUKEMIA | P01111 | GTPase NRas |
| 607785 | JUVENILE MYELOMONOCYTIC LEUKEMIA | Q9UNA1 | Rho GTPase-activating protein 26 |
| 607785 | JUVENILE MYELOMONOCYTIC LEUKEMIA | P21359 | Neurofibromin |
| 608089 | ENDOMETRIAL CANCER | P12830 | Epithelial cadherin |
| 608089 | ENDOMETRIAL CANCER | P40692 | DNA mismatch repair protein Mlh1 |

| | | | |
|---|---|---|---|
| 608089 | ENDOMETRIAL CANCER | P43246 | DNA mismatch repair protein Msh2 |
| 608089 | ENDOMETRIAL CANCER | P20585 | DNA mismatch repair protein Msh3 |
| 608089 | ENDOMETRIAL CANCER | P52701 | DNA mismatch repair protein Msh6 |
| 608089 | ENDOMETRIAL CANCER | P60484 | Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN |
| 608930 | MYASTHENIC SYNDROME, CONGENITAL, FAST-CHANNEL | P02708 | Acetylcholine receptor subunit alpha |
| 608930 | MYASTHENIC SYNDROME, CONGENITAL, FAST-CHANNEL | Q04844 | Acetylcholine receptor subunit epsilon |
| 608930 | MYASTHENIC SYNDROME, CONGENITAL, FAST-CHANNEL | Q07001 | Acetylcholine receptor subunit delta |
| 608931 | MYASTHENIC SYNDROME, CONGENITAL, ASSOCIATED WITH ACETYLCHOLINE RECEPTOR DEFICIENCY | P11230 | Acetylcholine receptor subunit beta |
| 608931 | MYASTHENIC SYNDROME, CONGENITAL, ASSOCIATED WITH ACETYLCHOLINE RECEPTOR DEFICIENCY | Q04844 | Acetylcholine receptor subunit epsilon |
| 608931 | MYASTHENIC SYNDROME, CONGENITAL, ASSOCIATED WITH ACETYLCHOLINE RECEPTOR DEFICIENCY | Q13702 | 43 kDa receptor-associated protein of the synapse |
| 608931 | MYASTHENIC SYNDROME, CONGENITAL, ASSOCIATED WITH ACETYLCHOLINE RECEPTOR DEFICIENCY | O15146 | Muscle, skeletal receptor tyrosine protein kinase |
| 608971 | SEVERE COMBINED IMMUNODEFICIENCY | P04234 | T-cell surface glycoprotein CD3 delta chain |
| 608971 | SEVERE COMBINED IMMUNODEFICIENCY | P16871 | Interleukin-7 receptor subunit alpha |
| 608971 | SEVERE COMBINED IMMUNODEFICIENCY | P08575 | Leukocyte common antigen |
| 609423 | SUSCEPTIBILITY TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 | P51681 | C-C chemokine receptor type 5 |
| 609423 | SUSCEPTIBILITY TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 | P49238 | CX3C chemokine receptor 1 |
| 609423 | SUSCEPTIBILITY TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 | P41597 | C-C chemokine receptor type 2 |
| 609830 | ABDOMINAL BODY FAT DISTRIBUTION | P37231 | Peroxisome proliferator-activated receptor gamma |
| 609830 | ABDOMINAL BODY FAT DISTRIBUTION | P18031 | Tyrosine-protein phosphatase non-receptor type 1 |
| 609830 | ABDOMINAL BODY FAT DISTRIBUTION | P01189 | Corticotropin-lipotropin |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | P01903 | HLA class II histocompatibility antigen, DR alpha chain |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | P16410 | Cytotoxic T-lymphocyte protein 4 |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | Q08334 | Interleukin-10 receptor beta chain |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | P48551 | Interferon-alpha/beta receptor beta chain |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | P11226 | Mannose-binding protein C |
| 610424 | SUSCEPTIBILITY TO HEPATITIS B VIRUS | P01375 | Tumor necrosis factor |

# References


[1] T. F. Smith and M. S. Waterman, "Identification of common molecular subsequences," *J Mol Biol,* vol. 147, pp. 195-7, Mar 25 1981.

[2] S. B. Needleman and C. D. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *J Mol Biol,* vol. 48, pp. 443-53, Mar 1970.

[3] O. Gotoh, "An improved algorithm for matching biological sequences," *J Mol Biol,* vol. 162, pp. 705-8, Dec 15 1982.

[4] S. F. Altschul and B. W. Erickson, "Optimal sequence alignment using affine gap costs," *Bull Math Biol,* vol. 48, pp. 603-16, 1986.

[5] D. J. Lipman, S. F. Altschul, and J. D. Kececioglu, "A tool for multiple sequence alignment," *Proc Natl Acad Sci U S A,* vol. 86, pp. 4412-5, Jun 1989.

[6] W. R. Pearson and D. J. Lipman, "Improved tools for biological sequence comparison," *Proc Natl Acad Sci U S A,* vol. 85, pp. 2444-8, Apr 1988.

[7] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Res,* vol. 22, pp. 4673-80, Nov 11 1994.

[8] C. Notredame, D. G. Higgins, and J. Heringa, "T-Coffee: A novel method for fast and accurate multiple sequence alignment," *J Mol Biol,* vol. 302, pp. 205-17, Sep 8 2000.

[9] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Res,* vol. 32, pp. 1792-7, 2004.

[10] C. Grasso and C. Lee, "Combining partial order alignment and progressive multiple sequence alignment increases alignment speed and scalability to very large alignment problems," *Bioinformatics,* vol. 20, pp. 1546-56, Jul 10 2004.

[11] R. Hughey and A. Krogh, "Hidden Markov models for sequence analysis: extension and analysis of the basic method," *Comput Appl Biosci,* vol. 12, pp. 95-107, Apr 1996.

[12] S. R. Eddy, "A probabilistic model of local sequence alignment that simplifies statistical significance estimation," *PLoS Comput Biol,* vol. 4, p. e1000069, May 2008.

[13] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *J Mol Biol,* vol. 215, pp. 403-10, Oct 5 1990.

[14] L. Holm and C. Sander, "Mapping the protein universe," *Science,* vol. 273, pp. 595-603, Aug 2 1996.

[15] I. N. Shindyalov and P. E. Bourne, "Protein structure alignment by incremental combinatorial extension (CE) of the optimal path," *Protein Eng,* vol. 11, pp. 739-47, Sep 1998.

[16] A. Guerler and E. W. Knapp, "Novel protein folds and their nonsequential structural analogs," *Protein Sci,* vol. 17, pp. 1374-82, Aug 2008.

[17] A. R. Ortiz, C. E. Strauss, and O. Olmea, "MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison," *Protein Sci,* vol. 11, pp. 2606-21, Nov 2002.

[18] R. Mosca and T. R. Schneider, "RAPIDO: a web server for the alignment of protein structures in the presence of conformational changes," *Nucleic Acids Res,* vol. 36, pp. W42-6, Jul 1 2008.

