TelNet - a database for human and yeast genes involved in telomere maintenance

Delia M. Braun[1], Inn Chung[1], Nick Kepper[1] & Karsten Rippe[1,*]

[1] German Cancer Research Center (DKFZ) and Bioquant, Research Group Genome Organization & Function, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

* Corresponding author: Karsten.Rippe@dkfz.de, Tel: +49-6221-54-51376, Fax: +49-6221-54-51487;

Keywords: telomerase, alternative telomere lengthening, pathway analysis

Abstract

Telomeres, the ends of linear chromosomes, consist of repetitive DNA sequences bound by the shelterin protein complex. Cancer cells need to extend these telomere repeats for an unlimited proliferation potential, either by reactivating the reverse transcriptase telomerase or by using the alternative lengthening of telomeres (ALT) pathway. The different telomere maintenance (TM) mechanisms are likely to involve hundreds of genes but their telomere repeat length related activities are only partly understood. Currently, a database that integrates information on TM relevant genes is missing. To provide a reference for studies that dissect TM features, we here introduce the *TelNet* database available via the http://www.cancertelsys.org/telnet link. It offers a comprehensive compilation of more than 1,900 human and over 1,100 yeast genes linked to telomere maintenance. The genes compiled in *TelNet* were annotated in terms of TM mechanism (telomerase or ALT), TM specific functions, validation of significance for TM including a calculated *TelNet* significance score, homology assignment between different organisms, and information from the literature. Different search modes to display the results are available that display TM information on a gene card, via a gene list view or as a statistics page of a TM pathway analysis. Thus, *TelNet* can be integrated into the annotation of genes identified from bioinformatic analysis pipelines to determine possible connections with TM networks. We anticipate that *TelNet* will be a helpful resource for researchers that study TM processes.

Database URL: http://www.cancertelsys.org/telnet

Introduction

Telomeres, the ends of linear chromosomes, consist of repetitive DNA sequences bound by the shelterin protein complex (1,2). This protein assembly protects the DNA ends from degradation and accidental recognition as DNA double-strand breaks (3-5). The progressive shortening of the telomere repeat sequences that accompanies normal replication limits the number of cell divisions. Thus, it needs to be circumvented by cancer cells in order to obtain the ability of unlimited proliferation. This is accomplished by activation of a telomere maintenance (TM) mechanism. It involves either the reactivation of the reverse transcriptase telomerase (6), which is normally repressed in somatic cells, or activation of the alternative lengthening of telomeres (ALT) pathway (7). For ALT in human cancer cells, it is known that DNA damage response, recombination and repair machineries play important roles but details on the mechanism remain elusive (8-11).

From recent studies, it is emerging that TM is a complex process that involves hundreds of different genes. Recent screening studies have revealed proteins that are located at telomere repeats (12,13) or in close proximity to a shelterin component (14). Several factors have been shown to regulate transcription and functional activity of the ribonucleoprotein telomerase (15,16). A characteristic feature of ALT is the presence ALT-associated PML nuclear bodies (APBs), which are complexes of PML nuclear bodies with telomeres that are associated with a variety of proteins (17-20). Furthermore, investigations of deregulating effects upon telomere shortening have related proteins to telomere crisis (21).

A well-studied model organism for telomere biology is the budding yeast *Saccharomyces cerevisiae (22)*. Several independent deletion screens with subsequent telomere length analysis have identified a comprehensive list of yeast genes involved in telomere length regulation (23-25). Since telomere structure and function are highly conserved between organisms, mammalian homologues exist for most of the genes identified in the various yeast screens. Thus, it can be informative to relate TM phenotypes found in yeast to human homologues (26). In *S. cerevisiae*, telomerase is constitutively active and its deletion leads to cellular senescence (27). Survivor cells that overcome cellular senescence use a mechanism based on homologous recombination for telomere elongation (28). Interestingly, similar to ALT in human cells, so-called type II survivors are characterized by heterogeneous telomere lengths (29,30).

Currently, the following databases have been published that provide some telomere-relevant information: The *TeloPIN* database (Telomeric Proteins Interaction Network, http://songyanglab.sysu.edu.cn/telopin/index.php) is a collection of interaction data in human and mouse cells from scientific literature and GEO database. It includes interactions between shelterin compounds and other proteins identified from protein-protein interactions

and protein interactions of the non-coding telomeric repeat-containing RNA (TERRA) (31). It also provides information on the methods used to identify the respective interaction. The *Telomerase database* (http://telomerase.asu.edu/overview.html) is a web-based tool for the study of structure, function, and evolution of the telomerase ribonucleoprotein (32). The objective of this database is to provide a comprehensive compilation of information on the telomerase enzyme and its DNA substrate. The collection of information includes sequences, alignments, secondary and tertiary structures, mutations known to cause human telomerase-deficiency diseases, and current telomerase researchers. The previously published *TeCK database* (http://www.bioinfosastra.com/services/teck/index.html) is a collection of telomeric and centromeric sequences as well as telomerase, centromere and kinetochore binding proteins This database appears to be no longer available (33). The *MiCroKiTS* (Midbody, Centrosome, Kinetochore, Telomere and Spindle) database focuses on information about proteins relevant for the molecular mechanisms during cell division and also includes telomere proteins (34). It contains information on the cellular localization of proteins in the course of cell cycle progression (http://microkit.biocuckoo.org).

