

Analysis of 173,303 exomes and genomes in the Pakistan Genome Resource

Corresponding Author: Dr Allan Gurtan

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.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

This paper presents exome or whole genome sequencing data from a total of ~175K individuals which allows them to identify human gene knockouts in about 34% of the protein coding genome. Having information about the human phenotypic effect of gene knockouts is a critical part of understanding gene function and in this paper the researchers have greatly expanded the landscape of this type of data. My comments are focused on improving the manuscript. The results particularly to inform clinical trials are incredible and I believe that the paper will be a landmark in the field and the knowledge of human phenotypic effects of knocking out 30% of the genes in the genome cannot be understated.

1) The population genetic analysis as they are presented provide little insight beyond the fact that these individuals are of South Asian ancestry (Fig 1 and 2). Would they be able to provide a PCA of just the individuals in their study, perhaps projected onto a West and East Eurasian basis? These were previously used to describe the Indian clinal populations first described in Reich et al, Nature, 2009 and provide a better view of the admixture history of these populations – you can use the references in Narasimhan Science 2019 to build these. In addition for fig 3a, would it be possible to separate the 1000 genomes SAS samples, to PJL to actually show the differential with the relevant 1000 genomes sub population?

2) What is interesting about the population genetics is not just the diversity of sampling and the ROH lengths but rather the how adding individuals of different ancestries within the PGR contribute to the overall number of knockouts seen. Could the authors report metrics like this sub-divided into different ancestries? For example, after adding several tens of thousands of genomes of Punjabi ancestry, how many new knockouts are seen in an additional genome of Kalash ancestry? This would help understanding how rare variation is distributed across subpopulations and how to prioritize future sampling efforts for this project as well as others, such as Genes and Health, and the UAE genome project

(<https://www.medrxiv.org/content/10.1101/2025.09.12.25334546v1>).

3) There are a number of individuals with extremely high autozygosity, almost 40% - is this real? I wonder if these are from the whole genome data, which need additional filters if running bcftools roh, as the error rates on the lower coverage sequencing maybe higher and this causes my code in bcftools roh to create "blips".

4) The association between Froh and reduced levels of MI but increased levels of TB and chronic kidney disease etc are interesting. But one issue is that Froh might be confounded with socioeconomic status or other cultural factors. Is there a way that they could split their populations into individual ancestry groups, Kashmiri, Punjabi etc. and rerun this analysis (hopefully they still have power with this enormous sample size)? As the relationship between Froh and environmental effects may be different across these groups this might provide additional evidence of this association (It is also possible that the association between Froh and environment maybe the same in all these groups).

5) In the suppression analysis, looking at the number of autozygous and non-autozygous regions and the burden of LoFs in them, are they restricting to LoFs that do not contain indels? This would make the comparison with missense and synonymous variants more "fair", as they cannot be generated by indels, though it is equally likely that an indel could fall into an autozygous and non-autozygous region.

6) Line 292 says that AAR can be used to identify genes subject to constraint. Are they able to show constraint at the level of individual genes with data of this size, or do they mean gene sets? If its gene sets they should modify the statement. If its individual genes, then this is indeed a powerful resource and they should make the list available publicly as this is currently now unavailable from datasets such as Gnomad.

7) Outside of AAR, there is an alternate way of computing suppressing involving subsampling of alleles at a particular

frequency which we used along with AAR in the Narasimhan Science 2016 paper. While not guaranteed that these two methods should arrive at the exact same point estimate of suppression – they should be in the same ballpark (particularly if using only SNPs and not indels). Can they also show this alternate approach for the many suppression analysis they carry out?

8) If they combine their data with the genes and health study and Gnomad or finnGen, how many total genes do they see knocked out recessively in humans? Rather than just restricting their results to what they observe in the paper it would be useful to know the total number of genes that could be tolerated with having knockouts.

This set of analysis involves follow up of individual genes not functionally, but just informatively using the data that they already should have or could get easily:

9) I saw that there are 4 PRDM9 knockouts. Are any of them male? If so, do they have children (regardless of male or female)? This is a simple query but would add a lot of value to the evolutionary genomics community, as this is one of the few known mammalian speciation genes, and also a gene for which the knockout in humans is different from expectation from mouse models and a knockout male individual with children would suggest major differences in how recombination works in humans and mice.

10) I see a huge number of individuals with knockouts in TLR genes – these have been identified to be under recent natural selection utilizing time series data from ancient DNA. The hypothesis is that these genes confer protection against pathogens and I wondered if these individuals have increased risk of tuberculosis or other disorders – is there power to look at this?

11) What is the eye color of the individual with a OCA2 knockout?

12) What is the skin pigmentation of individuals with an SLC24A5 knockout?

13) What is the BMI / glucose activity of the individuals with an OMA1 knockout? Mouse models

<https://pmc.ncbi.nlm.nih.gov/articles/PMC3343468/> indicate major changes to metabolic activity and body weight. Similarly do these individual also have low(er) mtDNA copy number?

14) For individuals with homozygous knockouts in the HLA, are there any indications of increased chronic infection?

Transcriptional adaptation

Recently meeting with Didier Stainier, I learned about a particular genetic mechanism that could explain the presence of some of these knockouts in humans – transcriptional adaptation (<https://www.nature.com/articles/s41586-024-08539-x#ref-CR3>). The idea is that mutations that do not completely ablate the protein but instead that cause NMD, actually can upregulate other paralogous proteins that may recover function. Thus for certain genes homozygous LoFs are better tolerated than homozygous missense. Is there a way that the authors could look at this phenomenon perhaps by computing ARR using missense vs lof in a gene by gene level?

Referee #2

(Remarks to the Author)

Thank you for the opportunity to review this paper – it is clearly a study of monumental scale with potential to further our understanding of autozygosity and new gene effects through population-specific and enriched LoF variants.

My co-reviewer and I have reviewed the manuscript and have the following major concerns as regards the content, synthesis and therefore conclusions.

1. In the abstract and throughout – what is the relevance of comparing variants to non-south Asians in gnomad? Given this is a deep variant-finding opportunity wouldn't it be more meaningful to include south Asians in gnomad (or perhaps present both)?

2. Given that 2 individuals contributed to the finding of LRRK2 KOs and that there was no deeper phenotyping of these individuals, the conclusion that KOs need renal monitoring is premature and conclusions should be more circumspect.

3. 115: "compendium of clinical data, including medical history, clinical measurements, and disease status for most participants"; this is a little vague; it would be useful to explicitly state what data is available and for whom. Typically this is contained in Table 1, describing the resource.

a. On reading the methods it's quite unclear how exactly a participant comes to be enrolled in any of these nested case/control cohorts, and for the international reader, how this interacts with provision of healthcare in Pakistan. Would participants accessing private healthcare be more or less likely to be enrolled? Urban vs rural? It would be useful in the discussion to comment on these recruitment factors and potential sources of bias in case/control ascertainment.

4. Results: A flow chart of which participants are included in which part of each analysis would be very helpful to orientate the reader (this feeds into comments above) – it is essential to orient the reader to what PROMIS-PGR contains in terms of recruitment, genetic data generation (different sources and technologies), omic data similarly and finally what clinical data is collected and contributes to the compendium described.

5. In the population genetics portion of the manuscript, it is not clear why self-reported ethnicity was used for FST as opposed to the just-projected PCs? Was it not possible to project the sub-populations within PGR with sufficient granularity? This could be expanded upon for clarity. Is it possible to clarify for how many individuals this was done?

a. I'm afraid as a non-expert in pop gen, the conclusions and take-away message from this portion of the manuscript were not apparent. Can this be better fleshed out and more explicit?

6. A major concern are the findings around Autozygosity

a. 165: Perhaps worth explaining why you tested per 0.0625, i.e. additional degree of relatedness from parents being 1st cousins – makes intuitive sense when elaborated upon but not much without, to the non-expert reader

b. 170-180: These are extremely interesting results, because they are at odds with existing literature e.g. Malawsky et al. in Cell 2023 who find quite the opposite associations with CVD, T2D etc. when using British Bangladeshis, British Pakistanis, UK Biobank British south Asians more generally, Europeans and finally in 23 & me.

<https://doi.org/10.1016/j.cell.2023.08.028> In this paper, the authors perform a sensitivity analysis in a subset of highly-consanguineous individuals, reasoning that cultural or environmental factors are less likely to vary within this stratum, and

that associations between increased FROH and traits restricted to this stratum are therefore more likely mediated by genetic (as opposed to environmental) effects. I wonder if a similar approach here might help to convince the reader further?

c. Are there some groups (defined by ethnicity or PCA) among whom FROH is higher? Are autozygosity analyses of trait associations adjusted for self-reported ethnicity or PCs? If not, this needs to be justified.

d. If rates of FROH and outcomes (eg T2D / MI) do vary between these ethnic groups, then demonstrating the association between T2D and FROH (for example) within each ancestry separately would help convince the reader that these associations are not driven by differences between these groups.

e. A major confounding factor is survivorship bias and competing risks. How are cases ascertained, and controls, and does this occur longitudinally? The risk of survivorship bias and how this has been handled needs to be addressed. Was there a depletion of individuals with very high FROH score before recruitment was even possible? As the current state of knowledge suggests that high autozygosity is linked to myriad developmental and health conditions, were individuals with the highest FROH able to participate, live long enough to be recruited? Competing risks analyses may also be necessary to detangle this.

f. 190-200: Again these are extremely interesting results. Why was age 40 chosen as a cutoff? This is a typical young-onset threshold for T2D, but is perhaps a little early for MI. It would be more interesting to see these effect sizes plotted out by 10-year age bin (ie 18-30, 30-40, etc) to see how the association varies across the full breadth of ages (accepting power will be limited here), as opposed to a binary $</>40$ bin which leaves the reader wondering where the effect arises from. These divergent effects across the age range are very surprising and a bit more granularity would be most interesting to the reader.

g. Is there some sort of artificial clumping driving this association? Is FROH associated with age in some way, i.e., is the population becoming more autozygous over time, more FROH in young people, who do not yet have T2D/MI etc?

h. In previous studies ROH has been called in imputed whole genotype data – it seems here ROH is called just in coding regions? This needs to be explicitly discussed and justified.

i. For these autozygosity analyses, did effects differ between males and females? It would be nice to see these effects plotted out sex-stratified

j. How are these outcomes defined? Is there ongoing surveying, or do participants have to access healthcare in order to receive a downstream diagnosis? If so, it seems autozygosity (associated with less education) may be associated with limited access to healthcare, wealth, etc – and although models are adjusted for this, perhaps these people simply do not obtain their dx of MI, and the association simply reflects healthcare engagement?

k. Finally, the discussion section is lacking depth as to why these findings diverge so entirely with previous studies. There is not the uncertainty on the effects of autozygosity as the authors suggest. There is in fact a consensus on the increased risk of ill-health. So the protective effects observed do require considered sensitivity testing and perhaps even external validation.

7. Biomarker burden

a. Again a flow-chart showing the extent of clinical data, proteomic data, etc, would be really helpful to orientate the reader.

b. 225: These are extremely interesting and well-presented results. To orientate the reader, it would be useful to have a short section explaining how many GENE*phenotype associations were tested, and what the associated Bonferroni-corrected p value would therefore be (for example, the “lower random glucose ($\beta=-0.05$, $P=2.3e-02$)” seems unlikely to be anything other than nominally significant, but it is not clear to the reader which other association would survive multiple testing correction.

c. It would also be useful to provide a summary table of how many individuals had available data for each phenotype discussed in the main text (this feeds back into a previous comment on study data presentation and flow charts)

d. It would be helpful in this section to compare to existing pLoF burden results identified either through eg AZphewas portal or GENEASS to orientate the reader to the novelty of these findings, as they are presented.

e. 229: It would be useful to compare to G&H recent effort characterising pLoF human KOs, recent preprint:

<https://doi.org/10.1101/2025.06.05.25329068>. For example the Genes and Health authors report 2991 homozygous carriers of pLoF in a sample size of ~44k individuals and this could be compared to findings from PROMIS-PGR for a better understanding of the landscape of genomic studies in SAS populations and sub-populations.

f. 217: Some of these associations with HBB, including iron and lipid traits, in addition to bilirubin, are reported in the Kim et al paper above and should be discussed in this context.

8. Analysis of LOF constraint

a. 261: it would be helpful to briefly define where these approved drugs are from (FDA? Some other authority?) to save the reader diving to the methods

b. 285-287: Are these AAFs significant? P values are missing, and no SEs are presented.

c. 297: Were synonymous variants “modestly depleted” if $p = 0.225$ (line 299)? And doesn't no depletion of synonymous variants in fact strengthen your argument, ie, as a negative control?

9. Human KO highlights

a. For recall samples, was sequencing repeated to confirm the initial sequencing/genotyping was correct?

b. 310: The LRRK2 findings are certainly most interesting, but I think the strength of assertion around need for renal monitoring should be couched firmly in the $N=2$ here, and perhaps more detail and deeper phenotyping on these 2. To my mind this argument would be considerably strengthened by checking for known genetic causes of CKD, eg PCKD, and perhaps calculating a CKD PGS. Ie, in the absence of other known genetic constructs causing CKD, we may be more inclined to believe it is really all mediated by LRRK2. Is further phenotyping possible (eg ACR, USS)? We appreciate this may not be feasible and are not suggesting this would be required for publication. However, the fullest picture of these individuals phenotypes would be needed to persuade readers of this journal that LRRK2 was the underlying cause here.

c. 326: Similarly, some more phenotypic information on the human KOs for SLC2A4 – if available – would be very interesting to the reader. Given they are first reported human KOs for this gene, there is likely to be enormous interest in knowing more about glycaemic measures in these individuals. The T2D knowledge portal reports findings with thyroid function at the gene-level, rare variants with ApoA1 and blood pressure. The authors have a unique opportunity to explore and characterise these rare individuals.

d. The findings in DENND1B KOs are interesting, but at odds with published summary stats from UKBiobank. How do

authors reconcile this?

- i. 336: It would be helpful to present some measure of range here (is this driven by a single BMI of 45?); or perhaps even all 4 BMIs, given the small number.
- ii. Given the novel finding, which is at odds with the AstraZeneca phewas portal findings for LoF DNND1B, it would be valuable to undertake replication with other data sets, UK Biobank, Genes & Health to confirm this findings.
- e. 338: Associations with RXFP1 have been demonstrated in UK Biobank, data available on AZphewas portal. The novelty here is unclear – but would benefit from highlighting and meta-analyses with UKB.
- f. 379: TRPM8 – bigger previously reported signals have been for lipid traits and MI, did this not replicate here?
- g. All these findings would be more impactful, if where possible, external validation or replication was undertaken. Coherent and directionally stable associations with ptns in UK Biobank, Genes & Health which is also likely enriched for LoF variants (albeit in a smaller sample size), PRECISE etc.
- h. Similar to the previous section, it would be helpful to know how many GENE*trait associations (*model type (additive / recessive)) etc were performed across this section, and an associated Bonferroni-corrected p value presented to the reader. Presumably not just the genes included in the vignettes were included in analysis, but it is difficult to get a sense of what a meaningful p value is without knowing the number of tests performed.

10. Discussion

- a. 391: as above, the actual median FROH is not within the range you describe
- b. 409: You describe the inconsistencies “Across previous studies”, but actually other studies that you describe were all directionally consistent with autozygosity being associated with deleterious phenotype (consistent with inbreeding depression).
- c. 411: It is unconvincing that the thrifty genotype hypothesis explains why these (age-stratified) results are at odds with the Genes & Health results in Malawsky et al – I think more fleshing out of the proposed mechanism is needed here to convince the reader.
- d. 416: I'm a little unclear on why the HBB-lipids associations have been determined to be causal, but other HBB-trait associations (e.g. MI) are not described as causal. Perhaps I have missed some key analysis step here? The HBB associations with iron, red cell biology and MI are strong and convincing, but I do not follow how you determine them to be causal?
- e. As below, it would be useful to include some comment on participant recruitment methods, and potential sources of bias associated with this.