[19] F. Teichert, U. Bastolla, and M. Porto, "SABERTOOTH: protein structural alignment based on a vectorial structure representation," *BMC Bioinformatics,* vol. 8, p. 425, 2007.

[20] W. R. Taylor, T. P. Flores, and C. A. Orengo, "Multiple protein structure alignment," *Protein Sci,* vol. 3, pp. 1858-70, Oct 1994.

[21] E. Krissinel and K. Henrick, "Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions," *Acta Crystallogr D Biol Crystallogr,* vol. 60, pp. 2256-68, Dec 2004.

[22] Y. Zhang and J. Skolnick, "TM-align: a protein structure alignment algorithm based on the TM-score," *Nucleic Acids Res,* vol. 33, pp. 2302-9, 2005.

[23] S. A. Canny, Y. Cruz, M. R. Southern, and P. R. Griffin, "PubChem promiscuity: a web resource for gathering compound promiscuity data from PubChem," *Bioinformatics,* vol. 28, pp. 140-1, Jan 1 2012.

[24] Y. Wang, E. Bolton, S. Dracheva, K. Karapetyan, B. A. Shoemaker, T. O. Suzek, *et al.*, "An overview of the PubChem BioAssay resource," *Nucleic Acids Res,* vol. 38, pp. D255-66, Jan 2010.

[25] S. Gunther, M. Kuhn, M. Dunkel, M. Campillos, C. Senger, E. Petsalaki, *et al*., "SuperTarget and Matador: resources for exploring drug-target relationships," *Nucleic Acids Res,* vol. 36, pp. D919-22, Jan 2008.

[26] M. Campillos, M. Kuhn, A. C. Gavin, L. J. Jensen, and P. Bork, "Drug target identification using side-effect similarity," *Science,* vol. 321, pp. 263-6, Jul 11 2008.

[27] D. B. Allison, X. Cui, G. P. Page, and M. Sabripour, "Microarray data analysis: from disarray to consolidation and consensus," *Nat Rev Genet,* vol. 7, pp. 55-65, Jan 2006.

[28] T. Mehta, M. Tanik, and D. B. Allison, "Towards sound epistemological foundations of statistical methods for high-dimensional biology," *Nat Genet,* vol. 36, pp. 943-7, Sep 2004.

[29] H. K. Lee, A. K. Hsu, J. Sajdak, J. Qin, and P. Pavlidis, "Coexpression analysis of human genes across many microarray data sets," *Genome Res,* vol. 14, pp. 1085-94, Jun 2004.

[30] D. Eisenberg, E. M. Marcotte, I. Xenarios, and T. O. Yeates, "Protein function in the post-genomic era," *Nature,* vol. 405, pp. 823-6, Jun 15 2000.

[31] M. R. Trusheim, B. Burgess, S. X. Hu, T. Long, S. D. Averbuch, A. A. Flynn, *et al*., "Quantifying factors for the success of stratified medicine," *Nat Rev Drug Discov,* vol. 10, pp. 817-33, 2011.

[32] Z. Dezso, Y. Nikolsky, T. Nikolskaya, J. Miller, D. Cherba, C. Webb, *et al*., "Identifying disease-specific genes based on their topological significance in protein networks," *BMC Syst Biol,* vol. 3, p. 36, 2009.

[33] J. B. Gibbs, "Mechanism-based target identification and drug discovery in cancer research," *Science,* vol. 287, pp. 1969-73, Mar 17 2000.

[34] K. A. Pattin and J. H. Moore, "Role for protein-protein interaction databases in human genetics," *Expert Rev Proteomics,* vol. 6, pp. 647-59, Dec 2009.

[35] P. Imming, C. Sinning, and A. Meyer, "Drugs, their targets and the nature and number of drug targets," *Nat Rev Drug Discov,* vol. 5, pp. 821-34, Oct 2006.

[36] D. C. Liebler and F. P. Guengerich, "Elucidating mechanisms of drug-induced toxicity," *Nat Rev Drug Discov,* vol. 4, pp. 410-20, May 2005.

[37] H. Najmabadi, H. Hu, M. Garshasbi, T. Zemojtel, S. S. Abedini, W. Chen, *et al*., "Deep sequencing reveals 50 novel genes for recessive cognitive disorders," *Nature,* vol. 478, pp. 57-63, Oct 6 2011.

[38] M. G. Tektonidou and M. M. Ward, "Validation of new biomarkers in systemic autoimmune diseases," *Nat Rev Rheumatol,* Nov 1 2011.

[39] E. C. Butcher, "Can cell systems biology rescue drug discovery?," *Nat Rev Drug Discov,* vol. 4, pp. 461-7, Jun 2005.

[40] J. Drews, "Strategic trends in the drug industry," *Drug Discovery Today,* vol. 8, pp. 411-420, May 1 2003.

[41] J. A. DiMasi, R. W. Hansen, and H. G. Grabowski, "The price of innovation: new estimates of drug development costs," *Journal of Health Economics,* vol. 22, pp. 151-185, Mar 2003.

[42] CBO. (2006, January 20). Research and Development in the Pharmaceutical Industry. *A CBO Study*. Available: http://www.cbo.gov/ftpdocs/76xx/doc7615/10-02-DrugR-D.pdf

[43] C. P. Adams and V. V. Brantner, "Estimating the cost of new drug development: Is it really $802 million?," *Health Affairs,* vol. 25, pp. 420-428, Mar-Apr 2006.

[44] E. A. Czerepak and S. Ryser, "Drug approvals and failures: implications for alliances," *Nature Reviews Drug Discovery,* vol. 7, pp. 197-198, Mar 2008.

[45] K. Wester, A. K. Jonsson, O. Spigset, H. Druid, and S. Hagg, "Incidence of fatal adverse drug reactions: a population based study," *British Journal of Clinical Pharmacology,* vol. 65, pp. 573-579, Apr 2008.

[46] J. Lazarou, B. H. Pomeranz, and P. N. Corey, "Incidence of adverse drug reactions in hospitalized patients - A meta-analysis of prospective studies," *Jama-Journal of the American Medical Association,* vol. 279, pp. 1200-1205, Apr 15 1998.

[47] J. T. Hanlon, C. F. Pieper, E. R. Hajjar, R. J. Sloane, C. I. Lindblad, C. M. Ruby, *et al*., "Incidence and predictors of all and preventable adverse drug reactions in frail elderly persons after hospital stay," *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* vol. 61, pp. 511-515, May 2006.

[48] A. E. Lobley, T. Nugent, C. A. Orengo, and D. T. Jones, "FFPred: an integrated feature-based function prediction server for vertebrate proteomes," *Nucleic Acids Res,* vol. 36, pp. W297-302, Jul 1 2008.

[49] E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman*, et al*., "A comprehensive catalogue of somatic mutations from a human cancer genome," *Nature,* vol. 463, pp. 191-6, Jan 14 2010.

[50] M. R. Stratton, P. J. Campbell, and P. A. Futreal, "The cancer genome," *Nature,* vol. 458, pp. 719-24, Apr 9 2009.

[51] M. Williams, "Target validation," *Curr Opin Pharmacol,* vol. 3, pp. 571-7, Oct 2003.

