Global Network Alignment In The Context Of Aging

Fazle Elahi Faisal, Han Zhao, and Tijana Milenković

Abstract—Analogous to sequence alignment, network alignment (NA) can be used to transfer biological knowledge across species between conserved network regions. NA faces two algorithmic challenges: 1) Which cost function to use to capture "similarities" between nodes in different networks? 2) Which alignment strategy to use to rapidly identify "high-scoring" alignments from all possible alignments? We "break down" existing state-of-the-art methods that use both different cost functions and different alignment strategies to evaluate each combination of their cost functions and alignment strategies. We find that a combination of the cost function of one method and the alignment strategy of another method beats the existing methods. Hence, we propose this combination as a novel superior NA method. Then, since human aging is hard to study experimentally due to long lifespan, we use NA to transfer aging-related knowledge from well annotated model species to poorly annotated human. By doing so, we produce novel human aging-related knowledge, which complements currently available knowledge about aging that has been obtained mainly by sequence alignment. We demonstrate significant similarity between topological and functional properties of our novel predictions and those of known aging-related genes. We are the first to use NA to learn more about aging.

Index Terms—Network Alignment; Aging; Protein Function Prediction.

1 Introduction

Bioinformatics research and genomic sequence alignment in particular have revolutionized our biological understanding. Sequence alignment finds regions of similarity that are a likely consequence of functional or evolutionary relationships between the sequences [1]. It has been extensively used, for example, to construct and interpret phylogenetic trees. Sequence alignment has been adopted from biology to other domains, e.g., natural language processing [2].

However, genes' protein products do not function in isolation. Instead, they interact with each other to keep us alive. And this is exactly what protein-protein interaction (PPI) networks model: in these networks, nodes correspond to proteins and edges indicate physical interactions between the proteins. Unlike genomic sequences, PPI networks enable the study of complex cellular processes that emerge from the collective behavior of the proteins. Note that even though we focus on PPI networks (for reasons explained above), our study is applicable to any type of biological networks, as well as to networks in other domains, such as social networks [3].

- *Fazle Elahi Faisal is with the Department of Computer Science and Engineering, Interdisciplinary Center for Network Science and Applications, and ECK Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: ffaisal@nd.edu.*
- *Han Zhao is with the Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, USA, and the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China. E-mail: zhaohan2009011312@gmail.com.*
- *Tijana Milenković is with the Department of Computer Science and Engineering, Interdisciplinary Center for Network Science and Applications, and ECK Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: tmilenko@nd.edu.*

High-throughput biotechologies (such as yeast two-hybrid (Y2H) assays or affinity purification coupled to mass spectrometry (AP/MS)) have produced PPI data for many model organisms and human [4], [5]. As more PPI data is becoming available, meaningful alignments of PPI networks of different species could be viewed as one of the foremost problems in computational and systems biology [6], [7], [8], [9]. Since network alignment aims to find regions of similarities between PPI networks of different species (see below), it could guide the transfer of biological knowledge across species between conserved network regions [6]. This is important, since many nodes in PPI networks are currently functionally unannotated even for well studied model species [10]. In particular, this is important when studying human aging: since human aging is hard to study experimentally due to long lifespan as well as ethical constraints, the knowledge about aging needs to be transferred from model species. Traditionally, the transfer of biological knowledge between species has been restricted to genomic sequence alignment. However, PPI networks and sequence data can give complementary biological insights [11], implying that PPI data can elucidate function that cannot be extracted from sequence data by current methods. Thus, restricting alignment to sequences may limit the knowledge transfer [12], [11].

In addition to across-species transfer of function, PPI network alignment can also be used to infer phylogenetic relationships of different species based on similarities between their PPI networks [7], [13], [8]. Further, network alignment has application outside of biological domain as well, with implications, e.g., on user privacy in online social networks [3].

Like sequence alignment, network alignment can be *local* and *global*. Initial methods have focused on local network alignment [14], [15], [16], [17], [18], [19], [20]. However, since these are generally unable to find *large* subgraphs that are conserved between the aligned networks, methods for global network alignment (GNA) (defined below) have been proposed [21], [22], [23], [24], [9], [7], [13], [8], [25], [3], [26], [27], [28], [29], [30], [31]. In this study, we focus on GNA, but our ideas are also applicable to local network alignment.

GNA is computationally hard due to the NP-completeness of the underlying subgraph isomorphism problem, which asks if a network (or graph) exists as an exact subgraph of another graph [32]. GNA is the more general problem of finding the best way to "fit" a graph into another graph even if the first graph is not an exact subgraph of the second one. Since GNA is computationally hard, heuristic methods must be sought. Existing GNA methods typically achieve an alignment by constructing a mapping between nodes of the compared networks that is expected to align topologically and functionally similar network regions [6]. They do so by considering two algorithmic challenges: 1) defining a cost function for capturing "similarities" between nodes in different networks and 2) presenting an alignment strategy for rapidly identifying from all possible alignments the high-scoring alignments with respect to some topological or biological alignment quality measure [8], [25], [9].

Different existing GNA methods typically use both different cost functions and different alignment strategies. So, when a method is superior to other methods, it is unclear whether the superiority comes from its cost function, its alignment strategy, or both. Hence, we aim to "break down" existing state-of-the-art GNA methods into the two components to fairly evaluate each combination of cost function and alignment strategy. This could result in a superior new GNA method if the combination of the cost function of one existing method and the alignment strategy of another existing method would beat the current methods.

Since the US is on average growing older because of $\sim$78 million of baby boomers who have begun turning 65 in 2011, and since susceptibility to diseases increases with age, studying molecular causes of aging gains importance. Hence, after we evaluate the GNA methods, we use them to transfer aging-related knowledge from well annotated network parts of some species to poorly annotated network parts of other species and human in particular. To our knowledge, we are the first to use network alignment to deepen our current knowledge about aging via across-species, network-based prediction of novel aging-related genes.

2 Our approach

Most methods define GNA as a *one-to-one* function that injectively maps nodes between *two* networks. MI-GRAAL [8] is a state-of-the-art method of this type. Some exceptions exist [22], [9]. For example, even though IsoRankN has been described as a GNA method [9], it allows for *many-to-many* node mapping as well as for simultaneous alignment of *multiple* networks. MI-GRAAL and IsoRankN are state-of-the-art. Their superiority has been demonstrated over other popular methods [8], [9], including a number of MI-GRAAL's [7], [13] and IsoRankN's [21], [23] predecessors, as well as Graemlin [16], [22] and NetworkBLAST-M [15].

As a proof of concept that it is indeed important to evaluate the relevance of each of two GNA components (cost function and alignment strategy) separately, we focus on a thorough and fair evaluation of these two state-of-the-art GNA algorithms, i.e., all combinations of their cost functions and alignment strategies. (Note that even though we focus on MI-GRAAL and IsoRankN, our study is applicable to any existing GNA method.) We aim to identify the best of the two cost functions when same alignment strategy is used. Ideally, we would also aim to identify the best of the two alignment strategies when same cost function is used. However, IsoRankN's alignment strategy that allows for many-to-many node mapping cannot be directly and fairly evaluated against MI-GRAAL's one-to-one mapping strategy because of their different nature. Nonetheless, we perform at least some indirect comparison of the two alignment strategies under same cost function, when we use them to predict new aging-related knowledge (see below).

Many existing GNA algorithms *by default* use in their cost function biological information *external* to network topology, such as sequence similarity [21], [9], [22], [24]. For each combination of cost function and alignment strategy, we first test how much of new biological knowledge can be uncovered *solely* from topology *before* integrating it with sequences [7], [13], [8]. Then, since both GNA algorithms do allow for adding sequence data to the cost function, we test whether using *both* topology and sequences improves alignment quality compared to using only topology (or only sequences).

We evaluate the different combinations of cost functions and alignment strategies (or "network aligners") as follows. We evaluate alignments resulting from MI-GRAAL's strategy both topologically (with respect to the amount of topology that is conserved across networks) and biologically (with respect to functional enrichment of the aligned nodes). Since IsoRankN's alignment strategy results in many-to-many node mappings, we cannot evaluate its alignments topologically. Instead, we evaluate the alignments biologically (with respect to functional enrichment).

We find that MI-GRAAL's cost function typically outperforms IsoRankN's cost function under both MI-GRAAL's and IsoRankN's alignment strategy, with respect to almost all alignment quality measures,

and independent on whether topology only or both topology and sequence are used in the cost function. Therefore, while the original IsoRankN, which combines IsoRankN's cost function and IsoRankN's alignment strategy, has been considered as a state-of-the-art method for *multiple* GNA, we show that the combination of MI-GRAAL's cost function and IsoRankN's alignment strategy outperforms the original IsoRankN. As such, this combination represents a new superior method for multiple GNA, which is an important contribution of our study.

When we study how well the different aligners uncover existing aging-related knowledge, we again find that MI-GRAAL's cost function dominates that of IsoRankN, especially under IsoRankN's alignment strategy. By "dominates", we mean that one method more often correctly aligns known aging-related network parts across species than another method. In this context, in an indirect comparison, we find that IsoRankN's alignment strategy is superior under MI-GRAAL's cost function, while MI-GRAAL's alignment strategy is superior under IsoRankN's cost function.