11. Methods

- a. 509-518: Did you set a specific minimum length for a ROH eg 1Mb? How was a ROH defined? I'm a little unclear after reading this section.
- b. 610: Your FROH distribution looks skewed (as would be expected) – did you log-transform it prior to logistic regression, and if so why/why not?

Minor comments ;

1. Intro 98: The first line feels a little at odds with the rest of the paper in setting the scene – much of this paper has no relevance to drug discovery?
2. After reaching the end of the introduction and starting results I was a little unclear on what exactly the aims / objectives of this study are – perhaps a brief sentence to summarise might help to orientate the reader ahead of the analyses being presented in the results
3. 103: I think for the non-expert reader it is necessary to expand on why autozygous populations are enriched for recessive / pLoF KO effects when explaining they can “uniquely power genetic discovery” ie spell out that same allele from common ancestor can arrive via each parent / IBD. A bit more detail on autozygosity for the non-expert reader, not necessarily obvious why they are more likely to have 2 copies, perhaps some sort of diagram and explanation of identity-by-descent
4. 160: A minor but concerning observation, but the median FROH value of 0.015 doesn't fall within your stated ranges 0.016 – 0.0625 as you say it does.

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

Referee #4

(Remarks to the Author)

Koch et al present a nice study of 173,303 individuals in Pakistan. They present interesting findings relevant for biomedicine using analyses of autozygosity, LOH variants and human knockouts, and the analyses are generally well done and innovative and the paper is well written.

I have few major concerns:

- For the FROH analysis, Malwasky et al 2023 clearly explain why such analyses can be confounded, and show that restricting analysis to offspring of first cousins is an effective way of reducing confounding due to social/environmental correlates of FROH. Why has no effort to do this been done? The authors should repeat this analysis using such an approach and including covariates for socioenvironmental variables that they have available. They also do not make any mention of multiple testing correction in this analysis and it is not clear if this has been done.
- The authors state that “Causal relationships were established between TIMD4 pLoF and circulating lipids, and between HBB and circulating lipids and apolipoproteins.” It's not clear why causal is the right word here and they should edit, they

have simply shown associations

- For this to be a resource, the availability of the dataset for researchers should be made clear, and it is not clear how researchers can access the dataset for other studies. This is a major concern. The authors state "Whole Genome Sequencing data is available for non-profit use at phs001569.v1.p1." but this only leads to the PROMIS dataset which is a tiny fraction of the data presented and analyzed. They also state that "All other academic requests to access relevant Center for Non-Communicable Diseases (CNCD) data or samples should be sent to info@cncdpk.com." How will it be ensured that the data is truly available to interested researchers and fair ethical practices will be applied for those requesting access?
- What plans do the authors have for community engagement and giving results back to participants?

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

Review by Vagheesh M. Narasimhan (I sign all my reviews).

My last concern is that the current population genetic figures (1, 2, and 3) present quite a lot of detail making them difficult to interpret. I would combine the relevant information into a single figure, highlight where the samples are from and use the same color coding on the PCA, admixture (the admixture plot doesn't separate the ancestries either) and location plots. I would also include figure 4a in this plot. and again follow the coloring with the other plots above. Otherwise the reviewers have addressed all my concerns - this is a remarkable paper and the end of an almost 10 year effort to understand the impact of gene knockouts on 1/3 of the human genome.

If at all possible performing sperm typing on the male prdm9 knockout would be a remarkable addition to the evolutionary genomics community, because we would understand how localization of recombination happens in humans in the absence of prdm9. The process is relatively inexpensive but I'm not sure if the individual will consent for additional sampling.

Referee #2

(Remarks to the Author)

Referee #2 (Remarks to the Author) - all new comments are prefaced by "Review 2" but the thread of comments is maintained and included to prevent confusion or contradictions in the course of review.

Thank you for the opportunity to review this paper – it is clearly a study of monumental scale with potential to further our understanding of autozygosity and new gene effects through population-specific and enriched LoF variants.

- We thank the reviewers for their positive overall assessment of the manuscript and provide answers to their questions below.

We thank the authors for another opportunity to review this excellent manuscript, which has been substantially improved during these revisions.

My co-reviewer and I have reviewed the manuscript and have the following major concerns as regards the content, synthesis and therefore conclusions.

2.1 In the abstract and throughout – what is the relevance of comparing variants to non-south Asians in gnomAD? Given this is a deep variant-finding opportunity wouldn't it be more meaningful to include south Asians in gnomAD (or perhaps present both)?

- We thank the reviewer for the suggestion. A focus on non-South Asians illustrates discovery power in South Asia. To address the reviewer's suggestion, which increases the significance and value of the comparisons, we have now revised Table 1 and Figure 6 to include South Asians in gnomAD.

Review2: We agree this improves the manuscript.

2.2 Given that 2 individuals contributed to the finding of LRRK2 KOs and that there was no deeper phenotyping of these individuals, the conclusion that KOs need renal monitoring is premature and conclusions should be more circumspect.

- We agree with the reviewer about the limitations of small numbers. To strengthen the observations from the two knockouts, we discovered a third independent KO through recontact and sequencing family members of an unrelated heterozygous loss-of-function variant carrier. Similarly to the two knockouts from the PGR biobank, the third knockout had low eGFR. As the reviewer rightfully suggested, we evaluated the exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, and APOL1). No damaging missense (REVEL score >0.7) or HC pLoF variants were found.

- We believe that our human data stands as a capstone to an extensive series of pre-clinical reports implicating LRRK2 in kidney function:

- o LRRK2 knockout mice exhibited impaired kidney function and hypertension (Tong et al., 2010; Herzig et al., 2011).
- o LRRK2 knockout rats exhibited renal dysfunction (Gu et al., 2023).
- o LRRK2 inhibitors impaired kidney function in rodents (Andersen et al., 2018).
- o LRRK2 inhibitors caused kidney pathology in non-human primates (Jennings et al., 2022).
- Our revised manuscript highlights these points (lines 349-359):

o Recall-by-genotype of LRRK2 pLoF carriers and consenting family members identified a 49-year-old female homozygous for LRRK2 splice donor variant, c.1418+1G>T (Extended Data Figure 10a). This individual was the sole identified KO in her family and had an eGFR of 51 ml/min, suggestive of kidney disease. Her eGFR was lower than all 13 of her recruited heterozygous and non-carrier family members (Extended Data Figure 10b). We evaluated the available exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, or APOL1) and found no damaging (REVEL score >0.7) or HC pLoF variants. Although the number of LRRK2 knockouts is low in our study, the observation of kidney dysfunction in these individuals is consistent with, and provides human data to support, the adverse renal phenotypes observed in rodent LRRK2 knockouts⁵¹⁻⁵³ and the preclinical renal findings of LRRK2 inhibitor administration in rodents⁵⁴ and non-human primates⁵⁵.

Review2:: We agree with the authors that the inclusion of an additional individual is noteworthy, and agree that the absence of pathogenic variants in known genes increases the believability of the findings. We also agree that the revised framing of the results, acknowledging the small sample size, is in keeping with the scientific robustness that would be expected from a manuscript submitted to this journal.

2.3.1 115: “compendium of clinical data, including medical history, clinical measurements, and disease status for most participants”; this is a little vague; it would be useful to explicitly state what data is available and for whom. Typically this is contained in Table 1, describing the resource.

• We agree with the reviewer. To address the reviewer’s point and fully describe the resource, we now list the specific clinical data that was available and analyzed in PGR. As suggested by the reviewer and for completeness, we fit the information into the newly expanded and revised Supplementary Table 1.

Review2: This is most helpful to orientate the reader and adds to the readability of the paper.

2.3.2 On reading the methods it’s quite unclear how exactly a participant comes to be enrolled in any of these nested case/control cohorts, and for the international reader, how this interacts with provision of healthcare in Pakistan. Would participants accessing private healthcare be more or less likely to be enrolled? Urban vs rural? It would be useful in the discussion to comment on these recruitment factors and potential sources of bias in case/control ascertainment.

• We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514):

• Recruitment. Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.

Review2: We agree this information helps orientate the reader in understanding how participants came to be recruited in the study.

2.4 Results: A flow chart of which participants are included in which part of each analysis would be very helpful to orientate the reader (this feeds into comments above) – it is essential to orient the reader to what PROMIS-PGR contains in terms of recruitment, genetic data generation (different sources and technologies), omic data similarly and finally what clinical data is collected and contributes to the compendium described.

• We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now additionally includes a list of all biomarkers and binary outcomes analyzed in this study, along with the number of individuals for which the data is available. Additionally, for each analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of which individuals were included.

Review2:: This is very helpful to orientate the reader. Please include the WES technology used and also list what other sources of genomic data exist other than WES – as the total number of PGR samples that contributed to coding variant counts was 173,303 whereas WES is available is 166,625. This can be added to the “Notes” section of the table.

2.5.1 In the population genetics portion of the manuscript, it is not clear why self-reported ethnicity was used for FST as opposed to the just-projected PCs? Was it not possible to project the sub-populations within PGR with sufficient granularity? This could be expanded upon for clarity. Is it possible to clarify for how many individuals this was done?

• We agree with the reviewer that self-reported ethnicity has limitations. We prepared a new Figure 3 to visualize clustering of South Asian populations. Individuals of the same ethnicity tend to be close to each other in PCA. This observation indicates that ethnicity was an appropriate input to the FST analysis, which included all participants.

Review2:: We agree with the authors that these changes help strengthen the paper. Can the authors explain what they mean by “ethnicity” because in PGR there is a group clubbed as “Urdu” which is a language rather than ethnic group.

2.5.2 I’m afraid as a non-expert in pop gen, the conclusions and take-away message from this portion of the manuscript were not apparent. Can this be better fleshed out and more explicit?

• We thank the reviewer for suggesting ways to make this section more insightful. To address the reviewer's point, we re-ordered the population analysis, which now presents Fst analysis, admixture, and PCA, for a clearer, stepwise presentation of South Asian ancestries and their genetic make-up. We also provided a PCA of the individuals in our study alongside South Asian 1000 Genomes (1KG-SAS) that we separated into its constituent ethnicities, as recommended by the reviewer. As the reviewer predicted, the new analysis projected PGR ethnicities along a geographic gradient of proximity to Europeans (PC1, Figure 3c,d). At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were the western populations of Pakistan, such as the Pathan and Balochi of PGR. The revised text includes the following information (lines 159-170):

o Principal Component Analysis (PCA) quantified the heterogeneity of PGR subpopulations relative to each other and to non-PGR populations from the 1KG (Figure 3). African, East Asian, European, and South Asian ancestries separated into clusters, with South Asians in PGR following a gradient along PC2 from European ancestries to East Asians (Figure 3a,b). Consistent with the admixture analyses, Pathan and Balochi ethnicities clustered more closely with Europeans than did other PGR ethnicities, while a subset of Balochis clustered more closely with Africans. Among individuals without recent African admixture, PC1 distinguished PGR ethnicities along a gradient of proximity to Europeans, consistent with previous observations (Figure 3c,d)^{19,20}. At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KGSAS, and at the other end were populations of western Pakistan, such as the Pathan and Balochi of PGR. PC2 and PC3 identified two additional clusters, one of which was predominantly Balochi and the other predominantly Punjabi, consistent with previous observations of substructure among the latter population²¹.

Review2:: We agree this adds some clarity to results for the non-expert reader. Can authors please differentiate between their use of ethnicity and genetically determined ancestry as these are not the same and, in many cases, it appears they are referencing the latter, certainly in Fig 2 and 3 which are PCA-reliant.

2.6 A major concern are the findings around Autozygosity

2.6.1 165: Perhaps worth explaining why you tested per 0.0625, i.e. additional degree of relatedness from parents being 1st cousins – makes intuitive sense when elaborated upon but not much without, to the non-expert reader

• We report in units of 0.0625 as this is the theoretical expectation of FROH for the offspring of first-cousin parents. To clarify, we have now added this point to the main text (lines 179182):

• We tested association of autozygosity (per FROH of 0.0625, the theoretical expectation of FROH for the offspring of first-cousin parents) with 16 quantitative and 18 binary phenotypes in PGR (Figure 5, Supplementary Table 3 and Supplementary Table 4).

Review2:: Thank you.

2.6.2 170-180: These are extremely interesting results, because they are at odds with existing literature e.g. Malawsky et al. in Cell 2023 who find quite the opposite associations with CVD, T2D etc. when using British Bangladeshi, British Pakistanis, UK Biobank British south Asians more generally, Europeans and finally in 23 & me. <https://doi.org/10.1016/j.cell.2023.08.028> In this paper, the authors perform a sensitivity analysis in a subset of highly-consanguineous individuals, reasoning that cultural or environmental factors are less likely to vary within this stratum, and that associations between increased FROH and traits restricted to this stratum are therefore more likely mediated by genetic (as opposed to environmental) effects. I wonder if a similar approach here might help to convince the reader further?

• Based on the reviewer's suggestion, and similar to Malawsky et al al, we have performed an additional subsetting analysis. We have added the following to the text (lines 214-219): o Subsetting to individuals with high autozygosity who were self-reported to have first-cousin parents eliminated the association between autozygosity and educational attainment, indicating that education as a trait may be confounded by social or economic factors (Extended Figure 7). For these individuals, however, the other associations were unchanged, suggesting that the observed effects are driven by autozygosity rather than socioeconomic confounders.

Review2:: This analysis is valuable and helps convince the reader that the basis of the FROH association may be genetic.

2.6.3 Are there some groups (defined by ethnicity or PCA) among whom FROH is higher? Are autozygosity analyses of trait associations adjusted for self-reported ethnicity or PCs? If not, this needs to be justified.

• All FROH association analyses are adjusted for genetic principal components. We now include analyses subset by ethnicity (Extended Data Figure 5) and have observed minimal heterogeneity in effect sizes across ethnicity. These results indicate that the associations are not driven by differences between groups.

Review2:: These additional analyses are very helpful and help convince the reader that some difference between groups defined by self-defined ethnicity (which may not have been adequately captured by PC adjustment) are not underlying the observed associations.

2.6.4 If rates of FROH and outcomes (eg T2D / MI) do vary between these ethnic groups, then demonstrating the association between T2D and FROH (for example) within each ancestry separately would help convince the reader that these associations are not driven by differences between these groups.

• All FROH association analyses are adjusted for genetic principal components. We now include analyses subset by ethnicity (Extended Data Figure 5) and have observed minimal heterogeneity in effect sizes across ethnicity. These results indicate that the associations are not driven by differences between groups.

Review2:: These results are most reassuring, and even more convincing than self-reported ethnicity-adjusted models in all participants.

2.6.5 A major confounding factor is survivorship bias and competing risks. How are cases ascertained, and controls, and does this occur longitudinally? The risk of survivorship bias and how this has been handled needs to be addressed. Was there a depletion of individuals with very high FROH score before recruitment was even possible? As the current state of

knowledge suggests that high autozygosity is linked to myriad developmental and health conditions, were individuals with the highest FROH able to participate, live long enough to be recruited? Competing risks analyses may also be necessary to detangle this.

- We agree with the reviewer that competing risks and survivorship bias analysis would have benefited this study. We currently do not have the longitudinal data to perform these analyses.

Review2:: We agree with the authors that these biases are critical to understanding the results observed in the study. We would suggest the authors would be well-served highlighting this to a greater extent in the discussion section as a critical limitation of their analyses.

Supplemental Table 4 is an excellent presentation of these data. Apologies if I have not been able to spot this, but is there a histogram that plots FROH (on Y axis) against age (on X)? There appears to be an age-based stratification of the effect of FROH on hypertension, CKD, and T2D and perhaps others. Wherein, if you were to meta-analyse the effects of FROH across these age-groups you would not find agreement in effect. Indeed this is captured in the p value for heterogeneity across groups. It could be, that individuals recruited at older ages have lower genome-wide/exome-wide runs of homozygosity, hence the question about distribution of FROH across age groups. There are also ~ 32,200 individuals recruited < 40 years whereas over 130k are over the age of 40 at recruitment.

