nature portfolio

Peer Review File

From genotype to phenotype with 1,086 near telomere-to-telomere yeast genomes

Corresponding Author: Professor Joseph Schacherer

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
Loegler et al. presents an extensive dataset of 1,086 Saccharomyces cerevisiae genomes, assembled nearly to telomere-to-telomere completeness, along with a large compendium of phenotypic measurements. Using long-read sequencing and assembly methods, they build a species-wide structural variant (SV) atlas, a gene-based pangenome, and a reference graph pangenome that incorporates 6,587 SVs. They then perform genome-wide association studies (GWAS) between this full catalog of genetic variants and 8,391 traits (molecular traits such as transcript, protein and metabolite levels, as well as organismal traits). The key findings include SVs and small InDels in the analysis and increases trait heritability estimates by ~14% (on average) compared to SNPs alone. SVs are more often associated with traits and appear more pleiotropic than SNPs and the genetic architectures of molecular vs. organismal traits show distinct patterns. Overall, the work is positioned as providing an unprecedented resource linking full-genome variation to phenotype, with implications for understanding genotype–phenotype relationships.

This manuscript presents a rich and valuable dataset on yeast genomic variation and its link to phenotype, and it addresses an important question in genetics. The scale of 1,086 near-complete genomes is impressive and could significantly advance the field. However, before publication, the authors should strengthen the methodological descriptions. In particular, clearer explanations of the assembly and SV-calling methods, and of the statistical thresholds used in GWAS, are needed. The Discussion should explicitly situate the work relative to recent pangenome and GWAS studies in yeast and other species to highlight what is novel. Given the broad interest of this topic, the manuscript should also more clearly articulate the general lessons about genotype–phenotype mapping (e.g. the utility of including SVs in association studies).

Major Comments
• Methodological rigor and clarity. The authors have undertaken an impressive scale of genome assembly, but the manuscript should clarify critical details of the assembly and variant-calling workflow. In particular, how many chromosomes were completely assembled and were any highly repetitive regions (rDNA arrays, telomeric repeats) unresolved? If multiple contigs per chromosome remain, the authors should explain how these gaps might affect SV calls. The sample set includes a mix of haploids and diploids (~75% diploid, with ~55% heterozygosity). Were heterozygous diploids phased or collapsed into consensus sequences? This choice could impact variant detection. More generally, the SV-calling pipeline should be described in full: which tools were used (assembly comparison vs. read mapping), what filtering criteria were applied, and how were false positives controlled for? Given the known challenges of detecting large insertions and complex rearrangements, some validation (e.g., by PCR or independent data) would strengthen confidence in the SV atlas. Similarly, the graph pangenome was built with Minigraph using 500 haplotypes – the authors should discuss whether this graph is fully consistent and robust, and how they ensured that all SV alleles were captured.
• Novelty and relation to prior work. The study's scale is unprecedented for yeast. The authors should explicitly compare their 1,086 genome SV catalog to the recent ScRAP study by O'Donnell et al. Nat Genet 55: 1390, 2023 and the study by Weller et al. Genome Res 33: 729, 2023; and explain how their larger size increases resolution and changes conclusions.
• Conceptual interpretation. The authors do note pleiotropy of SVs, but it would be useful to relate this to known examples. It is possible that SVs appear more pleiotropic simply because their larger genomic footprint makes them more likely to disrupt multiple genes. The authors should check whether the higher pleiotropy is robust after accounting for variant size and allele

frequency. Correlated traits could also inflate apparent pleiotropy. It would strengthen the claim to provide an analysis of whether SV-trait associations remain significant after conditioning on nearby SNPs (e.g. in a joint model). Likewise, the statement that “the genetic architecture of molecular and organismal traits differed markedly” needs more context. Are the differences driven by the type of trait (e.g., expression vs. growth) or by the source of the data? Adding a brief comparison to known patterns (e.g., omnigenic models, or prior yeast eQTL studies) would help interpret this point.
• Significance and interest. The work will be of high interest to yeast geneticists and to all scientists studying genotype–phenotype mapping. The large sample and multi-omic trait set make it a valuable resource. However, to appeal to Nature’s broad readership, the authors should highlight general insights – for example, the utility of incorporating SVs in any GWAS (consistent with studies in other organisms) and lessons for complex trait genetics more broadly. Is there evidence here that could inform human or plant genetics? Conversely, some readers may see this as a specialist yeast study unless the broader implications are made clear. If possible, the authors should emphasize one or two “take-home” messages that transcend yeast (e.g., SVs often harbor causative variation, so pangenomic approaches are generally warranted – Jeffares et al. Nat Comm 8: 14061, 2017). Stressing a conceptual advance would raise its impact.

Minor Comments
• Clarity of terminology. Define “near telomere-to-telomere” explicitly (how near? does it exclude rDNA?). Similarly, terms like “gene-based pangenome” should be briefly explained for non-specialists. The manuscript uses “SV” consistently, but ensure that each type of SV (deletion, duplication, inversion, translocation) is defined when first mentioned.
• Pleiotrophy definition. The term “pleiotropy” is used to mean “variant affects many traits.” The authors should state the criterion (e.g., number of trait associations) used to call a variant pleiotropic. Otherwise, readers may be confused as to whether, for example, any variant associated with more than one trait counts as pleiotropic, or if some threshold is applied.
• Figure legends and units. Ensure that all figures have clear legends. For Manhattan plots, confirm that all axes are labeled. If bar graphs have error bars or asterisks for significance, these should be explained. Check that any color-coding (e.g., trait types) has a key.

Statistics and Data Presentation
• Statistical methods. The use of linear mixed models (LMM) to control for population structure in GWAS is appropriate (Loegler et al. mention using FaST-LMM). However, the manuscript should provide more detail. For instance, how was the kinship (relatedness) matrix constructed – using all variants, or only pruned SNPs? It is noted that all variants (SNPs, InDels, SVs) were combined into one genotype matrix for LMM, which is reasonable. It would help to clarify the minor allele frequency (MAF) and linkage disequilibrium (LD) pruning thresholds used. If SVs tend to be rarer or in high LD, this could affect power and false-positive rates. The authors should specify how many variants of each type passed MAF filters, and whether allele frequency differences were accounted for.
• Heritability estimation. The claim that including SVs + InDels raises heritability by 14.3% should be clearly defined. Are these narrow-sense (additive) heritabilities estimated with LDAK/GCTA? The authors mention rank-based inverse normal transformation of traits and LDAK v5.2 for heritability. They should clarify whether heritability is on the original scale of each trait, and how significance of heritability was assessed. It would also be useful to discuss whether the increase in heritability is relative (e.g., heritability was 0.30 with SNPs only and 0.343 with SVs added, a 14.3% relative gain) or an absolute percentage point change. In similar graph-pangenome studies, notably in tomato, inclusion of SVs yielded a 24% relative increase in heritability (Zhou et al. Nature 606: 527, 2022). The yeast result is smaller but of the same order, which is plausible. The authors should ensure consistency with published definitions.
• Multiple testing and error bars. The manuscript should explicitly state how significance thresholds were set in the GWAS. With millions of SNPs plus thousands of SVs tested per trait, a genome-wide correction is needed. Did the authors use a strict Bonferroni or an FDR approach? For example, a 5% Bonferroni threshold would be extremely stringent given the variant counts. If an FDR (e.g., 5%) was used, this should be stated. It would also clarify how “SVs more often associated” was quantified – did an SV simply exceed the same p-value threshold as a SNP? If so, comparisons could be biased by the relative numbers of variants.
The figures appear to use error bars (e.g., in bar graphs of heritability or effect sizes). The legend or methods should define these error bars (standard error vs. confidence interval). All p-values mentioned in the text should specify whether they are one- or two-tailed and whether they are raw or adjusted. Throughout, the authors should maintain consistency (e.g., “$P<0.05$” should mean after correction if that is claimed to be “significant”).
• Data sufficiency and trait variation. Given that many traits may have been measured only once per strain, it would help to discuss the replication or measurement error. For instance, do organismal traits (growth rates, etc.) come from single experiments, or averages of replicates? Lack of replication could affect the accuracy of heritability estimates. If available, quoting residual variances or confidence intervals for heritabilities would demonstrate rigor. At minimum, the authors should acknowledge any assumptions (e.g., that environmental variance is similar across strains).
• Overall, the statistical approach appears sound (LMM, LDAK), but more detail on implementation and thresholding is needed to fully assess rigor.

Referee #2

(Remarks to the Author)
In this work, the authors present a catalog of structural variants (SVs) and their impact on genetic architecture in S. cerevisiae. The authors present near T2T genomes for >1,000 S. cerevisiae strains along with detailed analysis of SV classes, distributions, histories, and potential impacts. The manuscript will be an important resource for yeast biologists and contains insights of broad interest to geneticists.

The manuscript is generally well written and thorough; parts of the manuscript would be easier to digest with a little more detail in the main text, as outlined below. My biggest concerns were with a few of the methods.