We use all alignments in which the aligned network parts are statistically significantly enriched in known aging-related genes to predict novel aging-related genes in a given species based on their known aging-related aligned partners in other species. Therefore, another contribution of our study is a new data set of aging-related knowledge, obtained from a novel source of biological data, namely PPI networks. Thus, our study's output complements currently available aging-related knowledge that has been obtained mainly by genomic sequence alignment, especially in human.

We validate our novel aging-related predictions in human by: 1) showing that they significantly overlap with known aging-related genes from independent, external "ground truth" data sets; 2) demonstrating that their topological and functional properties are significantly similar to the properties of known aging-related genes, whereas their properties are different than the properties of genes that have not yet been associated to aging; 3) showing that our predictions are linked to aging-related biological pathways and diseases; and 4) associating almost all of our top-scoring predictions to aging via literature search.

3 METHODS

3.1 Data sets

We align PPI networks of four species: *S. cerevisiae* (yeast), *D. melanogaster* (fly), *C. elegans* (worm), and *H. sapiens* (human). Their networks have 3,321 proteins and 8,021 PPIs, 7,111 proteins and 23,376 PPIs, 2,582 proteins and 4,322 PPIs, and 6,167 proteins and 15,940 PPIs, respectively. They are largest connected components of Y2H networks from BioGRID 3.1.90 [4]. Of all PPIs, including "co-complex" AP/MS PPIs, we focus on "binary" Y2H PPIs because they are of higher quality than literature-curated PPIs supported by a single publication [33], [5]. What is important for a fair evaluation is that all aligners are tested on the same data, be it Y2H, AP/MS, or other network type.

For sequence similarity part of the cost function, we use BLAST bit-values [1] from NCBI database (http://www.ncbi.nlm.nih.gov/protein/).

To evaluate biological alignment quality with respect to functional enrichment of the aligned nodes, we use Gene Ontology (GO) data from July 2012 [34].

When we evaluate quality of alignments with respect to how well they uncover existing aging-related knowledge, we use GenAge data (version Build 16) [35] from September 2012. This data set contains 483, 79, 175, and 218 aging-related genes that are present in the yeast, fly, worm, and human PPI network, respectively. We refer to this set of aging-related genes as *GenAge*. Of the 218 human genes, only three have been linked to aging in humans *directly*. All others have been linked to aging in humans indirectly, e.g., via sequence-based homology from model species.

We also use GenAge to predict new aging-related genes from the alignments. Then, we validate the human predictions in two independent human aging-related "ground truth" data sets: 1) *ExpressionAge*, a set of 234 genes from the human PPI network that were predicted as aging-related because their expression levels significantly correlated with age [36], and 2) *DyNetAge*, a set of 394 genes from the human PPI network that were predicted as aging-related because their topological centralities in dynamic, age-specific PPI networks significantly changed with age [37]; the age-specific networks were obtained by integrating aging-related gene expression data with current static human PPI network data [37].

In all data sets, for consistency, gene names have been converted into gene IDs using DAVID tool [38].

3.2 Network aligners

MI-GRAAL. We previously designed a sensitive graphlet-based measure of topology, called *node graphlet degree vector (node-GDV)*, that captures up to 4-deep neighborhood of a node [39], [40]; a graphlet is a small induced subgraph of the network [41]. We designed a measure of topological similarity of such extended neighborhoods of two nodes, called *node-GDV-similarity*. Based on this measure, we developed GRAAL (GRAph ALigner) [7] and H-GRAAL (Hungarian-based GRAAL) [13] GNA algorithms that use topology *only* and can thus align networks in *any* domain, including social and computer networks [3]. Both use node-GDV-similarity as the cost function. They differ in their alignment strategies, as follows. GRAAL, a *greedy* "seed and extend" approach, like BLAST [1], first chooses the most node-GDV-similar "seed" node pair and then greedily expands the alignment radially outward around the seed to quickly

find *approximate* alignments. H-GRAAL finds *optimal* alignments with respect to the cost function by using computationally expensive Hungarian algorithm, a combinatorial optimization algorithm for finding a maximum weight matching in a weighted bipartite graph [32]. When used to align PPI networks of different species, both methods exposed regions of similarity *an order of magnitude larger* than other algorithms. The newer MI-GRAAL also uses node-GDV-similarity as the cost function while allowing for integration of additional biological data types, including sequence similarity. The cost function automatically adjusts the weights of nodes' topological similarity and their sequence similarity. MI-GRAAL's alignment strategy *combines* GRAAL's seed-and-extend alignment strategy with H-GRAAL's alignment strategy of solving the assignment problem, to further improve alignment quality at low computational cost [8].

IsoRankN. Its cost function is a spectral graph method which uses the intuition that two nodes should be matched only if their neighbors can also be matched [21]. The cost function can include sequence similarity in addition to topological similarity. The weights of topological similarity and sequence similarity are balanced by the parameter α, which is in [0,1] range. For example, if $\alpha = 0.6$, then the weight of topological similarity and sequence similarity is 60% and 40% of the total cost function, respectively. In our experiments, we set α to 0.7, as this is the recommended value in the original IsoRankN paper [9]. IsoRankN's alignment strategy relies on the notion of node-specific rankings and uses a method similar to PageRank-Nibble algorithm [9].

MI-GRAAL vs. IsoRankN. Cost functions of the two methods are conceptually similar: both consider as good matches those nodes whose extended network neighborhoods match well. However, the details of the two cost functions and their implementations are different [39], [8], [21], [9]. Consequently, as we will show, they result in alignments of different quality even under same alignment strategy.

Alignment strategies of the two methods are very different even conceptually. MI-GRAAL's alignment is a *one-to-one* function that injectively maps nodes between *two* networks. IsoRankN's alignment allows for a *many-to-many* node mapping (i.e., multiple nodes in one network can simultaneously be mapped to multiple nodes in another network), as well as for simultaneous alignment of *multiple* networks.

The different nature of the two alignment strategies makes their direct comparison hard. Hence, we compare the two cost functions under same alignment strategy, for each of the two alignment strategies. However, we cannot directly compare the two alignment strategies under same cost function. But, we still aim to compare them indirectly, in the context of across-species prediction of aging-related genes.

Combining each of the two cost functions with each of the two alignment strategies results in *four* network aligners. For each aligner, we align networks when using only topology in the cost function, as well as when using both topology and sequence in the cost function. Thus, we deal with two versions of each aligner, i.e., with *eight* aligners (Table 1). Also, for each of the two alignment strategies, we use sequence similarity information *only* in their cost function, resulting in *two additional* aligners (Table 1).

TABLE 1
Eight aligners resulting from two cost functions, two alignment strategies, and two data types used in the cost function (topology only ("T") or both topology and sequence ("T&S")), plus two aligners resulting from using sequence only ("S") as the cost function within a given alignment strategy.

Aligner	Cost function	Align. strategy	Data
MI-MI-T	MI-GRAAL	MI-GRAAL	T
MI-MI-TS	MI-GRAAL	MI-GRAAL	T&S
Iso-MI-T	IsoRankN	MI-GRAAL	T
Iso-MI-TS	IsoRankN	MI-GRAAL	T&S
MI-Iso-T	MI-GRAAL	IsoRankN	T
MI-Iso-TS	MI-GRAAL	IsoRankN	T&S
Iso-Iso-T	IsoRankN	IsoRankN	T
Iso-Iso-TS	IsoRankN	IsoRankN	T&S
X-MI-S	N/A (or "X")	MI-GRAAL	S
X-Iso-S	N/A (or "X")	IsoRankN	S

3.3 Evaluating different network aligners

We divide our evaluation into two steps. First, we evaluate the different cost functions under MI-GRAAL's alignment strategy. That is, we compare MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S, when each of them is used to produce six pairwise alignments of the four networks (Section 3.3.1). Second, we evaluate the cost functions under IsoRankN's alignment strategy. That is, we compare Iso-Iso-T, Iso-Iso-TS, MI-Iso-T, MI-Iso-TS, and X-Iso-S, when each of them is used to produce one simultaneous alignment of all four networks (Section 3.3.2). Since the two evaluation steps correspond to the two very different alignment strategies, in each step, we use measures of alignment quality that are tailored for the given alignment strategy, as follows.

3.3.1 Evaluation under MI-GRAAL's align. strategy

Given two networks $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$, where $|V_1| \leq |V_2|$, MI-GRAAL's alignment strategy results in a total injective function $f : V_1 \rightarrow V_2$ [8]. Function f is *total* if it maps all elements in V_1 to elements in V_2 and it is *injective* if it never maps different elements in V_1 to the same element in V_2. We denote by $V_2' = f(V_1)$ the set of nodes from G_2 that are aligned to nodes from V_1, by $E_2' = E_{f(V_1)}$ the set of edges from G_2 that exist between the nodes in V_2', and by $G_2'(V_2', E_2')$ the subgraph of G_2 *induced* on $f(V_1)$.

We evaluate MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S topologically *and* biologically.