All the above-mentioned databases cover telomere-related information but lack an annotation to the TM mechanism. Accordingly, we here introduce a compilation of information on telomere maintenance relevant genes via the *TelNet* database. *TelNet* comprises more than 1,900 human and over 1,100 yeast genes that are involved in TM pathways. The annotation of these genes includes the classification of telomere maintenance mechanisms along with a significance score as well as TM specific functions and homology assignments between different organisms. Furthermore, links to the respective literature sources are given. Thus, *TelNet* provides an integrative platform for dissecting TM networks and elucidating the alternative lengthening of telomeres pathway.

Construction and user interface

Database development and implementation

The *TelNet* database was constructed using Filemaker Pro Advanced version 13. It is available at http://www.cancertelsys.org/telnet and is distributed via Filemaker server version 15. It runs via the Filemaker webdirect software within a browser without the need to install additional software. The *TelNet* webpage provides general information about *TelNet* and the construction of the database as well instructions on how to use it. Links to other databases and contact information are provided. Currently, users can log into the *TelNet* database via selecting the guest account (no password needed).

TelNet user interface

On the front layout of *TelNet* an overview on the gene compilation for the two organisms *H. sapiens* and *S. cerevisiae* is presented (Fig. 1). Here, the total number of genes as well as their classification according to their TM significance is listed. The user can select either human or yeast genes (Fig. 2). A card view shows all information for an individual gene and scrollable list view displays selected information. The navigation menu can be found on the top left. Each gene was annotated in terms of general information from external databases and TM information from peer-reviewed literature. A short explanation of each annotation field is given by clicking on the corresponding button. The general part comprises the gene symbol, full name, synonyms, and the cellular function. Frequently used gene, transcript and protein identifiers were selected for each individual organism and are linked to the respective webpage. In the TM part, we provided information on the TM specific function of the selected gene and its role in TM mechanisms (Fig. 3). Genes were categorized according to their TM significance and the *TelNet* score is automatically calculated from provided information. Database entries from different organisms are connected via database hyperlinks. Literature information is linked for every single gene. Three different search modes enable convenient exploration of the *TelNet* database (Fig. 2). For a quick search throughout all fields a keyword can be entered into the search bar. If a user wishes to constrain the found set to a gene name, the ‘Find gene’ button can be pressed. By performing an advanced search, the user can enter specific search terms in corresponding fields. Furthermore, a complete list of gene identifiers can also be passed into the list search. The organism and identifier provided are mandatory to perform a list search. Genes found are then listed and can be selected to be included for further analysis. For a gene result list a database statistics page can be displayed that describes the performed search and gives a graphical overview over the distributions of various *TelNet* annotations. These results can be exported as a tab separated file for further analysis.

Compilation and annotation of human and yeast TM genes

To compile an initial set of TM relevant genes, we selected screening studies on genes or proteins that play a role in telomere biology (Table 1): (i) Proteins that bind to a telomere probe in an ALT- and a telomerase-positive cell line (12). (ii) Telomere proteins from the analysis of telomeric chromatin from telomerase-positive cells (13). (iii) Proteins binding in close proximity to at least one of the shelterin components (14). (iv) Proteins that affect ALT-associated promyelocytic leukemia (PML) nuclear bodies (20,35). (v) Deregulated proteins linked to telomere shortening caused by the telomerase inhibitor GRN163 (Geron) (21). (vi) Genes identified from a comparison of different published telomerase activity signatures derived from gene expression data used to predict telomerase activity (36). (vii) Potential telomerase regulators identified from a kinase library by a quantitative telomeric repeat amplification protocol (15). (viii) Telomerase regulators that act on the transcriptional level as reviewed recently (16). (ix) A gene set with potential relevance to telomeres and the ALT pathway (37).

The assignment of TM significance was as follows: Genes with a suggested role in telomere maintenance but lacking experimentally validation were ranked as ‘predicted’. The genes collected from the above-mentioned screening or review sources were classified as ‘screened’. Proteins with detailed experimental evidence for a connection to telomere maintenance were assigned as ‘validated’.

In budding yeast, deletion screens were performed that have directly evaluated if a given gene was linked to alterations in telomere length (23-25). Furthermore, telomere structures in post-senescent survivors of telomere-length-maintenance gene mutants have been investigated after telomerase knockout (38). Based on this analysis, human orthologues were included in the *TelNet* database as 'predicted', if a yeast gene’s TM significance was categorized as ‘screened’ or ‘validated’. Likewise, also yeast genes were included in cases where the respective human orthologue shows a TM significance higher than ‘predicted’. Moreover, all human and yeast genes with a GO annotation containing the term ‘telo’ (39) were included into the *TelNet* database. Finally, also transcription factors of telomerase were retrieved from Yeastract (40). In this manner, we compiled an initial list of human and yeast genes that was further curated and annotated manually.

General information from external databases

For a standardized nomenclature (41), we converted gene identifiers by the ID converter system from DAVID Bioinformatics Resources (https://david.ncifcrf.gov/) (42) or the BioMart tool from Ensembl (43). For each species, we individually selected identifiers that are

frequently used: Entrez Gene ID, Hugo Gene ID, Ensembl Gene ID, Locus tag, Refseq IDs and UniProt ID, and retrieved the information from Ensembl (44,45). To account for organism-specific differences such as the lack of splicing and therefore isoforms in yeast or the absence of locus tags in human, the identifiers were selected differentially for each species. General gene information was retrieved from designated external databases and repositories, such as the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) (46), HUGO Gene Nomenclature Committee (HGNC, http://www.genenames.org) (47), Ensembl (http://www.ensembl.org/index.html) (44,45), and the Saccharomyces Genome Database (SGD, http://www.yeastgenome.org) (48). The approved gene symbol, full name, and synonyms were also assigned from NCBI. UniProt (49) and Yeastmine (50) were consulted for the description of the cellular function in human and yeast, respectively. Moreover, we used YeastMine for the assignment of homologues. Based on the functional annotations from Gene Ontology (GO, http://www.geneontology.org) (39) and in line with biocuration guidelines (51) we generated a reasonable list of cellular functions similar to curation procedures at SGD (52). Every gene was manually annotated with the respective term that is most representative for its cellular function. In summary, general information for every single gene entry is applied from a variety of external databases.