1) Most importantly, the details on how SVs were called and classified is insufficient. The authors use the package Mum&Co but to those not familiar it is unclear what exactly it's doing. For example, the input compares each genome to "reference" but then what: are there size constraints or filters? Are the classifications into Presence-Absence Variation (PAV), CNV, etc. done by this package, or later by the authors? What distinguishes PAV from an indel? Are overlapping SVs or CNVs with different boundaries considered unique events, or are they related in any other way? Is aneuploidy included in the CNV class? Since this is the main focus and interest of this manuscript, more information in the methods is required. Furthermore, defining in the main text what an SV is here (e.g. difference in sequence stretch >1 bp compared to a reference strain) would be useful.

2) I did not understand the approach to calculate SV diversity (SV). This is the average number of SVs per site – averaged over the genome? Is each SV considered one event? Does the size of the SV matter at all? What is n in the formula? Perhaps it's just the # of SVs differences averaged over the number of bp in the genome, but if so I couldn't quite follow.

3) I interpreted PAV to mean difference in sequence content at a given locus between two strains, but later in the Methods there is a section describing gene presence and absence. Is PAV at the sequence level or expression level? Because it's ambiguous at what step and how these classifications are being made, this is unclear.

4) Perhaps gene presence and absence is something different, but it seems to be defined by expression – I fundamentally disagree with this. Gene presence and absence should be at the DNA / coding potential level, but the Methods imply that it is called based on RNA-seq reads (see also Fig S13). However, if expression regulation has evolved it's possible that expressing conditions haven't been sampled in some strains. The authors could distinguish between gene *sequence* presence / absence versus expression presence /absence, but these characteristics should be split into separate metrics.

Other more minor comments are below:

5) The paper frequently refers to "the reference" without any definition (aside of an R number with no citation deep in the methods). If this is the S288c reference published in SGD, then that should be explicit in the main text with a version number in the Methods and citations to SGD.

6) On page 3: how were unphased heterozygous genomes collapsed? I couldn't find information on that in the Methods.

7) There are many places where a little more detail in the results would improve digestibility. For example, at the top of page 5 it is stated that Ty elements "accounted for" X% of SVs, but it's impossible to know what that means without digging through the methods. Stating something like, "Ty elements were associated with (i.e. spanning 50% of the SV sequence) in X% …"

8) As above, it's not clear how a single SV diversity is calculated per site if SVs are different lengths.

9) Perhaps I missed it, but I did not see much analysis on SV length versus various features. For example, are longer SVs that likely affect more genes more likely to be pleiotropic than shorter ones? Is it possible to split each SV category into those that change coding sequence versus those that don't? I wondered if for example some of the minor insertion effects are because many are in noncoding regions, for example – would those within genes, and especially those that change frame, have larger effect? At any rate, testing how much the coding effects explain class differences would be useful.

10) Why is PAV length distribution bimodal in Fig S6?

11) Lines 132-137 were confusing to me: SV length is inversely proportional to heterozygosity, but the last line states that longer SV classes are enriched for being more heterozygous. Those statements seem conflicting to me.

12) Lines 153-155: the authors look at SV clade enrichments using a simple hypergeometric test, but this isn't appropriate because the strains are not independent of one another. The authors use this enrichment test as evidence for selection for SVs, but I don't buy this since it could easily be a bottleneck from a common ancestor. Any claims about selection are interesting but require better modeling using a phylogenetic perspective.

13) On the clade enrichment for different classes (around line 163). The manuscript says that translocations were enriched in the wine class, but it's unclear how this was calculated: is this a hypergeometric test on the number of unique SVs in the clade? Or is it the number of SVs across all genomes, including recurring and identical SVs? The latter doesn't seem appropriate because they could just be inherited neutrally from a common ancestor of the clade. Similarly, the numbers in Fig S8 on the right side are unclear: are those the number of unique SVs, or total SVs including redundant counts of the same SV in all of the strains?

14) Fig 2G was a little unclear: SNPs were called how, against a reference strain or somehow using the pan genome? A better description in the main text would help. The text states that Chinese strains had fewer SVs than expected, but really it looks like they have way more SNPs than other strains and an expectable number of SVs along the x-axis. Unless the SNP variants fit a clock expectation based on the tree, in which case that should be explained better.

15) Gene-based pangenome: it'd be useful to say in a sentence how the genes were annotated in the genomes. Likewise, a few words about the distinction between introgression and HGT would save the reader digging through the methods to find

out what the distinction is (namely, what species its closest ortholog was in).

16) Why is there a hump regarding gene family distribution in Fig 3G (also Fig S12), is that related to a distributed signal from the centromere? A quick check aligning on the CEN would reveal if something like that is happening. Either way, why the hump?

17) Line 219 cites that number of SVs – it might also be useful to cite the total number of basepairs spanned by those SVs.

18) I think the term “organismal” trait is overstated, since my understanding is that these are all colony-based growth rates. Minimally, that should be made clearer since I hesitate to use these as a representative of all organismal traits as a class.

19) Furthermore, I would bet that there is much more noise measuring the growth rates that are also subject to environmental variation. While it “makes sense” that organismal traits may have a more complex architecture, I’m not sure these data prove that. For example, is it true if the authors remove cis-acting e/pQTLs that likely have very local (and simple) effects? I wondered if molecular traits explained in trans have the same behaviors (maybe this is listed in one of the figures but I don’t think it is in this context).

20) Perhaps this is listed somewhere and I missed it, but in Figure 4A how do the authors control for SNPs versus SVs that span those SNPs (such that it is the SNP that’s causal)?

21) “Molecular” traits should be defined, I was initially thrown off by ‘local’ and ‘distant’ until I realized these must be eQTL from RNA-seq and pQTL from proteomics.

22) Several of the figure legends have alphabetical numbers that are not explained (e.g. Fig S5B, what are the numbers?) while others state that they refer to statistical significance but I could not find any key. In the latter case, just list the statistical significance on the figure or in the legend please so that each figure can stand on its own.

23) The availability statement is incomplete, furthermore I imagine there must be custom scripts for this work that should also be provided.

Referee #3

(Remarks to the Author)
In this study, the authors employ long-read sequencing to assemble the genomes of over 1000 Saccharomyces cerevisiae isolates, with a particular focus on structural variants (SVs) that are typically inaccessible via short-read sequencing approaches. This represents a technical achievement and extends previous large-scale efforts by the same group.
This work should be viewed in the context of two closely related prior studies:
1. The 2018 Nature paper (PMID: 29643504) from the same group, which provided analysis of SNPs and indels across the same set of strains using short-read sequencing, alongside a broad panel of phenotypic data. This foundational study offered a first view of the S. cerevisiae pangenome, including the definition of core and accessory genes.
2. A 2023 Nature Genetics paper (PMID: 37524789), involving members from the current team, reported telomere-to-telomere assemblies of 142 strains. That study focused on structural variation, however they integrated these only with genomic data, not with phenotypic traits.
The current manuscript builds directly on both studies, particularly by integrating the long-read-derived dataset with the phenotypic information collected in 2018. This provides valuable insights into the phenotypic relevance of SVs, an area that was previously unexplored at this scale in S. cerevisiae.

Comparisons with previous studies:
Variant types: while the 2018 study resolved SNPs and Indels and linked them to phenotypes, the current study adds SVs and connects these to phenotypic traits.
Ploidy: Ploidy determination was already possible in the 2018 dataset, though it is more precisely resolve here using long-read assemblies.
SV Catalog: The 2023 study identified 4809 SVs (across 142 strains). The current study, with its broader strain set, identifies 6587 SVs, increasing the median number of SVs per strain from 240 to 289.
Variable regions: Both the 2023 study and the current work conclude that SVs cluster in subtelomeric regions.
Gene content: The number of core and accessory genes increased from 4949 core/2856 accessory ORFs (2018) to 5047 core/3494 accessory genes (current), bringing the total gene count from 7805 to 8541.
Integration with Phenotypes: Unlike the 2023 study, which linked SVs to transcriptomic data only, the present work associates SVs with phenotypic traits recorded in 2018. This is an important contribution to the field.

Assessment of novelty and impact
While the study clearly required substantial effort and provides meaningful refinements over previous work, I am not fully convinced that the level of novelty justifies publication in Nature. The main findings represent a valuable extension of existing studies rather than a conceptual breakthrough. The integration of SVs with phenotype is important, but may be better suited for a high-impact specialist journal unless the manuscript is significantly reframed to highlight broader implications.

Clarity and Readability
The manuscript is currently difficult to read. Abbreviations are often undefined, and the structure lacks clarity in places.

Compared to the 2018 study, the presentation is less accessible to a broad audience. If the authors intend to reach a general Nature readership, substantial improvements in the writing and presentation are required.

Summary
This study provides a comprehensive analysis of long-read assemblies of >1,000 S. cerevisiae isolates and, for the first time at this scale, links SVs to phenotypic traits. While technically impressive and a valuable resource for the field, the conceptual advances appear incremental relative to earlier work. Moreover, the manuscript's readability must be improved if it is to be considered for publication in Nature.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
The revised manuscript has explicitly addressed my initial concerns and recommendations. The only remaining minor suggestions that I have are as follows:

1. The authors note that the graph pan genome limitations include reciprocal translocations. This is mentioned in the Methods and Materials section but should also be included somewhere in the Discussion section.

