



RESEARCH ARTICLE

Protein disorder reduced in *Saccharomyces cerevisiae* to survive heat shock [version 1; referees: 2 approved, 1 approved with reservations]

Esmeralda Vicedo^{1,3}, Zofia Gasik^{1,2}, Yu-An Dong^{1,4}, Tatyana Goldberg¹, Burkhard Rost^{1,5,6}

¹Department of Informatics, Bioinformatics & Computational Biology, TUM, Munich, Germany

²Graduate School of Information Science in Health, TUM, Munich, Germany

³Institute of Experimental Physics, Division of Biophysics, University of Warsaw, Warsaw, Poland

⁴Institute of Systems Biology, Shanghai University, Shanghai, China

⁵Institute of Advanced Study, TUM, Munich, Germany

⁶Institute for Food and Plant Sciences WZW, TUM, Freising, Germany

v1 First published: 06 Nov 2015, 4:1222 (doi: [10.12688/f1000research.7178.1](https://doi.org/10.12688/f1000research.7178.1))
 Latest published: 06 Nov 2015, 4:1222 (doi: [10.12688/f1000research.7178.1](https://doi.org/10.12688/f1000research.7178.1))

Abstract

Recent experiments established that a culture of *Saccharomyces cerevisiae* (baker's yeast) survives sudden high temperatures by specifically duplicating the entire chromosome III and two chromosomal fragments (from IV and XII). Heat shock proteins (HSPs) are not significantly over-abundant in the duplication. In contrast, we suggest a simple algorithm to "postdict" the experimental results: Find a small enough chromosome with minimal protein disorder and duplicate this region. This algorithm largely explains all observed duplications. In particular, all regions duplicated in the experiment reduced the overall content of protein disorder. The differential analysis of the functional makeup of the duplication remained inconclusive. Gene Ontology (GO) enrichment suggested over-representation in processes related to reproduction and nutrient uptake. Analyzing the protein-protein interaction network (PPI) revealed that few network-central proteins were duplicated. The predictive hypothesis hinges upon the concept of reducing proteins with long regions of disorder in order to become less sensitive to heat shock attack.

Open Peer Review

Referee Status:

	Invited Referees		
	1	2	3
version 1			
published 06 Nov 2015	report	report	report

- 1 **Melchor Sanchez-Martinez**, Mind the Byte Spain
- 2 **Paul Pavlidis**, University of British Columbia Canada
- 3 **Anuj Kumar**, University of Michigan USA

Discuss this article

Comments (0)

Corresponding author: Esmeralda Vicedo (vicedo@rostlab.org)

How to cite this article: Vicedo E, Gasik Z, Dong YA *et al.* **Protein disorder reduced in *Saccharomyces cerevisiae* to survive heat shock [version 1; referees: 2 approved, 1 approved with reservations]** *F1000Research* 2015, 4:1222 (doi: [10.12688/f1000research.7178.1](https://doi.org/10.12688/f1000research.7178.1))

Copyright: © 2015 Vicedo E *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: This work, and all authors were supported by the German Research Foundation (DFG) and the Technische Universität München within the funding program Open Access Publishing.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: No competing interests were disclosed.

First published: 06 Nov 2015, 4:1222 (doi: [10.12688/f1000research.7178.1](https://doi.org/10.12688/f1000research.7178.1))

Introduction

Saccharomyces cerevisiae (baker's yeast; for simplicity we mostly use yeast) was the first completely sequenced eukaryote¹. Being simple to handle and manipulate has rendered yeast a preferred model organism for genetics, biochemistry and systems biology²⁻⁴. It grows optimally within a narrow temperature range but tolerates moderate deviations, some of which impinge upon cell structure and function, often through rapid physiological adaptations. One such adaptation mechanism is the duplication of the whole genome or particular chromosomes (aneuploidy)⁵⁻⁷ that contain the genes necessary to rapidly cope with specific adverse conditions over the course of several generations of evolving yeast⁸⁻¹⁴. Such evolutionary adaptations imbalance the genome¹⁵, destabilize reactions or pathways^{16,17}, and cost energy^{18,19}. Aneuploidy, therefore, is a transient solution. Over many generations exposed to the same adverse conditions refined specific and less expensive solutions replace aneuploidy²⁰. Yeast cells can adapt to high-temperature stress by repeatedly duplicating chromosome III along with two other fragments (from chrIV and chrXII)²⁰. Why specifically copy these regions? Can particular biophysical features and/or functions of the proteins encoded in these regions explain the choice?

One simple biophysical feature is protein disorder: most proteins adopt well-defined three-dimensional (3D) structures²¹⁻²⁴, *i.e.* will largely remain identical at different times. In contrast, *disordered regions* do not adopt well-defined 3D structures in isolation²⁵, *i.e.* without binding substrates they will look very different at different time points. Proteins with long disordered regions encompass some unique biophysical characteristics²⁶⁻³⁴. Such regions are so difficult to characterize experimentally that there is no good experimental data set proxy for "all proteins with long regions of disorder in yeast". In contrast, acceptably accurate computational predictions are available for entire proteomes^{30,35,36}. Protein disorder seems one means for prokaryotes to adapt to extreme environments, *e.g.* halophiles have more proteins with long disorder than their closest phylogenetic relatives, while thermophiles tend to have fewer³⁷. Here, we hypothesized a similar effect to govern the response to high temperature-related duplication in yeast, namely that chromosomal regions duplicated under high temperature are depleted of proteins with long disorder.

Methods

Data

We downloaded the yeast (*S. cerevisiae*) proteome from UniProt (proteome ID: UP000002311)³⁸ as fasta files including only the reviewed proteins (UniProtKB/Swiss-Prot). Removal of duplicates applying the method Uniqueprot2³⁹ (with 100% pairwise sequence identity, keeping the longer sequence) left 5667 proteins (Table S1A). We considered the 16 nuclear chromosomes (matched through <http://www.yeastgenome.org>⁴⁰, the numbers of proteins per chromosome are given in Figure S1B). The yeastgenome.org resource also provided the annotations of heat-shock response proteins (HSR). Proteins known to interact with HSR proteins augmented this set of HSRs in the following way.

BioGRID (version 3.1.86) provided the data for experimental protein-protein interactions (PPIs) in yeast. After filtering out redundancy (a-b and b-a counted only once) and excluding self-interactions

(a-a), we based all subsequent analysis on the single largest connected component of the network. We focused on the most basic network features that allow the comparison and characterization of complex networks. The most elementary characteristic of a node is its degree or connectivity, defined as the number of interactions for a node (here protein), *i.e.* the number of interactions one protein has with all others. Another important parameter is the betweenness, *i.e.* the fraction of shortest paths between all other nodes that has to go through a given node. Additionally, we monitored the average degree of neighbors, which depends on the number of nodes and links in the network. These three parameters measured the importance of each node within the network.

Disorder predictions

We applied methods capturing different "flavors" of protein disorder^{29,41,42}. IUPred (version 1.0) is based on statistical contact potentials and exclusively uses single sequences^{28,43}. MD (Meta-Disorder)⁴² combines different original prediction methods through machine learning (neural network) with evolutionary profiles and predictions of solvent accessibility and protein flexibility. To some extent disorder is a gradual phenomenon, *i.e.* proteins may have more or less disorder⁴⁴. On the other hand, prediction methods distinguish between a 30-residue loop resembling "protein disorder" and another resembling a region with "regular structure"²⁹. Thus, protein disorder seems more a binary feature (it is there or not, or present/absent) than a gradual one²⁵. Unfortunately, no argument or data determines one single correct threshold for what constitutes present/absent for protein disorder. Typically, experts use a length threshold of the type: protein disorder is present when at least T consecutive residues in a protein are predicted to be disordered. If so, this protein is considered to contain a long region of disorder. More disorder in this model could imply, *e.g.* more than one region, or the entire protein. We analyzed many alternatives to choose the threshold for long disorder, and found most to be redundant. We included different views only if they provided relevant information. In particular, we largely focused on one threshold to define "long disorder": %long30, is the percentage of proteins with at least one region of ≥ 30 consecutive residues predicted as disordered (alternatives were: %long50 and %long80, *i.e.* with length thresholds at ≥ 50 and ≥ 80 , and completely disordered implying no region of 30 consecutive residues without any disorder).