Telomere maintenance annotation with literature information and scoring

Genes were further annotated with TM information from peer-reviewed literature for functional categorizing. The ‘TM annotation’ distinguishes how a gene is associated with the TM mechanism (telomerase or alternative lengthening of telomeres (ALT) pathway). Possible ways of association are ‘repressing’, ‘promoting’, or ‘conflicting’, with the latter describing cases where it was unclear if a gene has a repressing or promoting activity. The combination of assignments results in a dynamically determined predicted TM mechanism that is either ‘ALT’, ‘Telomerase-positive’ or ‘ambiguous’. The attribute of ‘ambiguous’ contains genes lacking any information on TM mechanism as well as genes with conflicting associations. For yeast genes, an additional classification for a role in survivor formation was annotated. We next assembled a list of ‘TM functions’ that comprises molecular functions as well as cellular processes and structures with regard to TM (Fig. 2). Up to five TM functions can be annotated according to the available information for the gene of interest. A knock-out or knock-down phenotype related to TM features such as alterations in telomere length, increased or decreased ALT hallmarks, or effects on telomerase was described as free-text in the field ‘TM phenotype’. All information from the literature were integratively summarized in the field called ‘TM comment’.

To quantify the significance of a given gene for TM we introduced the *TelNet* score that is automatically calculated from information entered into the *TelNet* database. Table 2 displays the scoring according to criteria like 'telomere biology' as a cellular function term, number and relevance of assigned TM functions, and the amount of experimental data associated with the TM function of a given gene. The resulting *TelNet* scores were rounded to integral numbers leading to a number from 1 to 23. Under 'statistics', a histogram for a given gene result set is automatically computed, providing the user with an overview on the *TelNet* score. Thus, genes were annotated with telomere relevant information extracted from peer-reviewed literature and TM significance can be assessed from the *TelNet* score.

TM pathway statistics of a gene result set

We implemented three main ways to search the *TelNet* database (Fig. 3). (i) The 'quick search' consults all database fields, whereas 'find gene' constrains the find to identifier and name fields. (ii) The 'advanced search' conducted with a TM function or predicted TMM for instance compiles a found set of genes for further analysis and TM network identification. (iii) A list of genes can be correlated with *TelNet* genes by using the 'list search' function. A given result set of genes is summarized on the statistics page, which can be found in the navigation menu. The statistics page consists of a summary of the performed search, a histogram of the *TelNet* score, and distribution of TM significance. Histograms displaying the TM annotation, and predicted TMM as well as cellular and TM functions are in the style of a conventional pathway analysis. These data can be exported as a tab separated file.

In order to further illustrate the use of our database, we correlated the *TelNet* genes with two gene lists from a current study on telomere length associations (36). In the latter study, reactome gene expression pathway analysis of the top 500 genes revealed an immunoreactive signature for genes associated with long telomeres (LT) and a proliferative signature for genes correlating with a short telomere (ST) phenotype. As concluded by the authors these findings could indicate differences in telomere maintenance. A *TelNet* list search with these two published gene lists revealed a number of genes that were in our database (Fig. 4A and B). Furthermore, *TelNet* statistics can be employed for a more detailed pathway analysis regarding telomere functions as well as indications for an associated telomere maintenance mechanism. This allows a further analysis of a gene result set, for example from a differential gene expression study or gene insertions/mutations, similar to existing pathway analyses. Thus, compilation of a result set can be achieved by three different ways of search with the integrated TM information being displayed on the statistics page together with the TM pathway analysis.

Discussion

The *TelNet* database was developed to provide a comprehensive collection of TM relevant genes. It is an ongoing research project and we will add new information on telomere maintenance to already existing gene entries or generate new entries as necessary. In addition, an extension of *TelNet* is planned to include also genes from other organisms such as *M. musculus*. We encourage other researchers working on telomeres to communicate their feedback to the curators to keep the information in the database accurate, current, and relevant via e-mail to telnet@dkfz.de. The annotations provided by *TelNet* allow the distinction between different types of TM mechanisms for a gene (set) of interest according to functional terms and a significance ranking. With these features *TelNet* supports the identification of TM networks in various ways. A gene set derived from a preceding bioinformatics analysis pipelines can be used for a *TelNet* list search to get more detailed insight on the corresponding TM associated genes. Possible TM links can be explored in an iterative manner. For example, current pan-cancer studies provide a wealth of information on mutated or deregulated genes that can be evaluated in terms of the associated TM mechanism (36). Some associations like mutations in *ATRX* and *DAXX* for ALT as well as *TERT* promoter mutations for telomerase-positive cells are well established. However, in many tumor samples mutations in these loci are absent. One would expect that for these cases the mutation status of a given cancer sample and its active TM are linked via other genes that might involve a combination of various factors. We anticipate that *TelNet* will prove to be a helpful analysis tool for revealing this type of correlations and to support the identification of active TM networks in different tumor entities.

Funding

This work was supported by the project CancerTelSys (01ZX1302) in the e:Med program of the German Federal Ministry of Education and Research (BMBF).

Acknowledgements

The authors thank all present and former members of the CancerTelSys Consortium especially Katharina Deeg, Brian Luke, André Maicher, Jan-Philipp Mallm, Alexandra Poos, and Rainer König for discussion and data contribution.