[52] M. Rask-Andersen, M. S. Almen, and H. B. Schioth, "Trends in the exploitation of novel drug targets," *Nat Rev Drug Discov,* vol. 10, pp. 579-90, Aug 2011.

[53] B. P. Zambrowicz and A. T. Sands, "Knockouts model the 100 best-selling drugs--will they model the next 100?," *Nat Rev Drug Discov,* vol. 2, pp. 38-51, Jan 2003.

[54] T. Takenaka, "Classical vs reverse pharmacology in drug discovery," *BJU Int,* vol. 88 Suppl 2, pp. 7-10; discussion 49-50, Sep 2001.

[55] C. Smith, "Drug target identification: a question of biology," *Nature,* vol. 428, pp. 225-31, Mar 11 2004.

[56] G. J. Crowther, D. Shanmugam, S. J. Carmona, M. A. Doyle, C. Hertz-Fowler, M. Berriman*, et al*., "Identification of attractive drug targets in neglected-disease pathogens using an in silico approach," *PLoS Negl Trop Dis,* vol. 4, p. e804, 2010.

[57] C. R. Caffrey, A. Rohwer, F. Oellien, R. J. Marhofer, S. Braschi, G. Oliveira*, et al*., "A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, Schistosoma mansoni," *PLoS One,* vol. 4, p. e4413, 2009.

[58] A. Schlicker and M. Albrecht, "FunSimMat: a comprehensive functional similarity database," *Nucleic Acids Res,* vol. 36, pp. D434-9, Jan 2008.

[59] A. Schlicker, F. S. Domingues, J. Rahnenfuhrer, and T. Lengauer, "A new measure for functional similarity of gene products based on Gene Ontology," *BMC Bioinformatics,* vol. 7, p. 302, 2006.

[60] A. Schlicker, T. Lengauer, and M. Albrecht, "Improving disease gene prioritization using the semantic similarity of Gene Ontology terms," *Bioinformatics,* vol. 26, pp. i561-7, Sep 15 2010.

[61] C. Pesquita, D. Faria, H. Bastos, A. E. Ferreira, A. O. Falcao, and F. M. Couto, "Metrics for GO based protein semantic similarity: a systematic evaluation," *BMC Bioinformatics,* vol. 9 Suppl 5, p. S4, 2008.

[62] G. Lerman and B. E. Shakhnovich, "Defining functional distance using manifold embeddings of gene ontology annotations," *Proc Natl Acad Sci U S A,* vol. 104, pp. 11334-9, Jul 3 2007.

[63] F. M. Couto, M\, \#225, r. J. Silva, and P. M. Coutinho, "Semantic similarity over the gene ontology: family correlation and selecting disjunctive ancestors," presented at the Proceedings of the 14th ACM international conference on Information and knowledge management, Bremen, Germany, 2005.

[64] F. M. Couto, M. J. Silva, and P. M. Coutinho, "Measuring semantic similarity between Gene Ontology terms," *Data & Knowledge Engineering,* vol. 61, pp. 137-152, 2007.

[65] A. del Pozo, F. Pazos, and A. Valencia, "Defining functional distances over gene ontology," *BMC Bioinformatics,* vol. 9, p. 50, 2008.

[66] S. Benabderrahmane, M. Smail-Tabbone, O. Poch, A. Napoli, and M. D. Devignes, "IntelliGO: a new vector-based semantic similarity measure including annotation origin," *BMC Bioinformatics,* vol. 11, p. 588, 2010.

[67] Y. R. Cho, W. Hwang, M. Ramanathan, and A. Zhang, "Semantic integration to identify overlapping functional modules in protein interaction networks," *BMC Bioinformatics,* vol. 8, p. 265, 2007.

[68] X. Guo, R. Liu, C. D. Shriver, H. Hu, and M. N. Liebman, "Assessing semantic similarity measures for the characterization of human regulatory pathways," *Bioinformatics,* vol. 22, pp. 967-73, Apr 15 2006.

[69] P. W. Lord, R. D. Stevens, A. Brass, and C. A. Goble, "Investigating semantic similarity measures across the Gene Ontology: the relationship between sequence and annotation," *Bioinformatics,* vol. 19, pp. 1275-83, Jul 1 2003.

[70] P. W. Lord, R. D. Stevens, A. Brass, and C. A. Goble, "Semantic similarity measures as tools for exploring the gene ontology," *Pac Symp Biocomput,* pp. 601-12, 2003.

[71] M. Mistry and P. Pavlidis, "Gene Ontology term overlap as a measure of gene functional similarity," *BMC Bioinformatics,* vol. 9, p. 327, 2008.

[72] C. Posse, A. Sanfilippo, B. Gopalan, R. Riensche, N. Beagley, and B. Baddeley, "Cross-Ontological Analytics: Combining Associative and Hierarchical Relations in the Gene Ontologies to Assess Gene Product Similarity," in *Computational Science – ICCS 2006*. vol. 3992, V. Alexandrov, G. van Albada, P. Sloot, and J. Dongarra, Eds., ed: Springer Berlin / Heidelberg, 2006, pp. 871-878-878.

[73] J. Z. Wang, Z. Du, R. Payattakool, P. S. Yu, and C. F. Chen, "A new method to measure the semantic similarity of GO terms," *Bioinformatics,* vol. 23, pp. 1274-81, May 15 2007.

[74] H. Wu, Z. Su, F. Mao, V. Olman, and Y. Xu, "Prediction of functional modules based on comparative genome analysis and Gene Ontology application," *Nucleic Acids Res,* vol. 33, pp. 2822-37, 2005.

[75] X. Wu, L. Zhu, J. Guo, D. Y. Zhang, and K. Lin, "Prediction of yeast protein-protein interaction network: insights from the Gene Ontology and annotations," *Nucleic Acids Res,* vol. 34, pp. 2137-50, 2006.

[76] H. Yu, L. Gao, K. Tu, and Z. Guo, "Broadly predicting specific gene functions with expression similarity and taxonomy similarity," *Gene,* vol. 352, pp. 75-81, Jun 6 2005.

[77] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, *et al*., "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium," *Nat Genet,* vol. 25, pp. 25-9, May 2000.

[78] B. Smith, J. Williams, and S. Schulze-Kremer, "The ontology of the gene ontology," *AMIA Annu Symp Proc,* pp. 609-13, 2003.

[79] Y. Tao, L. Sam, J. Li, C. Friedman, and Y. A. Lussier, "Information theory applied to the sparse gene ontology annotation network to predict novel gene function," *Bioinformatics,* vol. 23, pp. i529-38, Jul 1 2007.

[80] UniProt, "Ongoing and future developments at the Universal Protein Resource," *Nucleic Acids Res,* vol. 39, pp. D214-9, Jan 2011.

[81] H. Mcwilliam, F. Valentin, M. Goujon, W. Z. Li, M. Narayanasamy, J. Martin, *et al.*, "Web services at the European Bioinformatics Institute-2009," *Nucleic Acids Research,* vol. 37, pp. W6-W10, Jul 1 2009.