2. In the results section, please replace "... organismal traits (growth trait)..." with "...colony level growth traits (used here as organismal trait proxies)..." (line 232)

Referee #2

(Remarks to the Author)
The authors have done a good job adding substantial detail, clarifications, and revised analysis and have addressed my main comments. The methods in particular are now much more detailed. A few minor changes to the final draft could include:

Lines 175-177: "Wild clades, particularly the Chinese wild group, the species' ancestral population, deviates from this correlation, with fewer SVs than expected based on the number of SNPs, or
conversely (Fig. 2G)." "or conversely" is not clear. Perhaps the authors mean fewer SNPs compared to the rest of the distribution.

The authors nicely answered some of my questions in the response letter, but since other readers may have the same questions it may be worth adding those responses briefly For example:

"[R] The bimodal distribution of PAVs length (Fig S6) reflects the presence of SVs associated with Ty elements (~6 kb) and solo long terminal repeats (LTR, ~300 bp)." You could add this to the supplemental legend.

"[R] We believe that the reviewer is referring Fig. 3C and Fig. S12A. The hump observed here is not
associated with centromeres but originates from a large introgressed region of ~164 kb found in
chromosome 7 of one Mexican agave isolate. This region harbors 65 private genes, leading to a
localized enrichment in gene content in the pangenome, that appears as a peak on Fig. 3C" If space permits, the legend could add that a large introgression event in one strain (##) produces a signature at Xbp. That can be at the authors' discretion.

Referee #3

(Remarks to the Author)
The authors have substantially enhanced the clarity of the manuscript. In addition, they have rewritten the discussion, providing a much stronger rationale for the conceptual advances of this study compared to previous work.

Referee #1

Loegler et al. presents an extensive dataset of 1,086 *Saccharomyces cerevisiae* genomes, assembled nearly to telomere-to-telomere completeness, along with a large compendium of phenotypic measurements. Using long-read sequencing and assembly methods, they build a species-wide structural variant (SV) atlas, a gene-based pangenome, and a reference graph pangenome that incorporates 6,587 SVs. They then perform genome-wide association studies (GWAS) between this full catalog of genetic variants and 8,391 traits (molecular traits such as transcript, protein and metabolite levels, as well as organismal traits). The key findings include SVs and small InDels in the analysis and increases trait heritability estimates by ~14% (on average) compared to SNPs alone. SVs are more often associated with traits and appear more pleiotropic than SNPs and the genetic architectures of molecular vs. organismal traits show distinct patterns. Overall, the work is positioned as providing an unprecedented resource linking full-genome variation to phenotype, with implications for understanding genotype–phenotype relationships.

This manuscript presents a rich and valuable dataset on yeast genomic variation and its link to phenotype, and it addresses an important question in genetics. The scale of 1,086 near-complete genomes is impressive and could significantly advance the field. However, before publication, the authors should strengthen the methodological descriptions. In particular, clearer explanations of the assembly and SV-calling methods, and of the statistical thresholds used in GWAS, are needed. The Discussion should explicitly situate the work relative to recent pangenome and GWAS studies in yeast and other species to highlight what is novel. Given the broad interest of this topic, the manuscript should also more clearly articulate the general lessons about genotype–phenotype mapping (e.g. the utility of including SVs in association studies).

Major Comments

• Methodological rigor and clarity. The authors have undertaken an impressive scale of genome assembly, but the manuscript should clarify critical details of the assembly and variant-calling workflow. In particular, how many chromosomes were completely assembled and were any highly repetitive regions (rDNA arrays, telomeric repeats) unresolved? If multiple contigs per chromosome remain, the authors should explain how these gaps might affect SV calls.

[R] To address the reviewer's comment on assembly contiguity, we have added a supplementary note providing additional information on the contiguity and completeness of the assemblies.

Supplementary note 2 - Contiguity and completeness of genome assemblies

A total of 1,482 genome assemblies were generated from 1,086 natural isolates. Haplotype-resolved assemblies were obtained for 396 out of 456 non-polyploid heterozygous isolates. The median N50 per assembly is comparable to that of the S288c reference genome (910 vs. 924 kb).

To quantify the number of chromosomes assembled into a single contig, we considered a set of 22,635 chromosomes that do not carry reciprocal translocations, as these events complicate chromosome contiguity evaluation. A total of 22,008 (97.2%) chromosomes were assembled in a single contig (Fig. 1). The rDNA locus, located on chromosome 12, often caused discontiguity in the assemblies due to the significant copy number variation of this locus. This led to a reduced frequency of chromosome-scale contigs for chromosome 12, with 92.6% of the chromosomes assembled into a single contig.

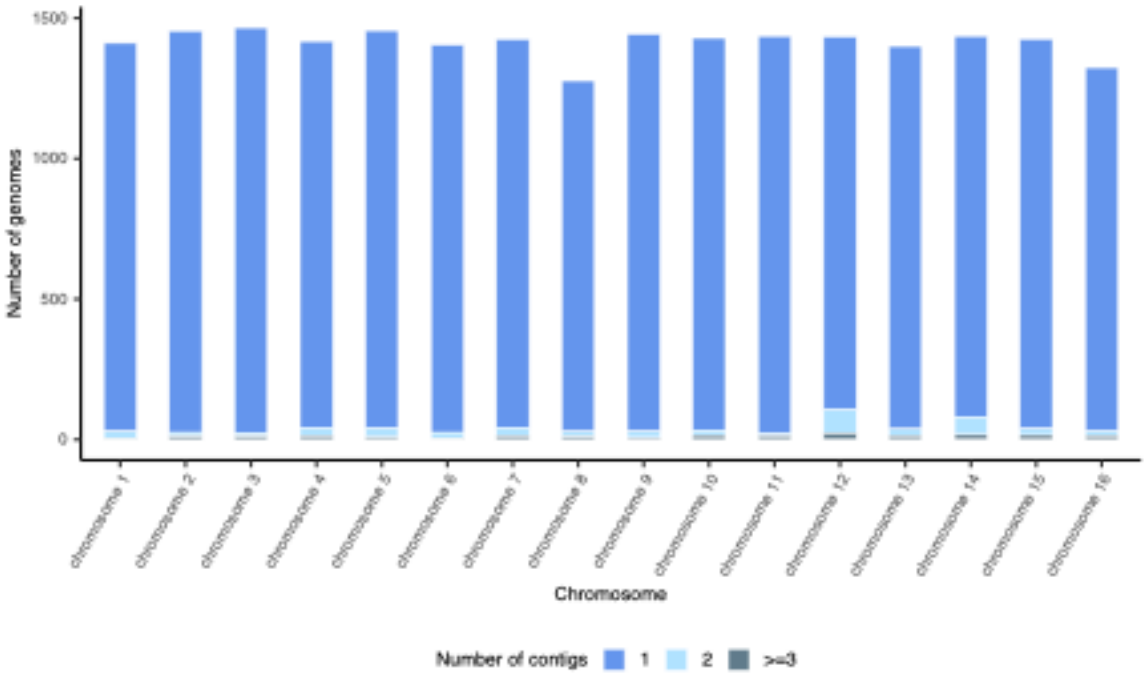

Fig. 1 | Number of chromosomes assembled. Only chromosome without reciprocal translocation were considered. The number of contigs associated to each assembled chromosome is color-coded.

Rigorous quality control measures were implemented to ensure the integrity of the genome assemblies, including the requirement that each assembly attain a minimum coverage of 95% of the S288c reference genome and include all chromosomes. However, this does not guarantee the achievement of telomere-to-telomere (T2T) status, which refers to a gapless assembly with the resolution of the 32 telomeric sequences present per haplotype. To assess telomeric completeness, we used Telofinder (O'Donnell et al., 2023, Nature Genetics, doi: 10.1038/s41588-023-01459-y) to identify resolved telomeres within the assembly collection. Of the 47,424 anticipated telomeres, 36,411 (76.8%) were successfully resolved. In total, 14,840 chromosomes (62.6%) were fully assembled into a single T2T contig, thereby capturing the entire chromosome sequence

Although our assemblies exhibit high contiguity and completeness, it would be inaccurate to describe them as telomere-to-telomere assemblies. It is therefore more accurate to describe them as near-T2T assemblies.

Since SV detection relies on assembly comparisons, a lack of assembly contiguity could result in SVs being overlooked. This is particularly relevant for regions that are difficult to resolve, such as subtelomeric sequences and the rDNA locus. The contiguity of genome assemblies (as measured by the N50 value) shows a weak correlation with the number of SVs detected per isolate (Fig. 2a; Pearson correlation, $R = 0.058$, $P = 0.026$). This weak correlation is primarily driven by outlier assemblies with an N50 value lower than 750 kb, which exhibit an average of 15% fewer SVs than other assemblies (215 vs. 253 SVs on average). Since the 27 outlier isolates are evenly distributed within the phylogeny (Fig. 2b), they are unlikely to affect downstream analyses, such as genome-wide associations, and were not discarded.