References

1. de Lange, T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev*, **19**, 2100-2110.
2. Blackburn, E.H., Epel, E.S. and Lin, J. (2015) Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science*, **350**, 1193-1198.
3. Denchi, E.L. and de Lange, T. (2007) Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature*, **448**, 1068-1071.
4. Palm, W. and de Lange, T. (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet*, **42**, 301-334.
5. Lazzerini-Denchi, E. and Sfeir, A. (2016) Stop pulling my strings - what telomeres taught us about the DNA damage response. *Nat Rev Mol Cell Biol*, **17**, 364-378.
6. Shay, J.W. (2016) Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov*, **6**, 584-593.
7. Cesare, A.J. and Reddel, R.R. (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet*, **11**, 319-330.
8. Zhong, Z.H., Jiang, W.Q., Cesare, A.J., Neumann, A.A., Wadhwa, R. and Reddel, R.R. (2007) Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J Biol Chem*, **282**, 29314-29322.
9. Draskovic, I. and Londono Vallejo, A. (2013) Telomere recombination and alternative telomere lengthening mechanisms. *Front Biosci (Landmark Ed)*, **18**, 1-20.
10. Gocha, A.R., Acharya, S. and Groden, J. (2014) WRN loss induces switching of telomerase-independent mechanisms of telomere elongation. *PLoS ONE*, **9**, e93991.
11. Dilley, R.L. and Greenberg, R.A. (2015) ALTernative Telomere Maintenance and Cancer. *Trends Cancer*, **1**, 145-156.
12. Dejardin, J. and Kingston, R.E. (2009) Purification of proteins associated with specific genomic Loci. *Cell*, **136**, 175-186.
13. Grolimund, L., Aeby, E., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M. and Lingner, J. (2013) A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun*, **4**, 2848.
14. Lee, O.H., Kim, H., He, Q., Baek, H.J., Yang, D., Chen, L.Y., Liang, J., Chae, H.K., Safari, A., Liu, D. *et al.* (2011) Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells. *Mol Cell Proteomics*, **10**, M110 001628.
15. Cerone, M.A., Burgess, D.J., Naceur-Lombardelli, C., Lord, C.J. and Ashworth, A. (2011) High-throughput RNAi screening reveals novel regulators of telomerase. *Cancer Res*, **71**, 3328-3340.
16. Ramlee, M.K., Wang, J., Toh, W.X. and Li, S. (2016) Transcription Regulation of the Human Telomerase Reverse Transcriptase (hTERT) Gene. *Genes (Basel)*, **7**, 50.
17. Yeager, T.R., Neumann, A.A., Englezou, A., Huschtscha, L.I., Noble, J.R. and Reddel, R.R. (1999) Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res*, **59**, 4175-4179.
18. Nabetani, A. and Ishikawa, F. (2011) Alternative lengthening of telomeres pathway: recombination-mediated telomere maintenance mechanism in human cells. *J Biochem*, **149**, 5-14.
19. Chung, I., Osterwald, S., Deeg, K.I. and Rippe, K. (2012) PML body meets telomere: the beginning of an ALTernate ending? *Nucleus*, **3**, 263-275.
20. Osterwald, S., Deeg, K.I., Chung, I., Parisotto, D., Wörz, S., Rohr, K., Erfle, H. and Rippe, K. (2015) PML induces compaction, TRF2 depletion and DNA damage signaling at telomeres and promotes their alternative lengthening. *J Cell Sci*, **128**, 1887-1900.
21. Uziel, O., Yosef, N., Sharan, R., Ruppin, E., Kupiec, M., Kushnir, M., Beery, E., Cohen-Diker, T., Nordenberg, J. and Lahav, M. (2015) The effects of telomere shortening on cancer cells: a network model of proteomic and microRNA analysis. *Genomics*, **105**, 5-16.
22. Kupiec, M. (2014) Biology of telomeres: lessons from budding yeast. *Fems Microbiol Rev*, **38**, 144-171.
23. Askree, S.H., Yehuda, T., Smolikov, S., Gurevich, R., Hawk, J., Coker, C., Krauskopf, A., Kupiec, M. and McEachern, M.J. (2004) A genome-wide screen for Saccharomyces cerevisiae deletion mutants that affect telomere length. *Proc Natl Acad Sci U S A*, **101**, 8658-8663.