The impact / heterogeneity across age groups makes it hard digest the results from the FROH analysis without a degree of scepticism. For example, if we were to accept the overall finding that higher FROH is protective from T2D, MI, what is the biological basis of this?

2.6.6 190-200: Again these are extremely interesting results. Why was age 40 chosen as a cutoff? This is a typical young-onset threshold for T2D, but is perhaps a little early for MI. It would be more interesting to see these effect sizes plotted out by 10-year age bin (ie 18-30, 30-40, etc) to see how the association varies across the full breadth of ages (accepting power will be limited here), as opposed to a binary </>40 bin which leaves the reader wondering where the effect arises from. These divergent effects across the age range are very surprising and a bit more granularity would be most interesting to the reader.

- We agree with the reviewer that the cutoff of 40 was arbitrary. We have now stratified our analyses into multiple age bins and describe the results in the revised text (lines 219-224): o To determine if the associations were age-dependent, we binned across age groups (Extended Figure 8). Significant heterogeneity of effect was observed for educational attainment ($P=3.4e-08$), number of children ($P=3.2e-05$), having had children ($P=6.2e-06$), diabetes ($P=9.0e-06$), and hypertension ($P=1.9e-08$), with autozygosity trending towards increased risk of hypertension and diabetes in individuals <40 years and protection from hypertension and diabetes in individuals 50 years old (Extended Figure 8).

This approach is much more informative to the reader and strengthens the manuscript

Review2:: agree. Thank you for making this table. See previous comment regarding acknowledging the lack of longitudinal data, recruitment of as many younger individuals <40 compared to >40 and >50, and heterogeneity of effects across age groups as a critical limitation of this analysis and generalisability of this finding.

2.6.7 Is there some sort of artificial clumping driving this association? Is FROH associated with age in some way, i.e., is the population becoming more autozygous over time, more FROH in young people, who do not yet have T2D/MI etc?

- We see less autozygosity with age in PGR ($P=8.4e-286$), albeit with a small effect size ($\beta=0.00026$ FROH per year of age). In Figure 1 of Malawsky et al., the authors report a similar magnitude of effect between age and autozygosity in British Pakistanis, but not in British Bangladeshis, arguing that survivorship bias or ascertainment bias are unlikely to explain the association between age and autozygosity. Additionally, we have included both age and age*age as covariates in all association analyses.

Review2:: We agree that through the lens the authors describe, survivorship bias does not necessarily explain these findings. But the age-based heterogeneity of effects is hard to ignore. There appears to be an age-independent but somehow intertwined artefact. Do those who survive to age 50, 60 and higher have runs of homozygosity around different loci that are protective compared to those who are younger and indeed theoretically, since you cannot measure this –in those who did not survive to participate in the study?

2.6.8 In previous studies ROH has been called in imputed whole genotype data – it seems here ROH is called just in coding regions? This needs to be explicitly discussed and justified.

- We used WES data for ROH calling to maximize our sample size. The original paper describing the BCFtools ROH algorithm demonstrated that the algorithm produces highly accurate results when applied to exome data (Figure S5 in Narasimhan et al. 2016) and is now commonly used in the field.

Review2:: Thank you for the clarification

2.6.9 For these autozygosity analyses, did effects differ between males and females? It would be nice to see these effects plotted out sex-stratified

- We now include analyses subset by sex and have observed minimal heterogeneity in effect sizes (Extended Data Figure 6).

Review2:: These results are valuable in confirming no sex heterogeneity of effects.

2.6.10 How are these outcomes defined? Is there ongoing surveying, or do participants have to access healthcare in order to receive a downstream diagnosis? If so, it seems autozygosity (associated with less education) may be associated with limited access to healthcare, wealth, etc – and although models are adjusted for this, perhaps these people simply do not obtain their dx of MI, and the association simply reflects healthcare engagement?

• We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514): o Recruitment. Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.

Review2:: As above this is helpful detail

2.6.11 Finally, the discussion section is lacking depth as to why these findings diverge so entirely with previous studies. There is not the uncertainty on the effects of autozygosity as the authors suggest. There is in fact a consensus on the increased risk of ill-health. So the protective effects observed do require considered sensitivity testing and perhaps even external validation.

• We agree with the reviewer that the lower diabetes and MI risk in our study contrast with conventional expectations. We do however observe many associations consistent with previous reports (Clark 2019): higher rates of childlessness, fewer children, lower educational attainment, lower tobacco usage, shorter stature, lower weight, and lower cholesterol. PGR is the largest consanguineous cohort in the world and can detect associations between autozygosity and traits that are not present in other cohorts in sufficient numbers, such as myocardial infarction. For these reasons, external validation will require the expansion of other highly autozygous cohorts to similar size and case/control composition. We have performed sensitivity analyses within PGR, such as subsetting by ethnicity.

Review2:: External validation of a finding as controversial as higher FROH being associated with lower risk of important health outcomes like MI and T2D etc. Is required if this finding is to be asserted. Why not utilise resources like G&H where there seem to be >20k individuals of pakistani ancestry with WES data available based on this preprint: <https://www.medrxiv.org/content/10.1101/2025.09.24.25336527v1.supplementary-material>

2.7 Biomarker burden

2.7.1 Again a flow-chart showing the extent of clinical data, proteomic data, etc, would be really helpful to orientate the reader.

• We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now additionally includes a list of all biomarkers and binary outcomes analyzed in this study, along with the number of individuals for which the data is available. Additionally, for each analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of which individuals were included.

Review2:: Thank you, this is very helpful

2.7.2 225: These are extremely interesting and well-presented results. To orientate the reader, it would be useful to have a short section explaining how many GENE*phenotype associations were tested, and what the associated Bonferroni-corrected p value would therefore be (for example, the “lower random glucose ($\beta=-0.05$, $P=2.3e-02$)” seems unlikely to be anything other than nominally significant, but it is not clear to the reader which other association would survive multiple testing correction.

• We thank the reviewer for the kind commentary. Regarding multiple testing, we perform an initial PheWAS where we report all gene-biomarker associations that are significant at an exome-wide Bonferroni threshold of $P<9.9e-08$ in Supplementary Table 5. Elsewhere, when following up on specific genes we report additional, nominal associations. We have specified in the text that these associations are nominally significant at this threshold.

Review2:: Thank you for specifying where associations are purely nominal. It would be helpful for the reader to understand how “specific genes” in this context were identified – on the basis of specifically interesting phenotypic associations, or on the basis of specifically interesting genes, if so how were these defined, and if not, so how were they selected?

My apologies for repeating the point, but in a PheWAS where Bonferroni correction requires a $p < 9.9e-8$ discussing as a major result, an association with a $P = 0.023$ seems far from statistically meaningful. If 1/3 of coding genes contain knockouts, as is stated in the title, then by chance one would expect many thousands of associations in the $p = 0.023$ threshold. Was there some other a priori definition of genes of interest which would confer a more relaxed P value threshold? I think some explicit justification of why results at relaxed P value thresholds are highlighted, given the scale of PheWAS association testing, is needed, or the authors should focus discussion on those associations with more convincing P values.

2.7.3 It would also be useful to provide a summary table of how many individuals had available data for each phenotype discussed in the main text (this feeds back into a previous comment on study data presentation and flow charts)

- Due to space constraints, we could not include this information in the main text. We have added this data to Supplementary Table 1.

Review2:: This is very helpful, many thanks

2.7.4 It would be helpful in this section to compare to existing pLoF burden results identified either through eg AZphewas portal or GENEASS to orientate the reader to the novelty of these findings, as they are presented.

- We agree with the reviewer that including UKBB data alongside PGR data would facilitate comparison. As stated in our revised text, include associations from the AstraZeneca PheWAS browser in Supplementary Table 5 (lines 236-237):
- For comparison, we include associations from the AstraZeneca PheWAS browser 26,27 (Supplementary Table 5).

Review2:: This is very helpful, thank you

2.7.5 229: It would be useful to compare to G&H recent effort characterising pLoF human KOs, recent preprint: <https://doi.org/10.1101/2025.06.05.25329068>. For example the Genes and Health authors report 2991 homozygous carriers of pLoF in a sample size of ~44k individuals and this could be compared to findings from PROMIS-PGR for a better understanding of the landscape of genomic studies in SAS populations and sub-populations.

- Based on the reviewer's feedback, we have included a comparison to South Asians in gnomAD using the same set of pLoF filters that were applied to PGR. gnomAD South Asians is comparable to Genes and Health in size and number of KOs: gnomAD has 45,546 South Asians, from which KOs are present in 3,022 genes. We have not analyzed the Genes & Health cohort as this preprint is not yet peer-reviewed and could be subject to change.

Review2:: We agree this is a pre-print and may be subject to change, but would also contest that referencing and learning from pre-printed manuscripts is increasingly the norm in the modern scientific era; particularly when making claims of particular novelty.

2.7.6 217: Some of these associations with HBB, including iron and lipid traits, in addition to bilirubin, are reported in the Kim et al paper above and should be discussed in this context. - Please see above.

Review2:: It is unlikely that an independent cohort reporting a consistent result with PGR is in any way problematic. If anything, a future reader of this paper would benefit from knowing that these findings are consistent and coherent with available literature from the limited other resources of south Asian data that exist.

2.8 Analysis of LOF constraint

2.8.1 261: it would be helpful to briefly define where these approved drugs are from (FDA? Some other authority?) to save the reader diving to the methods

- We downloaded the data used by Minikel et al 2024, which used data from Citeline Pharmaprojects. We added this citation to the text.

Response: This is very helpful

2.8.2 285-287: Are these AAFs significant? P values are missing, and no SEs are presented.

- We thank the reviewer for highlighting this oversight. We have performed additional statistical tests and added p-values to the main text. Please also see our response to 1.7 above.

Review2:: Many thanks

2.8.3 297: Were synonymous variants "modestly depleted" if $p = 0.225$ (line 299)? And doesn't no depletion of synonymous variants in fact strengthen your argument, ie, as a negative control?

- We thank the reviewer for this close reading. Our intention was to state that synonymous variants were only nominally significantly depleted in a small number of gene sets, but not gene sets that one would expect to be subject to recessive constraint – e.g. genes essential in vitro or associated with autosomal recessive conditions. We now describe a different metric in the main text and present AAR in supplemental analysis.

Review2:: We agree this is a sensible approach

2.9 Human KO highlights

2.9.1 For recall samples, was sequencing repeated to confirm the initial sequencing/genotyping was correct?

- Yes, confirmatory Sanger sequencing was conducted for recall samples

2.9.2 310: The LRRK2 findings are certainly most interesting, but I think the strength of assertion around need for renal monitoring should be couched firmly in the $N=2$ here, and perhaps more detail and deeper phenotyping on these 2. To my mind this argument would be considerably strengthened by checking for known genetic causes of CKD, eg PCKD, and perhaps calculating a CKD PGS. Ie, in the absence of other known genetic constructs causing CKD, we may be more inclined to believe it is really all mediated by LRRK2. Is further phenotyping possible (eg ACR, USS)? We appreciate this may not be feasible and are not suggesting this would be required for publication. However, the fullest picture of these individuals phenotypes would be needed to persuade readers of this journal that LRRK2 was the underlying cause here.

- We agree with the reviewer about the limitations of small numbers. To strengthen the observations from the two knockouts, we discovered a third independent KO through recontact and sequencing family members of an unrelated heterozygous

loss-of-function variant carrier. Similarly to the two knockouts from the PGR biobank, the third knockout had low eGFR. As the reviewer rightfully suggested, we evaluated the exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, and APOL1). No damaging missense (REVEL score >0.7) or HC pLoF variants were found.

- We believe that our human data stands as a capstone to an extensive series of pre-clinical reports implicating LRRK2 in kidney function:

- o LRRK2 knockout mice exhibited impaired kidney function and hypertension (Tong et al., 2010; Herzig et al., 2011).

- o LRRK2 knockout rats exhibited renal dysfunction (Gu et al., 2023).

- o LRRK2 inhibitors impaired kidney function in rodents (Andersen et al., 2018).

- o LRRK2 inhibitors caused kidney pathology in non-human primates (Jennings et al., 2022).

- Our revised manuscript highlights these points (lines 349-359):

- o Recall-by-genotype of LRRK2 pLoF carriers and consenting family members identified a 49-year-old female homozygous for LRRK2 splice donor variant,

c.1418+1G>T (Extended Data Figure 10a). This individual was the sole identified KO in her family and had an eGFR of 51 ml/min, suggestive of kidney disease. Her eGFR was lower than all 13 of her recruited heterozygous and non-carrier family members (Extended Data Figure 10b). We evaluated the available exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, or APOL1) and found no damaging (REVEL score >0.7) or HC pLoF variants. Although the number of LRRK2 knockouts is low in our study, the observation of kidney dysfunction in these individuals is consistent with, and provides human data to support, the adverse renal phenotypes observed in rodent LRRK2 knockouts⁵¹⁻⁵³ and the preclinical renal findings of LRRK2 inhibitor administration in rodents⁵⁴ and non-human primates⁵⁵.

Review2:: Thank you, as above, the literature is indeed convincing here. The main point I was making was really around framing, and I agree that describing these results as the capstone to what has gone before is appropriate – rather than couching the need for clinical renal monitoring having been identified in the present study.

2.9.3 326: Similarly, some more phenotypic information on the human KOs for SLC2A4 – if available – would be very interesting to the reader. Given they are first reported human KOs for this gene, there is likely to be enormous interest in knowing more about glycaemic measures in these individuals. The T2D knowledge portal reports findings with thyroid function at the gene-level, rare variants with ApoA1 and blood pressure. The authors have a unique opportunity to explore and characterise these rare individuals.

- We agree with the reviewer that this additional information would provide insights. We have added available glucose and HbA1c levels for SLC2A4 and the KOs in Supplementary Table 10. We were unable to recontact the SLC2A4 knockout for additional characterization.

Response: this is unfortunate but understandable. It would be good to include a statement saying so in text. Or simply, recontact was not possible for deeper phenotyping.

2.9.4 The findings in DENND1B KOs are interesting, but at odds with published summary stats from UK Biobank. How do authors reconcile this?

- We are similarly puzzled that the two new DENND1B KOs are not obese, in contrast to the report by Wallis et al. The difference may be due to variable penetrance, ancestral differences in weight gain, statistical limitations due to the rarity of DENND1B pLoF, or a combination of these factors.

- Wallis et al. does not show a formal burden analysis of the association between damaging or high confidence DENND1B loss-of-function variants and BMI with a corresponding effect size. Previously published analyses showed only nominally significant enrichment of combined damaging and protein truncating variants and a magnitude-free directionality of BMI (no reported effect size). Additionally, the paper provided supporting evidence of a single pediatric obesity case homozygous for a damaging missense variant (R501C)—the parents of this child were both heterozygous carriers and of normal BMI.

- In a formal burden analysis in PGR, we see no significant associations between DENND1B damaging/high confidence loss-of-function variants and BMI. Burden analyses in AZPheWAS also show no significant association between DENND1B rare damaging/loss of function variants and BMI. In addition, PGR contains two adult, high confidence loss-of-function homozygotes with normal BMI. Collectively, this suggests that DENND1B loss-of-function has little effect on BMI in adult humans.

Review2:: Thank you. As throughout the manuscript, I think external validation of your findings is reassuring and should be leveraged. Could authors please add the agreeing information from AZPheWAS portal to text to head off any concerns linked to disagreement with the Wallis paper.

2.9.5 336: It would be helpful to present some measure of range here (is this driven by a single BMI of 45?); or perhaps even all 4 BMIs, given the small number.

- All BMIs are included in Supplementary Table 10 and range from 28-42. We have now specifically included this information in the text for clarity.

Review2:: This is helpful, thank you

2.9.6 Given the novel finding, which is at odds with the AstraZeneca phewas portal findings for LoF DENND1B, it would be valuable to undertake replication with other data sets, UK Biobank, Genes & Health to confirm this findings.