GO term enrichment

We applied BINGO (Biological Networks Gene Ontology⁴⁴, version 2.44) to identify the enrichment of GO (Gene Ontology)⁴⁵ terms in subsets of experimentally annotated proteins. We focused on "biological process" and "molecular function". For two sets of proteins with annotated biological functions (more precisely: GO numbers) BINGO estimates to which extent their annotations differ in a statistically significant way. We visualized BINGO results with Cytoscape⁴⁶ platform (version 2.8). Our analysis focused on the hypergeometric test in BINGO, which accurately estimates p-values as it tests without replacement. Following the common procedure for BINGO, we considered p-values >0.05 as significant⁴⁶. Testing multiple hypotheses may give many false positives (Type I error: incorrect rejection of true null hypothesis^{47,48}). Using BINGO, we corrected for these through the Benjamini and Hochberg correction which provides strong control over the False Discovery Rate

(FDR, expected proportion of erroneous null hypothesis rejections among all rejections⁴⁸).

Results and discussion

Duplications in response to high temperature reduce protein disorder

In response to high temperature yeast (*S. cerevisiae*) duplicates the entire chromosome III (for brevity we use *chrN* to denote ‘yeast chromosome N’ with N as Roman numerals following convention) and fragments from chrIV and chrXII²⁰. The size of the 16 yeast chromosomes varies over six-fold (Table S1). The average protein length is similar between the 16 chromosomes (Figure S1, Table S1). The duplicated chrIII is the 3rd smallest with 183 genes, of which 153 are mapped and 132 constitute “verified ORFs”. Fewer genes are encoded only by chrI with 90, and chrVI with 125 proteins (Table S1). The relatively small number of genes on chrIII was one reason for choosing it as the first fully synthesized functional yeast chromosome⁴⁹. In contrast to protein length, the percentage of proteins with predicted long regions of disorder differed significantly between the 16 yeast chromosomes (Figure 1).

The least protein disorder was predicted for chrIII and chrX (Figure 1, Table S2). That means heat response duplicates one of the two chromosomes with the least disorder. In addition, the fragments of chrIV and chrXII that are duplicated along with the entire chrIII also clearly have less disorder than the chromosomes from which they were taken (Figure 1). This enhances the effect of protein disorder reduction in response to high temperatures.

The other low-disorder option is chrX: Why not duplicate chrX in response to high temperatures? ChrX is more than twice as large as chrIII (Figure S1). Thus duplicating chrX would “cost” twice as much. This might be prohibitive. ChrX might also not contain the cell activities important for coping with high temperature. Furthermore, as chrX and chrIII are similar in disorder content while chrX has twice the proteins of chrIII, the duplication of chrX would increase the overall level of proteins with disorder that might become unfolded and thereby “jam” cellular activity more than the duplication of chrIII.

Assume a certain amount of tolerable duplication were tolerable and that number were about 153 proteins (as for chrIII): where in the genome do we find a continuous stretch (within a chromosome) that has 153 proteins with the least disorder? Our results underscored the special role of chrIII (Figure S2): only 3% of all continuous genome fragments with 153 proteins have as little disorder as chrIII (corresponding numbers for chrX: 5%; 29-protein fragment from chrIV: 52%; 64-protein fragment from chrXII: 10%). These figures demonstrate that the duplication of chrIII might be THE optimal choice for a simple way to duplicate 153 proteins with as little disorder as possible.

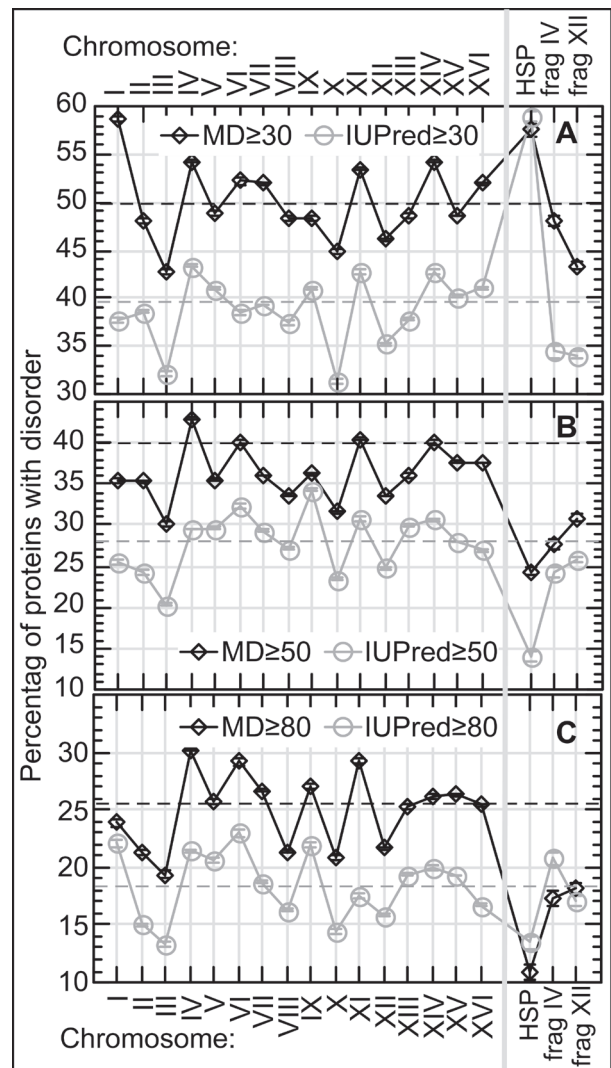


Figure 1. Protein disorder differs between yeast chromosomes.

The composition of proteins with long regions of disorder (y-axes) differed significantly between the 16 chromosomes of *S. cerevisiae* (x-axes) and also for the set of HSPs. The three rightmost marks on the x-axes describe: HSPs and the disorder predictions for the HSR-related duplicated fragments on chromosome IV and chromosome XII (frag IV and frag XII). The differences were similar for two different prediction methods (MD in black, IUPred in light gray), and for different thresholds with respect to the minimal length of a disordered region (A: ≥ 30 consecutive residues predicted in disorder, B: ≥ 50 , C: ≥ 80). Dashed horizontal lines mark the averages over all chromosomes. Error bars are too small to become visible on the scale chosen. The least disorder content was predicted for chromosome III and chromosome X. Overall, all duplications in response to heat shock treatment reduced the level of protein disorder in the offspring.

Heat-shock proteins do not explain the temperature-related duplication

Our results explained why duplicating 150–200 proteins from another chromosome might have been potentially more damaging than the duplication of chrIII in response to high temperatures. In other words, our model might suggest why the duplication of this particular region is better than other duplications. However, what is the selective benefit from the proteins on chrIII? We expected to find the answer to this question in proteins that actively help with coping with heat stress. The immediate suspects are heat-shock proteins (HSP) and the proteins known to interact with these HSPs (HSP-binders). The known HSPs and HSP-binders scatter over all 16 yeast chromosomes (Figure S3). All regions duplicated in response to heat shock contain only one known gene coding for HSPs (*HSP30*) and one known HSP-binder (*TAH1*). This implied that 1.3% of all known HSPs and HSP-binders were duplicated in an event that duplicated 0.5% of all genes, i.e. a 2.6-fold over-representation. This statistically insignificant finding that fewer than 1 in 50 of all HSPs and HSP-binders are duplicated might still be scientifically significant if *HSP30* and *TAH1/HSP90* were outstandingly important proteins for the given conditions. However, this is not the case, at least not given what is currently known about *HSP30*. Furthermore, introducing an extra copy of *HSP30* into wild-type cells did not increase the ability of the cells to cope with high temperature (Dahan & Pilpel, personal communication).