24. Gatbonton, T., Imbesi, M., Nelson, M., Akey, J.M., Ruderfer, D.M., Kruglyak, L., Simon, J.A. and Bedalov, A. (2006) Telomere length as a quantitative trait: genome-wide survey and genetic mapping of telomere length-control genes in yeast. *PLoS Genet*, **2**, e35.
25. Ungar, L., Yosef, N., Sela, Y., Sharan, R., Ruppin, E. and Kupiec, M. (2009) A genome-wide screen for essential yeast genes that affect telomere length maintenance. *Nucleic Acids Res*, **37**, 3840-3849.
26. Lippuner, A.D., Julou, T. and Barral, Y. (2014) Budding yeast as a model organism to study the effects of age. *Fems Microbiol Rev*, **38**, 300-325.
27. Lundblad, V. (2002) Telomere maintenance without telomerase. *Oncogene*, **21**, 522-531.
28. Lundblad, V. and Blackburn, E.H. (1993) An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell*, **73**, 347-360.
29. Chen, Q., Ijpma, A. and Greider, C.W. (2001) Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Mol Cell Biol*, **21**, 1819-1827.
30. Teng, S.C. and Zakian, V.A. (1999) Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in Saccharomyces cerevisiae. *Mol Cell Biol*, **19**, 8083-8093.
31. Luo, Z., Dai, Z., Xie, X., Feng, X., Liu, D., Songyang, Z. and Xiong, Y. (2015) TeloPIN: a database of telomeric proteins interaction network in mammalian cells. *Database (Oxford)*, **2015**, bav018-bav018.
32. Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X. and Chen, J.J. (2008) The telomerase database. *Nucleic Acids Res*, **36**, D339-343.
33. Gowthaman, R., Krishnamoorthy, S., Nandakumar, R.D. and Ayyarappan, V. (2007) TeCK database: a comprehensive collection of telomeric and centromeric sequences with their associated proteins. *Bioinformation*, **2**, 73-75.
34. Huang, Z., Ma, L., Wang, Y., Pan, Z., Ren, J., Liu, Z. and Xue, Y. (2015) MiCroKiTS 4.0: a database of midbody, centrosome, kinetochore, telomere and spindle. *Nucleic Acids Res*, **43**, D328-334.
35. Jiang, W.Q., Zhong, Z.H., Henson, J.D. and Reddel, R.R. (2007) Identification of candidate alternative lengthening of telomeres genes by methionine restriction and RNA interference. *Oncogene*, **26**, 4635-4647.
36. Barthel, F.P., Wei, W., Tang, M., Martinez-Ledesma, E., Hu, X., Amin, S.B., Akdemir, K.C., Seth, S., Song, X., Wang, Q. *et al.* (2017) Systematic analysis of telomere length and somatic alterations in 31 cancer types. *Nat Genet*, **49**, 349-357.
37. Lovejoy, C.A., Li, W., Reisenweber, S., Thongthip, S., Bruno, J., de Lange, T., De, S., Petrini, J.H., Sung, P.A., Jasin, M. *et al.* (2012) Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet*, **8**, e1002772.
38. Hu, Y., Tang, H.B., Liu, N.N., Tong, X.J., Dang, W., Duan, Y.M., Fu, X.H., Zhang, Y., Peng, J., Meng, F.L. *et al.* (2013) Telomerase-null survivor screening identifies novel telomere recombination regulators. *PLoS Genet*, **9**, e1003208.
39. Gene Ontology, C. (2015) Gene Ontology Consortium: going forward. *Nucleic Acids Res*, **43**, D1049-1056.
40. Teixeira, M.C., Monteiro, P.T., Guerreiro, J.F., Goncalves, J.P., Mira, N.P., dos Santos, S.C., Cabrito, T.R., Palma, M., Costa, C., Francisco, A.P. *et al.* (2014) The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae. *Nucleic Acids Res*, **42**, D161-166.
41. Klionsky, D.J., Bruford, E.A., Cherry, J.M., Hodgkin, J., Laulederkind, S.J. and Singer, A.G. (2012) In the beginning there was babble. *Autophagy*, **8**, 1165-1167.
42. Huang da, W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, **4**, 44-57.
43. Kinsella, R.J., Kahari, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A. *et al.* (2011) Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)*, **2011**, bar030.
44. Aken, B.L., Achuthan, P., Akanni, W., Amode, M.R., Bernsdorff, F., Bhai, J., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P. *et al.* (2017) Ensembl 2017. *Nucleic Acids Res*, **45**, D635-D642.
45. Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Fernandez Banet, J., Billis, K., Garcia Giron, C., Hourlier, T. *et al.* (2016) The Ensembl gene annotation system. *Database (Oxford)*, **2016**.

46. Brown, G.R., Hem, V., Katz, K.S., Ovetsky, M., Wallin, C., Ermolaeva, O., Tolstoy, I., Tatusova, T., Pruitt, K.D., Maglott, D.R. *et al.* (2015) Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res*, **43**, D36-42.
47. Yates, B., Braschi, B., Gray, K.A., Seal, R.L., Tweedie, S. and Bruford, E.A. (2017) Genenames.org: the HGNC and VGNC resources in 2017. *Nucleic Acids Res*, **45**, D619-D625.
48. Cherry, J.M. (2015) The Saccharomyces Genome Database: A Tool for Discovery. *Cold Spring Harb Protoc*, **2015**, pdb top083840.
49. The UniProt, C. (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res*, **45**, D158-D169.
50. Balakrishnan, R., Park, J., Karra, K., Hitz, B.C., Binkley, G., Hong, E.L., Sullivan, J., Micklem, G. and Cherry, J.M. (2012) YeastMine--an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit. *Database (Oxford)*, **2012**, bar062.
51. Poux, S. and Gaudet, P. (2017) Best Practices in Manual Annotation with the Gene Ontology. *Methods Mol Biol*, **1446**, 41-54.
52. Hong, E.L., Balakrishnan, R., Dong, Q., Christie, K.R., Park, J., Binkley, G., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G. *et al.* (2008) Gene Ontology annotations at SGD: new data sources and annotation methods. *Nucleic Acids Res*, **36**, D577-581.

Table 1. Screening studies included in *TelNet* for identification of TM genes

Method	Organism / cell line	# genes or proteins	Ref.
Proteomics of isolated chromatin segments (PICh)	human / Wi38-VA13 and HeLa 1.2.11	296	(12)
Quantitative telomeric chromatin isolation protocol (QTIP)	human / HeLa	34	(13)
Protein complementation assay (PCA/bimolecular fluorescent complementation (BiFC)) of shelterin compounds and ~12.000 candidate genes; GST-pulldown of FLAG-tagged genes	human / HTC75	339	(14)
siRNA mediated knockdown; APB formation	human / U2OS	29	(20,35)
Proteomic analysis of deregulated genes upon telomere shortening caused by the telomerase inhibitor GRN163	human / SK-N-MC	99	(21)
QTRAP: kinase library screen	human / HeLa	109	(15)
Telomerase regulators that act on transcriptional level	human / various	53	(16)
Gene set with potential relevance to telomeres and the ALT pathway	human / various	297	(37)
Telomerase activity signature	human	43	(36)
GO annotation containing the term 'telo'	human	245	(39)
Haploid deletion screen, telomere length	yeast	166	(23)
Telomere length-variation screen in deletion strains	yeast	138	(24)
Screen of DAmP collection	yeast	77	(25)
Telomerase null screen of yeast mutants	yeast	270	(38)
Telomerase regulators	yeast	35	(40)
GO annotation containing the term 'telo'	yeast	192	(39)

Selected screening studies and other sources with information on TM genes are shown. The method used for the respective investigation as well as the organism and if available the human cell line is indicated. Also, the number of genes, added from the given literature source is listed.