[82] Y. Arita, S. Nishimura, A. Matsuyama, Y. Yashiroda, T. Usui, C. Boone, *et al.*, "Microarray-based target identification using drug hypersensitive fission yeast expressing ORFeome," *Mol Biosyst,* vol. 7, pp. 1463-72, May 2011.

[83] M. T. Dittrich, G. W. Klau, A. Rosenwald, T. Dandekar, and T. Muller, "Identifying functional modules in protein-protein interaction networks: an integrated exact approach," *Bioinformatics,* vol. 24, pp. i223-31, Jul 1 2008.

[84] O. Bodenreider, M. Aubry, and A. Burgun, "Non-lexical approaches to identifying associative relations in the gene ontology," *Pac Symp Biocomput,* pp. 91-102, 2005.

[85] Y. R. Cho and A. Zhang, "Identification of functional hubs and modules by converting interactome networks into hierarchical ordering of proteins," *BMC Bioinformatics,* vol. 11 Suppl 3, p. S3, 2010.

[86] R. A. Craig and L. Liao, "Phylogenetic tree information aids supervised learning for predicting protein-protein interaction based on distance matrices," *BMC Bioinformatics,* vol. 8, p. 6, 2007.

[87] J. L. Sevilla, V. Segura, A. Podhorski, E. Guruceaga, J. M. Mato, L. A. Martinez-Cruz, *et al.*, "Correlation between gene expression and GO semantic similarity," *IEEE/ACM Trans Comput Biol Bioinform,* vol. 2, pp. 330-8, Oct-Dec 2005.

[88] A. Lewin and I. C. Grieve, "Grouping Gene Ontology terms to improve the assessment of gene set enrichment in microarray data," *BMC Bioinformatics,* vol. 7, p. 426, 2006.

[89] R. Jansen, H. Yu, D. Greenbaum, Y. Kluger, N. J. Krogan, S. Chung, *et al.*, "A Bayesian networks approach for predicting protein-protein interactions from genomic data," *Science,* vol. 302, pp. 449-53, Oct 17 2003.

[90] P. Resnik, "Using Information Content to Evaluate Semantic Similarity in a Taxonomy," *In Proceedings of the 14th International Joint Conference on Artificial Intelligence,* pp. 448-453, 1995.

[91] D. Lin, "An Information-Theoretic Definition of Similarity," in *In Proceedings of the 15th International Conference on Machine Learning*, 1998, pp. 296-304.

[92] B. Rost, "Enzyme function less conserved than anticipated," *J Mol Biol,* vol. 318, pp. 595-608, Apr 26 2002.

[93] S. E. Brenner, "Errors in genome annotation," *Trends Genet,* vol. 15, pp. 132-3, Apr 1999.

[94] D. Devos and A. Valencia, "Practical limits of function prediction," *Proteins,* vol. 41, pp. 98-107, Oct 1 2000.

[95] R. Gentleman. (2010). *Visualizing and Distances Using GO*. Available: http://www.bioconductor.org/packages/release/bioc/vignettes/GOstats/inst/doc/GOvis.pdf

[96] A. G. Murzin, S. E. Brenner, T. Hubbard, and C. Chothia, "SCOP: a structural classification of proteins database for the investigation of sequences and structures," *J Mol Biol,* vol. 247, pp. 536-40, Apr 7 1995.

[97] S. Hunter, P. Jones, A. Mitchell, R. Apweiler, T. K. Attwood, A. Bateman, *et al*., "InterPro in 2011: new developments in the family and domain prediction database," *Nucleic Acids Res,* vol. 40, pp. D306-D312, Jan 2012.

[98] J. J. Jiang and D. W. Conrath, "Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy," in *International Conference Research on Computational Linguistics (ROCLING X)*, 1997, p. 9008.

[99] R. M. Riensche, B. L. Baddeley, A. P. Sanfilippo, C. Posse, and B. Gopalan, "XOA: Web-Enabled Cross-Ontological Analytics," in *Services, 2007 IEEE Congress on*, 2007, pp. 99-105.

[100] A. Sanfilippo, C. Posse, B. Gopalan, R. Riensche, N. Beagley, B. Baddeley, *et al*., "Combining Hierarchical and Associative Gene Ontology Relations With Textual Evidence in Estimating Gene and Gene Product Similarity," *NanoBioscience, IEEE Transactions on,* vol. 6, pp. 51-59, 2007.

[101] P. Ye, B. D. Peyser, X. Pan, J. D. Boeke, F. A. Spencer, and J. S. Bader, "Gene function prediction from congruent synthetic lethal interactions in yeast," *Mol Syst Biol,* vol. 1, p. 2005 0026, 2005.

[102] V. Pekar and S. Staab, "Taxonomy learning: factoring the structure of a taxonomy into a semantic classification decision," in *Proceedings of the Nineteenth Conference on Computational Linguistics*, 2002, pp. 786-792.

[103] A. Nagar and H. Al-Mubaid, "Using path length measure for gene clustering based on similarity of annotation terms," in *Computers and Communications, 2008. ISCC 2008. IEEE Symposium on*, 2008, pp. 637-642.

[104] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing Images Using the Hausdorff Distance," *Ieee Transactions on Pattern Analysis and Machine Intelligence,* vol. 15, pp. 850-863, Sep 1993.

[105] X. X. Chi and J. Y. Hou, "An iterative approach of protein function prediction," *Bmc Bioinformatics,* vol. 12, Nov 10 2011.

[106] P. Khatri and S. Draghici, "Ontological analysis of gene expression data: current tools, limitations, and open problems," *Bioinformatics,* vol. 21, pp. 3587-3595, Sep 15 2005.

[107] S. Y. Rhee, V. Wood, K. Dolinski, and S. Draghici, "Use and misuse of the gene ontology annotations," *Nature Reviews Genetics,* vol. 9, pp. 509-515, Jul 2008.

[108] P. Khatri, B. Done, A. Rao, A. Done, and S. Draghici, "A semantic analysis of the annotations of the human genome," *Bioinformatics,* vol. 21, pp. 3416-21, Aug 15 2005.

[109] A. Nagar and H. Al-Mubaid, "A New Path Length Measure Based on GO for Gene Similarity with Evaluation using SGD Pathways," in *Computer-Based Medical Systems, 2008. CBMS '08. 21st IEEE International Symposium on*, 2008, pp. 590-595.

[110] M. Y. Galperin, D. R. Walker, and E. V. Koonin, "Analogous enzymes: independent inventions in enzyme evolution," *Genome Res,* vol. 8, pp. 779-90, Aug 1998.

[111] D. Devos and A. Valencia, "Intrinsic errors in genome annotation," *Trends Genet,* vol. 17, pp. 429-31, Aug 2001.

[112] R. D. Sleator and P. Walsh, "An overview of in silico protein function prediction," *Arch Microbiol,* vol. 192, pp. 151-5, Mar 2010.

[113] B. Stein, "Principles of hash-based text retrieval," in *Proceedings of the 30th annual international ACM SIGIR conference*, Amsterdam, Netherlands, 2007.

[114] X. W. Chen, J. C. Jeong, and P. Dermyer, "KUPS: constructing datasets of interacting and non-interacting protein pairs with associated attributions," *Nucleic Acids Res,* Oct 15 2010.