The set of known HSPs (Figure S3) slightly changed expression levels in response to heat stress during the fixation of the trisomy²⁰ only slightly but almost all HSPs were significantly up-regulated (arrows in Figure S3) when the “refined descendants” replaced the trisomy²⁰. This could imply that the duplicated genes are essential for survival under heat stress. Nevertheless, quite contrary to the naïve expectation, the HSPs and the HSP-binders by no means explained the heat-stress-specific duplications observed experimentally.

Incidentally, HSPs appeared particularly abundant in disorder regions of 30–50 consecutive residues (Figure 1A, in particular for IUPred). It has previously been argued that such disorder is required for proper function of HSPs⁵⁰. In contrast, HSPs are depleted of longer disorder (>50; Figure 1B,C).

Overall, we argue that HSPs could have explained the duplication of many other chromosomes, possibly even better than that of chrIII. Therefore, this explanation is not specific. Thus, we conclude that the duplication of known HSPs and HSP-binding proteins did not explain why chrIII was specifically duplicated. Many HSPs and HSP-binders might remain unknown. However, we have no scientific ground to suspect that the fraction of the unknown HSPs differs between the chromosomes, i.e. that there are particular HSPs on chrIII that remain undiscovered.

GO terms enriched for growth and reproduction in heat stress-duplication

Are any other proteins on chrIII important for growth under high temperature? Simply scanning GO⁴⁵ annotations is insufficient: the question is not whether proteins on chrIII have certain functions, but whether these are overrepresented enough to explain why

chrIII and not the other two small chromosomes (chrVI or chrI) are duplicated in response to high temperatures. In order to address this question, we need a GO term enrichment analysis of the duplicated regions⁵¹.

Growth and reproduction might be considered as the most important cell activities in the sense that the organism must grow and proliferate (cells that fail are not observed) even under stress. The GO enrichment analysis seemed to confirm this expectation (Figure 2): the two most abundant GO terms in the heat stress-duplicated regions were those related to (i) sexual reproduction (Figure 2 and Figure S2; “conjugation with cellular fusion”, “reproductive cellular process” and “response to pheromone”) and to (ii) sugar transport (hexose transport process as well as mannose, fructose and glucose transmembrane transporter activity; Figure S4).

The major energy source of yeast is sugar, in particular hexose monosaccharides (C₆H₁₂O₆; e.g. glucose, fructose, mannose). These nutrient sugars are imported into the cell through hexose transporters, which are encoded by *HXT* genes^{52,53}. The *HXT* yeast genes on the duplicated fragment of chromosome IV (*HXT3*, *HXT6* and *HXT7*) are almost five-fold over-represented with respect to random (yeast has 5667 N_{genY} genes, 243 N_{genD} are duplicated, 15 N_{genHXT} genes are in yeast; in a region with 243 N_{genD} genes we would find by chance 0.64 *HXT* genes in the duplicated regions $p_{\text{chance}} = \frac{N_{\text{genHXT}}}{N_{\text{genD}}} \cdot \frac{N_{\text{genD}}}{N_{\text{genY}}}$). Two *HXT* genes on the duplicated chrIV fragment (*HXT6* and *HXT7*) appear to encode high-affinity transporters required for growth at very low glucose concentrations (~0.1%⁵⁴), i.e. these two would become particularly important when yeast is cultured under glucose limitation⁵⁴. Interestingly, several works have detected duplication of these two genes (*HXT6* and *HXT7*) in yeast populations evolving under low nutrient availability^{8,55}. These numbers suggest that heat stress also puts strain upon obtaining the energy needed for growth and reproduction.

Sexual reproduction also appeared crucial for the survival of yeast cultured under heat stress^{56,57}. Seven of the ten molecular functions to be significantly overrepresented in the heat stress-duplicated chrIII (Table S3) by a standard GO-term enrichment analysis⁵¹ are involved in reproduction. Three of these seven molecular functions are related specifically to sexual reproduction; the others pertain to general reproductive processes (Figure 2). In particular, the reproduction-related processes involve cell fusion (*FUS1* and *FIG2*^{58–60}), pheromone response (*STE50* which is also required for optimal invasive growth and hyperosmotic stress signaling^{61,62} and *NOT1* that is also involved in several RNA regulation levels⁶³), nuclear fusion, chromosome disjunction, nuclear segregation after mating (*BIK1* which is involved in microtubule function during mitosis^{64,65}), fusion of haploid nuclei during mating; *KAR4* or *KARYOGAMY* plays a critical role in the choreography of the mating response⁶⁶), cytokinesis (division of cytoplasm and plasma membrane of a cell and its separation into two daughter cells which is also relevant for asexual mitotic growth: *CDC10*⁶⁷), specification of the site where the daughter cell will form (relevant for budding and asexual growth, also referred to as axial bud selection) and in the developmental process in which the size of a cell is generated and organized (also referred to as morphogenesis: *CDC10*^{67–69}). All these genes are also required for the correct