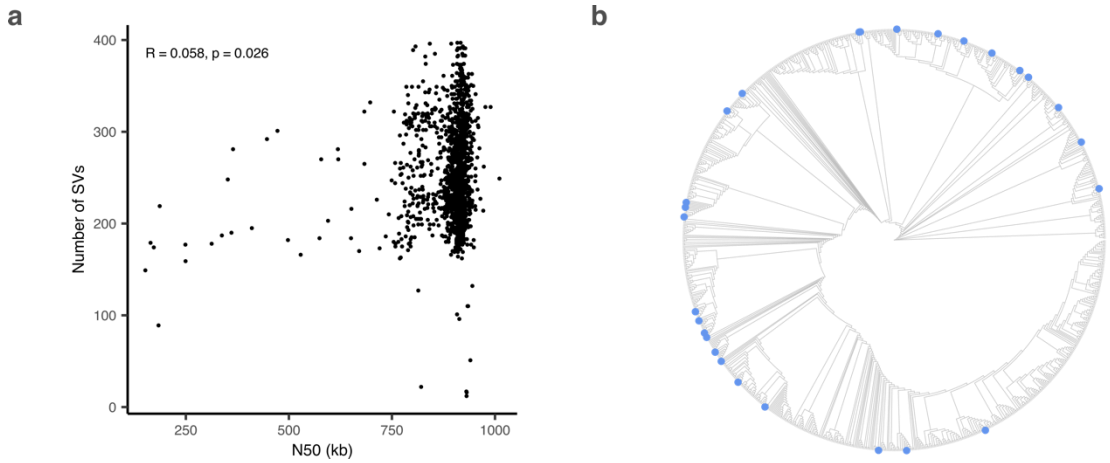

Fig. 2 | Assembly contiguity and structural variants (SVs). a. Number of SVs per assembly depending on the assembly N50. The correlation was computed using a Spearman correlation test. **b.** Neighbor-joining tree of the 1086 isolates that constitute the genome assembly collection. Isolates with a genome assembly of low contiguity (N50 lower than 750 kb) are indicated by a blue dot.

The sample set includes a mix of haploids and diploids (~75% diploid, with ~55% heterozygosity). Were heterozygous diploids phased or collapsed into consensus sequences? This choice could impact variant detection.

[R] The description of the sample and the assembly process was likely not precise enough in the manuscript. We have therefore added a note to provide more detail on this (see below). Briefly, for 396 out of 456 heterozygous diploids (~87%), we generated phased genome assemblies. The genome assemblies of the remaining 60 isolates were collapsed into a consensus sequence because the sequencing depth was insufficient to obtain a complete, contiguous, and phased assembly. SV calling was then performed on each haplotype for the phased genome assemblies, or on the consensus sequence for the other isolates. We acknowledge that this affects variant calling, as, in the case of collapsed consensus sequences, only one allele is randomly captured for each heterozygous locus. However, we prioritized genome contiguity and completeness over haplotype phasing, as we could not achieve both for heterozygous isolates with low sequencing depth.

Supplementary note 1 - Haplotype resolution among genome assemblies

Our genome assembly collection includes newly generated assemblies for 1,015 isolates, as well as previously generated assemblies for 71 strains (O'Donnell et al., 2023, Nature Genetics, doi: 10.1038/s41588-023-01459-y), reaching a total of 1,086 isolates (Table S2). Of these, 186 are haploids, 814 are diploids, 42 are triploids, 29 are tetraploids, and three are pentaploids. The ploidy of 12 isolates could not be determined due to their complete homozygosity, which precludes allele-frequency-based inference. In total, 527 isolates are homozygous, comprising seven haploids with a heterozygous aneuploid chromosome, 449 diploids, 40 triploids, 28 tetraploids, and three pentaploids.

Due to the variation in ploidy and the presence of heterozygosity in many isolates, we attempted to obtain haplotype-resolved (phased) genome assemblies when possible, to enable accurate structural variant genotyping. Collapsed assemblies, which represent one consensus haplotype capturing a random allele at each heterozygous locus, may miss some structural variants when haplotypes diverge. However, phased assemblies often exhibited reduced contiguity and completeness compared to collapsed assemblies. Poor assembly quality in phased assemblies can hinder SV detection, necessitating a trade-off between haplotype resolution and assembly quality.

Since polyploid genome phasing using ONT remains a considerable challenge, often at the expense of assembly contiguity, we opted for collapsed assemblies for the 71 heterozygous polyploid isolates. These were generated by assembling all sequencing reads from an isolate into a single collapsed

haplotype. For haploid and diploid heterozygous isolates, haplotype-resolved assemblies were considered when both haplotype assemblies met the quality threshold defined in the genome assembly pipeline (Methods). If one or both haplotypes failed to meet these criteria, the corresponding collapsed assembly was used instead. Therefore, phased assemblies were retained for six out of seven heterozygous haploids with aneuploid chromosomes and 390 out of 449 heterozygous diploids. Collapsed assemblies were retained for the remaining 60 isolates.

More generally, the SV-calling pipeline should be described in full: which tools were used (assembly comparison vs. read mapping), what filtering criteria were applied, and how were false positives controlled for?

[R] We have revised the paragraph in the Methods section describing SV detection to provide a clearer and more detailed explanation of the pipeline. Specifically, the filters applied to ensure a low false positive rate are now described in greater detail.

Given the known challenges of detecting large insertions and complex rearrangements, some validation (e.g., by PCR or independent data) would strengthen confidence in the SV atlas.

[R] We agree that validation is important given the complexity of detecting large insertions and rearrangements. To address this, we used an independent dataset (Peter et al., *Nature* 2018) of short-read sequences and we performed a new analysis on the long-read sequences. We have added an additional note on SV validation (see below), supported by a systematic examination of the mapping of both short and long reads to the reference genome. A sentence was also added to line 111 of the Results section.

Supplementary note 3 - Validation of SV calls

To assess the reliability of SV calls inferred from genome assembly comparisons, we performed validation using both Illumina and ONT sequencing reads mapped to the S288c reference genome. Illumina reads were aligned with bwa-mem2 (Vasimuddin et al., 2019, IPDPS, doi: 10.1109/IPDPS.2019.00041), and ONT reads with Minimap2 (Li, 2018, Bioinformatics, doi: 10.1093/bioinformatics/bty191).

We randomly selected 500 SVs for systematic inspection in IGV (Robinson et al., 2011, Nat. Biotechnology, doi: 10.1038/nbt.1754) (Fig. 1). Our validation set includes 252 insertions, 104 deletions, 79 duplications, 24 contractions, 16 inversions, and 25 translocations. These numbers are proportional to the fraction of each type of SV in the total set of SV calls.

Disrupted short-read alignments supported the presence of structural variation in 474 of 500 cases (**94.8%**). The 26 unsupported cases were predominantly insertions (19/26), consistent with the known limitations of short-read technologies in detecting such events (Delage et al., 2020, BMC Genomics, doi: 10.1186/s12864-020-07125-5).

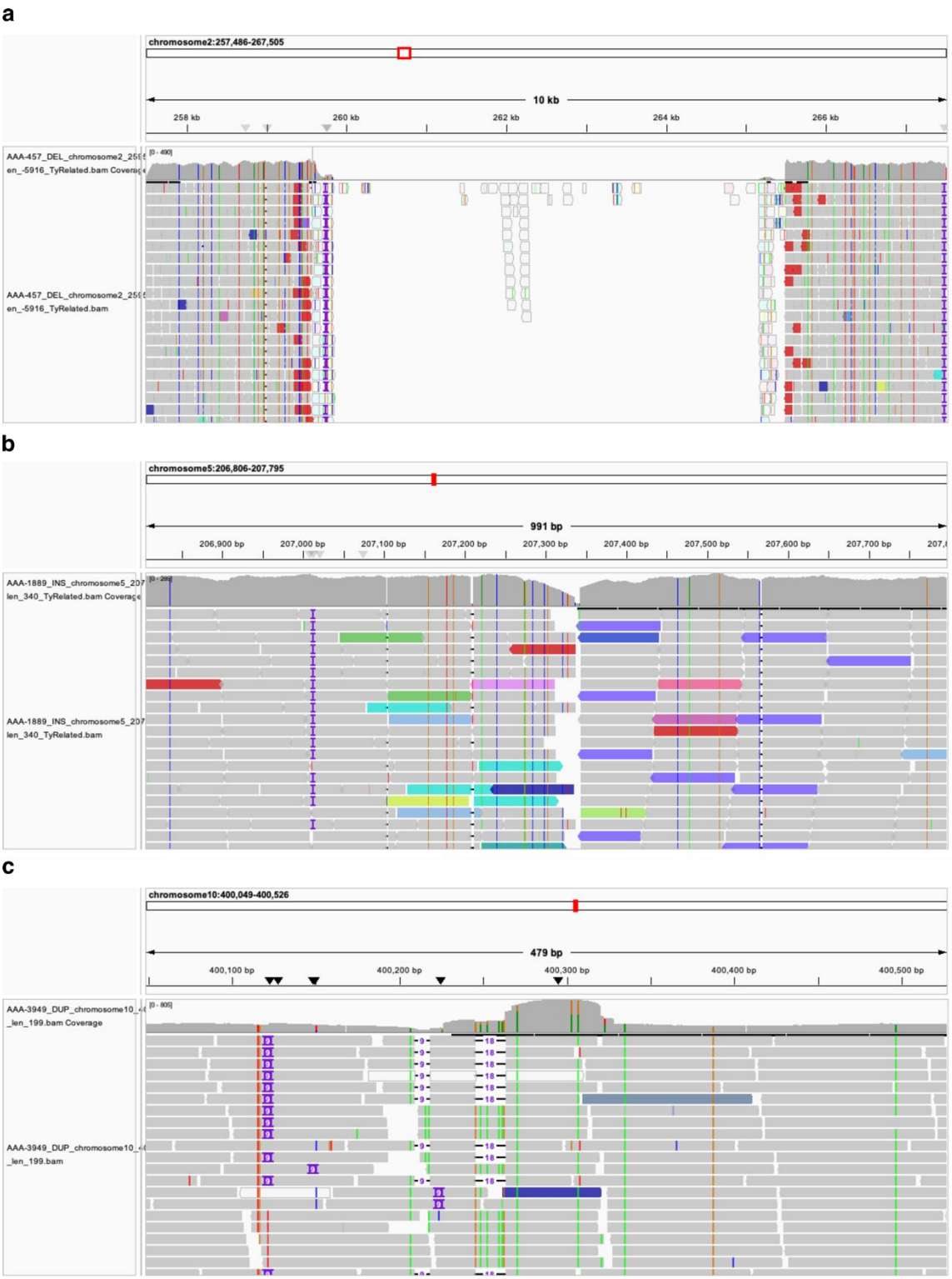

Fig. 1 | Disruption of short-read alignments in the presence of structural variants. Examples include a deletion (**a**), an insertion (**b**) and a duplication (**c**).