- With apologies to the reviewer if we misunderstood their comment, we do not see significant associations between DENND1B LoF and BMI in the AZPheWAS portal.

Review2:: With apologies to the authors, the results we were drawing their attention to were for the AZPheWAS “flexdmg” mask. In any case including AZPheWAS portal in Supp Table 5 helps orientate the reader to previously described results.

This is not cross-referenced in the results paragraph on DENND1B – could you please do so and reassure readers about consistency?

2.9.7 338: Associations with RXFP1 have been demonstrated in UK Biobank, data available on AZPheWAS portal. The novelty here is unclear – but would benefit from highlighting and meta-analyses with UKB.

- With apologies to the reviewer if we misunderstood their comment, we do not see significant associations in RXFP1 in the AZPheWAS portal.
- We believe that our report of RXFP1 KO is important because the healthy cardiovascular, metabolic, and reproductive profiles of these individuals strongly invalidate the longstanding pre-clinical observations that formed the foundation of past clinical trials. Decades of publications using rodent models suggested that RXFP1 is critical for nursing, fibrosis, and cardiac health, and metabolism. These rodent studies motivated many clinical trials testing RXFP1 activators in humans, including two large Phase 3 trials, to treat a broad spectrum of diseases, including impaired cervical ripening during childbirth, scleroderma, and acute heart failure. These trials have been unsuccessful.

Review2:: On the AZPheWAS portal RXFP1 rare dmgs variants are associated with scleroderma/systemic sclerosis ($p = 0.04$), heart failure/pulmonary oedema ($p = 5e-3$), and arterial strictures ($p = 1.2e-4$). We agree in principal with the author's contention that a healthy RXFP1 KO is of substantial interest, but the point remains that the p values the authors consider significant must be consistent across analyses – for example if they wish to highlight a p value of $p = 0.02$ in their results (as discussed in 2.7.2 above), then we would argue they should also draw the reader's attention to results for other findings elsewhere at a similar threshold and cannot dismiss these as not significant.

2.9.8 379: TRPM8 – bigger previously reported signals have been for lipid traits and MI, did this not replicate here?

- With apologies to the reviewer if we misunderstood their comment, we have not identified any previously published significant associations by GWAS or ExWAS between variants in TRPM8 and lipid traits or myocardial infarction. We also do not see any associations in the PGR biobank or in the recall by genotype study. However, there have been multiple studies reporting variants in the TRPM8 locus for migraine (Chasman et al., 2011; Freilinger et al., 2012), including variants associated with partial reductions in TRPM8 expression that reduce the risk for migraine and cold sensation (Gavva et al., 2019). While PGR lacks data on migraine-related phenotypes, our findings indicate that complete loss of TRPM8 is tolerated. Therefore, pharmacological inhibition of TRPM8 for migraine may be safe. We elaborate on these points in the revised text (lines 424-427):

- In GWAS studies, TRPM8 has been linked to migraine72-74, with the directionality suggesting that reduction in TRPM8 reduces risk of migraine68. While PGR lacks migraine-related phenotypes, our findings indicate that complete loss of TRPM8 is tolerated and, therefore, pharmacological inhibition of TRPM8 for migraine may be safe.

Review2:: Really this speaks to the point above around p-value thresholds- it is interesting the AZPheWAS portal result for ptxrare dmgs variants ~ angina (Union1208 other forms of angina pectoris, OR = 1.93, $p = 1.7e-4$) did not replicate, particularly given the strong associations with family history also reported on the AZPheWAS portal.

2.9.9 All these findings would be more impactful, if where possible, external validation or replication was undertaken. Coherent and directionally stable associations with ptxs in UK Biobank, Genes & Health which is also likely enriched for LoF variants (albeit in a smaller sample size), PRECISE etc.

- We have taken care to highlight examples of novel high confidence loss-of-function homozygotes that have not been previously reported or characterized by other cohorts. The absence of KOs in other cohorts precludes external validation of the phenotypes associated (or not associated) with these KOs.

Review2:: The logic of not replicating or meta-analysing novel findings where possible is not apparent to me. Particularly where p values are not entirely convincing.

If attempts have been made to identify LoF KOs in other cohorts and none were found this should be explicitly stated in text and will enhance the novelty of PGR.

2.9.10 Similar to the previous section, it would be helpful to know how many GENE*trait associations (*model type (additive / recessive)) etc were performed across this section, and an associated Bonferroni-corrected p value presented to the reader. Presumably not just the genes included in the vignettes were included in analysis, but it is difficult to get a sense of what a meaningful p value is without knowing the number of tests performed.

- The associations in this section are generally hypothesis-based, motivated by the observations of KOs in genes for which KOs have not been previously identified. Where appropriate, we have added additional descriptions of statistical tests.

Review2:: Our apologies, we struggled to find these in the text, please could the authors highlight which additional descriptions of tests have been included?

2.10 Discussion

2.10.1 391: as above, the actual median FROH is not within the range you describe

- We thank the reviewer for catching this error. It should state that the mean value (0.034) falls within the range. We have corrected this in the text (lines 176-178):
 - o The mean FROH in PGR was within the expected ranges of first (FROH = 0.0625) to second cousin (FROH = 0.016) level of parental relatedness and was consistent with the high percentage of individuals (30.6%) who self-report being born to parents who were first cousins.

2.10.2 409: You describe the inconsistencies "Across previous studies", but actually other studies that you describe were all

directionally consistent with autozygosity being associated with deleterious phenotype (consistent with inbreeding depression).

- We thank the reviewer for pointing out the lack of clarity in our text. We intended to highlight inconsistencies between PGR and some previous studies (e.g. diabetes in PGR vs. G&H), not across previous studies. We have revised the text (lines 455-456):

- The age-dependent effects of autozygosity on health may partially explain the inconsistencies for some of the findings in our studies relative to other studies.

Review2:: We agree this is helpful

2.10.3 411: It is unconvincing that the thrifty genotype hypothesis explains why these (age-stratified) results are at odds with the Genes & Health results in Malawsky et al – I think more fleshing out of the proposed mechanism is needed here to convince the reader.

- Given the retrospective nature of our observations, we offer only a hypothesis rather than a mechanism supported by data. In the revised text, we have clarified that the thrifty gene hypothesis and the possibility of past resource-limited environments may explain only a subset of the protective phenotypes, such as lower LDL-C, triglycerides, BMI, and MI (lines 456-460):

- o The mildly protective effects of autozygosity on some cardiometabolic biomarkers (LDL-C, triglycerides, BMI) and outcomes (MI) may represent reduced fitness of traits that were protective in past resource-limited environments, consistent with the “thrifty genotype” and related hypotheses that aim to explain high rates of obesity, diabetes, hypertension, and cardiovascular disease as a byproduct of adaptation to a divergent ancestral environment⁸³⁻⁸⁷.

Review2:: Given the age-based heterogeneity it would be useful to readers to understand how that aspect squares off with the mechanism proposed here.

2.10.4 416: I'm a little unclear on why the HBB-lipids associations have been determined to be causal, but other HBB-trait associations (e.g. MI) are not described as causal. Perhaps I have missed some key analysis step here? The HBB associations with iron, red cell biology and MI are strong and convincing, but I do not follow how you determine them to be causal?

- We thank the reviewer for pointing out that we may have overstated the causal relationship between HBB, iron, red cell biology, and MI. We have modified the language to suggest a relationship whose mechanism is yet to be established (lines 461-465):

- PGR is powered to reveal directional relationships between genetic variants and phenotypes, including circulating biomarkers. Such relationships were established between TIMD4 pLoF and circulating lipids, and between HBB, circulating lipids and apolipoproteins. In PGR, HBB pLoF is associated with lower risk of myocardial infarction. This suggests that HBB pLoF could modulate heart attack risk through hemoglobin or its constituent iron, or through lipids and lipoproteins⁸⁸.

Review2:: We agree this is much better framed as an association without discussion on causality

2.10.5 As below, it would be useful to include some comment on participant recruitment methods, and potential sources of bias associated with this.

- We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514): o Recruitment. Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.

2.11 Methods

2.11.1 509-518: Did you set a specific minimum length for a ROH eg 1Mb? How was a ROH defined? I'm a little unclear after reading this section.

- We identified regions of ROH using the widely used BCFtools roh algorithm (Narasimhan et al. 2016). The algorithm uses a hidden Markov model that models the probability of transitioning from ROH to non-ROH regions using parameters such as known position-specific recombination rates and population allele frequencies. As such, it does not require a user-defined minimum length threshold.

Review2:: Thank you for clarifying

2.11.2 610: Your FROH distribution looks skewed (as would be expected) – did you log-transform it prior to logistic regression, and if so why/why not?

- We thank the reviewer for their close attention to the data. For consistency with previous publications in the field (Clark et al. 2019, Malawsky et al. 2023), we have not transformed the FROH data prior to logistic regression. We do perform rank inverse normal transformation of quantitative traits prior to running linear regression.

Review2:: Thank you for clarifying

Minor comments ;

2.12 Intro 98: The first line feels a little at odds with the rest of the paper in setting the scene – much of this paper has no relevance to drug discovery?

- We agree with the reviewer that the paper does not focus exclusively on drug targets. At the same time, findings from human genetics are directly translatable to drug discovery.

Review2:: we will leave this to the editor's discretion.

2.13 After reaching the end of the introduction and starting results I was a little unclear on what exactly the aims / objectives of this study are – perhaps a brief sentence to summarise might help to orientate the reader ahead of the analyses being presented in the results

- We thank the reviewer for this observation. As the reviewer suggested, we have revised the text to summarize the aims/objectives of the study (lines 126-129):

- o This report summarizes the regional genetic diversity of the PGR, the impact of autozygosity on disease risk, the discovery of human knockouts for one-third of coding genes, and examples of recall-by-genotype studies that deepen insights into gene function.

2.14 103: I think for the non-expert reader it is necessary to expand on why autozygous populations are enriched for recessive / pLoF KO effects when explaining they can “uniquely power genetic discovery” ie spell out that same allele from common ancestor can arrive via each parent / IBD. A bit more detail on autozygosity for the non-expert reader, not necessarily obvious why they are more likely to have 2 copies, perhaps some sort of diagram and explanation of identity-by-descent

- We thank the reviewer for suggesting this improvement. As the reviewer suggested, we have revised the text to elaborate (lines 114-118):

- Consanguinity in PGR is driven by a high proportion of individuals who are the offspring of first- or second-cousin unions. These individuals are more likely to be homozygous for (have two alleles of) otherwise rare alleles. Therefore, relative to non-consanguineous cohorts, PGR is powered to discover individuals who are loss-of-function homozygotes, also known as naturally occurring human gene “knockouts”.

Thank you

2.15 160: A minor but concerning observation, but the median FROH value of 0.015 doesn't fall within your stated ranges 0.016 – 0.0625 as you say it does.

- We thank the reviewer for catching this error. It should state that the mean value (0.034) falls within the range. We have corrected this in the text (lines 176-179):

- o The mean FROH in PGR was within the expected ranges of first (FROH = 0.0625) to second cousin (FROH = 0.016) level of parental relatedness and was consistent with the high percentage of individuals (30.6%) who self-report being born to parents who were first cousins.

Thank you

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

Referee #4

(Remarks to the Author)

The authors have addressed most of my concerns. I still have two major concerns:

- human knockouts seems to be a ill-suited term to refer to humans and can lead to all kinds of stigma and miscommunication. I suggest the authors take the time to think of a better way to refer to said individuals and as such alter the title and use in the manuscript. The terminology we use is key to not go down a eugenics path again.
- The access to the dataset is still not clear, and generates a lot of concern if academics will truly be able to access the data for academic projects. If the data has to be homed at CNCD, how will the authors ensure a fair and transparent process post-publication?

Version 2:

Reviewer comments:

Referee #2

(Remarks to the Author)

Referee #2 (Remarks to the Author):

Thank you for the opportunity to review this paper – it is clearly a study of monumental scale with potential to further our understanding of autozygosity and new gene effects through population-specific and enriched LoF variants.

• We thank the reviewers for their positive overall assessment of the manuscript and provide answers to their questions below.

We thank the authors for another opportunity to review this excellent manuscript, which has been substantially improved during these revisions.

- We thank the reviewer for their positive assessment of our manuscript and for helping strengthen our paper. We also appreciate that the reviewer noted that most of our edits improved the manuscript. Below, we will focus only on the remaining questions.

2.4 Results: A flow chart of which participants are included in which part of each analysis would be very helpful to orientate the reader (this feeds into comments above) – it is essential to orient the reader to what PROMIS-PGR contains in terms of recruitment, genetic data generation (different sources and technologies), omic data similarly and finally what clinical data is collected and contributes to the compendium described.

• We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now additionally includes a list of all biomarkers and binary outcomes analyzed in this study, along with the number of individuals for which the data is available. Additionally, for each analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of which individuals were included.

- Review 2: This is very helpful to orientate the reader. Please include the WES technology used and also list what other sources of genomic data exist other than WES – as the total number of PGR samples that contributed to coding variant counts was 173,303 whereas WES is available is 166,625. This can be added to the “Notes” section of the table.

- We thank the reviewer for this suggestion. A total of 166,625 WES samples and 6,678 WGS samples were analyzed in this study. We have added a breakdown of the counts of WES and WGS samples used for each analysis to Supplementary Table 1 and have added more detail to the methods.

Review 3: This is most helpful, thank you.

2.5.1 In the population genetics portion of the manuscript, it is not clear why self-reported ethnicity was used for FST as opposed to the just-projected PCs? Was it not possible to project the sub-populations within PGR with sufficient granularity? This could be expanded upon for clarity. Is it possible to clarify for how many individuals this was done?

• We agree with the reviewer that self-reported ethnicity has limitations. We prepared a new Figure 3 to visualize clustering of South Asian populations. Individuals of the same ethnicity tend to be close to each other in PCA. This observation indicates that ethnicity was an appropriate input to the FST analysis, which included all participants.

- Review 2: We agree with the authors that these changes help strengthen the paper. Can the authors explain what they mean by “ethnicity” because in PGR there is a group clubbed as “Urdu” which is a language rather than ethnic group.

- The reviewer is correct that Urdu is a language and we thank them for raising this point. Individuals who identify themselves as Urdu represent a large, mixed group in Pakistan who have largely immigrated from North India during the post-colonial partition of the subcontinent. Individuals who identify as Urdu are ethnically distinct such that they don't recognize themselves as any other ethnic group in Pakistan. For instance, a person who recognizes himself or herself as “Urdu speaking” will not classify himself as “Punjabi”, “Sindhi”, “Pathan”, etc. Importantly, all ethnicity data in PGR is based on the self-reporting of participants. We state this information in the revised main text and revised Methods.

Review 3: This is helpful and now clear in the main text

2.5.2 I'm afraid as a non-expert in pop gen, the conclusions and take-away message from this portion of the manuscript were not apparent. Can this be better fleshed out and more explicit?

• We thank the reviewer for suggesting ways to make this section more insightful. To address the reviewer's point, we re-ordered the population analysis, which now presents Fst analysis, admixture, and PCA, for a clearer, stepwise presentation of South Asian ancestries and their genetic make-up. We also provided a PCA of the individuals in our study alongside South Asian 1000 Genomes (1KG-SAS) that we separated into its constituent ethnicities, as recommended by the reviewer. As the reviewer predicted, the new analysis projected PGR ethnicities along a geographic gradient of proximity to Europeans (PC1, Figure 3c,d). At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were the western populations of Pakistan, such as the Pathan and Balochi of PGR. The revised text includes the following information (lines 159-170):

o Principal Component Analysis (PCA) quantified the heterogeneity of PGR subpopulations relative to each other and to non-PGR populations from the 1KG (Figure 3). African, East Asian, European, and South Asian ancestries separated into clusters, with South Asians in PGR following a gradient along PC2 from European ancestries to East Asians (Figure 3a,b). Consistent with the admixture analyses, Pathan and Balochi ethnicities clustered more closely with Europeans than did other PGR ethnicities, while a subset of Balochis clustered more closely with Africans. Among individuals without recent African admixture, PC1 distinguished PGR ethnicities along a gradient of proximity to Europeans, consistent with previous observations (Figure 3c,d)^{19,20}. At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KGSAS, and at the other end were populations of western Pakistan, such as the Pathan and Balochi of PGR. PC2 and PC3 identified two additional clusters, one of which was predominantly Balochi and the other predominantly Punjabi, consistent with previous observations of substructure among the latter population²¹.