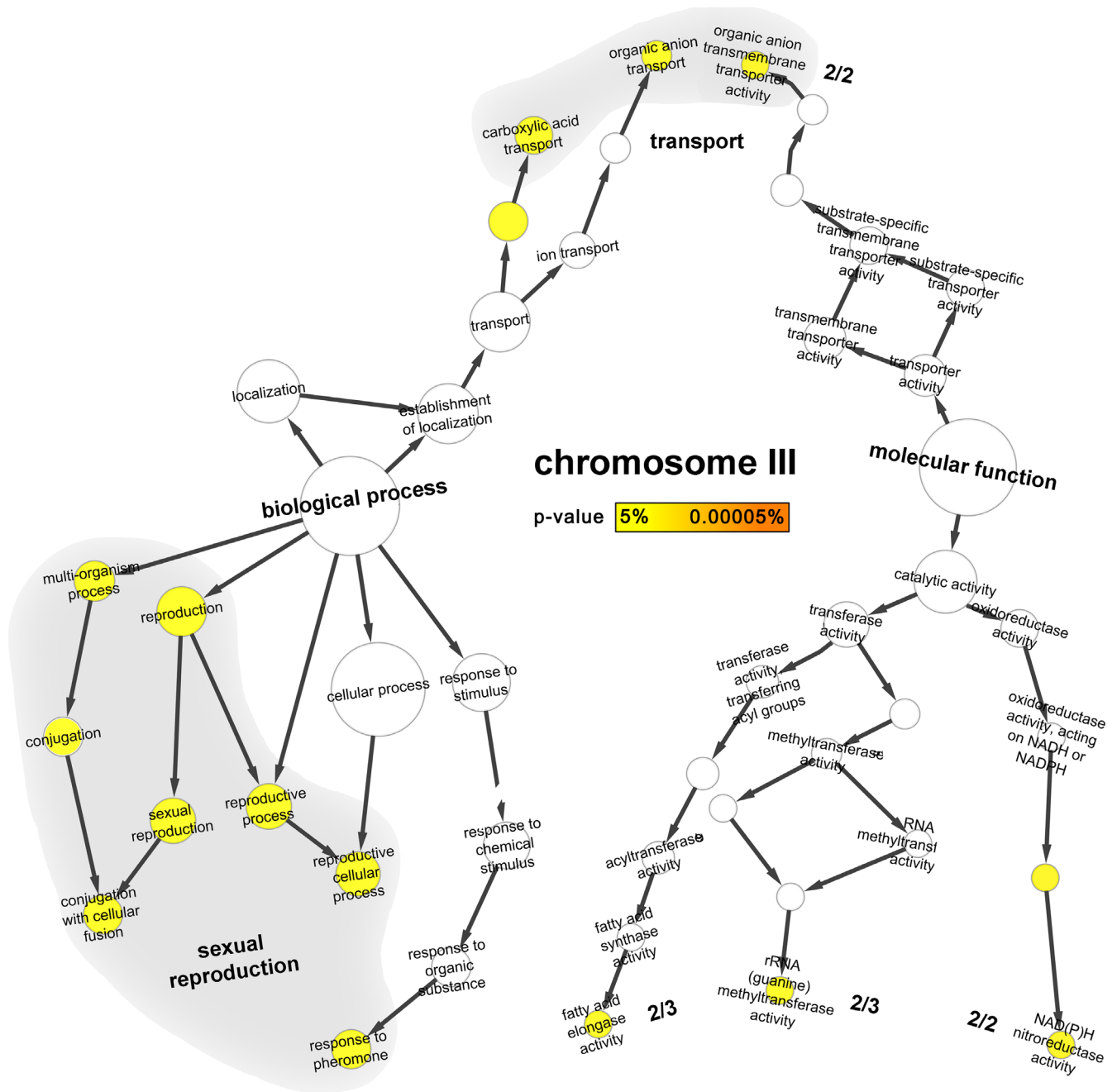


Figure 2. GO enrichment of sexual reproduction and nutrient uptake. The tree gives the complete set of all experimentally annotated GO-terms (Gene Ontology⁴⁵) for any of the proteins on chromosome III that describe *biological process* (left branch) and *molecular function* (right branch). The enrichment analysis⁵¹ describes how much chrIII GO-terms are enriched with respect to all other GO-terms from yeast: all terms marked by yellow circles are significantly enriched. Sexual reproduction (7 GO-terms on chrIII) and transport (carboxylic acid and organic anion 4 GO-terms on chrIII) mapped to most overrepresented GO terms on this chromosome.

localization of other proteins involved in cytokinesis and bud site selection^{67,70-73}. Other important processes and activities overrepresented on chrIII are related to the avoidance of oxidative stress (e.g. carboxylic acid transport – **Figure 2** - which may be important for the survival since during the vegetative asexual reproduction cells were exposed to oxidative stress) and NAD(P)H nitro-reductase activity (**Figure 2**). The only nitroreductase-related proteins in yeast – HBN1 and FRM2⁷⁴ - are only on chrIII (**Table S4**). The proteins involved in these two activities (carboxylic acid transport and NAD(P)H nitro-reductase activities) are also implicated in cellular detoxification⁷⁵, which is another task relevant for survival under stress.

All these data supported the view that chrIII is important for sexual reproduction. A seemingly convincing story, until we learned that the laboratory strains of yeast survived through asexual reproduction²⁰, *i.e.* apparently did not need what is so uniquely enriched in the heat stress duplication. The set of proteins known to be involved in reproduction on chrIII (**Table S3**) had more disorder than the average for chrIII (**Figure S5**). Some of these proteins with long disordered regions might not work correctly in heat.

Why duplicate proteins that fail? Not having found a convincing answer, we propose two conjectures: first, sexual reproduction might “frame” another cellular activity of the same protein that is more relevant to the growth conditions applied during the evolution in the laboratory experiment. For instance, *CDC10* is also required to maintain cell polarity (GO: 0030011), *BUD3* and *BUD5* are involved in axial cellular bud-site selection (GO: 0007120), *KCC4* a bud neck kinase involved in budding and cell bud growth (GO: 0007117) and *BIK1*, which is involved in microtubule function during mitosis. All of these activities are related to asexual reproduction. Our second proposition seems more far-fetched, namely that the set of proteins with the strongest GO-enrichment might have been duplicated coincidentally, *i.e.* the disorder-rich proteins related to sexual reproduction might have been duplicated because they happened to be on chrIII but not due their relevance for the survival in heat. If so, there must be something else we have not found yet on chrIII.

Several other processes were slightly enriched in the duplicated fragments with some relevance for yeast survival in heat but none of those gave a clear explanation (**Figure 2**): (i) fatty acid elongase, (ii) rRNA (guanine) methyltransferase, and (iii) the importin-alpha export receptor activities. We analyzed these in detail. (i) Fatty acid elongase: currently, only three proteins are known to be involved in lengthening fatty acids; two of those (*ELO2* and *APA1*; **Table S3**) are on chrIII. Fatty acid elongases are involved in sphingolipid biosynthesis. The sphingolipids are components of the cellular membrane and bioactive signaling molecules that contribute to heat tolerance as they are directly involved in organizational cellular structures (e.g. cell membrane)⁷⁶. (ii) rRNA methyltransferases: three yeast proteins are known to be involved in rRNA (guanine) methyltransferase activity; two of those (*BUD23* and *SPB1*) are on chrIII (**Table S4**). It is believed that the modification of

ribonucleotides optimizes the rRNA structure and represents a way to expand the topological potentials of RNA molecules. It is possible that the loss of modification affects fine-tuning of ribosome function that could give rise to the pronounced cold-sensitivity⁷⁷. (iii) Importin-alpha nuclear export: two yeast proteins contribute toward the importin-alpha export receptor activity; one of those (*MSN5*) is in the duplicated fragment of chromosome IV. *MSN5* knockout mutants show a variety of phenotypes, including carbon-source utilization, defects and sensitivity to high concentrations of ions, severe heat shock, and high pH⁷⁸. Moreover, these mutants are partially sterile⁷⁸. Therefore, this protein appears necessary for cell survival, especially under extreme conditions.

Only one cellular activity related to tRNA synthase appeared over-represented on the duplicated fragment of chrXII (*DUS3* and *DUS4* proteins; **Table S7**). In particular, to the tRNA dihydrouridine synthases, which are responsible for the reduction of the 5,6-double bond of a uridine residue on tRNA (one of the numerous modifications observed on tRNA cytoplasmatic⁷⁹). However, this particular finding appeared less relevant since the corresponding fragment was only duplicated in one of four growth experiments in response to high temperatures²⁰.

One crucial limitation for any functional enrichment study remains the incomplete experimental annotation even for an organism as intensively studied as yeast. It may be that all our speculation above missed the real causation because the functions of the proteins that are really relevant remain uncharacterized. Therefore, we complemented our analysis with one aspect of function for which we have a complete prediction, namely the prediction of sub-cellular localization of all yeast proteins. The experimental localization annotations for yeast are still cover at most 70% of all proteins⁸⁰. However, today's top prediction methods, such as LocTree3, are very reliable⁸⁰ and can make crucial differences for comparing ‘complete’ data sets⁸¹. We found nuclear proteins to be clearly depleted on chrIII (-4.6 percentage points with respect to the entire proteome; **Figure S7A**). Other abundant proteins found on chrIII were secreted (extra-cellular) or annotated as endoplasmic reticulum (ER) membrane proteins (each 3.2 percentage points higher than in the full yeast proteome). We also observed significantly more disorder in nuclear proteins (nuclear 77% vs. <40% for non-nuclear; **Figure S8**). This might explain the depletion of nuclear proteins on chrIII. While these findings were clear, they did not suggest a simple interpretation. The abundance of secreted proteins on chrIII (about 3.2 percentage points more on chrIII than in entire yeast; **Figure S7A**) implies that in the response to heat shock, more proteins are secreted into the ‘hot’ environment. Given the correlation between habitat and disorder³⁷, we expect that proteins are more likely to sustain high temperatures with less disorder. Unfortunately, a GO enrichment study of the secreted proteins also did not provide the answer we had been hoping for. However, the “secretome” alone could not explain the lower content of disordered proteins on chromosome III (disorder entire yeast-chrIII=50%-43%=7%>3% for secretome; **Table 1** and **Figure S7A**).