Topological evaluation. We use the following measures of topological alignment quality [8]:

1) *Edge correctness (EC)*, the percentage of edges from G_1, the smaller network (in terms of the number of nodes), which are aligned to edges from G_2, the larger network [7], [8]. Formally, $EC = \frac{|E_1 \cap E_2'|}{|E_1|} \times 100\%$, where the numerator is the number of "conserved" edges, i.e., edges that are aligned under the given node mapping. The larger the EC, the better the alignment.
2) *Induced conserved structure (ICS)*: $ICS = \frac{|E_1 \cap E_2'|}{|E_2'|} \times 100\%$. It has been argued that EC might fail to differentiate between alignments that one might intuitively consider to be of different topological quality [25], since it is defined with respect to edges in E_1. For example, aligning a k-node cycle in G_1 to a k-node cycle in G_2 would result in the same EC as aligning a k-node cycle in G_1 to a k-node clique (complete graph) in G_2. Clearly, the former is intuitively a better alignment than the latter, since no edges that exist between the k nodes in G_2 are left unaligned in the first case, whereas many edges are left unaligned in the second case. Since ICS is defined with respect to edges in E_2', it would have the maximum value of 100% when aligning a k-node cycle to a k-node cycle, and it would have a lower value when aligning a k-node cycle to a k-node clique. The larger the ICS, the better.
3) *Node coverage (NC)*, the percentage of $|V_1|$ aligned node pairs that participate in the conserved edges. The larger the NC, the better, as fewer nodes are mapped across networks that do not add to edge conservation.
4) The size of the *largest connected common subgraph (LCCS)* [7], [8], which we use because of two alignments with similar EC, ICS, or NC scores, one could expose large, contiguous, and topologically complex regions of network similarity, while the other could fail to do so. Thus, in addition to counting aligned edges (as EC and ICS do) or nodes that participate in the aligned edges (as NC does), it is important that the aligned edges cluster together to form large connected subgraphs rather than being isolated. Hence, we define a connected common subgraph (CCS) as a connected subgraph (not necessarily induced) that appears in both networks [13], [8]. We measure the size of the largest CCS (LCCS) in terms of the number of nodes as well as edges. Large LCCS are desirable.

Biological evaluation. Only alignments in which many aligned node pairs perform the same function should be used to transfer function from annotated parts of one network to unannotated parts of another network. Hence, we measure Gene Ontology (GO) [34] enrichment of aligned proteins pairs, i.e., the percentage of protein pairs in which the two proteins *share* at least one GO term, out of all aligned protein pairs in which both proteins are annotated with at least one GO term. We refer to this percentage as *GO correctness*. We do this with respect to complete GO annotation data, independent of GO evidence code. Also, since many GO annotations have been obtained via sequence comparison, and since some of the aligners use sequence information, we repeat the analysis considering only GO annotations with experimental evidence codes. In this case, we refer to GO correctness as *experimental GO correctness*. The higher the GO correctness, the better.

3.3.2 Evaluation under IsoRankN's align. strategy

With IsoRankN's *multiple* network alignment strategy, we align all PPI networks simultaneously.

Since IsoRankN's alignment strategy allows for multiple nodes in one network to be mapped to multiple nodes in another network, its output is a set of aligned clusters, where no two clusters overlap, but each cluster can contain multiple nodes from each network. Thus, IsoRankN's output cannot be quantified topologically with EC, ICS, NC, or LCCS, as one many-to-many node alignment can produce exponentially many one-to-one node alignments and enumerating all of them is infeasible [7].

Instead of using EC, ICS, NC, or LCCS, we use the original IsoRankN's measures [9] to evaluate MI-Iso-T, MI-Iso-TS, Iso-Iso-T, Iso-Iso-TS, and X-Iso-S. Intuitively, a good IsoRankN's alignment should produce aligned clusters such that genes in each cluster are functionally uniform or *consistent*. Also, it should produce many such clusters, so that it *covers* as many of the proteins from the aligned networks as possible. IsoRankN captures the notions of consistency and coverage with the following evaluation measures:

1) The *number of aligned clusters*, where the higher the number, the better the alignment.
2) *Exact cluster ratio*, the percentage of aligned clusters in which all proteins share at least one GO term. The higher the value, the better.
3) *Exact protein ratio*, the percentage of all proteins that are in the exact clusters (as defined above). The higher the value, the better.
4) *Mean entropy of alignment*. First, we compute the entropy of an aligned cluster S_v^* as: $H(S_v^*) = H(p_1, p_2, \ldots, p_d) = -\sum_{i=1}^{d} p_i \log p_i$, where p_i is the percentage of all proteins in S_v^* that have GO term i, and d is the total number of GO terms [9]. Then, the mean entropy of the alignment is obtained by averaging entropies across all clusters in the alignment. The lower the entropy of the alignment, the higher its average within-cluster GO term consistency, and consequently, the better its biological quality.

5) *Normalized mean entropy of alignment*. First, we compute the normalized entropy of an aligned cluster S_v^* as: $\bar{H}(S_v^*) = \frac{1}{\log d} H(S_v^*)$. Then, the mean normalized entropy is obtained by averaging normalized entropies across all aligned clusters. Low normalized mean entropy is desirable.

3.4 Network alignment in the context of aging

3.4.1 Aging under MI-GRAAL's alignment strategy

MI-GRAAL's alignment strategy produces pairwise GNA and one-to-one node mapping. Hence, we evaluate different aligners under MI-GRAAL's alignment strategy in the context of aging as follows. We denote a node pair as aging-related if both nodes are known aging-related genes. We count the number of aging-related node pairs in an alignment and determine the statistical significance of this result via the hypergeometric test. Let V_1 and V_2 be the sets of nodes in two networks to be aligned, and let A_1 and A_2 be the subsets of V_1 and V_2 that are known aging-related genes, respectively. Let E be the set of all node pairs that could be aligned: $|E| = |V_1| \times |V_2|$. Let G be the set of all aging-related node pairs that could be aligned: $|G| = |A_1| \times |A_2|$. Given an alignment, let H denote the set of aligned node pairs: $|H| = min(|V_1|, |V_2|)$. Let O be the subset of H containing all aligned aging-related node pairs, i.e., the intersection between G and H. Then, the probability p of observing $|O|$ or more aging-related node pairs in the alignment by chance is: $p = 1 - \sum_{i=0}^{|O|-1} \frac{\binom{|E|}{i}\binom{|E|-|H|}{|G|-i}}{\binom{|E|}{|G|}}$. Throughout the paper, we use p-value threshold of 0.05.

We focus on all statistically significant alignments. In such an alignment, we predict a gene as aging-related if the gene is aligned to a known aging-related gene. Then, we measure precision, recall, and F-score of all predictions in the given alignment. Precision is the percentage of the predictions that are known aging-related genes. Recall is the percentage of known aging-related genes that are among the predictions. F-score is the harmonic mean of precision and recall.

3.4.2 Aging under IsoRankN's alignment strategy

IsoRankN's alignment strategy produces multiple GNA and many-to-many node mapping. Hence, we evaluate different aligners under IsoRankN's alignment strategy in the context of aging as follows. We measure the enrichment in known aging-related genes of each aligned cluster, and we measure the statistical significance of this result via the hypergeometric test (Section 3.4.1), where now E is the set of nodes in the given alignment, i.e. nodes covered by all aligned clusters, G is the subset of nodes from E that are known aging-related genes, H is the set of nodes in the aligned cluster of interest, and O is the subset of nodes from H that are known aging-related genes.

We focus on all statistically significant clusters. We predict each gene in such a cluster as aging-related, since the gene is aligned to significantly many known aging-related genes. Then, we measure precision, recall, and F-score of all predictions made from all statistically significant clusters in the given alignment.

3.5 Validation of novel aging-related predictions

We aim to validate our novel aging-related predictions in human by: 1) computing their overlap with known aging-related genes from independent, external data sets, in hope to observe the significant overlap; 2) comparing their topological positions in the human PPI network against the positions of known aging-related genes as well as of genes that have not yet been implicated in aging, in hope to observe significant similarity of our predictions with the former but not the latter; 3) performing functional (Gene Ontology – GO) and disease-related (Disease Ontology – DO) enrichment analysis of our predictions, in hope to link them to aging-related biological processes; 4) computing the overlap between GO and DO terms enriched in our predictions and those enriched in known aging-related genes, in hope to observe the significant overlap; and 5) performing manual literature validation of our highest-confidence predictions, in hope to link them to aging in scientific articles.

3.5.1 Prediction overlap with the "ground truth" data

We measure the statistical significance of the overlap of genes in two data sets via the hypergeometric test (Section 3.4.1), where now E is the set of genes from the human network, G is the subset of genes from E that are in one of the data sets, H the subset of genes from E that are in the other data set, and O is the set of genes that are in the overlap between G and H.

3.5.2 Topological properties of our predictions

We analyze the topological position of a node in the network with respect to *seven* popular node centrality measures [42], [43], [44], [45], [46], as different measures capture different aspects of the network position of the node. The measures are: *degree centrality* (DEGC), *k-coreness centrality* (KC), *graphlet degree centrality* (GDC), *clustering coefficient centrality* (CLUSC), *betweenness centrality* (BETWC), *closeness centrality* (CLOSEC), and *eccentricity centrality* (ECC) (see Supplementary Section S1.1 for the definitions).