Long-reads alignments on the reference genome allowed to further validate the type and length of variant detected. Systematic inspection of ONT reads mapping confirmed 441 out of 500 SVs (**88.2%**). Validation was considered as read alignment disruption for long and complex SVs such as inversions and translocations, or within-read insertion or deletion for other SVs (Fig. 2). Unconfirmed cases were significantly enriched near chromosome ends (within 1 kb of termini; Fisher's exact test, P = 0.039), likely due to increased sequence complexity and reduced mapping accuracy in these regions.

A total of 9 out of 25 translocations were not confirmed by long read mapping, often because of read discrepancies: most of the unconfirmed translocation cases exhibit a few reads confirming the translocation, while many does not, resulting in difficult SV validation for these types of SVs.

Fig. 2 | Validation of SVs using long-read alignments. Examples include a deletion (**a**), an insertion (**b**) and a duplication (**c**).

Together, this systematic inspection of sequencing reads supports the presence of the majority (**94.8%**) of SVs in the selected cases.

Similarly, the graph pangenome was built with Minigraph using 500 haplotypes – the authors should discuss whether this graph is fully consistent and robust, and how they ensured that all SV alleles were captured.

[R] Due to the inherent method of graph-pangenome construction algorithms, not all structural variants can be captured in a graph-pangenome. All algorithms rely on alignments between homologous chromosomes. As a result, reciprocal translocations, which result in mixed chromosomes, are not handled by graph-pangenome algorithms and are missed in the resulting variant catalog. For this reason, we acknowledge that graph pangenomes do not fully capture the structural variation of a population. However, we ensure the inclusion of the most diverse haplotypes in the graph construction to capture the maximum amount of diversity possible. The 500 haplotypes selected capture all the SVs detected through assembly comparison. Because of this selection, any additional haplotype from our collection should have a limited impact on increasing the structural diversity captured by the graph pangenome.
A sentence highlighting the limitation of the graph pangenome in exhaustively capturing SVs has been added to line 339.

• Novelty and relation to prior work. The study's scale is unprecedented for yeast. The authors should explicitly compare their 1,086 genome SV catalog to the recent ScRAP study by O'Donnell et al. Nat Genet 55: 1390, 2023 and the study by Weller et al. Genome Res 33: 729, 2023; and explain how their larger size increases resolution and changes conclusions.

[R] We have added a paragraph in the Discussion section directly comparing previous preliminary studies in yeast (Istace et al. *GigaScience* 2017, Weller et al. *Genome Research* 2023, O'Donnell et al. *Nature Genetics* 2023). This new paragraph highlights the increased resolution afforded by our larger dataset and explains how our broader sampling refines and extends conclusions from these prior studies.

• Conceptual interpretation. The authors do note pleiotropy of SVs, but it would be useful to relate this to known examples. It is possible that SVs appear more pleiotropic simply because their larger genomic footprint makes them more likely to disrupt multiple genes. The authors should check whether the higher pleiotropy is robust after accounting for variant size and allele frequency. Correlated traits could also inflate apparent pleiotropy. It would strengthen the claim to provide an analysis of whether SV-trait associations remain significant after conditioning on nearby SNPs (e.g. in a joint model).

[R] We performed additional analyses to assess whether the elevated pleiotropy of SVs could be attributed to their larger size or a higher likelihood of gene disruption. To avoid size overestimation from gene copy number variants (CNVs), which do not accurately reflect physical SV length, we excluded these from the dataset, resulting in 95 trait-associated SVs ranging from 53 to 19,888 bp. Even with this adjustment, SVs remained significantly more pleiotropic than SNPs and InDels (two-sided Wilcoxon rank-sum test, $P = 0.00185$), and no correlation was observed between SV length and the number of associated traits (Spearman correlation, $P = 0.311$). We further corrected pleiotropy by the number of genes disrupted by each variant, including genes fully or partially overlapped by deletions or CNVs, embedded within insertions, or intersected by translocation breakpoints (±1 kb). After normalizing for gene disruption, SVs still showed significantly greater pleiotropy (two-sided Wilcoxon rank-sum test, $P = 0.00305$), suggesting that their heightened pleiotropic effects are not solely explained by size or gene count, but likely reflect distinct functional properties.
In addition, high linkage disequilibrium (LD) between associated SNPs and SVs can introduce ambiguity in pinpointing causal variants. To mitigate this issue, we excluded all associated linkage groups that contained multiple variant types. This filtering step removed 60 such groups, resulting in a final set of 7,385 linkage groups composed of a single variant type. Despite this more stringent classification, SVs continued to show significantly higher pleiotropy compared to both SNPs and InDels (two-sided Wilcoxon rank-sum test, $P = 6.57\times10^{-72}$ and $P = 2.54\times10^{-10}$, respectively).

Likewise, the statement that "the genetic architecture of molecular and organismal traits differed markedly" needs more context. Are the differences driven by the type of trait (e.g., expression vs. growth) or by the source of the data? Adding a brief comparison to known patterns (e.g., omnigenic models, or prior yeast eQTL studies) would help interpret this point.

[R] As mentioned in the text, the differences in genetic architecture between trait types are primarily reflected in the number of associated variants, their effect sizes, and the distribution of variant classes (Figure 5A–C). While it would be valuable to compare these results to known patterns from previous QTL studies, most existing work in yeast, particularly linkage mapping studies focusing on individual trait types (e.g., growth, transcript, or protein levels), has been limited to SNP-based analyses. As a result, direct comparisons are not possible. Only the comprehensive, multi-layered dataset generated in this study, which combines high-contiguity genome assemblies, extensive variant types, and diverse phenotypic measurements, enables such detailed conclusions about how genetic architecture varies across molecular and organismal traits.

• Significance and interest. The work will be of high interest to yeast geneticists and to all scientists studying genotype–phenotype mapping. The large sample and multi-omic trait set make it a valuable resource. However, to appeal to Nature's broad readership, the authors should highlight general insights – for example, the utility of incorporating SVs in any GWAS (consistent with studies in other organisms) and lessons for complex trait genetics more broadly. Is there evidence here that could inform human or plant genetics? Conversely, some readers may see this as a specialist yeast study unless the broader implications are made clear. If possible, the authors should emphasize one or two "take-home" messages that transcend yeast (e.g., SVs often harbor causative variation, so pangenomic approaches are generally warranted – Jeffares et al. Nat Comm 8: 14061, 2017). Stressing a conceptual advance would raise its impact.

[R] We appreciate the reviewer's insightful comments regarding the broader significance of our study. In response, we have revised the Discussion to explicitly highlight key conceptual insights that extend beyond yeast. In particular, we emphasize that structural variants (SVs) are disproportionately associated with complex traits and exhibit high pleiotropy. We also discuss how only a comprehensive, multi-layered dataset such as ours, integrating high-resolution SV maps, pangenomes, and multi-omic phenotyping, can fully resolve the contribution of different variant types to trait architecture. We hope these additions make the broader implications clearer and help position the work as a general framework for genotype–phenotype mapping across eukaryotes.

Minor Comments

• Clarity of terminology. Define "near telomere-to-telomere" explicitly (how near? does it exclude rDNA?). Similarly, terms like "gene-based pangenome" should be briefly explained for non-specialists. The manuscript uses "SV" consistently, but ensure that each type of SV (deletion, duplication, inversion, translocation) is defined when first mentioned.

[R] A supplementary note (SN2) has been added to provide detailed description of assembly contiguity and completeness. This note details the frequency of resolution of the rDNA locus, as well as the number of telomeres successfully resolved. We defined the near telomere-to-telomere (near-T2T) status as a level of contiguity and completeness analogous to the reference genome, while acknowledging that not all telomeres are necessarily resolved. For instance, the average N50 of our assemblies is close to that of the reference genome (910 vs. 924 kb), and 97.2% of the chromosomes were assembled in a single contig. Minimal assembly completeness has been guaranteed for each assembly by ensuring that at least 95% of the reference genome is covered. However, only 76.8% of the telomeric sequences are fully resolved, preventing the true T2T status. Nevertheless, this corresponds to near-T2T status. A sentence has been added to the main text to precise this definition (line 102).
In addition, a sentence has been added to the main text to define the term "gene-based pangenome" line 187.