Proteins from chrIII less implied in overall PPI network

As proteins cannot be understood without also considering their networks of interaction, we compared the network of experimentally characterized PPIs between the entire yeast and those fragments that are duplicated in heat evolving populations. As for the differential analysis of any experimental annotation, the limitation of such an approach lies in the incompleteness of the experimental data. In all 16 chromosomes, the degree (number of interactions per protein) was lowest for chrIII (average=16±2; **Figure 3A**). A similar trend was observed for betweenness (number of times that a protein acts as a bridge along the shortest path between two other proteins: average=1800±300; **Figure 3B**). Furthermore, chrIII is one of the chromosomes with the largest mean value for the average neighbor degree (average=380±40; **Figure 3C**). Our network analyses confirm chrIII as a good choice for a first line of defense against high temperature because the proteins encoded on this chromosome play less essential roles for the overall PPI network. However, once again, this portrays the duplication as a solution with least possible damage without positively suggesting causation.

Conclusions

Organisms can duplicate the whole genome or particular chromosomes (aneuploidy) in response to sudden dramatic changes in the environment. As such coarse-grained major changes are costly, aneuploidy tends to give way to more fine-tune focused solutions that require many generations to evolve. The entire chromosome III and two fragments from chromosomes IV and XII in a culture of budding yeast (*S. cerevisiae*) were duplicated as a “transient evolutionary solution” in response to high temperature - a “transition” that fostered the survival of between 400 and 2,000 generations. Here, we reported that while the proteins on all 16 main chromosomes from yeast have similar length, they differ substantially in the fraction of proteins with long regions predicted to contain protein disorder (≥ 30 –80 consecutive residues predicted as disordered by IUPred and MD). We found the regions duplicated under heat stress depleted of predicted disorder. In fact, chromosome III was one of the two chromosomes with the least disorder (**Figure 1**). The other (chromosome X) is twice as large, *i.e.* would cost twice to duplicate. Decreasing the overall content in protein disorder is likely an important strategy to protect against heat stress. A detailed analysis of the experimentally characterized PPI network in yeast revealed the duplicated proteins to be connected less than average (**Figure 3**). The PPI analysis, therefore, added to the explanation that the duplication causes minimal damage. However, why did the duplication create an advantage under heat stress? Surprisingly, we found no sustained evidence for a significant over-representation of HSPs in the duplication *i.e.* of proteins that usually help out under such stress. Instead, a Gene Ontology (GO) enrichment analysis suggested that the duplicated regions were enriched in processes related to reproduction and to the import of nutrients (**Figure 2**). The enrichment was strongest for proteins related to sexual reproduction although the heat stress survival was maintained through budding, *i.e.* through asexual reproduction. Nevertheless, the set of GO enriched proteins appeared so important that they were duplicated although high in disorder. This might point to where the explanation for the duplication might be found. Overall, our data suggested a very simple

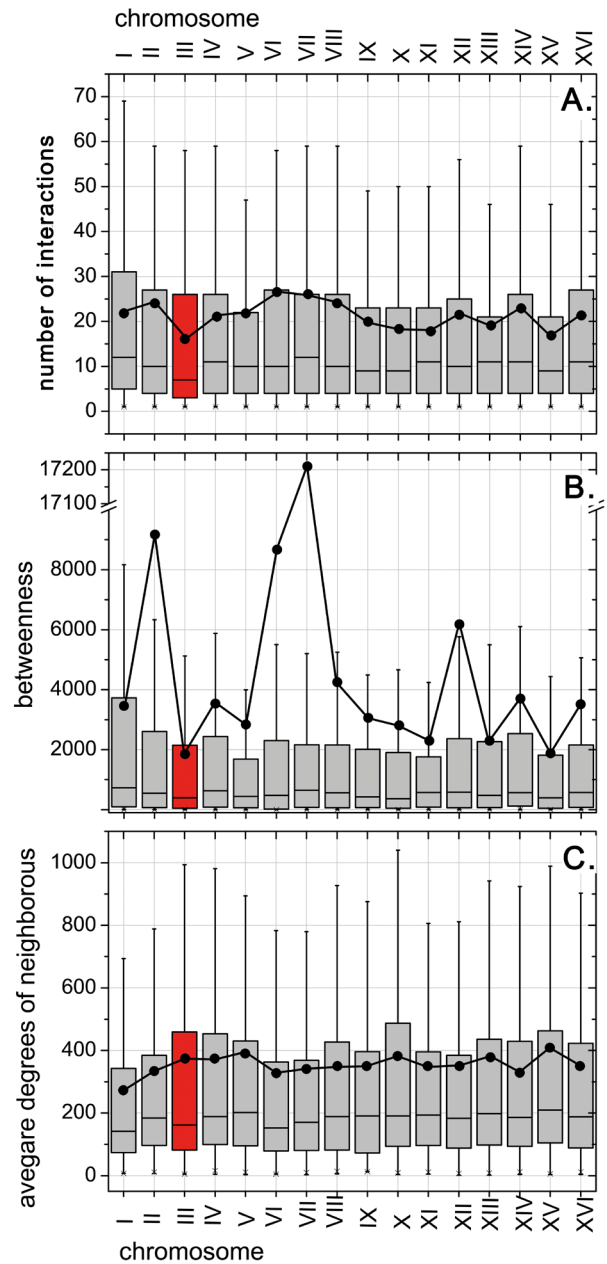


Figure 3. PPI network differs between yeast chromosomes.

We began with the entire network of all PPIs with experimental annotations in yeast (Methods), and then differentially analyzed major network features: **(A) Degree:** The number of PPIs per protein (degree) was minimal for the proteins from chrIII (box in red; lowest mean - black dot and lowest median - black line in the box). **(B) Betweenness:** betweenness (number of times that a node acts as a bridge along the shortest path between two other nodes) was also lowest for chrIII. **(C) Average neighbor degree:** plotting the average degree for all network neighbors of all proteins on chrIII (*i.e.* all those proteins in direct PPI with proteins on chrIII), we observed a much less differentiated view. For this network feature, the proteins from chrIII had one of the highest means (black dot), but one of the lowest medians. Clearly, the proteins from the HSR-duplicated chromosome appeared less involved in the yeast network than expected by chance.

algorithm: identify the region with lowest protein disorder that is large enough, yet not too large and duplicate it along with possibly other fragments that are also depleted of disorder in order to cope with heat stress.

Author contributions

EV and BR conceived the study and designed the experiments. EV, ZG, YD, TG and MJ carried out the research. YD and TG provided expertise in protein-protein interactions and protein localization prediction respectively. EV prepared the first draft of the manuscript. EV, ZG and TG contributed to the graphics and preparation of the manuscript. All authors were involved in the revision of the draft manuscript and have agreed to the final content.

Competing interests

No competing interests were disclosed.

Grant information

This work, and all authors were supported by the German Research Foundation (DFG) and the Technische Universität München within the funding program Open Access Publishing.

I confirm that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

Thanks to Tim Karl and Laszlo Kajan (TUM) for invaluable help with hardware and software; to Marlena Drabik and Inga Weise (TUM) for administrative support; to Tobias Hamp and Edda Kloppmann for helpful comments on the manuscript. A special thanks goes to the Pilpel-lab (Weizman Inst. Rehovot, Israel), in particular to Yitzhak Pilpel (Weizman) and Orna Dahan (Weizman) for the data and crucial help. Last, not least, thanks to all those who deposit their experimental data in public databases, and to those who maintain these databases.

Supplementary material

Supplementary material for Vicedo *et al.*, 2015 'Protein disorder reduced in *Saccharomyces cerevisiae* to survive heat shock'.

Table of Contents for Supplementary Material

Figure S1: Number of genes per chromosome in *Saccharomyces cerevisiae* (yeast).

Figure S2: Fragments with less disorder than proteins duplicated during heat shock response (HSR).

Figure S3: Distribution of Heat Shock proteins (HSP) along chromosomes.

Figure S4: Complete set of known GO (Gene Ontology) terms for a fragment of the HSR-duplicated chromosome IV.

Figure S5: Disordered proteins differentiate by chromosomes and roles.

Figure S6: Disorder difference between chromosome III proteins and their paralogs.

Figure S7: Distribution comparison of chromosome III proteins and the complete yeast proteome across localization classes.