Using each of the above measures, we compute the centrality of each node in the human PPI network. We group the nodes into those that are: 1) our novel predictions; 2) known aging-related genes; and 3) currently not associated with aging. Then, we perform Wilcoxon Rank-Sum test between centrality values of each pair of the groups to test whether the two node groups have statistically different centrality values, i.e., topological positions in the network.

3.5.3 Functional properties of our predictions

We study the enrichment of a gene set in biological process GO terms [47]. We use: 1) all 4,137 GO terms that annotate (independent on the evidence code) at least two genes from the human PPI network and 2) 1,656 GO terms that annotate (with respect to an experimental evidence code only) at least two genes from the network. For GO term g, we compute the statistical significance of its enrichment via the hypergeometric test (Section 3.4.1), where now E is the set of genes from the network that are annotated by any GO term, G is the gene set in which we are measuring GO term enrichment, H is the subset of genes from E that are annotated by GO term g, and O is the set of genes in the overlap between G and H.

3.5.4 Functional overlap with the "ground truth" data

We measure the statistical significance of the overlap of GO terms enriched in one gene set and GO terms enriched in another gene set via the hypergeometric test (Section 3.4.1), where now E is the set of GO terms that annotate at least two genes from the human PPI network, G is the set of GO terms enriched in any one of the two gene sets, H is the set of GO terms enriched in the other gene set, and O is the set of GO terms that are in the overlap between G and H.

3.5.5 Disease-related properties of our predictions

We study the enrichment of a gene set in all 467 DO terms that annotate at least two genes from the human network [48] just as we study GO term enrichments.

3.5.6 Disease overlap with the "ground truth" data

We study the overlap of DO terms from two gene sets in the same way as when we study GO term overlaps.

3.5.7 Literature validation of our predictions

Since automatic literature validation is prone to errors, we manually search for our novel predictions in PubMed articles for an evidence that a prediction is (in)directly linked to aging. As such manual search is time-consuming, we perform it only on our highest-confidence predictions. By highest-confidence, we mean the following. We score each novel prediction based on the minimum p-value over all alignment(s) supporting the prediction (Sections 3.4.1 and 3.4.2), as well as on the number of aligners supporting the prediction, so that the lower the p-value and the more aligners, the higher the confidence score of the prediction. Specifically, we convert the p-value into the "aging confidence score" (CS_{AG}) by taking its negative logarithm. We refer to the number of aligners that support the prediction as the "aligner confidence score" (CS_{AL}). After we normalize each of the two scores by dividing it with the maximum score over all novel predictions, we compute the "total confidence score" (CS_T) of a gene as the arithmetic mean of the normalized CS_{AG} and CS_{AL} scores. Then, we search in the literature for the highest-scoring genes.

4 RESULTS

4.1 Evaluating different network aligners

4.1.1 MI-GRAAL's vs. IsoRankN's cost function under MI-GRAAL's alignment strategy

Topological evaluation. We compare MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S with respect to EC, ICS, NC, and LCCS (Section 3.3.1). MI-GRAAL's cost function outperforms IsoRankN's (and sequence-only) cost function within MI-GRAAL's alignment strategy, for each measure (Fig. 1).

One would expect that using topology only in the cost function would result in alignments that are superior in terms of topological quality to alignments produced when using both topology and sequence in the cost function, as sequence would favor aligning network parts which are functionally but not necessarily topologically similar. Surprisingly, this is not the case for any of the cost functions: using both sequence and topology results in alignments of almost the same topological quality as using topology only (Fig. 1). Yet, using some topology in the (MI-GRAAL's) cost function significantly improves topological alignment quality compared to using sequence only (Fig. 1).

Biological evaluation. We compare MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S with respect to (experimental) GO correctness (Section 3.3.1). Now, quality of the different aligners is nearly indistinguishable (Fig. 2). Thus, the choice of the cost function under MI-GRAAL's alignment strategy does not affect biological alignment quality.

One would expect that using both topology and sequence (or only sequence) would result in alignments that are biologically superior to alignments produced when using topology only, as the latter would favor aligning network parts which are topologically but not necessarily functionally similar. Surprisingly, this is not the case: using topology only results in alignments of the same biological quality as using both sequence and topology (or only sequence) (Fig. 2).

4.1.2 MI-GRAAL's vs. IsoRankN's cost function under IsoRankN's alignment strategy

We compare MI-Iso-T, MI-Iso-TS, Iso-Iso-T, Iso-Iso-TS, and X-Iso-S with respect to: the number of aligned clusters, exact cluster ratio, exact protein ratio, mean entropy, and mean normalized entropy (Section 3.3.2). Using both topology and sequence within MI-GRAAL's cost function significantly outperforms using topology only as well as using both topology and sequence within IsoRankN's cost function, with respect to each measure; also, it outperforms using sequence-only cost function (Table 2). Further, using topology only within MI-GRAAL's cost function outperforms using topology only as well as using both topology and sequence within IsoRankN's cost function for three of the five measures; also, it outperforms using sequence only for two of the five

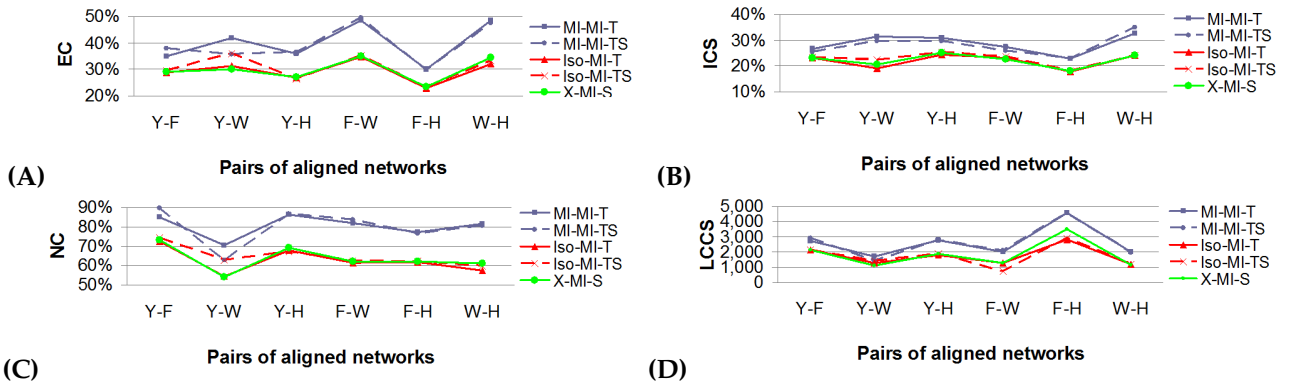


Fig. 1. Topological alignment quality of different cost functions under MI-GRAAL's alignment strategy. **(A)** EC, **(B)** ICS, **(C)** NC, and **(D)** LCCS scores of MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S when they are used on each pair of networks of yeast (Y), fly (F), worm (W), and human (H). The size of the LCCS is given in terms of the number of its nodes; results are equivalent with respect to the number of edges in the LCCS (results not shown). In all panels, the higher the values, the better the alignment quality.

measures (Table 2). Hence, MI-GRAAL's cost function dominates IsoRankN's (as well as sequence-only) cost function under IsoRankN's alignment strategy.

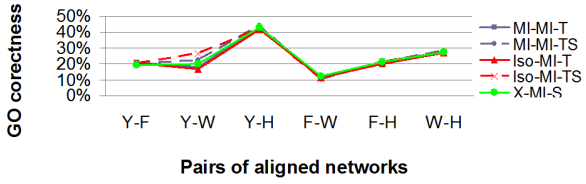


Fig. 2. Biological alignment quality of different cost functions under MI-GRAAL's alignment strategy. GO correctness scores of MI-MI-T, MI-MI-TS, Iso-MI-T, Iso-MI-TS, and X-MI-S are shown, when they are used on each pair of networks of yeast (Y), fly (F), worm (W), and human (H). The higher the score, the better the alignment quality. The results are consistent for *experimental* GO correctness (Supplementary Fig. S1).

4.1.3 *Evaluation: summary and significance*

MI-GRAAL's cost function is superior to IsoRankN's (and sequence-only) cost function, independent on the alignment strategy or evaluation measure. We recommend using MI-GRAAL's cost function under both MI-GRAAL's and IsoRankN's alignment strategy, and we recommend using both topology and sequence in MI-GRAAL's cost function (especially under IsoRankN's alignment strategy). Since the sequence-only aligners (X-MI-S and X-Iso-S) do not add to alignment quality compared to using some topology in the cost function, we do not consider them in Section 4.2.

The significance of our results is as follows. First, we show that it is indeed important to consider *separately* the contribution of a method's cost function and its alignment strategy to its alignment quality, which the research community has typically failed to do so far. Second, while the original IsoRankN, which combines IsoRankN's cost function and IsoRankN's alignment strategy, has been considered as a state-of-the-art method for *multiple* GNA, we demonstrate that combing MI-GRAAL's cost function with IsoRankN's alignment strategy outperforms the original IsoRankN. Therefore, an important contribution of our study is a new superior method for multiple GNA, namely MI-Iso-TS.