Finally, a revised version of the SV calling paragraph in the method section now ensure the precise definition of the different SV types (line 467 & 485).

• Pleiotrophy definition. The term “pleiotropy” is used to mean “variant affects many traits.” The authors should state the criterion (e.g., number of trait associations) used to call a variant pleiotropic. Otherwise, readers may be confused as to whether, for example, any variant associated with more than one trait counts as pleiotropic, or if some threshold is applied.

[R] We used the term pleiotropy to refer to the average number of traits associated with each variant type, as stated on line 242. In this context, higher pleiotropy indicates that, on average, variants of that type are associated with a greater number of traits. In our study, we define a pleiotropic variant as one that is associated with more than one trait. To clarify this, we have added a sentence to the main text that explicitly defines our use of the term and reports the proportion of pleiotropic QTL for each variant type (line 243).

• Figure legends and units. Ensure that all figures have clear legends. For Manhattan plots, confirm that all axes are labeled. If bar graphs have error bars or asterisks for significance, these should be explained. Check that any color-coding (e.g., trait types) has a key.

[R] We revised figures and figure legends that lacked information about *P* values, color keys, error bars, and explanation of boxplots. Asterisks used for significance were explicitly defined in the legend.

Statistics and Data Presentation

• Statistical methods. The use of linear mixed models (LMM) to control for population structure in GWAS is appropriate (Loegler et al. mention using FaST-LMM). However, the manuscript should provide more detail. For instance, how was the kinship (relatedness) matrix constructed – using all variants, or only pruned SNPs? It is noted that all variants (SNPs, InDels, SVs) were combined into one genotype matrix for LMM, which is reasonable. It would help to clarify the minor allele frequency (MAF) and linkage disequilibrium (LD) pruning thresholds used. If SVs tend to be rarer or in high LD, this could affect power and false-positive rates. The authors should specify how many variants of each type passed MAF filters, and whether allele frequency differences were accounted for.

[R] The paragraph on GWAS in the method section (line 681) has been fully revised to answer all the reviewer’s comments. The kinship used was detailed. The number of variants retained after MAF filtering and LD pruning was indicated. There is a higher proportion of common SVs (MAF > 0.05, 10.7%) than common SNPs (6.4%). This difference was not accounted for, and this has been clearly stated in the revised paragraph. A similar fraction of SVs and SNPs were discarded through LD pruning (58.9% vs. 60.7%, respectively), indicating a similar degree of linkage disequilibrium.

• Heritability estimation. The claim that including SVs + InDels raises heritability by 14.3% should be clearly defined. Are these narrow-sense (additive) heritabilities estimated with LDAK/GCTA? The authors mention rank-based inverse normal transformation of traits and LDAK v5.2 for heritability. They should clarify whether heritability is on the original scale of each trait, and how significance of heritability was assessed. It would also be useful to discuss whether the increase in heritability is relative (e.g., heritability was 0.30 with SNPs only and 0.343 with SVs added, a 14.3% relative gain) or an absolute percentage point change. In similar graph-pangenome studies, notably in tomato, inclusion of SVs yielded a 24% relative increase in heritability (Zhou et al. Nature 606: 527, 2022). The yeast result is smaller but of the same order, which is plausible. The authors should ensure consistency with published definitions.

[R] Narrow-sense heritability was estimated using LDAK. Heritability estimates are scaled from 0 to 1, not at the original trait scale. After discussion with LDAK developer, variants in high correlation with the phenotype may bias the heritability results obtained. To ensure reliability of our estimates, we

discarded any traits with highly significant associations ($P < 10^{-20}$). A sentence has been added to the method section to justify this choice (line 672).

The increase of 14.3% in heritability correspond to a relative gain, from an average of 0.36 with SNPs, to an average of 0.41 with all types of variants considered. Average numbers have been added to the result section for clarification (line 234).

• Multiple testing and error bars. The manuscript should explicitly state how significance thresholds were set in the GWAS. With millions of SNPs plus thousands of SVs tested per trait, a genome-wide correction is needed. Did the authors use a strict Bonferroni or an FDR approach? For example, a 5% Bonferroni threshold would be extremely stringent given the variant counts. If an FDR (e.g., 5%) was used, this should be stated. It would also clarify how "SVs more often associated" was quantified – did an SV simply exceed the same p-value threshold as a SNP? If so, comparisons could be biased by the relative numbers of variants.

[R] The establishment of GWAS significance thresholds is now clearly described in the revised Methods section. Thresholds were defined using a permutation-based approach, controlling the false discovery rate (FDR) at 5%.

The statement "more often associated" refers to the proportion of tested variants that were significantly associated with traits, as reported in line 239. This measure is normalized by the total number of tested SNPs, InDels, and SVs, and is therefore not biased by differences in the absolute number of variants across categories.

The figures appear to use error bars (e.g., in bar graphs of heritability or effect sizes). The legend or methods should define these error bars (standard error vs. confidence interval). All p-values mentioned in the text should specify whether they are one- or two-tailed and whether they are raw or adjusted. Throughout, the authors should maintain consistency (e.g., "P<0.05" should mean after correction if that is claimed to be "significant").

[R] The figure legends have been revised to precise the signification of all elements in the bar graphs. We revised the main text to precise one- or two-tailed for all statistical tests used. When multiple testing was performed, p-value adjustment method was indicated.

• Data sufficiency and trait variation. Given that many traits may have been measured only once per strain, it would help to discuss the replication or measurement error. For instance, do organismal traits (growth rates, etc.) come from single experiments, or averages of replicates? Lack of replication could affect the accuracy of heritability estimates. If available, quoting residual variances or confidence intervals for heritabilities would demonstrate rigor. At minimum, the authors should acknowledge any assumptions (e.g., that environmental variance is similar across strains).

[R] As mentioned in Peter et al. (*Nature*, 2018), Caudal et al. (*Nature Genetics*, 2024), Teyssonnière et al. (*PNAS*, 2024), and Muenzner et al. (*Nature*, 2024), most phenotypic traits (growth, transcriptome, proteome) were indeed measured once per isolate. However, the robustness of these datasets was validated using replicates for a subset of ~100 representative strains, which showed high reproducibility with correlation coefficients ($R > 0.9$) across all data types. These results confirm the robustness of the phenotypic measurements and therefore should not impact heritability estimates.

• Overall, the statistical approach appears sound (LMM, LDAK), but more detail on implementation and thresholding is needed to fully assess rigor.

[R] As mentioned above, we have now added detailed descriptions of the statistical approach and thresholding procedures in the revised Methods section.

Referee #2

In this work, the authors present a catalog of structural variants (SVs) and their impact on genetic architecture in *S. cerevisiae*. The authors present near T2T genomes for >1,000 S. cerevisiae strains along with detailed analysis of SV classes, distributions, histories, and potential impacts. The manuscript will be an important resource for yeast biologists and contains insights of broad interest to geneticists.

The manuscript is generally well written and thorough; parts of the manuscript would be easier to digest with a little more detail in the main text, as outlined below. My biggest concerns were with a few of the methods.

1) Most importantly, the details on how SVs were called and classified is insufficient. The authors use the package Mum&Co but to those not familiar it is unclear what exactly it's doing. For example, the input compares each genome to "reference" but then what: are there size constraints or filters? Are the classifications into Presence-Absence Variation (PAV), CNV, etc. done by this package, or later by the authors? What distinguishes PAV from an indel? Are overlapping SVs or CNVs with different boundaries considered unique events, or are they related in any other way? Is aneuploidy included in the CNV class? Since this is the main focus and interest of this manuscript, more information in the methods is required. Furthermore, defining in the main text what an SV is here (e.g. difference in sequence stretch >1 bp compared to a reference strain) would be useful.

[R] We have revised the Methods sections describing SV calling and the construction of the exhaustive variant catalog. These revisions improve clarity on the definition of structural variants, including the different SV types detected and how aneuploidies are captured. We now also explicitly describe how our strategy ensures completeness of the variant catalog, albeit with high redundancy, and how this redundancy is accounted for in downstream analyses.

2) I did not understand the approach to calculate SV diversity (πSV). This is the average number of SVs per site – averaged over the genome? Is each SV considered one event? Does the size of the SV matter at all? What is n in the formula? Perhaps it's just the # of SVs differences averaged over the number of bp in the genome, but if so I couldn't quite follow.

[R] The paragraph on structural diversity in the method section has been revised to enhance clarity. The method to define the number of SVs in a window is now better described line 501. To briefly answer the reviewer's comment, each SV is considered as a single event, regardless of its size.

3) I interpreted PAV to mean difference in sequence content at a given locus between two strains, but later in the Methods there is a section describing gene presence and absence. Is PAV at the sequence level or expression level? Because it's ambiguous at what step and how these classifications are being made, this is unclear.

4) Perhaps gene presence and absence is something different, but it seems to be defined by expression – I fundamentally disagree with this. Gene presence and absence should be at the DNA / coding potential level, but the Methods imply that it is called based on RNA-seq reads (see also Fig S13). However, if expression regulation has evolved it's possible that expressing conditions haven't been sampled in some strains. The authors could distinguish between gene *sequence* presence / absence versus expression presence /absence, but these characteristics should be split into separate metrics.