- Review 2: We agree this adds some clarity to results for the non-expert reader. Can authors please differentiate between their use of ethnicity and genetically determined ancestry as these are not the same and, in many cases, it appears they are referencing the latter, certainly in Fig 2 and 3 which are PCA-reliant.

- The reviewer is correct that the admixture and PCA plots each provide a view of genetically determined ancestry. Within these plots, the grouping (admixture plots) and labelling (PCA plots) are done by self-reported ethnicity to show the genetic variation that exists both within and across these self-reported ethnic groups.

Review 3: We thank the authors for addressing this concern

2.6 Supplemental Table 4 is an excellent presentation of these data. Apologies if I have not been able to spot this, but is there a histogram that plots FROH (on Y axis) against age (on X)? There appears to be an age-based stratification of the effect of FROH on hypertension, CKD, and T2D and perhaps others. Wherein, if you were to meta-analyse the effects of FROH across these age-groups you would not find agreement in effect. Indeed this is captured in the p value for

heterogeneity across groups. It could be, that individuals recruited at older ages have lower genome-wide/exome-wide runs of homozygosity, hence the question about distribution of FROH across age groups. There are also ~ 32,200 individuals recruited < 40 years whereas over 130k are over the age of 40 at recruitment.

The impact / heterogeneity across age groups makes it hard digest the results from the FROH analysis without a degree of scepticism. For example, if we were to accept the overall finding that higher FROH is protective from T2D, MI, what is the biological basis of this?

- We thank the reviewer for these comments and suggestions. As proposed, we have now added additional considerations about the possibility of these biases to the discussion (lines 431-433):

- o "Future longitudinal studies may elucidate these age-dependent associations and exclude the possibility of ascertainment or survivorship bias."

- Regarding the biological basis of our findings, we offer an evolutionary hypothesis in the discussion. Future studies will focus on understanding possible biological mechanisms.

Review 3: It is helpful that these limitations have been added, but I would still argue the overall framing of certainty around this finding which is (1) in conflict with the existing literature and (2) not replicated is too strong

2.6.6 190-200: Again these are extremely interesting results. Why was age 40 chosen as a cutoff? This is a typical young-onset threshold for T2D, but is perhaps a little early for MI. It would be more interesting to see these effect sizes plotted out by 10-year age bin (ie 18-30, 30-40, etc) to see how the association varies across the full breadth of ages (accepting power will be limited here), as opposed to a binary </>40 bin which leaves the reader wondering where the effect arises from. These divergent effects across the age range are very surprising and a bit more granularity would be most interesting to the reader.

- We agree with the reviewer that the cutoff of 40 was arbitrary. We have now stratified our analyses into multiple age bins and describe the results in the revised text (lines 219-224): o To determine if the associations were age-dependent, we binned across age groups (Extended Figure 8). Significant heterogeneity of effect was observed for educational attainment ($P=3.4e-08$), number of children ($P=3.2e05$), having had children ($P=6.2e-06$), diabetes ($P=9.0e-06$), and hypertension ($P=1.9e-08$), with autozygosity trending towards increased risk of hypertension and diabetes in individuals <40 years and protection from hypertension and diabetes in individuals 50 years old (Extended Figure 8). This approach is much more informative to the reader and strengthens the manuscript

- Review2: agree. Thank you for making this table. See previous comment regarding acknowledging the lack of longitudinal data, recruitment of as many younger individuals <40 compared to >40 and >50, and heterogeneity of effects across age groups as a critical limitation of this analysis and generalisability of this finding. - Please see our response above.

2.6.7 Is there some sort of artificial clumping driving this association? Is FROH associated with age in some way, i.e., is the population becoming more autozygous over time, more FROH in young people, who do not yet have T2D/MI etc?

- We see less autozygosity with age in PGR ($P=8.4e-286$), albeit with a small effect size ($\beta=0.00026$ FROH per year of age). In Figure 1 of Malawsky et al., the authors report a similar magnitude of effect between age and autozygosity in British Pakistanis, but not in British Bangladeshis, arguing that survivorship bias or ascertainment bias are unlikely to explain the association between age and autozygosity. Additionally, we have included both age and age*age as covariates in all association analyses.

- Review 2: We agree that through the lens the authors describe, survivorship bias does not necessarily explain these findings. But the age-based heterogeneity of effects is hard to ignore. There appears to be an age-independent but somehow intertwined artefact. Do those who survive to age 50, 60 and higher have runs of homozygosity around different loci that are protective compared to those who are younger and indeed theoretically, since you cannot measure this –in those who did not survive to participate in the study? - Please see our response above.

2.6.11 Finally, the discussion section is lacking depth as to why these findings diverge so entirely with previous studies. There is not the uncertainty on the effects of autozygosity as the authors suggest. There is in fact a consensus on the increased risk of ill-health. So the protective effects observed do require considered sensitivity testing and perhaps even external validation.

- We agree with the reviewer that the lower diabetes and MI risk in our study contrast with conventional expectations. We do however observe many associations consistent with previous reports (Clark 2019): higher rates of childlessness, fewer children, lower educational attainment, lower tobacco usage, shorter stature, lower weight, and lower cholesterol. PGR is the largest consanguineous cohort in the world and can detect associations between autozygosity and traits that are not present in other cohorts in sufficient numbers, such as myocardial infarction. For these reasons, external validation will require the expansion of other highly autozygous cohorts to similar size and case/control composition. We have performed sensitivity analyses within PGR, such as subsetting by ethnicity.

- Review 2: External validation of a finding as controversial as higher FROH being associated with lower risk of important health outcomes like MI and T2D etc. Is required if this finding is to be asserted. Why not utilise resources like G&H where there seem to be >20k individuals of pakistani ancestry with WES data available based on this preprint:

<https://www.medrxiv.org/content/10.1101/2025.09.24.25336527v1.supplementary-material>

- We agree with the reviewer that external validation, wherever possible, would be beneficial. We cite Malawsky et al. (2023) which finds an association between FROH and increased risk of T2D using the G&H data described by the reviewer. For many outcomes, such as MI, PGR is uniquely powered due not only to its large number of highly consanguineous individuals, but its case/control breakdown. Additional large, highly consanguineous cohorts with similar phenotypic distributions are not available.

Review 3: Thank you for your response. Do the authors have, for example, power calculations from other available cohorts to support their assertion that PROMIS-PGR is uniquely powered to detect these effects, even if only at a nominally or borderline significant (with consistent direction of effect) level? I am not suggesting this should be in the main text, but it seems difficult to believe an effect such as that observed for MI ~ FROH that is so well-powered for detection in a cohort of ~150k individuals would be entirely absent in a highly consanguineous cohort of ~55k individuals (~20k Pakistani) like Genes & Health. Indeed this association was tested and not significant in Malawsky et al.

2.7.2 225: These are extremely interesting and well-presented results. To orientate the reader, it would be useful to have a short section explaining how many GENE*phenotype associations were tested, and what the associated Bonferroni-corrected p value would therefore be (for example, the “lower random glucose ($\beta=-0.05$, $P=2.3e-02$)” seems unlikely to be anything other than nominally significant, but it is not clear to the reader which other association would survive multiple testing correction.

• We thank the reviewer for the kind commentary. Regarding multiple testing, we perform an initial PheWAS where we report all gene-biomarker associations that are significant at an exome-wide Bonferroni threshold of $P<9.9e-08$ in Supplementary Table 5. Elsewhere, when following up on specific genes we report additional, nominal associations. We have specified in the text that these associations are nominally significant at this threshold.

- Review 2: Thank you for specifying where associations are purely nominal. It would be helpful for the reader to understand how “specific genes” in this context were identified – on the basis of specifically interesting phenotypic associations, or on the basis of specifically interesting genes, if

so how were these defined, and if not, so how were they selected?

My apologies for repeating the point, but in a PheWAS where Bonferroni correction requires a $p < 9.9e-8$ discussing as a major result, an association with a $P = 0.023$ seems far from statistically meaningful. If 1/3 of coding genes contain knockouts, as is stated in the title, then by chance one would expect many thousands of associations in the $p = 0.023$ threshold. Was there some other a priori definition of genes of interest which would confer a more relaxed P value threshold? I think some explicit justification of why results at relaxed P value thresholds are highlighted, given the scale of PheWAS association testing, is needed, or the authors should focus discussion on those associations with more convincing P values.

- We thank the reviewer for these comments. We report all gene-biomarker associations that are significant after multiple test correction ($p<9.9e-8$). For select genes of clinical importance that are discovered through this approach, such as HBB, we show all nominal associations to provide a full picture of the gene, specifying that these additional associations are nominal ($p<0.05$).

Review 3: Apologies for reiterating this point, but it is not clear how these “select” genes were defined or selected, whether it was a priori or not, and if not, what the relevance of nominally significant p values of ~ 0.05 is in the context of a PheWAS – the reader will be much less convinced if they suspect these associations were selectively picked out post-hoc. It is also not entirely clear how the $p < 9.9e-8$ is arrived at, as above, the authors describe this as exome-wide significant, but it is unclear if this is corrected for the multiple phenotypes tested? Shouldn't the value of interest be genome-wide significant?

This is of course a relatively minor point on which we disagree, in the broader context of a well-run study and fabulous manuscript, and I will be more than content to leave it to the editor's expertise to determine whether this comment should be discounted or needs further consideration.

2.7.5 229: It would be useful to compare to G&H recent effort characterising pLoF human KOs, recent preprint:

<https://doi.org/10.1101/2025.06.05.25329068>. For example the Genes and Health authors report 2991 homozygous carriers of pLoF in a sample size of $\sim 44k$ individuals and this could be compared to findings from PROMIS-PGR for a better understanding of the landscape of genomic studies in SAS populations and sub-populations.

• Based on the reviewer's feedback, we have included a comparison to South Asians in gnomAD using the same set of pLoF filters that were applied to PGR. gnomAD South Asians is comparable to Genes and Health in size and number of KOs: gnomAD has 45,546 South Asians, from which KOs are present in 3,022 genes. We have not analyzed the Genes & Health cohort as this preprint is not yet peer-reviewed and could be subject to change.

- Review 2: We agree this is a pre-print and may be subject to change, but would also contest that referencing and learning from pre-printed manuscripts is increasingly the norm in the modern scientific era; particularly when making claims of particular novelty.

- We recognize that Genes & Health is a valuable resource and appreciate the reviewer's suggestion to include a direct comparison of the homLoF genes present in PGR and Genes & Health. We do not have access to variant-level data in G&H that would allow us to apply the additional loss-of-function variant filters applied to PGR—including correction of errors in LOFTEE—which we describe in the methods, supplementary text, and Supplementary Figure 2, and would allow for a relevant head-to-head comparison of homLoF genes between these two cohorts.

Review 3: While we were suggesting that obtaining access to G&H for simple analyses could substantially strengthen this study and work, we appreciate limitations of time and the advanced stage of review may preclude this for this body of work.

2.7.6 217: Some of these associations with HBB, including iron and lipid traits, in addition to bilirubin, are reported in the Kim et al paper above and should be discussed in this context. - Please see above.

- Review 2: It is unlikely that an independent cohort reporting a consistent result with PGR is in any way problematic. If anything, a future reader of this paper would benefit from knowing that these findings are consistent and coherent with available literature from the limited other resources of south Asian data that exist.

- We thank the reviewer for the suggestion. We have updated the manuscript and added this citation (lines 234-235):

• “HBB pLoF variants are associated with reduced hematocrit, increased protection against malaria parasite, increased risk of β -thalassemia, bilirubin levels, and various lipid traits³⁰⁻³².”

Review 3: We agree this strengthens the manuscript

2.9.3 326: Similarly, some more phenotypic information on the human KOs for SLC2A4– if available – would be very interesting to the reader. Given they are first reported human KOs for this gene, there is likely to be enormous interest in knowing more about glycaemic measures in these individuals. The T2D knowledge portal reports findings with thyroid function at the gene-level, rare variants with ApoA1 and blood pressure. The authors have a unique opportunity to explore and characterise these rare individuals. • We agree with the reviewer that this additional information would provide insights. We have added available glucose and HbA1c levels for SLC2A4 and the KOs in Supplementary Table 10. We were unable

to recontact the SLC2A4 knockout for additional characterization.

- Review 2: this is unfortunate but understandable. It would be good to include a statement saying so in text. Or simply, recontact was not possible for deeper phenotyping.

- We agree with the reviewer and added a statement to the text (lines 348-349):

• “As predicted by preclinical models⁵⁷, this participant, who was not available for RBG, has type 2 diabetes.”

Review 3: We agree this strengthens the manuscript

2.9.4 The findings in DENND1B KO are interesting, but at odds with published summary stats from UKBiobank. How do authors reconcile this?

• We are similarly puzzled that the two new DENND1B KO are not obese, in contrast to the report by Wallis et al. The difference may be due to variable penetrance, ancestral differences in weight gain, statistical limitations due to the rarity of DENND1B pLoF, or a combination of these factors.

• Wallis et al. does not show a formal burden analysis of the association between damaging or high confidence DENND1B loss-of-function variants and BMI with a corresponding effect size. Previously published analyses showed only nominally significant enrichment of combined damaging and protein truncating variants and a magnitude-free directionality of BMI (no reported effect size). Additionally, the paper provided supporting evidence of a single pediatric obesity case homozygous for a damaging missense variant (R501C)—the parents of this child were both heterozygous carriers and of normal BMI.

• In a formal burden analysis in PGR, we see no significant associations between DENND1B damaging/high confidence loss-of-function variants and BMI. Burden analyses in AZPheWAS also show no significant association between DENND1B rare damaging/loss of function variants and BMI. In addition, PGR contains two adult, high confidence loss-of-function homozygotes with normal BMI. Collectively, this suggests that DENND1B loss-of-function has little effect on BMI in adult humans.

- Review 2: Thank you. As throughout the manuscript, I think external validation of your findings is reassuring and should be leveraged. Could authors please add the agreeing information from AZPheWAS portal to text to head off any concerns linked to disagreement with the Wallis paper.

- We thank the reviewer for this suggestion. We have added the following to the text (lines 352 – 353):

o “In the UK Biobank, predicted damaging variants in DENND1B are nominally associated with higher BMI in European ancestries²⁶”

Review 3: We agree this acknowledgement of externally supportive associations strengthens the manuscript overall, and adds to the findings being reported here

2.9.6 Given the novel finding, which is at odds with the AstraZeneca phewas portal findings for LoF

DENND1B, it would be valuable to undertake replication with other data sets, UK Biobank, Genes & Health to confirm this findings.

• With apologies to the reviewer if we misunderstood their comment, we do not see significant associations between DENND1B LoF and BMI in the AZPheWAS portal.

- Review 2: With apologies to the authors, the results we were drawing their attention to were for the AZPheWAS “flexdmg” mask. In any case including AZPheWAS portal in Supp Table 5 helps orientate the reader to previously described results. This is not cross-referenced in the results paragraph on DENND1B – could you please do so and reassure readers about consistency?

- We thank the reviewer for this comment. We have checked the manuscript throughout for consistency in referencing, including Supplemental Table 5.

Review 3: This is much clearer and strengthens the manuscript

2.9.7 338: Associations with RXFP1 have been demonstrated in UK Biobank, data available on AZPheWAS portal. The novelty here is unclear – but would benefit from highlighting and meta-analyses with UKB.