4.2 Network alignment in the context of aging

Since human aging is hard to study experimentally, the aging-related knowledge needs to be transferred from model species. This transfer has mostly been restricted to sequence alignment [49], [35]. But, since topology and sequence can give complementary biological insights [12], [11], and since not all aging-related genes implicated in aging in model species have sequence orthologs in human [50], restricting comparison to sequence may limit the knowledge transfer. Network alignment can help by transferring aging-related knowledge between conserved network regions of different species.

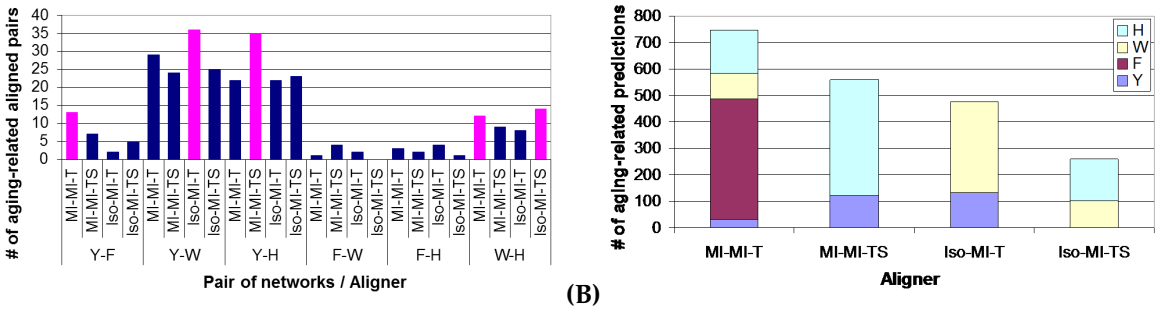
Hence, we evaluate the eight combinations of cost functions and alignment strategies from Table 1 (without considering X-MI-S and X-Iso-S sequence-only aligners) by measuring how well they can uncover existing aging-related knowledge, in the sense that they align known aging-related network parts of one species to known aging-related network parts of other species. Then, from the alignments that achieve this with statistically significantly high accuracy, we predict new aging-related knowledge in currently unannotated network regions whenever such regions have been aligned to known aging-related network regions.

4.2.1 *Aging under MI-GRAAL's alignment strategy*

In four of six pairs of species, at least one of MI-MI-T, MI-MI-TS, Iso-MI-T, and Iso-MI-TS aligns a sta-

TABLE 2
Biological alignment quality of different cost functions under IsoRankN's alignment strategy. MI-Iso-T, MI-Iso-TS, Iso-Iso-T, Iso-Iso-TS, and X-Iso-S are evaluated when aligning all networks simultaneously with respect to five measures. The higher the values for the first three measures and the lower the values for the last two measures, the better the alignment quality. The best aligner for each measure is shown in bold.

Measure	MI-Iso-T	MI-Iso-TS	Iso-Iso-T	Iso-Iso-TS	X-Iso-S
Number of clusters	1923	**3095**	79	99	1627
Exact cluster ratio	10.3%	**41.9%**	7.6%	10.1%	12.7%
Exact protein ratio	6.1%	**40.4%**	5.2%	5.9%	11.5%
Mean entropy	1.31	**0.66**	1.29	1.23	1.06
Mean normalized entropy	0.84	**0.54**	0.86	0.82	0.84



(A) **(B)**

Fig. 3. Performance of different cost functions under MI-GRAAL's alignment strategy in the context of aging. **(A)** The number of aligned node pairs in which both nodes are known aging-related genes, when each pair of the four networks (yeast (Y), fly (F), worm (W), and human (H)) is aligned by each of the four aligners (MI-MI-T, MI-MI-TS, Iso-MI-T, and Iso-MI-TS). Pink color indicates that the observed number of aging-related aligned pairs is statistically significant, whereas blue color indicates that it is not. **(B)** The number of *novel* aging-related genes predicted by each aligner in each species from the statistically significant alignments (in pink in panel (A)); only predictions that are currently not associated with aging are shown.

TABLE 3
MI-GRAAL's vs. IsoRankN's cost function with respect to *known* aging-related knowledge (in terms of precision, recall, and F-score) across all species under: 1) MI-GRAAL's alignment strategy when using topology only, 2) MI-GRAAL's alignment strategy when using both topology and sequence, 3) IsoRankN's alignment strategy when using topology only, and 4) IsoRankN's alignment strategy when using both topology and sequence.

	Aligners	Precision	Recall	F-score
1	MI-MI-T vs. Iso-MI-T	6.3% vs. 13.2%	5.2% vs. 7.5%	5.7% vs. 9.6%
2	MI-MI-TS vs. Iso-MI-TS	11.1% vs. 9.8%	7.3% vs. 2.9%	8.8% vs. 4.5%
3	MI-Iso-T vs. Iso-Iso-T	48.4% vs. 47.6%	7.9% vs. 1.0%	13.5% vs. 2.0%
4	MI-Iso-TS vs. Iso-Iso-TS	52.9% vs. 50.0%	18.4% vs. 0.6%	27.3% vs. 1.2%

tistically significantly large number of known aging-related genes across species (Section 3.4.1). MI-MI-T achieves this when aligning two pairs of species, whereas MI-MI-TS, Iso-MI-T, and Iso-MI-TS each achieve this when aligning one pair of species (Fig. 3 (A)). Thus, all aligners perform comparably. We use the five significant alignments (in pink in Fig. 3 (A)) to predict aging-related genes (Section 3.4.1).

In terms of the ability of the aligners to correctly predict *known* aging-related knowledge, MI-GRAAL's cost function is superior when using both topology and sequence, while IsoRankN's cost function is superior when using topology only (Table 3). Hence, again, the two are comparable. In terms of the ability of the aligners to predict *new* aging-related knowledge, MI-MI-T makes the most predictions, followed by MI-MI-TS, Iso-MI-T, and Iso-MI-TS, respectively (Fig. 3 (B)). In *human*, MI-MI-TS makes the most predictions, followed by MI-MI-T and Iso-MI-TS, whereas Iso-MI-T fails to predict any new aging-related genes (Fig. 3 (B)). Hence, in this context, MI-GRAAL's cost function is in general superior to IsoRankN's cost function.

Some overlap between new predictions of the different aligners under MI-GRAAL's strategy is encouraging (Supplementary Fig. S2). We provide the list of the predictions (http://nd.edu/∼cone/netal/ST1.xlsx).

4.2.2 *Aging under IsoRankN's alignment strategy*

Compared to Iso-Iso-T and Iso-Iso-TS, MI-Iso-T and MI-Iso-TS result in more aligned clusters that are statistically significantly enriched in known aging-related genes, and the enriched clusters cover more proteins (Section 3.4.2 and Fig. 4 (A) and (B)). Hence, MI-GRAAL's cost function is superior under IsoRankN's strategy. We use the significant clusters to predict aging-related genes (Section 3.4.2).

In terms of the ability of the aligners to correctly predict *known* aging-related knowledge, MI-GRAAL's cost function is significantly superior to IsoRankN's (Table 3). In terms of the ability of the aligners to predict *new* aging-related knowledge, MI-Iso-TS makes the most predictions, followed by MI-Iso-T, Iso-Iso-T, and Iso-Iso-TS, respectively (Fig. 4 (C)). Hence, MI-GRAAL's cost function is again dominant.

Unlike under MI-GRAAL's alignment strategy (Supplementary Fig. S2), there is no overlap in any species between new predictions of the different aligners under IsoRankN's strategy. Only in fly, two predictions overlap between MI-Iso-T and MI-Iso-TS. We provide the list of the predictions (http://nd.edu/∼cone/netal/ST2.xlsx).

4.2.3 MI-GRAAL's vs. IsoRankN's alignment strategy under same cost function

Thus far, we were unable to directly compare MI-GRAAL's and IsoRankN's alignment strategies under same cost function due to their different nature. However, here, we can indirectly compare the two, by comparing precision, recall, and F-score values of their *known* aging-related predictions, as well as by evaluating the amount of *novel* predictions that they can make. Note, however, that aging-related predictions, based on which this comparison is made, have been derived differently for the two alignment strategies (Sections 3.4.1 and 3.4.2). Plus, the comparison is done only on the biological level, ignoring any topological aspects, since IsoRankN's alignments cannot be evaluated topologically. Hence, this comparison may be unfair to MI-GRAAL.

With respect to known aging-related knowledge, interestingly, IsoRankN's alignment strategy is superior to MI-GRAAL's strategy under MI-GRAAL's cost function, while MI-GRAAL's alignment strategy is superior to IsoRankN's strategy under IsoRankN's cost function (Table 4). Overall, the best performing combination is MI-Iso-TS, as it results in the highest precision as well as recall.

With respect to novel aging-related knowledge, alignment strategy of MI-GRAAL's is always superior, independent of the cost function (Table 5). The overlap between novel predictions of the two alignment strategies under same cost function is relatively low, and only MI-MI-T and MI-Iso-T result in a statistically significant overlap (Table 5).