[R] Response for comments 3 & 4

Thanks for pointing this out, we realize this part of the Methods may have caused some confusion.

First, gene presence/absence in our study is strictly defined at the DNA level, based on short-read sequencing depth over gene coding sequences. It is not inferred from RNA expression levels, and gene expression plays no role in the final determination of presence or absence in a given isolate.

Second, to determine whether a gene is present in a given genome, we mapped Illumina reads to the representative coding sequences (CDS) of each gene family. We then computed a normalized depth of coverage for each gene, which we used to assess presence or absence. To account for variable ploidy across isolates, we introduced a correction factor (raising the ploidy to a power x) in our gene presence formula. The value of x was empirically optimized using a precision-recall framework.

Third, the precision-recall optimization step did make use of transcriptome data, but only to build a gold standard matrix of gene presence/absence across 553 isolates, where both genome assemblies and transcriptomic data were available. This step was necessary to determine the best threshold and ploidy correction factor for inferring gene presence based on sequencing depth. Specifically, we considered a gene present in the gold standard when it was both annotated in the genome assembly and expressed at ≥2 TPM in RNA-seq data, and absent when it was neither annotated nor expressed. Ambiguous cases were excluded. This gold standard was used only for model calibration, not for determining presence/absence in the actual dataset.

In summary, gene presence/absence is determined from DNA-level evidence only (short-read sequencing depth), and expression data was used only in the training of our thresholding method, not in any final gene calls. We clarified this in the method section to avoid future confusion (line 573-598).

Other more minor comments are below:

5) The paper frequently refers to "the reference" without any definition (aside of an R number with no citation deep in the methods). If this is the S288c reference published in SGD, then that should be explicit in the main text with a version number in the Methods and citations to SGD.

[R] The reference we used corresponds to the S288c reference published in SGD. Explicit mention of this reference and version has been added at the first occurrence of the term "reference genome" in the method section (line 409), and at the first occurrence of the reference annotations (line 548). Citation to SGD has been added as well.

6) On page 3: how were unphased heterozygous genomes collapsed? I couldn't find information on that in the Methods.

[R] Collapsed genome assemblies for heterozygous isolates were obtained by assembling all sequencing reads into a single haplotype. A supplementary note (SN1) has been added to clarify haplotype resolution within genome assemblies, a precise how collapsed assemblies were obtained.

7) There are many places where a little more detail in the results would improve digestibility. For example, at the top of **page 5** it is stated that Ty elements "accounted for" X% of SVs, but it's impossible to know what that means without digging through the methods. Stating something like, "Ty elements were associated with (i.e. spanning 50% of the SV sequence) in X% …"

[R] We agree that this enhances clarity in the result section. The revised this sentence accordingly (line 115).

8) As above, it's not clear how a single SV diversity is calculated per site if SVs are different lengths.

[R] SVs are considered as unique events, regardless of their size. However, large deletions, CNVs or inversions will be considered across multiple sliding windows along the genome. The paragraph about structural diversity has been clarified, as asked in comment 2.

9) Perhaps I missed it, but I did not see much analysis on SV length versus various features. For example, are longer SVs that likely affect more genes more likely to be pleiotropic than shorter ones?

Is it possible to split each SV category into those that change coding sequence versus those that don't? I wondered if for example some of the minor insertion effects are because many are in noncoding regions, for example – would those within genes, and especially those that change frame, have larger effect? At any rate, testing how much the coding effects explain class differences would be useful.

[R] We agree that these assumptions would make sense. To check for the impact of SV length, we excluded gene CNVs (which do not accurately reflect original SV size), which reduced the dataset to 95 associated SVs ranging from 53 to 19,888 bp. We did not see any significant correlation between the length of SV-QTL and the number of traits associated (Spearman correlation test, S = 127,858, P = 0.311). However, the sample size remains limited to be conclusive on this question.
To investigate how gene disruption affects trait variation, we focused on large insertions. Of the 73 associations involving insertion-QTL, 32 involve an insertion within a CDS, and 41 involve an insertion outside a CDS. The difference in effect size is minimal between these categories (0.018 vs. 0.019, two-sided Mann-Whitney-Wilcoxon test, P = 0.0175). Once again, the limited sample size precludes conclusions.

10) Why is PAV length distribution bimodal in Fig S6?

[R] The bimodal distribution of PAVs length reflects the presence of SVs associated with Ty elements (~6 kb) and solo long terminal repeats (LTR, ~300 bp).

11) Lines 132-137 were confusing to me: SV length is inversely proportional to heterozygosity, but the last line states that longer SV classes are enriched for being more heterozygous. Those statements seem conflicting to me.

[R] We thank the reviewer for pointing this out. This was a misstatement in the text. SV length is proportional to heterozygosity, not inversely proportional, as correctly shown in Fig. S6F. This has now been corrected in the main text (line 139).

12) Lines 153-155: the authors look at SV clade enrichments using a simple hypergeometric test, but this isn't appropriate because the strains are not independent of one another. The authors use this enrichment test as evidence for selection for SVs, but I don't buy this since it could easily be a bottleneck from a common ancestor. Any claims about selection are interesting but require better modeling using a phylogenetic perspective.

[R] We agree with the reviewer that clade enrichment for SVs may both originate from selection, or from population bottleneck. This was changed and specified in the main text line 161.

13) On the clade enrichment for different classes (around line 163). The manuscript says that translocations were enriched in the wine class, but it's unclear how this was calculated: is this a hypergeometric test on the number of unique SVs in the clade? Or is it the number of SVs across all genomes, including recurring and identical SVs? The latter doesn't seem appropriate because they could just be inherited neutrally from a common ancestor of the clade. Similarly, the numbers in Fig S8 on the right side are unclear: are those the number of unique SVs, or total SVs including redundant counts of the same SV in all of the strains?

[R] The enrichment for translocation in wine isolates were calculated using a hypergeometric test on the number of wine-specific SVs against all other clade-specific SVs. Each SV event is counted once, and multiple occurrences of a single event is not accounted for.
On Fig. S8, the numbers on the right side represent the total number of SV events enriched per clade (*i.e.*, the clade SV signature). The legend of the figure has been modified to specify that.

14) Fig 2G was a little unclear: SNPs were called how, against a reference strain or somehow using the pan genome? A better description in the main text would help. The text states that Chinese strains had fewer SVs than expected, but really it looks like they have way more SNPs than other strains and

an expectable number of SVs along the x-axis. Unless the SNP variants fit a clock expectation based on the tree, in which case that should be explained better.

[R] As pinpointed by the reviewer, the paragraph about SNPs and InDels calling was missing from the method section. It has now been added. Briefly, small variants were called using the reference genome with a traditional mapping strategy. To ensure high-quality variant calls, we retained only sites with confidently called genotypes in at least 99% of the isolates. This filter excludes regions affected by presence/absence variation, as such regions would result in missing genotypes for some strains. Therefore, the use of a pangenome graph would not substantially improve small variant detection under this filtering criteria.
Regarding the interpretation of Fig. 2G, we agree with the reviewer that it is not totally accurate to state that Chinese isolate have fewer SVs than expected. However, we also pinpoint that stating they have more SNPs than expected would be speculative without additional analyses. Therefore, we have revised the main text (line 176) to clarify that the Chinese isolates deviate from the overall correlation between SNPs and SVs observed across the broader dataset, without implying a specific direction for this deviation.

15) Gene-based pangenome: it'd be useful to say in a sentence how the genes were annotated in the genomes. Likewise, a few words about the distinction between introgression and HGT would save the reader digging through the methods to find out what the distinction is (namely, what species its closest ortholog was in).

[R] The sentence stating gene origin in the result section has been revised to enhance clarity (line 205). It is now clearly said that introgressions were defined as genes with highest similarity to a *Saccharomyces* gene, while HGTs have highest similarity to a non-*Saccharomyces* gene.

16) Why is there a hump regarding gene family distribution in Fig 3G (also Fig S12), is that related to a distributed signal from the centromere? A quick check aligning on the CEN would reveal if something like that is happening. Either way, why the hump?

[R] We believe that the reviewer is referring Fig. 3C and Fig. S12A. The hump observed here is not associated with centromeres but originates from a large introgressed region of ~164 kb found in chromosome 7 of one Mexican agave isolate. This region harbors 65 private genes, leading to a localized enrichment in gene content in the pangenome, that appears as a peak on Fig. 3C.

17) Line 219 cites that number of SVs – it might also be useful to cite the total number of basepairs spanned by those SVs.

[R] As suggested, we have now included the total number of base pairs spanned by structural variants in the revised manuscript. Specifically, we report that SVs (excluding translocations) span a total of 27.3 Mb of sequence. This information has been added to the main text (line 113).

18) I think the term "organismal" trait is overstated, since my understanding is that these are all colony-based growth rates. Minimally, that should be made clearer since I hesitate to use these as a representative of all organismal traits as a class.

[R] We agree with the reviewer that the growth traits used in this work are based on colony-level measurements. We assumed that these measurements represent a proxy for organismal traits. We have clarified this point in the main text (line 232). However, we decided to retain the term "organismal traits" in the text to maintain readability.