[115] J. Ling, B. R. So, S. S. Yadavalli, H. Roy, S. Shoji, K. Fredrick, *et al*., "Resampling and editing of mischarged tRNA prior to translation elongation," *Mol Cell,* vol. 33, pp. 654-60, Mar 13 2009.

[116] H. S. Zaher and R. Green, "Fidelity at the Molecular Level: Lessons from Protein Synthesis," *Cell,* vol. 136, pp. 746-762, Feb 20 2009.

[117] E. N. Kersh, A. S. Shaw, and P. M. Allen, "Fidelity of T cell activation through multistep T cell receptor zeta phosphorylation," *Science,* vol. 281, pp. 572-575, Jul 24 1998.

[118] T. Xu, L. F. Du, and Y. Zhou, "Evaluation of GO-based functional similarity measures using S. cerevisiae protein interaction and expression profile data," *BMC Bioinformatics,* vol. 9, Nov 6 2008.

[119] C. Chothia, J. Gough, C. Vogel, and S. A. Teichmann, "Evolution of the protein repertoire," *Science,* vol. 300, pp. 1701-3, Jun 13 2003.

[120] B. Long, P. S. Yu, and Z. Zhang, "A General Model for Multiple View Unsupervised Learning," in *In Proceedings of the 2008 SIAM International Conference on Data Mining*, Atlanta, Georgia, USA, 2008, pp. 822-833.

[121] S. Rüping and T. Scheffer, "Learning With Multiple Views," in *Proceedings of the ICML 2005 Workshop*, Bonn, Germany, 2005.

[122] A. Andreeva, D. Howorth, J. M. Chandonia, S. E. Brenner, T. J. Hubbard, C. Chothia, *et al*., "Data growth and its impact on the SCOP database: new developments," *Nucleic Acids Res,* vol. 36, pp. D419-25, Jan 2008.

[123] C. A. Orengo, A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells, and J. M. Thornton, "CATH--a hierarchic classification of protein domain structures," *Structure,* vol. 5, pp. 1093-108, Aug 15 1997.

[124] S. Hunter, P. Jones, A. Mitchell, R. Apweiler, T. K. Attwood, A. Bateman, *et al*., "InterPro in 2011: new developments in the family and domain prediction database," *Nucleic Acids Research,* vol. 40, pp. D306-D312, Jan 2012.

[125] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, *et al*., "The Protein Data Bank," *Nucleic Acids Research,* vol. 28, pp. 235-242, Jan 1 2000.

[126] E. M. Marcotte, M. Pellegrini, H. L. Ng, D. W. Rice, T. O. Yeates, and D. Eisenberg, "Detecting protein function and protein-protein interactions from genome sequences," *Science,* vol. 285, pp. 751-3, Jul 30 1999.

[127] S. Letovsky and S. Kasif, "Predicting protein function from protein/protein interaction data: a probabilistic approach," *Bioinformatics,* vol. 19, pp. i197-i204, Jul 2003.

[128] A. Vazquez, A. Flammini, A. Maritan, and A. Vespignani, "Global protein function prediction from protein-protein interaction networks," *Nat Biotechnol,* vol. 21, pp. 697-700, Jun 2003.

[129] J. Muller and A. Hemphill, "Drug Target Identification in Intracellular and Extracellular Protozoan Parasites," *Curr Top Med Chem,* May 26 2011.

[130] X. W. Chen, J. C. Jeong, and P. Dermyer, "KUPS: constructing datasets of interacting and non-interacting protein pairs with associated attributions," *Nucleic Acids Res,* vol. 39, pp. D750-4, Jan 2011.

[131] D. Wilson, R. Pethica, Y. D. Zhou, C. Talbot, C. Vogel, M. Madera*, et al.*, "SUPERFAMILY-sophisticated comparative genomics, data mining, visualization and phylogeny," *Nucleic Acids Research,* vol. 37, pp. D380-D386, Jan 2009.

[132] D. Lee, O. Redfern, and C. Orengo, "Predicting protein function from sequence and structure," *Nat Rev Mol Cell Biol,* vol. 8, pp. 995-1005, Dec 2007.

[133] L. Stein, "Genome annotation: from sequence to biology," *Nat Rev Genet,* vol. 2, pp. 493-503, Jul 2001.

[134] Y. Loewenstein, D. Raimondo, O. C. Redfern, J. Watson, D. Frishman, M. Linial*, et al.*, "Protein function annotation by homology-based inference," *Genome Biol,* vol. 10, p. 207, 2009.

[135] J. M. Chandonia, G. Hon, N. S. Walker, L. Lo Conte, P. Koehl, M. Levitt*, et al.*, "The ASTRAL Compendium in 2004," *Nucleic Acids Res,* vol. 32, pp. D189-92, Jan 1 2004.

[136] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer*, et al.*, "BLAST+: architecture and applications," *BMC Bioinformatics,* vol. 10, p. 421, 2009.

[137] M. Pellegrini, E. M. Marcotte, M. J. Thompson, D. Eisenberg, and T. O. Yeates, "Assigning protein functions by comparative genome analysis: protein phylogenetic profiles," *Proc Natl Acad Sci U S A,* vol. 96, pp. 4285-8, Apr 13 1999.

[138] C. Chothia and A. M. Lesk, "The Relation between the Divergence of Sequence and Structure in Proteins," *Embo Journal,* vol. 5, pp. 823-826, Apr 1986.

[139] H. Hasegawa and L. Holm, "Advances and pitfalls of protein structural alignment," *Curr Opin Struct Biol,* vol. 19, pp. 341-8, Jun 2009.

[140] T. J. Ley, E. R. Mardis, L. Ding, B. Fulton, M. D. McLellan, K. Chen, *et al*., "DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome," *Nature,* vol. 456, pp. 66-72, Nov 6 2008.

[141] X. W. Chen and J. C. Jeong, "Sequence-based prediction of protein interaction sites with an integrative method," *Bioinformatics,* vol. 25, pp. 585-91, Mar 1 2009.

[142] R. R. Crichton and M. Charloteaux-Wauters, "Iron transport and storage," *Eur J Biochem,* vol. 164, pp. 485-506, May 4 1987.

[143] E. C. Theil, "Ferritin: structure, gene regulation, and cellular function in animals, plants, and microorganisms," *Annu Rev Biochem,* vol. 56, pp. 289-315, 1987.

[144] J. E. Dutton, L. J. Rogers, B. G. Haslett, I. A. H. Takruri, J. T. Gleaves, and D. Boulter, "Comparative Studies on the Properties of Two Ferredoxins from Pisum sativum L," *Journal of Experimental Botany,* vol. 31, pp. 379-391, April 1, 1980 1980.

[145] J. R. Mason and R. Cammack, "The electron-transport proteins of hydroxylating bacterial dioxygenases," *Annu Rev Microbiol,* vol. 46, pp. 277-305, 1992.

[146] L. Hedstrom, "Serine protease mechanism and specificity," *Chem Rev,* vol. 102, pp. 4501-24, Dec 2002.