We further compare the two alignment strategies by identifying novel aging-related predictions in human by each alignment strategy, regardless of the cost function, and by computing the overlap of the resulting predictions with each of two external, independent sets of known human aging-related genes: ExpressionAge and DyNetAge (Section 3.1). By doing so, we find that MI-GRAAL's alignment strategy is superior to IsoRankN's. Specifically, of 720 human aging-related genes predicted under MI-GRAAL's alignment strategy, 36 are present in ExpressionAge (p-value of 0.072) and 79 are present in DyNetAge (p-value of 0.014). Hence, it is encouraging that under MI-GRAAL's alignment strategy, a statistically significantly large number of the predictions are found in independent sources of aging-related data. Of 87 human aging-related genes predicted under IsoRankN's alignment strategy, only one is present in ExpressionAge and only 11 are present in DyNetAge. None of these two overlaps under IsoRankN's alignment strategy is statistically significant. Yet, the existence of some overlap is still encouraging, since 1) IsoRankN, a *state-of-the-art* method for multiple GNA, has been used to produce the predictions, 2) the overlap could be low due to the *noisiness* of all aging-related data sets (including not just the predicted set but also ExpressionAge and DyNetAge), and 3) statistically non-significant results may still be *biologically significant* [51], [52]. Note that we further investigate topological and functional properties of the novel human predictions in Section 4.3, in order to validate them.

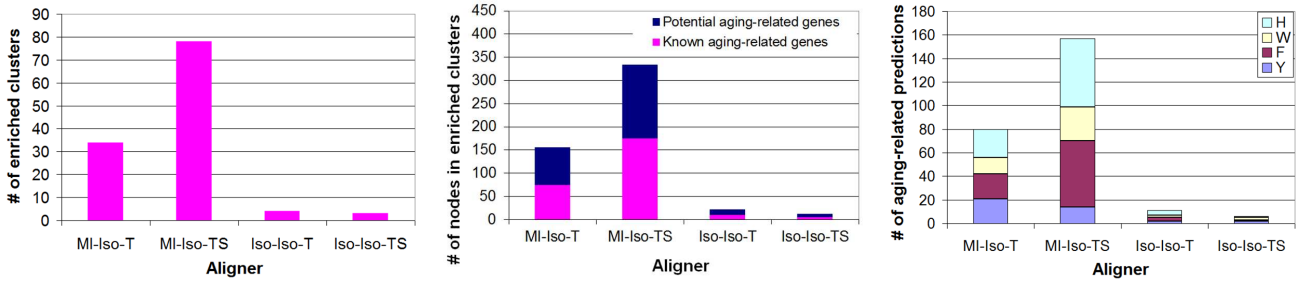
TABLE 4
MI-GRAAL's vs. IsoRankN's alignment strategy with respect to *known* aging knowledge (in terms of precision, recall, and F-score) under: 1) MI-GRAAL's cost function when using topology only, 2) MI-GRAAL's cost function when using topology and sequence, 3) IsoRankN's cost function when using topology only, and 4) IsoRankN's cost function when using topology and sequence.

	Aligners	Precision	Recall	F-score
1	MI-MI-T vs. MI-Iso-T	6.3% vs. 48.4%	5.2% vs. 7.9%	5.7% vs. 13.5%
2	MI-MI-TS vs. MI-Iso-TS	11.1% vs. 52.9%	7.3% vs. 18.4%	8.8% vs. 27.3%
3	Iso-MI-T vs. Iso-Iso-T	13.2% vs. 47.6%	7.5% vs. 1.0%	9.6% vs. 2.0%
4	Iso-MI-TS vs. Iso-Iso-TS	9.8% vs. 50.0%	2.9% vs. 0.6%	4.5% vs. 1.2%

4.2.4 Aging through GNA: summary and significance

First, we compare the two cost functions under same alignment strategy. MI-GRAAL's cost function is overall superior over that of IsoRankN independent on the alignment strategy, both when extracting known and predicting new aging-related knowledge, especially when using both topology and sequence in the cost function. Hence, consistent to our evaluation results (Section 4.1.3), we again recommend using MI-GRAAL's cost function under both alignment strategies, and we recommend using both topology and sequence information in MI-GRAAL's cost function, especially under IsoRankN's alignment strategy.

Second, we compare the two alignment strategies under same cost function. IsoRankN's alignment strategy is superior under MI-GRAAL's cost function,



(A) **(B)** **(C)**

Fig. 4. Performance of different cost functions under IsoRankN's alignment strategy in the context of aging. **(A)** The number of clusters that are statistically significantly enriched in known aging-related genes. **(B)** The number of proteins in the statistically significantly enriched clusters, broken down into known aging-related genes and genes that are currently not associated with aging. **(C)** The number of *novel* aging-related genes predicted by each aligner in each of the species (yeast (Y), fly (F), worm (W), and human (H)) from the statistically significant clusters; only predictions that are currently not associated with aging are shown.

while MI-GRAAL's alignment strategy is superior under IsoRank's cost function. The highest prediction accuracy is achieved by MI-Iso-TS. Thus, if one's goal is to increase the prediction accuracy, we recommend using this aligner. If instead one's goal is to increase topological similarity between networks while still producing statistically significant aging-related predictions, we recommend using MI-GRAAL's alignment strategy. MI-GRAAL's strategy is also superior with respect to the overlap of novel predictions with independent existing aging-related knowledge.

We produce new human aging-related knowledge via network alignment, complementing current knowledge obtained mainly via sequence alignment.

TABLE 5
MI-GRAAL's vs. IsoRankN's alignment strategy with respect to *new* aging-related knowledge predicted under: 1) MI-GRAAL's cost function when using topology only, 2) MI-GRAAL's cost function when using topology and sequence, 3) IsoRankN's cost function when using topology only, and 4) IsoRankN's cost function when using topology and sequence. The table lists the number of predictions ("N") for each aligner, the size of the overlap between predictions of two given aligners ("O"), and p-value of the overlap ("p"). "N/A" means that p-value could not be computed due to no overlap.

	Aligners	N	O	p
1	MI-MI-T vs. MI-Iso-T	747 vs. 80	9	0.005
2	MI-MI-TS vs. MI-Iso-TS	558 vs. 157	2	0.951
4	Iso-MI-T vs. Iso-Iso-T	475 vs. 11	0	N/A
4	Iso-MI-TS vs. Iso-Iso-TS	259 vs. 6	0	N/A

4.3 Validation of novel aging-related predictions

Next, we aim to validate our novel aging-related predictions in human by studying their topological and functional properties, in hope to link the properties to current knowledge about aging (Sections 4.3.1–4.3.7). For this purpose, we denote each of the eight sets of novel predictions in human produced by one of the eight aligners as *Novel-"aligner"*. For example, *Novel-MI-Iso-TS*, would correspond to the set of novel human predictions by MI-Iso-TS. Also, we denote the set of all predictions in human (the union over all aligners) as *Novel-All*. Finally, we denote by *Complement* the set of all genes from the human PPI network minus any gene from Novel-All or GenAge. That is, Complement contains what can be considered the least likely aging-related candidates. We use Complement as a negative control data in the following sections, in hope that properties of our novel predictions will be significantly similar to properties of known aging-related genes, whereas they will be significantly dissimilar to properties of genes from Complement. Indeed, this is exactly what we observe.

4.3.1 Prediction overlap with the "ground truth" data

To validate our novel predictions (Novel-All), we first compute their overlap with each of ExpressionAge and DyNetAge (Section 3.5.1). Since the novel predictions are disjoint with GenAge, to compute the overlaps fairly, here we exclude genes from ExpressionAge and DyNetAge that are also in GenAge. It is highly encouraging that Novel-All overlaps significantly with DyNetAge (with 70 genes in the overlap; p-value of 1.83×10^{-3}) and marginally significantly with ExpressionAge (with 36 genes in the overlap; p-value of 0.08), whereas its overlap with Complement is non-significant (p-value of above 0.5).

4.3.2 Topological properties of our predictions

Since it has been argued that aging-related genes play a central role in the network [53], we evaluate topological positions of our novel predictions in the human PPI network with respect to seven centrality measures (Section 3.5.2). We hope that topological positions of our predictions are significantly different

that those of genes in Complement, whereas they are similar to topological positions of known aging-related genes (Section 3.5.2). And this is exactly what we observe (Fig. 5 and Supplementary Table S1).

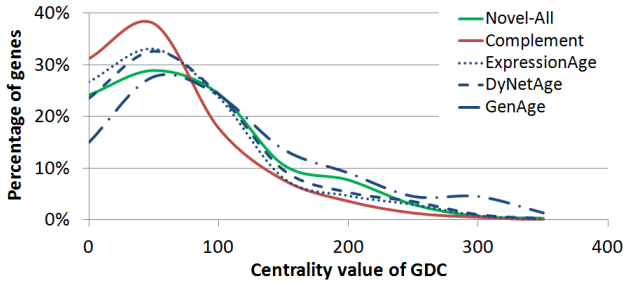


Fig. 5. GDC distribution of genes in Novel-All, ExpressionAge, DyNetAge, GenAge, and Complement. Complement is significantly less central compared to all other data sets, including Novel-All, whereas Novel-all behaves similarly as the three known aging-related data sets. Similar trends are observed for other centralities as well (Supplementary Table S1).

4.3.3 *Functional properties of our predictions*

We compute enrichment of our novel predictions in biological process GO terms (Section 3.5.3), in hope that these GO terms are significantly similar to those enriched in known aging-related genes, while they are dissimilar to those enriched in Complement. And this is exactly what we observe (Section 4.3.4). Supplementary Table S2 summarizes the number of GO terms enriched in each of the data sets.