19) Furthermore, I would bet that there is much more noise measuring the growth rates that are also subject to environmental variation. While it "makes sense" that organismal traits may have a more complex architecture, I'm not sure these data prove that. For example, is it true if the authors remove cis-acting e/pQTLs that likely have very local (and simple) effects? I wondered if molecular traits

explained in trans have the same behaviors (maybe this is listed in one of the figures but I don't think it is in this context).

[R] To address this concern about the bias of local e/pQTL on the analysis of traits complexity, we discarded all molecular traits affected by a local QTL. This led to the removal of 1,091 traits, resulting in a total of 176 growth traits and 2,450 molecular traits considered. On average, growth traits were associated with 2.34 QTL, compared to only 1.45 for the remaining molecular traits. This indicates a significantly higher complexity for organismal traits (two-sided Mann-Whitney-Wilcoxon test, $P = 7.7\times10^{-27}$), even after removing locally affected molecular traits. This result support our previous findings.

20) Perhaps this is listed somewhere and I missed it, but in Figure 4A how do the authors control for SNPs versus SVs that span those SNPs (such that it is the SNP that's causal)?

[R] To control for variants in linkage disequilibrium that could lead to false positive associations, we defined groups of linkage for variants associated with a same trait. For each group, only the most significant variant was retained. This step has been described more clearly in the revised version of the manuscript (line 699).

21) "Molecular" traits should be defined, I was initially thrown off by 'local' and 'distant' until I realized these must be eQTL from RNA-seq and pQTL from proteomics.

[R] We agree that the use of the term molecular trait was not well defined. To prevent confusion, we explicitly mentioned transcriptomic and proteomic measurements in the revised version of the main text (line 232).

22) Several of the figure legends have alphabetical numbers that are not explained (e.g. Fig S5B, what are the numbers?) while others state that they refer to statistical significance but I could not find any key. In the latter case, just list the statistical significance on the figure or in the legend please so that each figure can stand on its own.

[R] All figure legends have been revised to include full description of the statistical tests, statistical significance key and alphabetical numbers explanation.

23) The availability statement is incomplete, furthermore I imagine there must be custom scripts for this work that should also be provided.

[R] The availability statement has been completed by adding reference to the Zenodo repository and adding all custom scripts used in the Zenodo repository and in a GitHub repository. An early access link to the zenodo repository is provided.

Referee #3

In this study, the authors employ long-read sequencing to assemble the genomes of over 1000 *Saccharomyces cerevisiae* isolates, with a particular focus on structural variants (SVs) that are typically inaccessible via short-read sequencing approaches. This represents a technical achievement and extends previous large-scale efforts by the same group.
This work should be viewed in the context of two closely related prior studies:
1. The 2018 Nature paper (PMID: 29643504) from the same group, which provided analysis of SNPs and indels across the same set of strains using short-read sequencing, alongside a broad panel of phenotypic data. This foundational study offered a first view of the *S. cerevisiae* pangenome, including the definition of core and accessory genes.

2. A 2023 Nature Genetics paper (PMID: 37524789), involving members from the current team, reported telomere-to-telomere assemblies of 142 strains. That study focused on structural variation, however they integrated these only with genomic data, not with phenotypic traits.
The current manuscript builds directly on both studies, particularly by integrating the long-read-derived dataset with the phenotypic information collected in 2018. This provides valuable insights into the phenotypic relevance of SVs, an area that was previously unexplored at this scale in *S. cerevisiae*.

Comparisons with previous studies
Variant types: while the 2018 study resolved SNPs and Indels and linked them to phenotypes, the current study adds SVs and connects these to phenotypic traits.
Ploidy: Ploidy determination was already possible in the 2018 dataset, though it is more precisely resolve here using long-read assemblies.
SV Catalog: The 2023 study identified 4809 SVs (across 142 strains). The current study, with its broader strain set, identifies 6587 SVs, increasing the median number of SVs per strain from 240 to 289.
Variable regions: Both the 2023 study and the current work conclude that SVs cluster in subtelomeric regions.
Gene content: The number of core and accessory genes increased from 4949 core/2856 accessory ORFs (2018) to 5047 core/3494 accessory genes (current), bringing the total gene count from 7805 to 8541.
Integration with Phenotypes: Unlike the 2023 study, which linked SVs to transcriptomic data only, the present work associates SVs with phenotypic traits recorded in 2018. This is an important contribution to the field.

[R] We agree with the reviewer that the essential advance of this study lies in the comprehensive capture of all major genetic variant types, including structural variants, across a population-scale set of near telomere-to-telomere genomes, and in linking this variation to a broad range of molecular and organismal traits. This integration has not been achieved in any other eukaryotic species to date, making the dataset uniquely comprehensive in terms of both genomic resolution and phenotypic breadth. As such, it goes well beyond yeast and provides a framework for studying genotype-phenotype relationships in other systems. We have clarified this point in the revised Discussion to emphasize the broader impact of the work.

Assessment of novelty and impact
While the study clearly required substantial effort and provides meaningful refinements over previous work, I am not fully convinced that the level of novelty justifies publication in Nature. The main findings represent a valuable extension of existing studies rather than a conceptual breakthrough. The integration of SVs with phenotype is important, but may be better suited for a high-impact specialist journal unless the manuscript is significantly reframed to highlight broader implications.

[R] We thank the reviewer for this perspective. As noted in our response to Reviewer 1, we have revised the Discussion to better articulate the conceptual advances of our work and to highlight its broader relevance beyond yeast. In particular, we show that structural variants (SVs) are disproportionately associated with complex traits and exhibit high pleiotropy. Importantly, our study is the first to integrate population-scale telomere-to-telomere assemblies, a complete SV and gene content catalog, and multilayered phenotypic data across thousands of traits in any eukaryote. This uniquely comprehensive framework enables resolution of trait architecture at an unprecedented scale and serves as a model for future genotype-phenotype studies in other systems. We hope that the revised framing more clearly conveys the novelty and broad impact of the work.

Clarity and Readability
The manuscript is currently difficult to read. Abbreviations are often undefined, and the structure lacks clarity in places. Compared to the 2018 study, the presentation is less accessible to a broad audience. If the authors intend to reach a general Nature readership, substantial improvements in the writing and presentation are required.

[R] We agree with the reviewer that some abbreviations and terms were not clearly defined, a point also raised by two other reviewers. In response, we have carefully revised the manuscript to ensure that all abbreviations and key terms are defined upon first use and have tried to improve the overall clarity. We have also added supplementary notes.

Summary

This study provides a comprehensive analysis of long-read assemblies of >1,000 *S. cerevisiae* isolates and, for the first time at this scale, links SVs to phenotypic traits. While technically impressive and a valuable resource for the field, the conceptual advances appear incremental relative to earlier work. Moreover, the manuscript's readability must be improved if it is to be considered for publication in Nature.

[R] We have responded to these comments above.

Reviewer comments

Referee #1:

The revised manuscript has explicitly addressed my initial concerns and recommendations. The only remaining minor suggestions that I have are as follows:

1. The authors note that the graph pan genome limitations include reciprocal translocations. This is mentioned in the Methods and Materials section but should also be included somewhere in the Discussion section.

[R] In fact, the limitations of graph pangenomes (including translocations) are already mentioned in the last paragraph of the Results section, not in the Methods and Materials section. Therefore, we think that it would be redundant to repeat this part in the Discussion section.

2. In the results section, please replace "... organismal traits (growth trait)..." with "...colony level growth traits (used here as organismal trait proxies)..." (line 232)

[R] The sentence "… colony growth traits (used here as organismal traits proxies) …" was added in the revised manuscript.

Referee #2:

The authors have done a good job adding substantial detail, clarifications, and revised analysis and have addressed my main comments. The methods in particular are now much more detailed. A few minor changes to the final draft could include:

Lines 175-177: "Wild clades, particularly the Chinese wild group, the species' ancestral population, deviates from this correlation, with fewer SVs than expected based on the number of SNPs, or conversely (Fig. 2G)." "or conversely" is not clear. Perhaps the authors mean fewer SNPs compared to the rest of the distribution.

[R] The sentence has been revised to clarify the meaning of "conversely":
"Wild clades, particularly the Chinese wild group, the species' ancestral population, deviates from this correlation, with fewer SVs than expected based on the number of SNPs, or more SNPs than expected based on the number of SVs".

The authors nicely answered some of my questions in the response letter, but since other readers may have the same questions it may be worth adding those responses briefly For example:

"[R] The bimodal distribution of PAVs length (Fig S6) reflects the presence of SVs associated with Ty elements (~6 kb) and solo long terminal repeats (LTR, ~300 bp)." You could add this to the supplemental legend.

[R] The sentence has been added to the legend of Fig. S6G.

"[R] We believe that the reviewer is referring Fig. 3C and Fig. S12A. The hump observed here is not associated with centromeres but originates from a large introgressed region of ~164 kb found in chromosome 7 of one Mexican agave isolate. This region harbors 65 private genes, leading to a localized enrichment in gene content in the pangenome, that appears as a peak on Fig. 3C" If space permits, the legend could add that a large introgression event in one strain (##) produces a signature at Xbp. That can be at the authors' discretion.

[R] A sentence has been added to the legend of Figure 3C.

Referee #3:

The authors have substantially enhanced the clarity of the manuscript. In addition, they have rewritten the discussion, providing a much stronger rationale for the conceptual advances of this study compared to previous work.

[R] We thank the reviewer for the positive feedback