[147] N. D. Rawlings and A. J. Barrett, "Families of serine peptidases," *Methods Enzymol,* vol. 244, pp. 19-61, 1994.

[148] E. S. Alnemri, D. J. Livingston, D. W. Nicholson, G. Salvesen, N. A. Thornberry, W. W. Wong, *et al*., "Human ICE/CED-3 protease nomenclature," *Cell,* vol. 87, p. 171, Oct 18 1996.

[149] E. A. Stura, G. G. Fieser, and I. A. Wilson, "Crystallization of Antibodies and Antibody-Antigen Complexes," *ImmunoMethods,* vol. 3, pp. 164-179, 1993.

[150] M. G. Marinus and N. R. Morris, "Isolation of deoxyribonucleic acid methylase mutants of Escherichia coli K-12," *J Bacteriol,* vol. 114, pp. 1143-50, Jun 1973.

[151] M. R. Badger and G. D. Price, "The Role of Carbonic-Anhydrase in Photosynthesis," *Annual Review of Plant Physiology and Plant Molecular Biology,* vol. 45, pp. 369-392, 1994.

[152] K. H. Liu, J. J. E. Dore, M. P. Roberts, R. Krishnan, F. M. Hopkins, and J. D. Godkin, "Expression and Cellular-Localization of Retinol-Binding Protein Messenger-Ribonucleic-Acid in Bovine Blastocysts and Extraembryonic Membranes," *Biology of Reproduction,* vol. 49, pp. 393-400, Aug 1993.

[153] J. Hurwitz, "The discovery of RNA polymerase," *J Biol Chem,* vol. 280, pp. 42477-85, Dec 30 2005.

[154] S. Takahashi and T. Koyama, "Structure and function of cis-prenyl chain elongating enzymes," *Chem Rec,* vol. 6, pp. 194-205, 2006.

[155] A. E. Chung and J. H. Law, "Cyclopropane Fatty Acid Synthetase: Partial Purification and Properties," *Biochemistry,* vol. 3, pp. 967-74, Jul 1964.

[156] P. Dinadayala, F. Laval, C. Raynaud, A. Lemassu, M. A. Laneelle, G. Laneelle, *et al*., "Tracking the putative biosynthetic precursors of oxygenated mycolates of Mycobacterium tuberculosis. Structural analysis of fatty acids of a mutant strain deviod of methoxy- and ketomycolates," *J Biol Chem,* vol. 278, pp. 7310-9, Feb 28 2003.

[157] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, "Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders," *Nucleic Acids Res,* vol. 33, pp. D514-7, Jan 1 2005.

[158] H. J. Bandelt, Q. P. Kong, W. Parson, and A. Salas, "More evidence for non-maternal inheritance of mitochondrial DNA?," *J Med Genet,* vol. 42, pp. 957-60, Dec 2005.

[159] M. Pronicki, E. Matyja, D. Piekutowska-Abramczuk, T. Szymanska-Debinska, A. Karkucinska-Wieckowska, E. Karczmarewicz, *et al*., "Light and electron microscopy characteristics of the muscle of patients with SURF1 gene mutations associated with Leigh disease," *Journal of Clinical Pathology,* vol. 61, pp. 460-466, Apr 2008.

[160] M. Hirano and S. G. Pavlakis, "Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes (MELAS): current concepts," *J Child Neurol,* vol. 9, pp. 4-13, Jan 1994.

[161] O. Gimm, M. Armanios, H. Dziema, H. P. Neumann, and C. Eng, "Somatic and occult germ-line mutations in SDHD, a mitochondrial complex II gene, in nonfamilial pheochromocytoma," *Cancer Res,* vol. 60, pp. 6822-5, Dec 15 2000.

[162] D. Cotesta, C. Caliumi, P. Alo, L. Petramala, M. G. Reale, R. Masciangelo, *et al*., "High plasma levels of human chromogranin A and adrenomedullin in patients with pheochromocytoma," *Tumori,* vol. 91, pp. 53-8, Jan-Feb 2005.

[163] R. McFarland, D. M. Kirby, K. J. Fowler, A. Ohtake, M. T. Ryan, D. J. Amor, *et al*., "De novo mutations in the mitochondrial ND3 gene as a cause of infantile mitochondrial encephalopathy and complex I deficiency.," *American Journal of Human Genetics,* vol. 73, pp. 462-462, Nov 2003.

[164] A. D. Sheftel, O. Stehling, A. J. Pierik, H. P. Elsasser, U. Muhlenhoff, H. Webert, *et al*., "Humans possess two mitochondrial ferredoxins, Fdx1 and Fdx2, with distinct roles in steroidogenesis, heme, and Fe/S cluster biosynthesis," *Proc Natl Acad Sci U S A,* vol. 107, pp. 11775-80, Jun 29 2010.

[165] S. J. Hoefs, C. E. Dieteren, F. Distelmaier, R. J. Janssen, A. Epplen, H. G. Swarts, *et al*., "NDUFA2 complex I mutation leads to Leigh disease," *American Journal of Human Genetics,* vol. 82, pp. 1306-15, Jun 2008.

[166] J. L. Loeffen, R. H. Triepels, L. P. van den Heuvel, M. Schuelke, C. A. Buskens, R. J. Smeets, *et al*., "cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: human complex I cDNA characterization completed," *Biochem Biophys Res Commun,* vol. 253, pp. 415-22, Dec 18 1998.

[167] M. Bugiani, F. Invernizzi, S. Alberio, E. Briem, E. Lamantea, F. Carrara, *et al*., "Clinical and molecular findings in children with complex I deficiency," *Biochim Biophys Acta,* vol. 1659, pp. 136-47, Dec 6 2004.

[168] E. Ostergaard, R. J. Rodenburg, M. van den Brand, L. L. Thomsen, M. Duno, M. Batbayli, *et al*., "Respiratory chain complex I deficiency due to NDUFA12 mutations as a new cause of Leigh syndrome," *Journal of Medical Genetics,* vol. 48, pp. 737-740, Nov 2011.

[169] H. Lu and X. M. Cao, "GRIM-19 is essential for maintenance of mitochondrial membrane potential," *Molecular Biology of the Cell,* vol. 19, pp. 1893-1902, May 2008.

[170] M. P. Puissegur, N. M. Mazure, T. Bertero, L. Pradelli, S. Grosso, K. Robbe-Sermesant, *et al*., "miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity," *Cell Death and Differentiation,* vol. 18, pp. 465-478, Mar 2011.

[171] T. Marui, I. Funatogawa, S. Koishi, K. Yamamoto, H. Matsumoto, O. Hashimoto, *et al*., "The NADH-ubiquinone oxidoreductase 1 alpha subcomplex 5 (NDUFA5) gene variants are associated with autism," *Acta Psychiatrica Scandinavica,* vol. 123, pp. 118-124, Feb 2011.

[172] G. D. Mellick, P. A. Silburn, J. A. Prince, and A. J. Brookes, "A novel screen for nuclear mitochondrial gene associations with Parkinson's disease," *Journal of Neural Transmission,* vol. 111, pp. 191-199, Feb 2004.