Figure S8: Distribution of disordered/ordered nuclear proteins and other yeast proteins.

Table S1: General information about yeast chromosomes.

Table S2: Disorder abundance content.

Table S3: Overrepresented GO terms for molecular function analysis of chromosome III.

Table S4: Overrepresented GO terms for biological process analysis of chromosome III.

Table S5: Overrepresented GO terms for molecular function of the duplicated fragment of chromosome IV.

Table S6: Overrepresented GO terms for biological process of the duplicated fragment of chromosome IV.

Table S7: Overrepresented GO terms for biological process of the duplicated fragment of chromosome XII.

Table S8: Heat shock proteins distribution on chromosomes.

[Click here to access the data.](#)

References

1. Goffeau A, Barrell BG, Bussey H, *et al.*: **Life with 6000 genes.** *Science*. 1996; **274**(5287): 546, 563–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Alberghina L, Mavelli G, Drovandi G, *et al.*: **Cell growth and cell cycle in *Saccharomyces cerevisiae*: basic regulatory design and protein-protein interaction network.** *Biotechnol Adv*. 2012; **30**(1): 52–72.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Kitano H: **Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology.** *Curr Genet*. 2002; **41**(1): 1–10.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Westerhoff HV, Palsson BO: **The evolution of molecular biology into systems biology.** *Nat Biotechnol*. 2004; **22**(10): 1249–52.
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Torres EM, Sokolsky T, Tucker CM, *et al.*: **Effects of aneuploidy on cellular physiology and cell division in haploid yeast.** *Science*. 2007; **317**(5840): 916–24.
[PubMed Abstract](#) | [Publisher Full Text](#)
6. Pavelka N, Rancati G, Zhu J, *et al.*: **Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast.** *Nature*. 2010; **468**(7321): 321–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Gordon DJ, Resio B, Pellman D: **Causes and consequences of aneuploidy in cancer.** *Nat Rev Genet*. 2012; **13**(3): 189–203.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Gresham D, Desai MM, Tucker CM, *et al.*: **The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast.** *PLoS Genet*. 2008; **4**(12): e1000303.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Hughes TR, Roberts CJ, Dai H, *et al.*: **Widespread aneuploidy revealed by DNA microarray expression profiling.** *Nat Genet*. 2000; **25**(3): 333–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Rancati G, Pavelka N, Fleharty B, *et al.*: **Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor.** *Cell*. 2008; **135**(5): 879–93.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Selmecki A, Gerami-Nejad M, Paulson C, *et al.*: **An isochromosome confers drug resistance *in vivo* by amplification of two genes, *ERG11* and *TAC1*.** *Mol Microbiol*. 2008; **68**(3): 624–41.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Dunham MJ, Badrane H, Ferea T, *et al.*: **Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*.** *Proc Natl Acad Sci U S A*. 2002; **99**(25): 16144–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Polakova S, Blume C, Zarate JA, *et al.*: **Formation of new chromosomes as a virulence mechanism in yeast *Candida glabrata*.** *Proc Natl Acad Sci U S A*. 2009; **106**(8): 2688–93.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Fawcett JA, Maere S, Van de Peer Y: **Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event.** *Proc Natl Acad Sci U S A*. 2009; **106**(14): 5737–42.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Torres EM, Williams BR, Amon A: **Aneuploidy: cells losing their balance.** *Genetics*. 2008; **179**(2): 737–46.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Sheltzer JM, Blank HM, Pfau SJ, *et al.*: **Aneuploidy drives genomic instability in yeast.** *Science*. 2011; **333**(6045): 1026–30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Torres EM, Dephoure N, Panneerselvam A, *et al.*: **Identification of aneuploidy-tolerating mutations.** *Cell*. 2010; **143**(1): 71–83.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Sheltzer JM, Amon A: **The aneuploidy paradox: costs and benefits of an incorrect karyotype.** *Trends Genet*. 2011; **27**(11): 446–53.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Pavelka N, Rancati G, Li R: **Dr Jekyll and Mr Hyde: role of aneuploidy in cellular adaptation and cancer.** *Curr Opin Cell Biol*. 2010; **22**(6): 809–15.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Yona AH, Manor YS, Herbst RH, *et al.*: **Chromosomal duplication is a transient evolutionary solution to stress.** *Proc Natl Acad Sci U S A*. 2012; **109**(51): 21010–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Anfinsen CB, Scheraga HA: **Experimental and theoretical aspects of protein folding.** *Adv Protein Chem*. 1975; **29**: 205–300.
[PubMed Abstract](#) | [Publisher Full Text](#)
22. Karplus M, Weaver DL: **Protein-folding dynamics.** *Nature*. 1976; **260**(5550): 404–6.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. Levitt M, Chothia C: **Structural patterns in globular proteins.** *Nature*. 1976; **261**(5561): 552–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
24. Levitt M, Warshel A: **Computer simulation of protein folding.** *Nature*. 1975; **253**(5494): 694–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Schlessinger A, Schaefer C, Vicedo E, *et al.*: **Protein disorder—a breakthrough invention of evolution?** *Curr Opin Struct Biol*. 2011; **21**(3): 412–8.
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Wright PE, Dyson HJ: **Linking folding and binding.** *Curr Opin Struct Biol*. 2009; **19**(1): 31–8.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Dunker AK, Cortese MS, Romero P, *et al.*: **Flexible nets. The roles of intrinsic disorder in protein interaction networks.** *FEBS J*. 2005; **272**(20): 5129–48.
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Dosztányi Z, Csizmok V, Tompa P, *et al.*: **IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.** *Bioinformatics*. 2005; **21**(16): 3433–4.
[PubMed Abstract](#) | [Publisher Full Text](#)
29. Schlessinger A, Liu J, Rost B: **Natively unstructured loops differ from other loops.** *PLoS Comput Biol*. 2007; **3**(7): e140.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Dunker AK, Silman I, Uversky VN, *et al.*: **Function and structure of inherently disordered proteins.** *Curr Opin Struct Biol*. 2008; **18**(6): 756–64.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Singh GP, Dash D: **Intrinsic disorder in yeast transcriptional regulatory network.** *Proteins*. 2007; **68**(3): 602–5.
[PubMed Abstract](#) | [Publisher Full Text](#)
32. Fuxreiter M, Tompa P, Simon I, *et al.*: **Malleable machines take shape in eukaryotic transcriptional regulation.** *Nat Chem Biol*. 2008; **4**(12): 728–37.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Liu J, Tan H, Rost B: **Loopy proteins appear conserved in evolution.** *J Mol Biol*. 2002; **322**(1): 53–64.
[PubMed Abstract](#) | [Publisher Full Text](#)
34. Devos D, Dokudovskaya S, Williams R, *et al.*: **Simple fold composition and modular architecture of the nuclear pore complex.** *Proc Natl Acad Sci U S A*. 2006; **103**(7): 2172–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Radivojac P, Iakoucheva LM, Oldfield CJ, *et al.*: **Intrinsic disorder and functional proteomics.** *Biophys J*. 2007; **92**(5): 1439–56.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Dunker AK, Gough J: **Sequences and topology: intrinsic disorder in the evolving universe of protein structure.** *Curr Opin Struct Biol*. 2011; **21**(3): 379–81.
[PubMed Abstract](#) | [Publisher Full Text](#)
37. Vicedo E, Schlessinger A, Rost B: **Environmental Pressure May Change the Composition Protein Disorder in Prokaryotes.** *PLoS One*. 2015; **10**(8): e0133990.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
38. UniProt Consortium: **Reorganizing the protein space at the Universal Protein Resource (UniProt).** *Nucleic Acids Res*. 2012; **40**(Database issue): D71–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
39. Mika S, Rost B: **UniqueProt: Creating representative protein sequence sets.** *Nucleic Acids Res*. 2003; **31**(13): 3789–91.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Cherry JM, Hong EL, Amundsen C, *et al.*: **Saccharomyces Genome Database: the genomics resource of budding yeast.** *Nucleic Acids Res*. 2012; **40**(Database issue): D700–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
41. Schlessinger A, Punta M, Rost B: **Natively unstructured regions in proteins identified from contact predictions.** *Bioinformatics*. 2007; **23**(18): 2376–84.
[PubMed Abstract](#) | [Publisher Full Text](#)
42. Schlessinger A, Punta M, Yachdav G, *et al.*: **Improved disorder prediction by combination of orthogonal approaches.** *PLoS One*. 2009; **4**(2): e4433.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Dosztányi Z, Csizmok V, Tompa P, *et al.*: **The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins.** *J Mol Biol*. 2005; **347**(4): 827–39.
[PubMed Abstract](#) | [Publisher Full Text](#)
44. Maere S, Heymans K, Kuiper M: **BINGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics*. 2005; **21**(16): 3448–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
45. Ashburner M, Ball CA, Blake JA, *et al.*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet*. 2000; **25**(1): 25–9.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
46. Shannon P, Markiel A, Ozier O, *et al.*: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res*. 2003; **13**(11): 2498–504.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
47. Farcomeni A: **A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion.** *Stat Methods Med Res*. 2008; **17**(4): 347–88.
[PubMed Abstract](#) | [Publisher Full Text](#)
48. Benjamini Y, Yekutieli D: **The control of the false discovery rate in multiple testing under dependency.** *Ann Statist*. 2001; **29**(4): 1165–88.
[Publisher Full Text](#)