Table 2. Calculation of the *TelNet* score

Condition	Score
If a gene is included in the *TelNet* database	1 point
If the **TM significance** is 'validated'	5 points
If the protein/ gene has been identified in a **screen.** If it has been found in several screens, 1 point per screen is assigned.	1 point
If the **TM significance** is 'predicted'. Here, 1/10 of the homologs *TelNet* Score is added.	0-2 points
If the **cellular function** contains the term 'telomere'	3 points
For each **TM function**	1 point
If the **TM function** contains the terms ‘telomere’, ‘telomerase’, ‘telomeric repeat’, ‘TERRA’, or ‘APB’	1 point

The computation of the *TelNet* score is shown. The *TelNet* score is calculated as a sum of information that was entered into the *TelNet* database. In particular, TM significance and TM functions are employed for calculation.

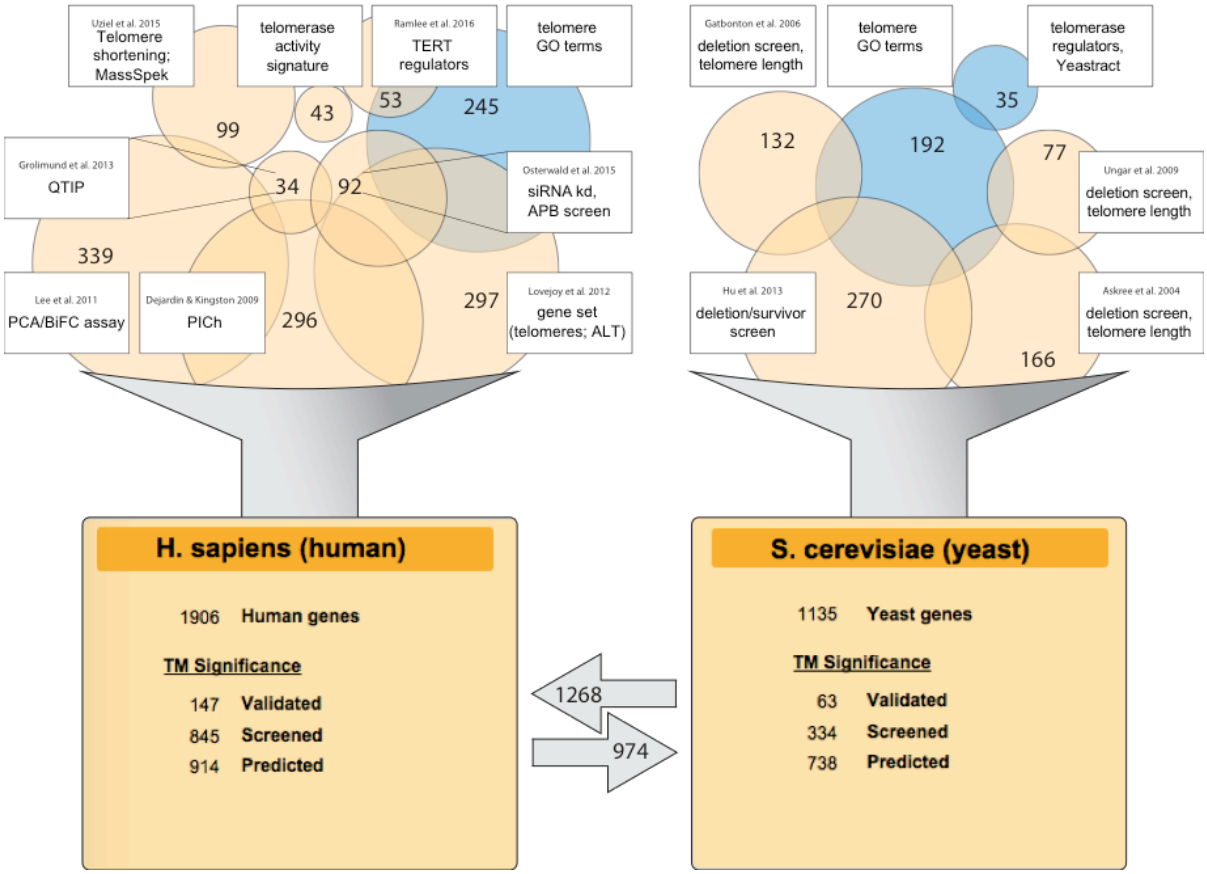

Figure 1. Data sources of TM genes for *TelNet*

Selected screening studies and other references that serve as sources for TM genes are indicated. In total over 1,900 human genes and more than 1,100 yeast genes were included in the *TelNet* database. More than 600 human and yeast genes were added to the gene list of the other organism, respectively, based on homology assignments.