• With apologies to the reviewer if we misunderstood their comment, we do not see significant associations in RXFP1 in the AZPheWAS portal.

• We believe that our report of RXFP1 KO is important because the healthy cardiovascular, metabolic, and reproductive profiles of these individuals strongly invalidate the longstanding pre-clinical observations that formed the foundation of past clinical trials. Decades of publications using rodent models suggested that RXFP1 is critical for nursing, fibrosis, and cardiac health, and metabolism. These rodent studies motivated many clinical trials testing RXFP1 activators in humans, including two large Phase 3 trials, to treat a broad spectrum of diseases, including impaired cervical ripening during childbirth, scleroderma, and acute heart failure. These trials have been unsuccessful.

- Review 2: On the AZPheWAS portal RXFP1 rare dmg variants are associated with scleroderma/systemic sclerosis ($p = 0.04$), heart failure/pulmonary oedema ($p = 5e-3$), and arterial stricture ($p = 1.2e-4$). We agree in principal with the author's contention that a healthy RXFP1 KO is of substantial interest, but the point remains that the p values the authors consider significant must be consistent across analyses – for example if they wish to highlight a p value of $p = 0.02$ in their results (as discussed in 2.7.2 above), then we would argue they should also draw the reader's attention to results for other findings elsewhere at a similar threshold and cannot dismiss these as not significant.

- We agree with the reviewer, including the point that the RXFP1 homLoF is of substantial interest. Space considerations preclude adding supplementary association statistics.

Review 3: This point really echoes that made above about the exome-wide Bonferroni significance threshold and we are of course happy to leave this to the editor's expertise and discretion

2.9.9 All these findings would be more impactful, if where possible, external validation or replication was undertaken.

Coherent and directionally stable associations with ptns in UK Biobank, Genes & Health which is also likely enriched for LoF variants (albeit in a smaller sample size), PRECISE etc.

• We have taken care to highlight examples of novel high confidence loss-of-function homozygotes that have not been

previously reported or characterized by other cohorts. The absence of KOs in other cohorts precludes external validation of the phenotypes associated (or not associated) with these KOs. Review2:: The logic of not replicating or meta-analysing novel findings where possible is not apparent to me. Particularly where p values are not entirely convincing.

- Review 2: If attempts have been made to identify LoF KOs in other cohorts and none were found this should be explicitly stated in text and will enhance the novelty of PGR.

- We agree that meta-analysis of novel findings is valuable wherever possible. In the text, we clarified that homLoF individuals for these genes have not previously been described.

Review 3: We agree this strengthens the manuscript

2.9.10 Similar to the previous section, it would be helpful to know how many GENE*trait associations (*model type (additive / recessive)) etc were performed across this section, and an associated Bonferroni-corrected p value presented to the reader. Presumably not just the genes included in the vignettes were included in analysis, but it is difficult to get a sense of what a meaningful p value is without knowing the number of tests performed.

• The associations in this section are generally hypothesis-based, motivated by the observations of KOs in genes for which KOs have not been previously identified. Where appropriate, we have added additional descriptions of statistical tests.

- Review 2: Our apologies, we struggled to find these in the text, please could the authors highlight which additional descriptions of tests have been included?

- Apologies to the reviewer. We intended to state that we added additional information, such as Bonferroni thresholds, to the other sections where they are relevant.

2.10.3 411: It is unconvincing that the thrifty genotype hypothesis explains why these (age-stratified) results are at odds with the Genes & Health results in Malawsky et al – I think more fleshing out of the proposed mechanism is needed here to convince the reader.

• Given the retrospective nature of our observations, we offer only a hypothesis rather than a mechanism supported by data. In the revised text, we have clarified that the thrifty gene hypothesis and the possibility of past resource-limited environments may explain only a subset of the protective phenotypes, such as lower LDL-C, triglycerides, BMI, and MI (lines 456-460):

o The mildly protective effects of autozygosity on some cardiometabolic biomarkers (LDL-C, triglycerides, BMI) and outcomes (MI) may represent reduced fitness of traits that were protective in past resource-limited environments, consistent with the “thrifty genotype” and related hypotheses that aim to explain high rates of obesity, diabetes, hypertension, and cardiovascular disease as a byproduct of adaptation to a divergent ancestral environment⁸³⁻⁸⁷.

- Review 2: Given the age-based heterogeneity it would be useful to readers to understand how that aspect squares off with the mechanism proposed here. - Please see our response above.

Referee #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

Referee #4

(Remarks to the Author)

Thank you for addressing concerns further.

I think article titles and labels used in the article should be considered carefully especially when addressing humans, but obviously in all cases.

homozygous loss-of-function humans is a very strange way to refer to humans, and can easily be misconstrued in all kinds of political and social narratives. This is a term used for certain genetic variants and to apply it to humans like this is not reasonable, as it explicitly makes it seem like these humans have lost functions (extremely eugenic terms) when their own analyses show that certain variant profiles are not always correlated to traits.

I think it is important that the authors think of a more thoughtful title for the study, and also a better label throughout the article. I see this as a major concern to fix, they should likely consult with a bioethicist or historian of science (for e.g. Anna Lewis, Tayyaba Jiwani).

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

POINT-BY-POINT RESPONSES

We thank the editor for helpful suggestions, instructions, and summarizing reviewers' comments.

We are grateful to the reviewers for their kind appraisal of our paper, as well as for many insightful and helpful comments. We have addressed every point they raised, and believe the manuscript is now much improved.

To elicit the most impact and broadest appeal of the manuscript's findings in the scientific community, particularly as highlighted by the reviewers, we propose a revised title for consideration:

Translational Insights from Human Knockouts in One-Third of Coding Genes

Referee #1 (Remarks to the Author):

This paper presents exome or whole genome sequencing data from a total of ~175K individuals which allows them to identify human gene knockouts in about 34% of the protein coding genome. Having information about the human phenotypic effect of gene knockouts is a critical part of understanding gene function and in this paper the researchers have greatly expanded the landscape of this type of data. My comments are focused on improving the manuscript. The results particularly to inform clinical trials are incredible and I believe that the paper will be a landmark in the field and the knowledge of human phenotypic effects of knocking out 30% of the genes in the genome cannot be understated.

- **We kindly thank the reviewer for the encouraging appraisal of our study and insightful suggestions on improvements.**

1.1 The population genetic analysis as they are presented provide little insight beyond the fact that these individuals are of South Asian ancestry (Fig 1 and 2). Would they be able to provide a PCA of just the individuals in their study, perhaps projected onto a West and East Eurasian basis? These were previously used to describe the Indian clinal populations first described in Reich et al, Nature, 2009 and provide a better view of the admixture history of these populations – you can use the references in Narasimhan Science 2019 to build these. In addition for fig 3a, would it be possible to separate the 1000 genomes SAS samples, to PJI to actually show the differential with the relevant 1000 genomes sub population?

- **We thank the reviewer for suggesting ways to make this section more insightful. To address the reviewer's point, we re-ordered the population analysis, which now presents Fst analysis, admixture, and PCA, for a clearer, stepwise presentation of South Asian ancestries and their genetic make-up. We also provided a PCA of the individuals in our study alongside South Asian 1000 Genomes (1KG-SAS) that we separated into its constituent ethnicities, as recommended by the reviewer. As the reviewer predicted, the new analysis projected PGR ethnicities along a geographic gradient of proximity to Europeans (PC1, Figure 3c,d). At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were the western populations of Pakistan, such as the Pathan and Balochi of PGR. The revised text includes the following information (lines 159-170):**

- o **Principal Component Analysis (PCA) quantified the heterogeneity of PGR subpopulations relative to each other and to non-PGR populations from the 1KG (Figure 3). African, East Asian, European, and South Asian ancestries separated into clusters, with South Asians in PGR following a gradient along PC2 from European ancestries to East Asians (Figure 3a,b). Consistent with the admixture analyses, Pathan and Balochi ethnicities clustered more closely with Europeans than did other PGR ethnicities, while a subset of Balochis clustered more closely with Africans. Among individuals without recent African admixture, PC1 distinguished PGR ethnicities along a gradient of proximity to Europeans, consistent with previous observations (Figure 3c,d)^{19,20}. At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-**

SAS, and at the other end were populations of western Pakistan, such as the Pathan and Balochi of PGR. PC2 and PC3 identified two additional clusters, one of which was predominantly Balochi and the other predominantly Punjabi, consistent with previous observations of substructure among the latter population²¹. We have added an additional panel to our figure (now Figure 4) that compares the distribution of F_{ROH} in PGR populations to 1KG SAS populations. Punjabis in PGR have slightly higher, albeit similar, levels of F_{ROH} compared to 1KG Punjabis.

1.2 What is interesting about the population genetics is not just the diversity of sampling and the ROH lengths but rather the how adding individuals of different ancestries within the PGR contribute to the overall number of knockouts seen. Could the authors report metrics like this sub-divided into different ancestries? For example, after adding several tens of thousands of genomes of Punjabi ancestry, how many new knockouts are seen in an additional genome of Kalash ancestry? This would help understanding how rare variation is distributed across subpopulations and how to prioritize future sampling efforts for this project as well as others, such as Genes and Health, and the UAE genome project (<https://www.medrxiv.org/content/10.1101/2025.09.12.25334546v1>).

- **The reviewer raises an interesting point that may inform future biobank strategies. To address this point, we provide the analyses and figures below. The new analyses indicate that the discovery power in PGR is greatest among the Sindhi, Punjabi, Balochi, and Kashmiri, and lowest in among the Urdu. New analyses include:**
 - o **Samples per gene KO by ethnicity (Extended Figure 9c).**
 - o **Number of KO genes by ethnicity (Extended Figure 9d).**
 - o **Number of Unique KO genes by ethnicity (Extended Figure 9e).**
 - o **Comparison of KO discovery by ethnicity controlling for sample size (Extended Figure 9f).**

1.3 There are a number of individuals with extremely high autozygosity, almost 40% - is this real? I wonder if these are from the whole genome data, which need additional filters if running bcftools roh, as the error rates on the lower coverage sequencing maybe higher and this causes my code in bcftools roh to create "blips".

- **We agree with the reviewer's observation about the autozygosity level. All the autozygosity calculations are from WES. The maximum autozygosity observed here (38.5%) is similar to the highest values observed for other cohorts (Narasimhan et al. 2016 Fig S4A; Ziyatdinov et al. 2023 Extended Data Figure 5).**

1.4 The association between Froh and reduced levels of MI but increased levels of TB and chronic kidney disease etc are interesting. But one issue is that Froh might be confounded with socioeconomic status or other cultural factors. Is there a way that they could split their populations into individual ancestry groups, Kashmiri, Punjabi etc. and rerun this analysis (hopefully they still have power with this enormous sample size)? As the relationship between Froh and environmental effects may be different across these groups this might provide additional evidence of this association (It is also possible that the association between Froh and environment maybe the same in all these groups).

- **We now include analyses subset by ethnicity and have observed minimal heterogeneity in effect sizes across ethnicity (Extended Data Figure 5).**

1.5 In the suppression analysis, looking at the number of autozygous and non-autozygous regions and the burden of LoFs in them, are they restricting to LoFs that do not contain indels? This would make the comparison with missense and synonymous variants more "fair", as they cannot be generated by indels, though it is equally likely that an indel could fall into an autozygous and non-autozygous region.

- **As the reviewer points out, indel and substitution rates may differ. For this reason, we compare rates of LoF in autozygous positions to rates of LoF in non-autozygous positions. Additionally, we now provide AAR calculations by variant type in Supplementary Table 9, showing that frameshift and stop gain variants generally show signs of recessive constraint in the same gene sets.**

1.6 Line 292 says that AAR can be used to identify genes subject to constraint. Are they able to show constraint at the level of individual genes with data of this size, or do they mean gene sets? If its gene sets they should modify the statement. If its individual genes, then this is indeed a powerful resource and they should make the list available publicly as this is currently now unavailable from datasets such as Gnomad.

- **The reviewer raises an important point: Demonstrating depletion at the level of individual genes would inform many domains of biology and therapeutics. Even with the vast number of KOs identified here, we are not powered to calculate constraint for individual genes. Our analysis is focused on gene sets. We clarify this point in the revised text (lines 309-312):**
 - o **To determine recessive constraint at the level of gene sets, we calculated percent depletion or enrichment of variants with at least one homozygous individual relative to allele frequency-matched synonymous variants, as described previously³⁸.**

1.7 Outside of AAR, there is an alternate way of computing suppressing involving subsampling of alleles at a particular frequency which we used along with AAR in the Narasimhan Science 2016 paper. While not guaranteed that these two methods should arrive at the exact same point estimate of suppression – they should be in the same ballpark (particularly if using only SNPs and not indels). Can they also show this alternate approach for the many suppression analysis they carry out?

- **At the reviewer's suggestion, we repeated all depletion analyses using the metric suggested by the reviewer. The results were indeed very similar to AAR. Because the reviewer's suggested analysis is more intuitive to understand, we feature it in the main text and have moved the AAR analysis entirely to supplemental. We continue to include both methods because the convergence of the results strengthens the conclusions, as the reviewer indicates.**
- **We have revised the text accordingly:**
 - o **To determine recessive constraint at the level of gene sets, we calculated percent depletion or enrichment of variants with at least one homozygous individual relative to allele frequency-matched synonymous variants, as described previously³⁸.**

1.8 If they combine their data with the genes and health study and Gnomad or finnGen, how many total genes do they see knocked out recessively in humans? Rather than just restricting their results to what they observe in the paper it would be useful to know the total number of genes that could be tolerated with having knockouts.

- **We thoroughly agree with the suggestion and its usefulness. By combining data from PGR and gnomAD, including South Asians, we identify knockouts in 7,954 genes. We have added this information to Supplementary Table 8 and the main text (lines 260-263):**
 - o **Of these 6,476 KO genes, 3,323 (51%) were unique to PGR relative to 761,616 non-South Asian individuals in gnomAD; 2,298 (35%) were unique to PGR relative to all 807,162 individuals in gnomAD (Figure 6b, Supplementary Table 8). There were 7,954 KO genes in PGR and gnomAD combined (Supplementary Table 8).**

1.9 I saw that there are 4 PRDM9 knockouts. Are any of them male? If so, do they have children (regardless of male or female)? This is a simple query but would add a lot of value to the evolutionary genomics community, as this is one of the few known mammalian speciation genes, and also a gene for which the knockout in humans is different from expectation from mouse models and a knockout male individual with children would suggest major differences in how recombination works in humans and mice.

- **We thank the reviewer for highlighting this example. One of the PRDM9 knockouts is male and has had children. The other 3 KOs are female and have had children. The significant discovery by Narasimhan et al. (Science, 2016) that recombination mechanisms differ substantially across species is further enhanced by our findings. We added these points to the revised text (lines 406-413):**
 - o **Reproduction. PRDM9 is a DNA-binding histone methyltransferase that determines sites of DNA recombination during meiosis across many mammalian species,**

including mice and humans⁶⁸⁻⁷⁰. *PRDM9* knockout mice exhibit errant patterns of recombination along with male and female infertility, illustrating the importance of this gene to the process of murine meiosis. A previous report described a female *PRDM9* KO who has had children³⁸. Extending this observation, PGR contains 3 females, all of whom have had children (range = 2-7 children; Supplementary Table 10). PGR also contains a male *PRDM9* KO who has had 4 children. The existence of both male and female human *PRDM9* KOs who were fertile indicates that *PRDM9*'s role in meiosis and fertility markedly diverges between humans and mice.

1.10 I see a huge number of individuals with knockouts in TLR genes – these have been identified to be under recent natural selection utilizing time series data from ancient DNA. The hypothesis is that these genes confer protection against pathogens and I wondered if these individuals have increased risk of tuberculosis or other disorders – is there power to look at this?