49. Annaluru N, Muller H, Mitchell LA, *et al.*: **Total synthesis of a functional designer eukaryotic chromosome.** *Science*. 2014; **344**(6179): 55–8.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Tompa P, Csermely P: **The role of structural disorder in the function of RNA and protein chaperones.** *FASEB J*. 2004; **18**(11): 1169–75.
[PubMed Abstract](#) | [Publisher Full Text](#)
51. Tarca AL, Bhatti G, Romero R: **A comparison of gene set analysis methods in terms of sensitivity, prioritization and specificity.** *PLoS One*. 2013; **8**(11): e79217.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
52. Boles E, Hollenberg CP: **The molecular genetics of hexose transport in yeasts.** *FEMS Microbiol Rev*. 1997; **21**(1): 85–111.
[PubMed Abstract](#) | [Publisher Full Text](#)
53. Ozcan S, Dover J, Johnston M: **Glucose sensing and signaling by two glucose receptors in the yeast *Saccharomyces cerevisiae*.** *EMBO J*. 1998; **17**(9): 2566–73.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
54. Reifenberger E, Freidel K, Ciriacy M: **Identification of novel HXT genes in *Saccharomyces cerevisiae* reveals the impact of individual hexose transporters on glycolytic flux.** *Mol Microbiol*. 1995; **16**(1): 157–67.
[PubMed Abstract](#) | [Publisher Full Text](#)
55. Brown CJ, Todd KM, Rosenzweig RF: **Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment.** *Mol Biol Evol*. 1998; **15**(8): 931–42.
[PubMed Abstract](#) | [Publisher Full Text](#)
56. Tannenbaum E: **A comparison of sexual and asexual replication strategies in a simplified model based on the yeast life cycle.** *Theory Biosci*. 2008; **127**(4): 323–33.
[PubMed Abstract](#) | [Publisher Full Text](#)
57. Zeyl C, Curtin C, Karnap K, *et al.*: **Antagonism between sexual and natural selection in experimental populations of *Saccharomyces cerevisiae*.** *Evolution*. 2005; **59**(10): 2109–15.
[PubMed Abstract](#) | [Publisher Full Text](#)
58. Nelson B, Parsons AB, Evangelista M, *et al.*: **Fus1p interacts with components of the Hog1p mitogen-activated protein kinase and Cdc42p morphogenesis signaling pathways to control cell fusion during yeast mating.** *Genetics*. 2004; **166**(1): 67–77.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
59. Zhang M, Bennett D, Erdman SE: **Maintenance of mating cell integrity requires the adhesin Fig2p.** *Eukaryot Cell*. 2002; **1**(5): 811–22.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
60. Erdman S, Lin L, Malczynski M, *et al.*: **Pheromone-regulated genes required for yeast mating differentiation.** *J Cell Biol*. 1998; **140**(3): 461–83.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
61. Ramezani-Rad M: **The role of adaptor protein Ste50-dependent regulation of the MAPKKK Ste11 in multiple signalling pathways of yeast.** *Curr Genet*. 2003; **43**(3): 161–70.
[PubMed Abstract](#) | [Publisher Full Text](#)
62. Truckses DM, Bloomekatz JE, Thorne J: **The RA domain of Ste50 adaptor protein is required for delivery of Ste11 to the plasma membrane in the filamentous growth signaling pathway of the yeast *Saccharomyces cerevisiae*.** *Mol Cell Biol*. 2006; **26**(3): 912–28.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
63. Sandler H, Kreth J, Timmers HT, *et al.*: **Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin.** *Nucleic Acids Res*. 2011; **39**(10): 4373–86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
64. Berlin V, Styles CA, Fink GR: **BIK1, a protein required for microtubule function during mating and mitosis in *Saccharomyces cerevisiae*, colocalizes with tubulin.** *J Cell Biol*. 1990; **111**(6 Pt 1): 2573–86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. Miller RK, Rose MD: **Kar9p is a novel cortical protein required for cytoplasmic microtubule orientation in yeast.** *J Cell Biol*. 1998; **140**(2): 377–90.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
66. Lahav R, Gammie A, Tavazoie S, *et al.*: **Role of transcription factor Kar4 in regulating downstream events in the *Saccharomyces cerevisiae* pheromone response pathway.** *Mol Cell Biol*. 2007; **27**(3): 818–29.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
67. Field CM, Kellogg D: **Septins: cytoskeletal polymers or signalling GTPases?** *Trends Cell Biol*. 1999; **9**(10): 387–94.
[PubMed Abstract](#) | [Publisher Full Text](#)
68. Cid VJ, Adamiková L, Cenamor R, *et al.*: **Cell integrity and morphogenesis in a budding yeast septin mutant.** *Microbiology*. 1998; **144**(Pt 12): 3463–74.
[PubMed Abstract](#) | [Publisher Full Text](#)
69. Longtine MS, DeMarini DJ, Valencik ML, *et al.*: **The septins: roles in cytokinesis and other processes.** *Curr Opin Cell Biol*. 1996; **8**(1): 106–19.
[PubMed Abstract](#) | [Publisher Full Text](#)
70. Madden K, Snyder M: **Cell polarity and morphogenesis in budding yeast.** *Annu Rev Microbiol*. 1998; **52**: 687–744.
[PubMed Abstract](#) | [Publisher Full Text](#)
71. Carroll CW, Altman R, Schieltz D, *et al.*: **The septins are required for the mitosis-specific activation of the Gin4 kinase.** *J Cell Biol*. 1998; **143**(3): 709–17.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
72. Barral Y, Parra M, Bidlingmaier S, *et al.*: **Nim1-related kinases coordinate cell cycle progression with the organization of the peripheral cytoskeleton in yeast.** *Genes Dev*. 1999; **13**(2): 176–87.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
73. Park HO, Sanson A, Herskowitz I: **Localization of Bud2p, a GTPase-activating protein necessary for programming cell polarity in yeast to the presumptive bud site.** *Genes Dev*. 1999; **13**(15): 1912–7.
[PubMed Abstract](#) | [Free Full Text](#)
74. de Oliveira IM, Henriques JA, Bonatto D: **In silico identification of a new group of specific bacterial and fungal nitroreductases-like proteins.** *Biochem Biophys Res Commun*. 2007; **355**(4): 919–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
75. Forestier C, Frangne N, Eggmann T, *et al.*: **Differential sensitivity of plant and yeast MRP (ABCC)-mediated organic anion transport processes towards sulfonylureas.** *FEBS Lett*. 2003; **554**(1–2): 23–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
76. Chen PW, Fonseca LL, Hannun YA, *et al.*: **Coordination of rapid sphingolipid responses to heat stress in yeast.** *PLoS Comput Biol*. 2013; **9**(5): e1003078.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. Peifer C, Sharma S, Watzinger P, *et al.*: **Yeast Rrp8p, a novel methyltransferase responsible for m1A 645 base modification of 25S rRNA.** *Nucleic Acids Res*. 2013; **41**(2): 1151–63.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
78. Alepuz PM, Matheos D, Cunningham KW, *et al.*: **The *Saccharomyces cerevisiae* RanGTP-binding protein msn5p is involved in different signal transduction pathways.** *Genetics*. 1999; **153**(3): 1219–31.
[PubMed Abstract](#) | [Free Full Text](#)
79. Boyle EI, Weng S, Gollub J, *et al.*: **GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes.** *Bioinformatics*. 2004; **20**(18): 3710–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
80. Goldberg T, Hecht M, Hamp T, *et al.*: **LocTree3 prediction of localization.** *Nucleic Acids Res*. 2014; **42**(Web Server issue): W350–5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
81. Ramilowski JA, Goldberg T, Harshbarger J, *et al.*: **A draft network of ligand-receptor-mediated multicellular signalling in human.** *Nat Commun*. 2015; **6**: 7866.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 20 November 2015