4.3.4 *Functional overlap with the "ground truth" data*

Indeed, we observe high overlap between GO terms enriched in our novel predictions and those enriched in known aging-related data sets, whereas their overlap with GO terms enriched in Complement is extremely low (Fig. 6 (A)), all of which validates our predictions. When we focus on GO terms based on any evidence code, GO terms of Novel-All significantly overlap with GO terms of DyNetAge (p-value of 4.23×10^{-4}); it is highly encouraging that GO terms enriched in the aging-related predictions from two different network-based approaches overlap significantly. Also, 5 GO terms overlap between Novel-All and ExpressionAge, and 12 GO terms overlap between Novel-All and SequenceAge (Fig. 6 (A)), although these overlaps are not statistically significant.

When we focus on the novel predictions by MI-Iso-TS (the best aligner in the context aging with respect to F-score; Table 3), it is highly encouraging that GO terms enriched in this set significantly overlap with GenAge (p-value of 0.02), as the two sets are completely disjoint. Furthermore, we observe significant GO term overlaps between Novel-MI-MI-T and ExpressionAge (p-value of 2.89×10^{-3}), Novel-MI-MI-T and DyNetAge (p-value of 0.03), Novel-MI-MI-TS and DyNetAge (p-value of 4.25×10^{-5}), Novel-Iso-MI-TS and DyNetAge (p-value of 2.32×10^{-4}), and Novel-Iso-MI-TS and GenAge (p-value of 0.02).

When we focus on GO data based on experimental evidence codes only, the number of enriched GO terms is significantly reduced for each data set (Supplementary Table S2). Therefore, GO overlaps between the data sets are expected to reduce as well. Yet, we still observe some (marginally) significant overlaps: between Novel-MI-MI-TS and DyNetAge (p-value of 9.65×10^{-3}) and between Novel-All and DyNetAge (p-value of 0.07).

Importantly, none of our novel prediction sets significantly overlap with Complement.

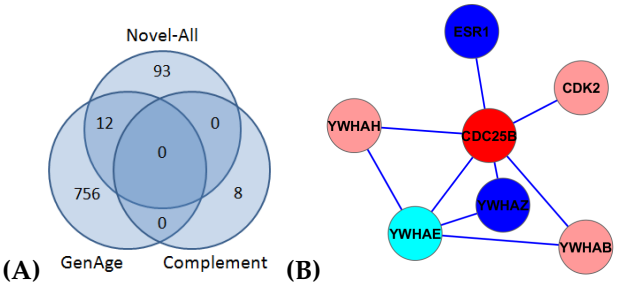


Fig. 6. **(A)** GO term overlap between Novel-All, GenAge, and Complement. Trends are similar for DyNetAge and ExpressionAge. **(B)** Our novel prediction CDC25B and its induced neighborhood in the human network. The neighbors in dark and light blue are from GenAge and ExpressionAge, respectively.

4.3.5 *Disease-related properties of our predictions*

We compute enrichment of our novel predictions in DO terms as well to further investigate whether the enriched diseases have any connection to aging. We find that this *is* the case. *Mental retardation*, enriched in Novel-All (p-value of 0.03), has been linked to aging (PubMed ID (PMID): 15823058). Mental retardation is also enriched in Novel-MI-MI-T (p-value of 0.04). Furthermore, *brain tumor* is enriched in MI-MI-T (p-value of 0.04), *prostate cancer* is enriched in Iso-MI-TS (p-value of 0.02), and *cancer* is enriched in MI-Iso-TS (p-value of 0.03). Note that tumors and cancers are known to be linked to aging (PMID: 17942417).

4.3.6 *Disease overlap with the "ground-truth" data*

It is encouraging that DO terms enriched in our novel predictions overlap with those enriched in known aging-related data sets. *Pancreas disease* is enriched in both Novel-MI-MI-T and GenAge. This disease is associated with aging because telomere shortening, an important risk factor for this disease (PMID: 23093543), is directly linked to aging (PMID: 24246679). *Prostate cancer* is enriched in both Novel-Iso-MI-T and GenAge, while cancers are known to be linked with aging (PMID: 17942417). For the same

reason, it is encouraging that *cancer* is enriched in both Novel-MI-Iso-TS and GenAge. *Kidney disease* is enriched in both Novel-MI-Iso-T and GenAge, while aging has been found to be a vital risk factor for this disease (PMID: 15888561). We note that very few DO terms are enriched in our novel predictions, and therefore, in most cases, statistically it makes no sense to compute the significance of the overlaps.

4.3.7 Literature validation of our predictions

We aim to manually validate our top-scoring novel predictions in the literature. After ranking all of the predictions (Section 3.5.7), there is a single top-scoring gene, but there are already 33 second top-scoring genes that are tied, which might be too many for manual literature search. In this case, we give priority to a gene identified by the best aligner in terms of F-score (Table 3) over a tied gene identified by a different aligner. This results in the total of 22 novel predictions with top two confidence scores, which we then attempt to link to aging in the literature. We successfully find aging-related evidence for 20 (91%) of the 22 genes (Table 6). As an illustration, we describe the aging-related evidence for several of the 20 genes, as follows. For example, KLHDC5, TANK, and TMPRSS3 have been linked to Alzheimer's disease (PMID:21705112,PMID: 9300664), ATP5B, MVK, and MYOD1 have been linked to Hutchinson-Gilfold progeria syndrome (PMID:21346760), and DSC3 has been linked to Parkinson disease (PMID: 23184149), whereas all of these diseases have been linked to aging. For the remaining predictions, see Table 6.

TABLE 6
PubMed ID (PMID) evidence of 20 of our predictions.

Gene	PMID	Gene	PMID
RGPD5	[54]	DSC3	23184149
AP2B1, MRPL4, STON2, TAC3, TNS1, ZBTB8A	18832152	ATPAF1	12783983
		CALM2	22719074
		CDKN2D	11103932
KLHDC5, TANK	21705112	FANCC	23303816
ATP5B, MVK, MYOD1	21346760	LRP8	19519777
JUP	18582489	TMPRSS3	9300664

4.3.8 Novel predictions: summary and significance

We have validated our novel predictions in a number of ways, demonstrating that they share topological and functional properties with known aging-related genes, whereas they differ from unlikely aging-related candidates. We finalize the analysis by *illustrating* the interconnectedness of one of our novel predictions in human PPI network with other genes, many of which are known aging-related genes (Fig. 6 (B)), which further validates potential involvement of this prediction in an aging-associated biological pathway.

5 Conclusions

We comprehensively evaluate combinations of cost functions and alignment strategies of two state-of-the-art GNA methods. The existing combination of MI-GRAAL's cost function and its alignment strategy remains superior for pairwise GNA. But our new combination of MI-GRAAL's cost function and IsoRankN's alignment strategy beats the existing combination of IsoRankN's cost function and its alignment strategy, which has been state-of-the-art for multiple GNA. Thus, we propose it as a new superior method.

We show that GNA can align with statistically significantly high accuracy known aging-related network parts across species. Thus, we transfer known aging-related knowledge from well annotated species to poorly annotated species (including human) between aligned network regions, hence producing novel and valuable aging-related knowledge. We validate our novel predictions in a number of ways.

Acknowledgments

This work was supported by NSF CCF-1319469 and NSF EAGER CCF-1243295 grants.

References

[1] S. Altschul, W. Gish, W. Miller, and D. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, pp. 403–410, 1990.
[2] M. Yandell and W. Majoros, "Genomics and Natural Language Processing," *Nature Reviews Genetics*, vol. 3, no. 8, pp. 601–610, 2002.
[3] A. Narayanan, E. Shi, and B. Rubinstein, "Link prediction by de-anonymization: How we won the Kaggle social network challenge," in *Proceedings of the 2011 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2011, pp. 1825–1834.
[4] B. J. Breitkreutz *et al.*, "The BioGRID Interaction Database: 2008 update," *Nucleic Acids Research*, vol. 36, pp. D637–D640, 2008.
[5] R. Solava and T. Milenković, "Revealing missing parts of the interactome," *arXiv:1307.3329 [q-bio.MN]*, 2013.
[6] R. Sharan and T. Ideker, "Modeling cellular machinery through biological network comparison," *Nature Biotechnology*, vol. 24, no. 4, pp. 427–433, 2006.
[7] O. Kuchaiev *et al.*, "Topological network alignment uncovers biological function and phylogeny," *Journal of the Royal Society Interface*, vol. 7, pp. 1341–1354, 2010.
[8] O. Kuchaiev and N. Pržulj, "Integrative network alignment reveals large regions of global network similarity in yeast and human," *Bioinformatics*, vol. 27, no. 10, pp. 1390–1396, 2011.
[9] C. Liao *et al.*, "IsoRankN: spectral methods for global alignment of multiple protein networks," *Bioinformatics*, vol. 25, no. 12, pp. i253–258, 2009.
[10] R. Sharan, I. Ulitsky, and R. Shamir, "Network-based prediction of protein function," *Molecular Systems Biology*, vol. 3, no. 88, pp. 1–13, 2007.
[11] V. Memisević, T. Milenković, and N. Pržulj, "Complementarity of network and sequence information in homologous proteins," *Journal of Integrative Bioinformatics*, vol. 7(3):135, 2010.
[12] D. Gautheret, F. Major, and R. Cedergren, "Pattern searching/alignment with RNA primary and secondary structures: an effective descriptor for tRNA," *Computer Applications in the Biosciences*, vol. 6, no. 4, pp. 325–331, 1990.
[13] T. Milenković, W. Ng, W. Hayes, and N. Pržulj, "Optimal network alignment with graphlet degree vectors," *Cancer Informatics*, vol. 9, pp. 121–137, 2010.
[14] B. Kelley *et al.*, "PathBLAST: a tool for alignment of protein interaction networks," *Nucleic Acids Research*, vol. 32, pp. 83–88, 2004.
[15] R. Sharan *et al.*, "Conserved patterns of protein interaction in multiple species," *Proceedings of the National Academy of Sciences*, vol. 102, no. 6, pp. 1974–1979, 2005.