- **Following the reviewer's suggestion, we performed this analysis. We did not observe significant association between tuberculosis and pLoF in any of the TLR genes in PGR.**

1.11 What is the eye color of the individual with a *OCA2* knockout?

- **Eye color and skin pigmentation were not captured in PGR. We were unable to recontact this individual.**

1.12 What is the skin pigmentation of individuals with an *SLC24A5* knockout?

- **Please see above.**

1.13 What is the BMI / glucose activity of the individuals with an *OMA1* knockout? Mouse models <https://pmc.ncbi.nlm.nih.gov/articles/PMC3343468/> indicate major changes to metabolic activity and body weight. Similarly do these individual also have low(er) mtDNA copy number?

- **There are two *OMA1* knockouts. One has a BMI of 30 and random glucose of 60 mg/dl. The other knockout has a random glucose of 87 mg/dl, and BMI is not available. We have not analyzed mtDNA copy number, which is outside of the scope of this study.**

1.14 For individuals with homozygous knockouts in the HLA, are there any indications of increased chronic infection?

- **Upon the reviewer's suggestion, we analyzed infection risk and did not find significant association between infection and pLoF in HLA.**

1.15 Transcriptional adaptation

Recently meeting with Didier Stainier, I learned about a particular genetic mechanism that could explain the presence of some of these knockouts in humans – transcriptional adaptation (<https://www.nature.com/articles/s41586-024-08539-x#ref-CR3>). The idea is that mutations that do not completely ablate the protein but instead that cause NMD, actually can upregulate other paralogous proteins that may recover function. Thus for certain genes homozygous LoFs are better tolerated than homozygous missense. Is there a way that the authors could look at this phenomenon perhaps by computing ARR using missense vs lof in a gene by gene level?

- **We agree that human KO datasets will eventually present an opportunity to investigate the hypothesized mechanism. Studies would require sufficient KOs across members of paralogous genes (like the *DMD*-UTRN example) alongside mRNA expression data from relevant tissues to test concomitant and concordant changes in NMD and paralog expression. As PGR expands its size and scope of 'omics datasets, these critical biological questions may become tractable in the future.**

Referee #2 (Remarks to the Author):

Thank you for the opportunity to review this paper – it is clearly a study of monumental scale with potential to further our understanding of autozygosity and new gene effects through population-specific and enriched LoF variants.

- 223 - **We thank the reviewers for their positive overall assessment of the manuscript and**
224 **provide answers to their questions below.**

225
226 My co-reviewer and I have reviewed the manuscript and have the following major concerns as regards
227 the content, synthesis and therefore conclusions.

228
229 2.1 In the abstract and throughout – what is the relevance of comparing variants to non-south Asians in
230 gnomad? Given this is a deep variant-finding opportunity wouldn't it be more meaningful to include south
231 Asians in gnomad (or perhaps present both)?

- 232 - **We thank the reviewer for the suggestion. A focus on non-South Asians illustrates**
233 **discovery power in South Asia. To address the reviewer's suggestion, which increases the**
234 **significance and value of the comparisons, we have now revised Table 1 and Figure 6 to**
235 **include South Asians in gnomAD.**

236
237 2.2 Given that 2 individuals contributed to the finding of LRRK2 KOs and that there was no deeper
238 phenotyping of these individuals, the conclusion that KOs need renal monitoring is premature and
239 conclusions should be more circumspect.

- 240 - **We agree with the reviewer about the limitations of small numbers. To strengthen the**
241 **observations from the two knockouts, we discovered a third independent KO through**
242 **recontact and sequencing family members of an unrelated heterozygous loss-of-function**
243 **variant carrier. Similarly to the two knockouts from the PGR biobank, the third knockout**
244 **had low eGFR. As the reviewer rightfully suggested, we evaluated the exomes of the two**
245 **PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2,**
246 **COL4A3, COL4A4, COL4A5, HNF1B, UMOD, and APOL1). No damaging missense (REVEL**
247 **score >0.7) or HC pLoF variants were found.**
248 - **We believe that our human data stands as a capstone to an extensive series of pre-clinical**
249 **reports implicating LRRK2 in kidney function:**
 - 250 ○ **LRRK2 knockout mice exhibited impaired kidney function and hypertension (Tong**
251 **et al., 2010; Herzig et al., 2011).**
 - 252 ○ **LRRK2 knockout rats exhibited renal dysfunction (Gu et al., 2023).**
 - 253 ○ **LRRK2 inhibitors impaired kidney function in rodents (Andersen et al., 2018).**
 - 254 ○ **LRRK2 inhibitors caused kidney pathology in non-human primates (Jennings et al.,**
255 **2022).**
- 256 - **Our revised manuscript highlights these points (lines 349-359):**
 - 257 ○ **Recall-by-genotype of LRRK2 pLoF carriers and consenting family members**
258 **identified a 49-year-old female homozygous for LRRK2 splice donor variant,**
259 **c.1418+1G>T (Extended Data Figure 10a). This individual was the sole identified**
260 **KO in her family and had an eGFR of 51 ml/min, suggestive of kidney disease. Her**
261 **eGFR was lower than all 13 of her recruited heterozygous and non-carrier family**
262 **members (Extended Data Figure 10b). We evaluated the available exomes of the**
263 **two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1,**
264 **PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, or APOL1) and found no**
265 **damaging (REVEL score >0.7) or HC pLoF variants. Although the number of LRRK2**
266 **knockouts is low in our study, the observation of kidney dysfunction in these**
267 **individuals is consistent with, and provides human data to support, the adverse**
268 **renal phenotypes observed in rodent LRRK2 knockouts⁵¹⁻⁵³ and the preclinical**
269 **renal findings of LRRK2 inhibitor administration in rodents⁵⁴ and non-human**
270 **primates⁵⁵.**

271
272 2.3.1 115: "compendium of clinical data, including medical history, clinical measurements, and disease
273 status for most participants"; this is a little vague; it would be useful to explicitly state what data is
274 available and for whom. Typically this is contained in Table 1, describing the resource.

- 275 - **We agree with the reviewer. To address the reviewer's point and fully describe the**
276 **resource, we now list the specific clinical data that was available and analyzed in PGR. As**
277 **suggested by the reviewer and for completeness, we fit the information into the newly**
278 **expanded and revised Supplementary Table 1.**

2.3.2 On reading the methods it's quite unclear how exactly a participant comes to be enrolled in any of these nested case/control cohorts, and for the international reader, how this interacts with provision of healthcare in Pakistan. Would participants accessing private healthcare be more or less likely to be enrolled? Urban vs rural? It would be useful in the discussion to comment on these recruitment factors and potential sources of bias in case/control ascertainment.

- **We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514):**

- o ***Recruitment.* Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.**

2.4 Results: A flow chart of which participants are included in which part of each analysis would be very helpful to orientate the reader (this feeds into comments above) – it is essential to orient the reader to what PROMIS-PGR contains in terms of recruitment, genetic data generation (different sources and technologies), omic data similarly and finally what clinical data is collected and contributes to the compendium described.

- **We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now additionally includes a list of all biomarkers and binary outcomes analyzed in this study, along with the number of individuals for which the data is available. Additionally, for each analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of which individuals were included.**

2.5.1 In the population genetics portion of the manuscript, it is not clear why self-reported ethnicity was used for FST as opposed to the just-projected PCs? Was it not possible to project the sub-populations within PGR with sufficient granularity? This could be expanded upon for clarity. Is it possible to clarify for how many individuals this was done?

- **We agree with the reviewer that self-reported ethnicity has limitations. We prepared a new Figure 3 to visualize clustering of South Asian populations. Individuals of the same ethnicity tend to be close to each other in PCA. This observation indicates that ethnicity was an appropriate input to the FST analysis, which included all participants.**

2.5.2 I'm afraid as a non-expert in pop gen, the conclusions and take-away message from this portion of the manuscript were not apparent. Can this be better fleshed out and more explicit?

- **We thank the reviewer for suggesting ways to make this section more insightful. To address the reviewer's point, we re-ordered the population analysis, which now presents Fst analysis, admixture, and PCA, for a clearer, stepwise presentation of South Asian ancestries and their genetic make-up. We also provided a PCA of the individuals in our**

study alongside South Asian 1000 Genomes (1KG-SAS) that we separated into its constituent ethnicities, as recommended by the reviewer. As the reviewer predicted, the new analysis projected PGR ethnicities along a geographic gradient of proximity to Europeans (PC1, Figure 3c,d). At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were the western populations of Pakistan, such as the Pathan and Balochi of PGR. The revised text includes the following information (lines 159-170):

- Principal Component Analysis (PCA) quantified the heterogeneity of PGR subpopulations relative to each other and to non-PGR populations from the 1KG (Figure 3). African, East Asian, European, and South Asian ancestries separated into clusters, with South Asians in PGR following a gradient along PC2 from European ancestries to East Asians (Figure 3a,b). Consistent with the admixture analyses, Pathan and Balochi ethnicities clustered more closely with Europeans than did other PGR ethnicities, while a subset of Balochis clustered more closely with Africans. Among individuals without recent African admixture, PC1 distinguished PGR ethnicities along a gradient of proximity to Europeans, consistent with previous observations (Figure 3c,d)^{19,20}. At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were populations of western Pakistan, such as the Pathan and Balochi of PGR. PC2 and PC3 identified two additional clusters, one of which was predominantly Balochi and the other predominantly Punjabi, consistent with previous observations of substructure among the latter population²¹.

2.6 A major concern are the findings around Autozygosity

2.6.1 165: Perhaps worth explaining why you tested per 0.0625, i.e. additional degree of relatedness from parents being 1st cousins – makes intuitive sense when elaborated upon but not much without, to the non-expert reader

- **We report in units of 0.0625 as this is the theoretical expectation of F_{ROH} for the offspring of first-cousin parents. To clarify, we have now added this point to the main text (lines 179-182):**

- **We tested association of autozygosity (per F_{ROH} of 0.0625, the theoretical expectation of F_{ROH} for the offspring of first-cousin parents) with 16 quantitative and 18 binary phenotypes in PGR (Figure 5, Supplementary Table 3 and Supplementary Table 4).**

2.6.2 170-180: These are extremely interesting results, because they are at odds with existing literature e.g. Malawsky et al. in Cell 2023 who find quite the opposite associations with CVD, T2D etc. when using British Bangladeshis, British Pakistanis, UK Biobank British south Asians more generally, Europeans and finally in 23 & me. <https://doi.org/10.1016/j.cell.2023.08.028> In this paper, the authors perform a sensitivity analysis in a subset of highly-consanguineous individuals, reasoning that cultural or environmental factors are less likely to vary within this stratum, and that associations between increased FROH and traits restricted to this stratum are therefore more likely mediated by genetic (as opposed to environmental) effects. I wonder if a similar approach here might help to convince the reader further?

- **Based on the reviewer's suggestion, and similar to Malawsky et al al, we have performed an additional subsetting analysis. We have added the following to the text (lines 214-219):**

- **Subsetting to individuals with high autozygosity who were self-reported to have first-cousin parents eliminated the association between autozygosity and educational attainment, indicating that education as a trait may be confounded by social or economic factors (Extended Figure 7). For these individuals, however, the other associations were unchanged, suggesting that the observed effects are driven by autozygosity rather than socioeconomic confounders.**

2.6.3 Are there some groups (defined by ethnicity or PCA) among whom FROH is higher? Are autozygosity analyses of trait associations adjusted for self-reported ethnicity or PCs? If not, this needs to be justified.

- **All FROH association analyses are adjusted for genetic principal components. We now include analyses subset by ethnicity (Extended Data Figure 5) and have observed minimal heterogeneity in effect sizes across ethnicity. These results indicate that the associations are not driven by differences between groups.**

2.6.4 If rates of FROH and outcomes (eg T2D / MI) do vary between these ethnic groups, then demonstrating the association between T2D and FROH (for example) within each ancestry separately would help convince the reader that these associations are not driven by differences between these groups.

- **All FROH association analyses are adjusted for genetic principal components. We now include analyses subset by ethnicity (Extended Data Figure 5) and have observed minimal heterogeneity in effect sizes across ethnicity. These results indicate that the associations are not driven by differences between groups.**

2.6.5 A major confounding factor is survivorship bias and competing risks. How are cases ascertained, and controls, and does this occur longitudinally? The risk of survivorship bias and how this has been handled needs to be addressed. Was there a depletion of individuals with very high FROH score before recruitment was even possible? As the current state of knowledge suggests that high autozygosity is linked to myriad developmental and health conditions, were individuals with the highest FROH able to participate, live long enough to be recruited? Competing risks analyses may also be necessary to detangle this.

- **We agree with the reviewer that competing risks and survivorship bias analysis would have benefited this study. We currently do not have the longitudinal data to perform these analyses.**

2.6.6 190-200: Again these are extremely interesting results. Why was age 40 chosen as a cutoff? This is a typical young-onset threshold for T2D, but is perhaps a little early for MI. It would be more interesting to see these effect sizes plotted out by 10-year age bin (ie 18-30, 30-40, etc) to see how the association varies across the full breadth of ages (accepting power will be limited here), as opposed to a binary $</>40$ bin which leaves the reader wondering where the effect arises from. These divergent effects across the age range are very surprising and a bit more granularity would be most interesting to the reader.

- **We agree with the reviewer that the cutoff of 40 was arbitrary. We have now stratified our analyses into multiple age bins and describe the results in the revised text (lines 219-224):**
 - o **To determine if the associations were age-dependent, we binned across age groups (Extended Figure 8). Significant heterogeneity of effect was observed for educational attainment ($P=3.4e-08$), number of children ($P=3.2e-05$), having had children ($P=6.2e-06$), diabetes ($P=9.0e-06$), and hypertension ($P=1.9e-08$), with autozygosity trending towards increased risk of hypertension and diabetes in individuals <40 years and protection from hypertension and diabetes in individuals ≥ 50 years old (Extended Figure 8).**

2.6.7 Is there some sort of artificial clumping driving this association? Is FROH associated with age in some way, i.e., is the population becoming more autozygous over time, more FROH in young people, who do not yet have T2D/MI etc?

- **We see less autozygosity with age in PGR ($P=8.4e-286$), albeit with a small effect size ($\beta=-0.00026 F_{ROH}$ per year of age). In Figure 1 of Malawsky et al., the authors report a similar magnitude of effect between age and autozygosity in British Pakistanis, but not in British Bangladeshis, arguing that survivorship bias or ascertainment bias are unlikely to explain the association between age and autozygosity. Additionally, we have included both age and age*age as covariates in all association analyses.**

2.6.8 In previous studies ROH has been called in imputed whole genotype data – it seems here ROH is called just in coding regions? This needs to be explicitly discussed and justified.

- **We used WES data for ROH calling to maximize our sample size. The original paper describing the BCFtools ROH algorithm demonstrated that the algorithm produces highly**

accurate results when applied to exome data (Figure S5 in Narasimhan et al. 2016) and is now commonly used in the field.

2.6.9 For these autozygosity analyses, did effects differ between males and females? It would be nice to see these effects plotted out sex-stratified

- **We now include analyses subset by sex and have observed minimal heterogeneity in effect sizes (Extended Data Figure 6).**

2.6.10 How are these outcomes defined? Is there ongoing surveying, or do participants have to access healthcare in order to receive a downstream diagnosis? If so, it seems autozygosity (associated with less education) may be associated with limited access to healthcare, wealth, etc – and although models are adjusted for this, perhaps these people simply do not obtain their dx of MI, and the association simply reflects healthcare engagement?

- **We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514):**

- o **Recruitment.** Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.

2.6.11 Finally, the discussion section is lacking depth as to why these findings diverge so entirely with previous studies. There is not the uncertainty on the effects of autozygosity as the authors suggest. There is in fact a consensus on the increased risk of ill-health. So the protective effects observed do require considered sensitivity testing and perhaps even external validation.