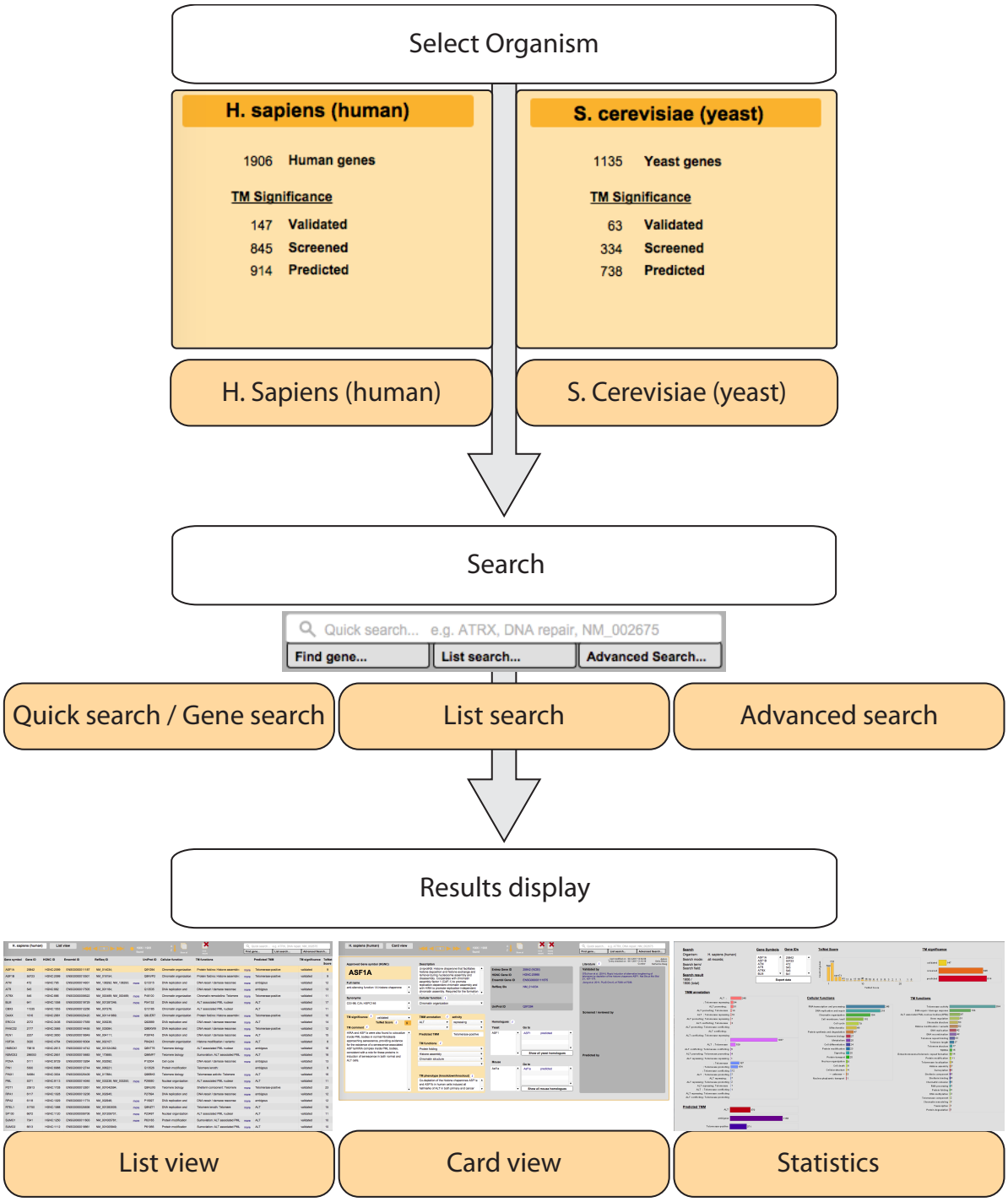

Figure 2. Typical *TelNet* workflow

Top: On the start page, the organism is selected. *Middle:* Three different kinds of search tools, namely quick search, list search and advanced search, are available to retrieve a set of genes. *Bottom:* The resulting genes can be displayed as a scrollable list or as a series of single gene cards. An overview of the associated TM annotations is obtained from the statistics page.