[173] M. Gu, J. M. Cooper, J. W. Taanman, and A. H. V. Schapira, "Mitochondrial DNA transmission of the mitochondrial defect in Parkinson's disease," *Annals of Neurology,* vol. 44, pp. 177-186, Aug 1998.

[174] R. H. Swerdlow, J. K. Parks, S. W. Miller, J. B. Tuttle, P. A. Trimmer, J. P. Sheehan, *et al*., "Origin and functional consequences of the complex I defect in Parkinson's disease," *Annals of Neurology,* vol. 40, pp. 663-671, Oct 1996.

[175] B. J. C. van den Bosch, M. Gerards, W. Sluiter, A. P. A. Stegmann, E. L. C. Jongen, D. M. E. I. Hellebrekers, *et al*., "Defective NDUFA9 as a novel cause of neonatally fatal complex I disease," *Journal of Medical Genetics,* vol. 49, pp. 10-15, Jan 2012.

[176] R. Janssen, J. Smeitink, R. Smeets, and L. van Den Heuvel, "CIA30 complex I assembly factor: a candidate for human complex I deficiency?," *Hum Genet,* vol. 110, pp. 264-70, Mar 2002.

[177] E. Fassone, J. W. Taanman, I. P. Hargreaves, N. J. Sebire, M. A. Cleary, M. Burch, *et al*., "Mutations in the mitochondrial complex I assembly factor NDUFAF1 cause fatal infantile hypertrophic cardiomyopathy," *J Med Genet,* vol. 48, pp. 691-7, Oct 2011.

[178] S. E. Calvo, A. G. Compton, S. G. Hershman, S. C. Lim, D. S. Lieber, E. J. Tucker, *et al*., "Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing," *Sci Transl Med,* vol. 4, p. 118ra10, Jan 25 2012.

[179] T. B. Haack, F. Madignier, M. Herzer, E. Lamantea, K. Danhauser, F. Invernizzi, *et al*., "Mutation screening of 75 candidate genes in 152 complex I deficiency cases identifies pathogenic variants in 16 genes including NDUFB9," *Journal of Medical Genetics,* vol. 49, pp. 83-89, Feb 2012.

[180] R. Szklarczyk, B. F. J. Wanschers, S. B. Nabuurs, J. Nouws, L. G. Nijtmans, and M. A. Huynen, "NDUFB7 and NDUFA8 are located at the intermembrane surface of complex I," *Febs Letters,* vol. 585, pp. 737-743, Mar 9 2011.

[181] S. Loublier, A. Bayot, M. Rak, R. El-Khoury, P. Benit, and P. Rustin, "The NDUFB6 subunit of the mitochondrial respiratory chain complex I is required for electron transfer activity: A proof of principle study on stable and controlled RNA interference in human cell lines," *Biochemical and Biophysical Research Communications,* vol. 414, pp. 367-372, Oct 22 2011.

[182] C. Ugalde, R. J. Janssen, L. P. van den Heuvel, J. A. Smeitink, and L. G. Nijtmans, "Differences in assembly or stability of complex I and other mitochondrial OXPHOS complexes in inherited complex I deficiency," *Hum Mol Genet,* vol. 13, pp. 659-67, Mar 15 2004.

[183] D. M. Kirby, R. Salemi, C. Sugiana, A. Ohtake, L. Parry, K. M. Bell, *et al*., "NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency," *J Clin Invest,* vol. 114, pp. 837-45, Sep 2004.

[184] S. Granata, G. Zaza, S. Simone, G. Villani, D. Latorre, P. Pontrelli, *et al*., "Mitochondrial dysregulation and oxidative stress in patients with chronic kidney disease," *Bmc Genomics,* vol. 10, Aug 21 2009.

[185] Y. A. Su, J. Wu, L. Zhang, Q. Y. Zhang, D. M. Su, P. He, *et al*., "Dysregulated Mitochondrial Genes and Networks with Drug Targets in Postmortem Brain of Patients with Posttraumatic Stress Disorder (PTSD) Revealed by Human Mitochondria-Focused cDNA Microarrays," *International Journal of Biological Sciences,* vol. 4, pp. 223-235, 2008.

[186] E. J. Tucker, S. G. Hershman, C. Kohrer, C. A. Belcher-Timme, J. Patel, O. A. Goldberger, *et al*., "Mutations in MTFMT underlie a human disorder of formylation causing impaired mitochondrial translation," *Cell Metab,* vol. 14, pp. 428-34, Sep 7 2011.

[187] M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, and T. Ideker, "Cytoscape 2.8: new features for data integration and network visualization," *Bioinformatics,* vol. 27, pp. 431-2, Feb 1 2011.

[188] D. Jaquet, J. Leger, C. Levy-Marchal, J. F. Oury, and P. Czernichow, "Ontogeny of leptin in human fetuses and newborns: effect of intrauterine growth retardation on serum leptin concentrations," *J Clin Endocrinol Metab,* vol. 83, pp. 1243-6, Apr 1998.

[189] L. A. Tartaglia, M. Dembski, X. Weng, N. Deng, J. Culpepper, R. Devos, *et al.*, "Identification and expression cloning of a leptin receptor, OB-R," *Cell,* vol. 83, pp. 1263-71, Dec 29 1995.

[190] R. S. Jackson, J. W. Creemers, S. Ohagi, M. L. Raffin-Sanson, L. Sanders, C. T. Montague, *et al.*, "Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene," *Nat Genet,* vol. 16, pp. 303-6, Jul 1997.

[191] L. Yaswen, N. Diehl, M. B. Brennan, and U. Hochgeschwender, "Obesity in the mouse model of pro-opiomelanocortin deficiency responds to peripheral melanocortin," *Nat Med,* vol. 5, pp. 1066-70, Sep 1999.

[192] W. H. Dietz, "Health consequences of obesity in youth: childhood predictors of adult disease," *Pediatrics,* vol. 101, pp. 518-25, Mar 1998.

[193] H. Hampel, N. S. Abraham, and H. B. El-Serag, "Meta-analysis: obesity and the risk for gastroesophageal reflux disease and its complications," *Ann Intern Med,* vol. 143, pp. 199-211, Aug 2 2005.

[194] D. L. Roberts, C. Dive, and A. G. Renehan, "Biological mechanisms linking obesity and cancer risk: new perspectives," *Annu Rev Med,* vol. 61, pp. 301-16, 2010.

[195] E. E. Calle and M. J. Thun, "Obesity and cancer," *Oncogene,* vol. 23, pp. 6365-78, Aug 23 2004.

[196] H. Masuya, Y. Makita, N. Kobayashi, K. Nishikata, Y. Yoshida, Y. Mochizuki, *et al.*, "The RIKEN integrated database of mammals," *Nucleic Acids Res,* vol. 39, pp. D861-70, Jan 2011.

[197] I. E. Scheffer and S. F. Berkovic, "Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes," *Brain,* vol. 120 ( Pt 3), pp. 479-90, Mar 1997.