doi:10.5256/f1000research.7734.r11125



Anuj Kumar

Department of Molecular, Cellular & Developmental Biology, University of Michigan, Ann Arbor, MI, USA

The manuscript by Vicedo and colleagues presents an interesting observation: the authors have examined chromosomal regions (all of chromosome III and fragments of chromosomes IV and XII) that are duplicated in *Saccharomyces cerevisiae* in response to sudden exposure to high temperature and find that these chromosomal sequences are significantly decreased for genes encoding proteins with long disordered regions. The authors further analyzed these duplicated regions and the encompassed genes for any enrichment in annotated GO terms, as well as for encoded protein positioning in interaction networks. The results do not indicate significant GO term enrichment and reveal that the encoded proteins exhibit a decreased number of interactions per protein. The biological advantage to this duplication remains unclear.

Comments/suggestions:

The main conclusion presented here is interesting, but as the authors themselves attest, this observation does not explain a biological advantage behind the duplication.

On p. 4, the authors state that introducing an extra copy of *HSP30* into wild-type yeast does not modify the ability of the cells to cope with high temperature. The inclusion of laboratory data considering the effect of adding an extra copy of genes or chromosomal regions corresponding to some of the duplicated sequences would strengthen the paper significantly. This seems to be the easiest way to address a biological effect from duplication of a given gene.

In regards to the analysis, are the observed GO function annotations enriched with respect to other chromosomes/segments as opposed to being enriched against the genome as a whole? If the advantage to the cell centered on the functions associated with the genes in the duplicated regions, then these regions relative to other regions may be enriched for a function. If I'm thinking of this correctly, that would be slightly different than comparing a region for enrichment against the whole genome. Maybe the authors could compare enrichment in one chromosome versus another or utilize a sliding window corresponding to the size of a duplicated fragment to identify regions that would be most enriched for some potentially interesting functions. That might be a more sensitive means of identifying a functional enrichment for the duplicated regions.

Typos/stylistic suggestions:

- on p.2, first line under Introduction: I think it would be sufficient to state “The baker’s yeast *Saccharomyces cerevisiae*” rather than the text in parentheses.
- on p. 3, fourth paragraph under “Duplications in response to high temperature reduce protein disorder”: the first sentence in this paragraph (“Assume a certain amount ...”) needs to be reworded.
- on p. 4, first paragraph, line 16: delete “the” from “insignificant the finding”

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 17 November 2015

doi:10.5256/f1000research.7734.r11121



Paul Pavlidis

Centre for High-Throughput Biology and Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

Vicedo *et al.* report a computational analysis sparked by the interesting findings of Yona *et al.* (2012) of yeast duplication of chrIII having a selective advantage in the face of heat stress. Yona *et al.* did not fully mechanistically explain the reason chrIII aneuploidy is the one selected for, so Vicedo *et al.* have proposed a hypothesis: that chrIII has a substantially lower number of disordered proteins. They test this computationally, followed by some additional bioinformatics and “by hand” characterization of chrIII genes (and some other regions of interest following from Yona *et al.*).

The difficulty here is assigning cause vs. “permissive”. As Vicedo *et al.* report, the disorder hypothesis has limited predictive value because chrX genes also have a low disorder (on average), so the size of the chromosome is posed as the other important variable. However, Vicedo *et al.* seem to be proposing that “low disorder” is good for heat resistance per se (I grant them this) – and that overexpression of low disorder proteins is even better. I have difficulty with this second step, because the way the experiment of Yona *et al.* was done, it could easily be that there are “heat resistance proteins” on chrIII and that the overall duplication of chrIII is tolerated in the context of the advantage of overexpression those genes. But if this was the end of the story it would be hard to make a determination of whether this is a viable hypothesis.

However, there is an obvious problem: the work of Yona *et al.* identified 17 genes on chrIII that appear to be the main culprits for the heat resistance (at least most of them). I see no mention of these 17 in Vicedo *et al.* nor of the 22 control genes tested by Yona *et al.* If Vicedo *et al.* are right then there should be a difference in the disorder of these two sets of proteins. Otherwise, the observations might still be relevant, but that the orderedness of chrIII proteins might be permissive for overexpression of the actual heat-resistance genes via aneuploidy. In that case it might be the rest of the proteins on chrIII that have the orderedness properties, not the 17. (Note that I was not familiar with the Yona work before this review and I have not checked to see if Yona *et al.* or others have done any followup.)

Given the omission of discussion of the 17, the sections of this paper on network analysis, GO and

localization cannot be interpreted with confidence. While I have some quibbles about them I would rather wait to see the response to the comments above.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 17 November 2015

doi:10.5256/f1000research.7734.r11122



Melchor Sanchez-Martinez

Mind the Byte, Barcelona, Spain

The research article entitled 'Protein disorder reduced in *Saccharomyces cerevisiae* to survive heat shock' by Vicedo, Rost and co-workers shows how *Saccharomyces cerevisiae* reduces protein disorder to survive heat shock. It constitutes an interesting example about the usage of bioinformatic techniques to analyze protein disorder and its implications at a whole proteome level. In the article, there is a comprehensive explanation of study design, methods and analysis. The conclusions are well explained and justified on the basis of the results.

Consequently, the manuscript is recommended for approval. It is a good piece of science that meets the indexation requirements of F1000Research.

However I have some comments that the authors may consider and/or answer.

1. As far as I know except some rare exceptions the protein disorder is reduced as temperature increases, oppositely as happens with ordered proteins or protein regions. With increasing temperature, disordered proteins and regions tend to adopt a transitory structure. Commonly this transitory structure is necessary for proteins to perform its biological function. In other recent works that the authors have published have published (Reference 37 in the References section), they stated that "protein disorder appeared as a possible building block to bring about evolutionary changes such as the adaptation to different habitats" and in that sense seems that more disorder should imply a better response to heat shock.

Thus is surprising for me that in response to heat there is a protein disorder reduction, whereas I expect a disorder increment. Why does it happens? Maybe the answer is so easy as that the disordered proteins do not help to "fight" against heat shock or as the authors said "...Some of these proteins with long disordered regions might not work correctly in heat...", but I am curious about that. Do you have any evidence or supported hypothesis to explain that?

2. Regarding to authors statement "...Some of these proteins with long disordered regions might not work correctly in heat...", a plausible way to study that and obtain a more conclusive answer could be to perform a molecular simulation. Maybe a Replica Exchange Molecular Dynamics or Monte Carlo simulations could give a better understanding of what happens with these protein at high temperatures.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.