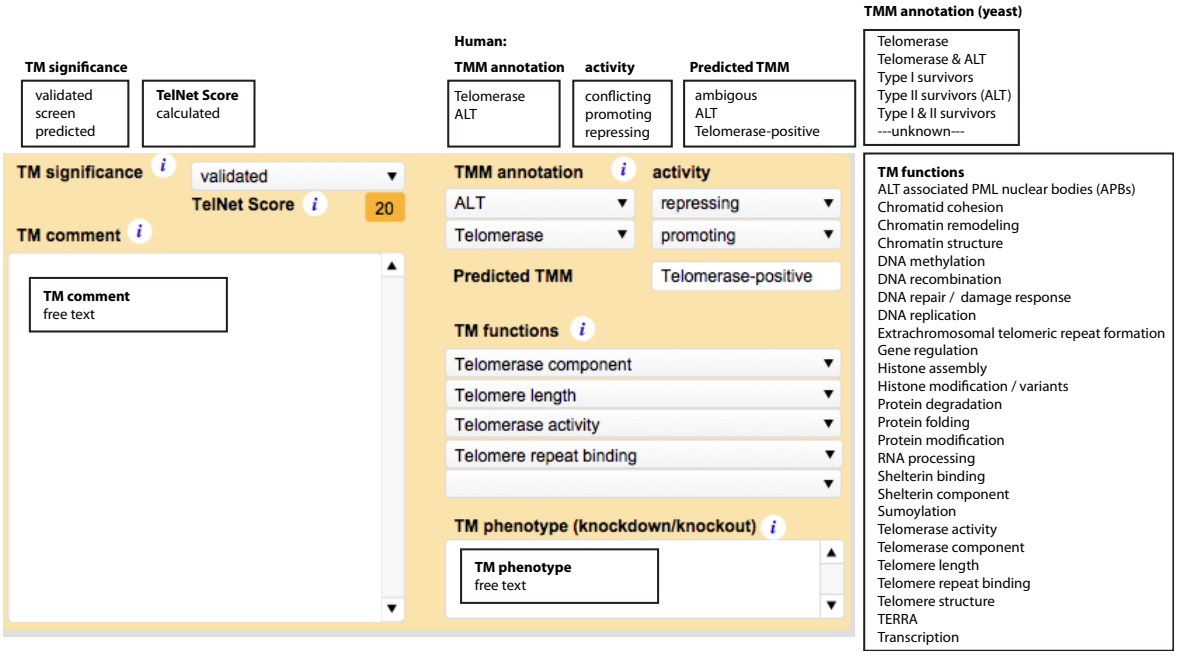

Figure 3. Gene card view of the telomere maintenance information

Possible entry terms are listed in the respective annotation field for TM significance associated with the *TelNet* score, TM comment, details on the regulation of TM ('TMM annotation'), TM functions, and TM phenotype.

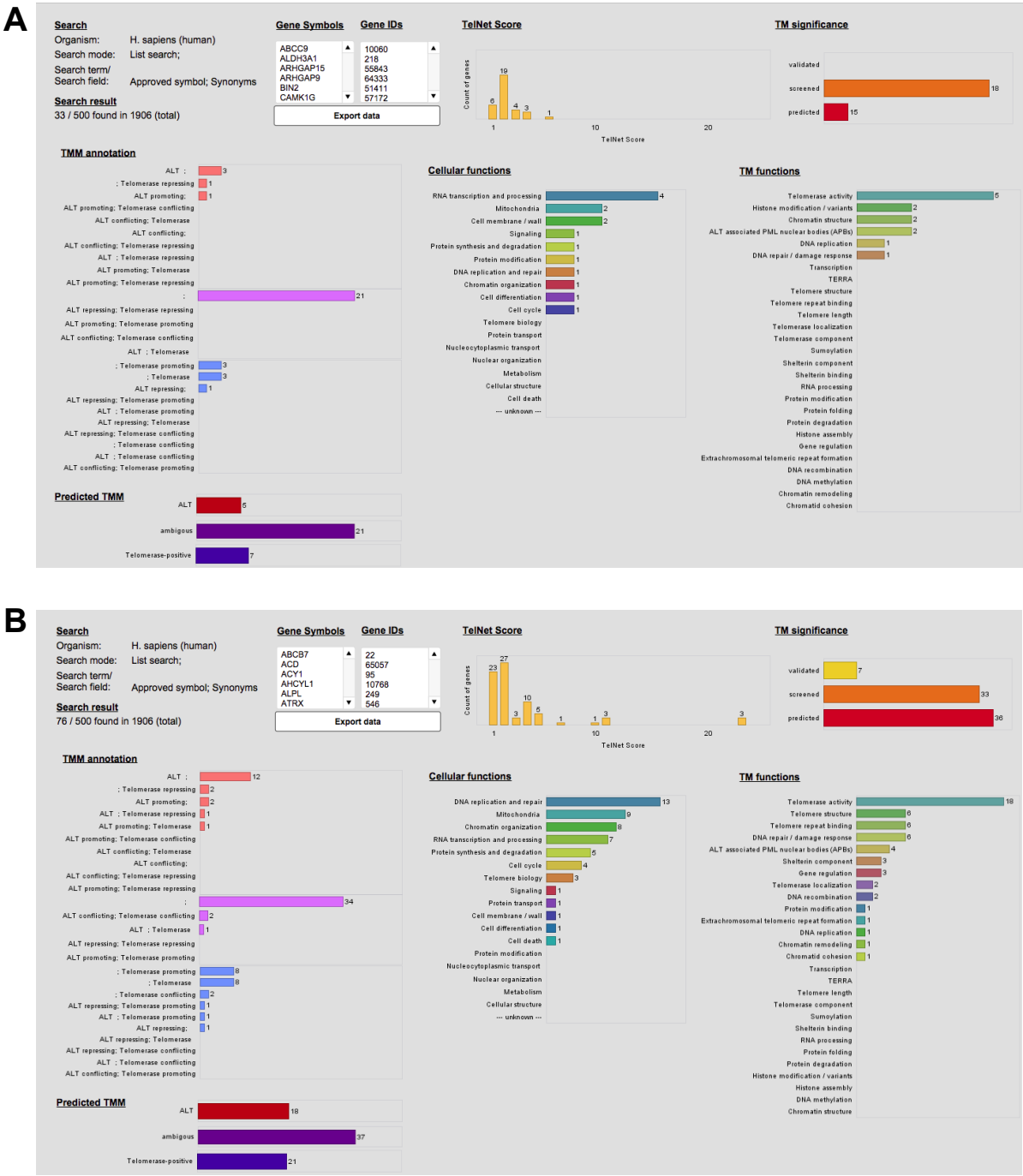

Figure 4. Statistics page for analysis of a gene list

Two gene lists associated with long or short telomeres were obtain from a study on telomere length associations (36). The top 500 genes of each group were submitted to *TelNet* to obtain statistics and histograms for the TelNet Score, the TM significance, the TMM annotation, predicted TMM, cellular functions, as well as TM functions. (A) Long telomere gene set that resulted in 33 *TelNet* hits for which the statistics page is displayed. (B) Same as panel A but for the short telomere gene list that yielded 76 hits.