[16] J. Flannick *et al.*, "Graemlin: General and robust alignment of multiple large interaction networks," *Genome Research*, vol. 16, no. 9, pp. 1169–1181, 2006.
[17] M. Koyuturk *et al.*, "Pairwise alignment of protein interaction networks," *Journal of Computational Biology*, vol. 13, no. 2, 2006.
[18] J. Berg and M. Lassig, "Local graph alignment and motif search in biological networks," *Proceedings of the National Academy of Sciences*, vol. 101, pp. 14 689–14 694, 2004.
[19] Z. Liang, M. Xu, M. Teng, and L. Niu, "NetAlign: a web-based tool for comparison of protein interaction networks," *Bioinformatics*, vol. 22, no. 17, pp. 2175–2177, 2006.
[20] J. Berg and M. Lassig, "Cross-species analysis of biological networks by Bayesian alignment," *Proceedings of the National Academy of Sciences*, vol. 103, no. 29, pp. 10 967–10 972, 2006.
[21] R. Singh, J. Xu, and B. Berger, "Pairwise global alignment of protein interaction networks by matching neighborhood topology," in *Research in Computational Molecular Biology (RECOMB)*. Springer, 2007, pp. 16–31.
[22] J. Flannick *et al.*, "Automatic parameter learning for multiple network alignment," in *Research in Computational Molecular Biology (RECOMB)*, 2008, pp. 214–231.
[23] R. Singh, J. Xu, and B. Berger, "Global alignment of multiple protein interaction networks," *Proceedings of Pacific Symposium on Biocomputing*, pp. 303–314, 2008.
[24] M. Zaslavskiy, F. Bach, and J. P. Vert, "Global alignment of protein-protein interaction networks by graph matching methods," *Bioinformatics*, vol. 25, no. 12, pp. i259–i267, 2009.
[25] R. Patro and C. Kingsford, "Global network alignment using multiscale spectral signatures," *Bioinformatics*, vol. 28, no. 23, pp. 3105–3114, 2012.
[26] X. Guo and A. Hartemink, "Domain-oriented edge-based alignment of protein interaction networks." *Bioinformatics*, vol. 25, no. 12, pp. i240–1246, 2009.
[27] G. Klau, "A new graph-based method for pairwise global network alignment." *BMC Bioinformatics*, vol. 10, no. Suppl 1, p. S59, 2009.
[28] V. Memišević and N. Pržulj, "C-GRAAL: Common-neighbors-based global graph alignment of biological networks," *Integrative Biology*, vol. 4, no. 7, pp. 734–743, 2012.
[29] L. Chindelevitch, C.-Y. Ma, C.-S. Liao, and B. Berger, "Optimizing a global alignment of protein interaction networks," *Bioinformatics*, vol. 29, no. 21, pp. 2765–2773, 2013.
[30] B. Neyshabur, A. Khadem, S. Hashemifar, and S. S. Arab, "Netal: a new graph-based method for global alignment of proteinprotein interaction networks," *Bioinformatics*, vol. 29, no. 13, pp. 1654–1662, 2013.
[31] V. Saraph and T. Milenković, "Magna: Maximizing accuracy in global network alignment," *arXiv:1311.2452 [q-bio.MN]*, 2013.
[32] D. West, *Introduction to Graph Theory*, 2nd ed. Prentice Hall, Upper Saddle River, NJ, 2001.
[33] K. Venkatesan *et al.*, "An empirical framework for binary interactome mapping." *Nature Methods*, vol. 6, no. 1, pp. 83–90, 2009.
[34] T. G. O. Consortium, "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, pp. 25–29, 2000.
[35] J. de Magalhães *et al.*, "The Human Ageing Genomic Resources: online databases and tools for biogerontologists." *Aging Cell*, vol. 8, no. 1, pp. 65–72, 2009.
[36] T. Lu *et al.*, "Gene regulation and DNA damage in the ageing human brain," *Nature*, vol. 429, no. 6994, pp. 883–891, 2004.
[37] F. E. Faisal and T. Milenković, "Dynamic networks reveal key players in aging," *arXiv:1307.3388 [cs.CE]*, 2013.
[38] D. W. Huang *et al.*, "David gene id conversion tool," *Bioinformatics*, vol. 2, no. 10, pp. 428–430, 2008.
[39] T. Milenković and N. Pržulj, "Uncovering biological network function via graphlet degree signatures," *Cancer Informatics*, vol. 6, pp. 257–273, 2008.
[40] R. Solava, R. Michaels, and T. Milenković, "Graphlet-based edge clustering reveals pathogen-interacting proteins," *Bioinformatics*, vol. 18, no. 28, pp. i480–i486, 2012, also, in Proceedings of the *11th European Conference on Computational Biology* (ECCB), Basel, Switzerland, September 9-12, 2012 (acceptance rate: 14%).
[41] N. Pržulj, D. G. Corneil, and I. Jurisica, "Modeling interactome: Scale-free or geometric?" *Bioinformatics*, vol. 20, no. 18, pp. 3508–3515, 2004.
[42] A. Barabási and Z. Oltvai, "Network biology: Understanding the cell's functional organization," *Nature Reviews*, vol. 5, pp. 101–113, 2004.
[43] R. Sharan and T. Ideker, "Protein networks in disease," *Genome Research*, vol. 18, pp. 644–652, 2008.
[44] T. Milenković, V. Memišević, A. Bonato, and N. Pržulj, "Dominating biological networks," *PLOS ONE*, vol. 6, no. 8, p. e23016, 2011.
[45] P. Radivojac *et al.*, "An integrated approach to inferring gene-disease associations in humans," *Proteins*, vol. 72, no. 3, pp. 1030–1037, 2008.
[46] V. Janjić and N. Pržulj, "The core diseasome," *Molecular bioSystems*, vol. 8, no. 10, pp. 2614–25, 2012.
[47] M. Ashburner *et al.*, "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
[48] P. Du *et al.*, "From disease ontology to disease-ontology lite: statistical methods to adapt a general-purpose ontology for the test of gene-ontology associations," *Bioinformatics*, vol. 25, no. 12, pp. i63–68, 2009.
[49] D. Wieser, I. Papatheodorou, M. Ziehm, and J. Thornton, "Computational biology for ageing," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1561, pp. 51–63, 2011.
[50] R. Tacutu, A. Budovsky, and V. Fraifeld, "The netage database: a compendium of networks for longevity, age-related diseases and associated processes," *Biogerontology*, vol. 11, no. 4, 2010.
[51] H. Motulsky, *Intuitive Biostatistics*, 1st ed. Oxford University Press, 1995.
[52] H. Ho *et al.*, "Protein interaction network uncovers melanogenesis regulatory network components within functional genomics datasets," *BMC Systems Biology*, vol. 4, no. 84, 2010.
[53] P. Langfelder, P. S. Mischel, and S. Horvath, "When is hub gene selection better than standard meta-analysis?" *PLOS ONE*, vol. 8, no. 4, p. e61505, 2013.
[54] R. Tacutu *et al.*, "Human ageing genomic resources: Integrated databases and tools for the biology and genetics of ageing," *Nucleic Acids Research*, vol. 41, no. D1, pp. D1027–D1033, 2013.

Fazle Elahi Faisal has been a Ph.D. student in the Department of Computer Science and Engineering at the University of Notre Dame since 2012. He earned his M.Sc. in Bioinformatics from the University of Memphis in 2011, and his M.Sc and B.Sc. in Computer Science and Engineering from Bangladesh University of Engineering and Technology in 2008 and 2004, respectively. His research is in network-based computational biology.

Han Zhao got his B.E. degree in Computer Science from Tsinghua University in 2013. He has been a Masters candidate in David R. Cheriton School of Computer Science at the University of Waterloo since 2013. He was visiting the Department of Computer Science and Engineering at the University of Notre Dame in 2012. His research is in machine learning and its applications to natural language processing and network mining.

Dr. Tijana Milenković is an Assistant Professor in the Department of Computer Science and Engineering at the University of Notre Dame. She earned her Ph.D. and M.Sc. in Computer Science from the University of California, Irvine in 2010 and 2008, respectively, and her B.Sc. in Electrical Engineering and Computer Science from the University of Sarajevo in 2005. Her research is in complex networks and computational (systems) biology. She is funded by NSF, NIH, and Google.