[198] R. Singh, I. E. Scheffer, K. Crossland, and S. F. Berkovic, "Generalized epilepsy with febrile seizures plus: a common childhood-onset genetic epilepsy syndrome," *Annals of Neurology,* vol. 45, pp. 75-81, Jan 1999.

[199] L. Zinman, E. Ng, and V. Bril, "IV immunoglobulin in patients with myasthenia gravis: a randomized controlled trial," *Neurology,* vol. 68, pp. 837-41, Mar 13 2007.

[200] M. Giraud, R. Taubert, C. Vandiedonck, X. Ke, M. Levi-Strauss, F. Pagani*, et al*., "An IRF8-binding promoter variant and AIRE control CHRNA1 promiscuous expression in thymus," *Nature,* vol. 448, pp. 934-7, Aug 23 2007.

[201] H. J. Garchon, F. Djabiri, J. P. Viard, P. Gajdos, and J. F. Bach, "Involvement of human muscle acetylcholine receptor alpha-subunit gene (CHRNA) in susceptibility to myasthenia gravis," *Proc Natl Acad Sci U S A,* vol. 91, pp. 4668-72, May 24 1994.

[202] H. L. Wang, M. Milone, K. Ohno, X. M. Shen, A. Tsujino, A. P. Batocchi*, et al*., "Acetylcholine receptor M3 domain: stereochemical and volume contributions to channel gating," *Nat Neurosci,* vol. 2, pp. 226-33, Mar 1999.

[203] A. Masuda, X. M. Shen, M. Ito, T. Matsuura, A. G. Engel, and K. Ohno, "hnRNP H enhances skipping of a nonfunctional exon P3A in CHRNA1 and a mutation disrupting its binding causes congenital myasthenic syndrome," *Hum Mol Genet,* vol. 17, pp. 4022-35, Dec 15 2008.

[204] O. Uchitel, A. G. Engel, T. J. Walls, A. Nagel, M. Z. Atassi, and V. Bril, "Congenital myasthenic syndromes: II. Syndrome attributed to abnormal interaction of acetylcholine with its receptor," *Muscle Nerve,* vol. 16, pp. 1293-301, Dec 1993.

[205] K. Ohno, H. L. Wang, M. Milone, N. Bren, J. M. Brengman, S. Nakano*, et al*., "Congenital myasthenic syndrome caused by decreased agonist binding affinity due to a mutation in the acetylcholine receptor epsilon subunit," *Neuron,* vol. 17, pp. 157-70, Jul 1996.

[206] R. Croxen, C. Hatton, C. Shelley, M. Brydson, G. Chauplannaz, H. Oosterhuis, *et al*., "Recessive inheritance and variable penetrance of slow-channel congenital myasthenic syndromes," *Neurology,* vol. 59, pp. 162-8, Jul 23 2002.

[207] A. G. Engel, K. Ohno, M. Milone, H. L. Wang, S. Nakano, C. Bouzat, *et al*., "New mutations in acetylcholine receptor subunit genes reveal heterogeneity in the slow-channel congenital myasthenic syndrome," *Hum Mol Genet,* vol. 5, pp. 1217-27, Sep 1996.

[208] C. M. Gomez, R. Maselli, J. Gammack, J. Lasalde, S. Tamamizu, D. R. Cornblath, *et al*., "A beta-subunit mutation in the acetylcholine receptor channel gate causes severe slow-channel syndrome," *Annals of Neurology,* vol. 39, pp. 712-23, Jun 1996.

[209] C. M. Gomez, R. A. Maselli, B. P. Vohra, M. Navedo, J. R. Stiles, P. Charnet, *et al*., "Novel delta subunit mutation in slow-channel syndrome causes severe weakness by novel mechanisms," *Annals of Neurology,* vol. 51, pp. 102-12, Jan 2002.

[210] E. Hund, R. P. Linke, F. Willig, and A. Grau, "Transthyretin-associated neuropathic amyloidosis. Pathogenesis and treatment," *Neurology,* vol. 56, pp. 431-5, Feb 27 2001.

[211] J. N. Buxbaum and N. Reixach, "Transthyretin: the servant of many masters," *Cell Mol Life Sci,* vol. 66, pp. 3095-101, Oct 2009.

[212] M. A. Gertz and R. A. Kyle, "Myopathy in primary systemic amyloidosis," *J Neurol Neurosurg Psychiatry,* vol. 60, pp. 655-60, Jun 1996.

[213] J. Keith, Z. Afshar-Ghotli, R. Roussev, B. Ernst, B. Young, and J. M. Bilbao, "Myopathy as the initial manifestation of primary amyloidosis," *Can J Neurol Sci,* vol. 38, pp. 161-4, Jan 2011.

[214] Y. Friedman, J. T. Paul, J. Turley, L. N. Hazrati, and D. Munoz, "Axial myopathy due to primary amyloidosis," *Muscle Nerve,* vol. 36, pp. 542-6, Oct 2007.

[215] D. Karacostas, M. Soumpourou, I. Mavromatis, G. Karkavelas, I. Poulios, and I. Milonas, "Isolated myopathy as the initial manifestation of primary systemic amyloidosis," *J Neurol,* vol. 252, pp. 853-4, Jul 2005.

[216] P. Bielefeld, L. Leveque, A. Turcu, L. Popitean, E. Justrabo, and J. F. Besancenot, "[Myopathy revealing AL amyloidosis]," *Presse Med,* vol. 31, p. 361, Mar 2 2002.

[217] N. Nadkarni, M. Freimer, and J. R. Mendell, "Amyloidosis causing a progressive myopathy," *Muscle Nerve,* vol. 18, pp. 1016-8, Sep 1995.

[218] H. Tuomaala, M. Karppa, H. Tuominen, and A. M. Remes, "Amyloid myopathy: a diagnostic challenge," *Neurol Int,* vol. 1, p. e7, 2009.

[219] D. Petrey, M. Fischer, and B. Honig, "Structural relationships among proteins with different global topologies and their implications for function annotation strategies," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 106, pp. 17377-17382, Oct 13 2009.

[220] H. Y. Yu, N. M. Luscombe, H. X. Lu, X. W. Zhu, Y. Xia, J. D. J. Han, *et al.*, "Annotation transfer between genomes: Protein-protein interologs and protein-DNA regulogs," *Genome Research,* vol. 14, pp. 1107-1118, Jun 2004.

[221] D. Petrey and B. Honig, "Is protein classification necessary? Toward alternative approaches to function annotation," *Current Opinion in Structural Biology,* vol. 19, pp. 363-368, Jun 2009.

[222] M. Osadchy and R. Kolodny, "Maps of protein structure space reveal a fundamental relationship between protein structure and function," *Proc Natl Acad Sci U S A,* vol. 108, pp. 12301-6, Jul 26 2011.

[223] M. A. Yildirim, K. I. Goh, M. E. Cusick, A. L. Barabasi, and M. Vidal, "Drug-target network," *Nat Biotechnol,* vol. 25, pp. 1119-26, Oct 2007.

[224] T. U. Consortium, "The Universal Protein Resource (UniProt) in 2010," *Nucleic Acids Res,* vol. 38, pp. D142-8, Jan 2010.