- **We agree with the reviewer that the lower diabetes and MI risk in our study contrast with conventional expectations. We do however observe many associations consistent with previous reports (Clark 2019): higher rates of childlessness, fewer children, lower educational attainment, lower tobacco usage, shorter stature, lower weight, and lower cholesterol. PGR is the largest consanguineous cohort in the world and can detect associations between autozygosity and traits that are not present in other cohorts in sufficient numbers, such as myocardial infarction. For these reasons, external validation will require the expansion of other highly autozygous cohorts to similar size and case/control composition. We have performed sensitivity analyses within PGR, such as subsetting by ethnicity.**

2.7 Biomarker burden

2.7.1 Again a flow-chart showing the extent of clinical data, proteomic data, etc, would be really helpful to orientate the reader.

- **We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now**

501 additionally includes a list of all biomarkers and binary outcomes analyzed in this study,
502 along with the number of individuals for which the data is available. Additionally, for each
503 analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of
504 which individuals were included.
505

506 2.7.2 225: These are extremely interesting and well-presented results. To orientate the reader, it would be
507 useful to have a short section explaining how many GENE*phenotype associations were tested, and what
508 the associated Bonferroni-corrected p value would therefore be (for example, the “lower random glucose
509 ($\beta=-0.05$, $P=2.3e-02$)” seems unlikely to be anything other than nominally significant, but it is not clear to
510 the reader which other association would survive multiple testing correction.

- 511 - **We thank the reviewer for the kind commentary. Regarding multiple testing, we perform an**
512 **initial PheWAS where we report all gene-biomarker associations that are significant at an**
513 **exome-wide Bonferroni threshold of $P<9.9e-08$ in Supplementary Table 5. Elsewhere, when**
514 **following up on specific genes we report additional, nominal associations. We have**
515 **specified in the text that these associations are nominally significant at this threshold.**
516

517 2.7.3 It would also be useful to provide a summary table of how many individuals had available data for
518 each phenotype discussed in the main text (this feeds back into a previous comment on study data
519 presentation and flow charts)

- 520 - **Due to space constraints, we could not include this information in the main text. We have**
521 **added this data to Supplementary Table 1.**
522

523 2.7.4 It would be helpful in this section to compare to existing pLoF burden results identified either
524 through eg AZphewas portal or GENEASS to orientate the reader to the novelty of these findings, as
525 they are presented.

- 526 - **We agree with the reviewer that including UKBB data alongside PGR data would facilitate**
527 **comparison. As stated in our revised text, include associations from the AstraZeneca**
528 **PheWAS browser in Supplementary Table 5 (lines236-237):**
 - 529 ○ **For comparison, we include associations from the AstraZeneca PheWAS**
530 **browser^{26,27} (Supplementary Table 5).**

531 532 2.7.5 229: It would be useful to compare to G&H recent effort characterising pLoF human KOs, recent
533 preprint: <https://doi.org/10.1101/2025.06.05.25329068>. For example the Genes and Health authors report
534 2991 homozygous carriers of pLoF in a sample size of ~44k individuals and this could be compared to
535 findings from PROMIS-PGR for a better understanding of the landscape of genomic studies in SAS
536 populations and sub-populations.

- 537 - **Based on the reviewer’s feedback, we have included a comparison to South Asians in**
538 **gnomAD using the same set of pLoF filters that were applied to PGR. gnomAD South**
539 **Asians is comparable to Genes and Health in size and number of KOs: gnomAD has**
540 **45,546 South Asians, from which KOs are present in 3,022 genes. We have not analyzed**
541 **the Genes & Health cohort as this preprint is not yet peer-reviewed and could be subject to**
542 **change.**
543

544 2.7.6 217: Some of these associations with HBB, including iron and lipid traits, in addition to bilirubin, are
545 reported in the Kim et al paper above and should be discussed in this context.

- 546 - **Please see above.**
547

548 2.8 Analysis of LOF constraint

549 2.8.1 261: it would be helpful to briefly define where these approved drugs are from (FDA? Some other
550 authority?) to save the reader diving to the methods

- 551 - **We downloaded the data used by Minikel et al 2024, which used data from Citeline**
552 **Pharmaprojects. We added this citation to the text.**
553

554 2.8.2 285-287: Are these AAFs significant? P values are missing, and no SEs are presented.

- **We thank the reviewer for highlighting this oversight. We have performed additional statistical tests and added p-values to the main text. Please also see our response to 1.7 above.**

2.8.3 297: Were synonymous variants “modestly depleted” if $p = 0.225$ (line 299)? And doesn’t no depletion of synonymous variants in fact strengthen your argument, ie, as a negative control?

- **We thank the reviewer for this close reading. Our intention was to state that synonymous variants were only nominally significantly depleted in a small number of gene sets, but not gene sets that one would expect to be subject to recessive constraint – e.g. genes essential *in vitro* or associated with autosomal recessive conditions. We now describe a different metric in the main text and present AAR in supplemental analysis.**

2.9 Human KO highlights

2.9.1 For recall samples, was sequencing repeated to confirm the initial sequencing/genotyping was correct?

- **Yes, confirmatory Sanger sequencing was conducted for recall samples**

2.9.2 310: The LRRK2 findings are certainly most interesting, but I think the strength of assertion around need for renal monitoring should be couched firmly in the $N=2$ here, and perhaps more detail and deeper phenotyping on these 2. To my mind this argument would be considerably strengthened by checking for known genetic causes of CKD, eg PCKD, and perhaps calculating a CKD PGS. Ie, in the absence of other known genetic constructs causing CKD, we may be more inclined to believe it is really all mediated by LRRK2. Is further phenotyping possible (eg ACR, USS)? We appreciate this may not be feasible and are not suggesting this would be required for publication. However, the fullest picture of these individuals phenotypes would be needed to persuade readers of this journal that LRRK2 was the underlying cause here.

- **We agree with the reviewer about the limitations of small numbers. To strengthen the observations from the two knockouts, we discovered a third independent KO through recontact and sequencing family members of an unrelated heterozygous loss-of-function variant carrier. Similarly to the two knockouts from the PGR biobank, the third knockout had low eGFR. As the reviewer rightfully suggested, we evaluated the exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (PKD1, PKD2, COL4A3, COL4A4, COL4A5, HNF1B, UMOD, and APOL1). No damaging missense (REVEL score >0.7) or HC pLoF variants were found.**
- **We believe that our human data stands as a capstone to an extensive series of pre-clinical reports implicating LRRK2 in kidney function:**
 - o LRRK2 knockout mice exhibited impaired kidney function and hypertension (Tong et al., 2010; Herzig et al., 2011).
 - o LRRK2 knockout rats exhibited renal dysfunction (Gu et al., 2023).
 - o LRRK2 inhibitors impaired kidney function in rodents (Andersen et al., 2018).
 - o LRRK2 inhibitors caused kidney pathology in non-human primates (Jennings et al., 2022).
- **Our revised manuscript highlights these points (lines 349-359):**
 - o Recall-by-genotype of *LRRK2* pLoF carriers and consenting family members identified a 49-year-old female homozygous for *LRRK2* splice donor variant, c.1418+1G>T (Extended Data Figure 10a). This individual was the sole identified KO in her family and had an eGFR of 51 ml/min, suggestive of kidney disease. Her eGFR was lower than all 13 of her recruited heterozygous and non-carrier family members (Extended Data Figure 10b). We evaluated the available exomes of the two PGR KOs for variants (with MAF <1%) in known kidney disease genes (*PKD1*, *PKD2*, *COL4A3*, *COL4A4*, *COL4A5*, *HNF1B*, *UMOD*, or *APOL1*) and found no damaging (REVEL score >0.7) or HC pLoF variants. Although the number of *LRRK2* knockouts is low in our study, the observation of kidney dysfunction in these individuals is consistent with, and provides human data to support, the adverse renal phenotypes observed in rodent *LRRK2* knockouts⁵¹⁻⁵³ and the preclinical

610 **renal findings of LRRK2 inhibitor administration in rodents⁵⁴ and non-human**
611 **primates⁵⁵.**
612

613 2.9.3 326: Similarly, some more phenotypic information on the human KOs for SLC2A4– if available –
614 would be very interesting to the reader. Given they are first reported human KOs for this gene, there is
615 likely to be enormous interest in knowing more about glycaemic measures in these individuals. The T2D
616 knowledge portal reports findings with thyroid function at the gene-level, rare variants with ApoA1 and
617 blood pressure. The authors have a unique opportunity to explore and characterise these rare individuals.

- 618 - **We agree with the reviewer that this additional information would provide insights. We**
619 **have added available glucose and HbA1c levels for SLC2A4 and the KOs in Supplementary**
620 **Table 10. We were unable to recontact the SLC2A4 knockout for additional**
621 **characterization.**
622

623 2.9.4 The findings in DENND1B KOs are interesting, but at odds with published summary stats from
624 UKBiobank. How do authors reconcile this?

- 625 - **We are similarly puzzled that the two new DENND1B KOs are not obese, in contrast to the**
626 **report by Wallis et al. The difference may be due to variable penetrance, ancestral**
627 **differences in weight gain, statistical limitations due to the rarity of DENND1B pLoF, or a**
628 **combination of these factors.**
629 - **Wallis et al. does not show a formal burden analysis of the association between damaging**
630 **or high confidence DENND1B loss-of-function variants and BMI with a corresponding**
631 **effect size. Previously published analyses showed only nominally significant enrichment**
632 **of combined damaging and protein truncating variants and a magnitude-free directionality**
633 **of BMI (no reported effect size). Additionally, the paper provided supporting evidence of a**
634 **single pediatric obesity case homozygous for a damaging missense variant (R501C)—the**
635 **parents of this child were both heterozygous carriers and of normal BMI.**
636 - **In a formal burden analysis in PGR, we see no significant associations between DENND1B**
637 **damaging/high confidence loss-of-function variants and BMI. Burden analyses in**
638 **AZPheWAS also show no significant association between DENND1B rare damaging/loss of**
639 **function variants and BMI. In addition, PGR contains two adult, high confidence loss-of-**
640 **function homozygotes with normal BMI. Collectively, this suggests that DENND1B loss-of-**
641 **function has little effect on BMI in adult humans.**
642

643 2.9.5 336: It would be helpful to present some measure of range here (is this driven by a single BMI of
644 45?); or perhaps even all 4 BMIs, given the small number.

- 645 - **All BMIs are included in Supplementary Table 10 and range from 28-42. We have now**
646 **specifically included this information in the text for clarity.**
647

648 2.9.6 Given the novel finding, which is at odds with the AstraZeneca phewas portal findings for LoF
649 DNND1B, it would be valuable to undertake replication with other data sets, UK Biobank, Genes & Health
650 to confirm this findings.

- 651 - **With apologies to the reviewer if we misunderstood their comment, we do not see**
652 **significant associations between DENND1B LoF and BMI in the AZPheWAS portal.**
653

654 2.9.7 338: Associations with RXFP1 have been demonstrated in UK Biobank, data available on
655 AZphewas portal. The novelty here is unclear – but would benefit from highlighting and meta-analyses
656 with UKB.

- 657 - **With apologies to the reviewer if we misunderstood their comment, we do not see**
658 **significant associations in RXFP1 in the AZPheWAS portal.**
659 - **We believe that our report of RXFP1 KOs is important because the healthy cardiovascular,**
660 **metabolic, and reproductive profiles of these individuals strongly invalidate the**
661 **longstanding pre-clinical observations that formed the foundation of past clinical trials.**
662 **Decades of publications using rodent models suggested that RXFP1 is critical for nursing,**
663 **fibrosis, and cardiac health, and metabolism. These rodent studies motivated many**
664 **clinical trials testing RXFP1 activators in humans, including two large Phase 3 trials, to**

treat a broad spectrum of diseases, including impaired cervical ripening during childbirth, scleroderma, and acute heart failure. These trials have been unsuccessful.

2.9.8 379: TRPM8 – bigger previously reported signals have been for lipid traits and MI, did this not replicate here?

- **With apologies to the reviewer if we misunderstood their comment, we have not identified any previously published significant associations by GWAS or ExWAS between variants in TRPM8 and lipid traits or myocardial infarction. We also do not see any associations in the PGR biobank or in the recall by genotype study. However, there have been multiple studies reporting variants in the TRPM8 locus for migraine (Chasman et al., 2011; Freilinger et al., 2012), including variants associated with partial reductions in TRPM8 expression that reduce the risk for migraine and cold sensation (Gavva et al., 2019). While PGR lacks data on migraine-related phenotypes, our findings indicate that complete loss of TRPM8 is tolerated. Therefore, pharmacological inhibition of TRPM8 for migraine may be safe. We elaborate on these points in the revised text (lines 424-427):**

- o **In GWAS studies, *TPRM8* has been linked to migraine⁷²⁻⁷⁴, with the directionality suggesting that reduction in TRPM8 reduces risk of migraine⁶⁸. While PGR lacks migraine-related phenotypes, our findings indicate that complete loss of TRPM8 is tolerated and, therefore, pharmacological inhibition of TRPM8 for migraine may be safe.**

2.9.9 All these findings would be more impactful, if where possible, external validation or replication was undertaken. Coherent and directionally stable associations with ptns in UK Biobank, Genes & Health which is also likely enriched for LoF variants (albeit in a smaller sample size), PRECISE etc.

- **We have taken care to highlight examples of novel high confidence loss-of-function homozygotes that have not been previously reported or characterized by other cohorts. The absence of KOs in other cohorts precludes external validation of the phenotypes associated (or not associated) with these KOs.**

2.9.10 Similar to the previous section, it would be helpful to know how many GENE*trait associations (*model type (additive / recessive)) etc were performed across this section, and an associated Bonferroni-corrected p value presented to the reader. Presumably not just the genes included in the vignettes were included in analysis, but it is difficult to get a sense of what a meaningful p value is without knowing the number of tests performed.

- **The associations in this section are generally hypothesis-based, motivated by the observations of KOs in genes for which KOs have not been previously identified. Where appropriate, we have added additional descriptions of statistical tests.**

2.10 Discussion

2.10.1 391: as above, the actual median FROH is not within the range you describe

- **We thank the reviewer for catching this error. It should state that the mean value (0.034) falls within the range. We have corrected this in the text (lines 176-178):**
 - o **The mean F_{ROH} in PGR was within the expected ranges of first ($F_{ROH} = 0.0625$) to second cousin ($F_{ROH} = 0.016$) level of parental relatedness and was consistent with the high percentage of individuals (30.6%) who self-report being born to parents who were first cousins.**

2.10.2 409: You describe the inconsistencies “Across previous studies”, but actually other studies that you describe were all directionally consistent with autozygosity being associated with deleterious phenotype (consistent with inbreeding depression).

- **We thank the reviewer for pointing out the lack of clarity in our text. We intended to highlight inconsistencies between PGR and some previous studies (e.g. diabetes in PGR vs. G&H), not across previous studies. We have revised the text (lines 455-456):**
 - o **The age-dependent effects of autozygosity on health may partially explain the inconsistencies for some of the findings in our studies relative to other studies.**

2.10.3 411: It is unconvincing that the thrifty genotype hypothesis explains why these (age-stratified) results are at odds with the Genes & Health results in Malawsky et al – I think more fleshing out of the proposed mechanism is needed here to convince the reader.

- **Given the retrospective nature of our observations, we offer only a hypothesis rather than a mechanism supported by data. In the revised text, we have clarified that the thrifty gene hypothesis and the possibility of past resource-limited environments may explain only a subset of the protective phenotypes, such as lower LDL-C, triglycerides, BMI, and MI (lines 456-460):**

- o **The mildly protective effects of autozygosity on some cardiometabolic biomarkers (LDL-C, triglycerides, BMI) and outcomes (MI) may represent reduced fitness of traits that were protective in past resource-limited environments, consistent with the “thrifty genotype” and related hypotheses that aim to explain high rates of obesity, diabetes, hypertension, and cardiovascular disease as a byproduct of adaptation to a divergent ancestral environment⁸³⁻⁸⁷.**

2.10.4 416: I'm a little unclear on why the HBB-lipids associations have been determined to be causal, but other HBB-trait associations (e.g. MI) are not described as causal. Perhaps I have missed some key analysis step here? The HBB associations with iron, red cell biology and MI are strong and convincing, but I do not follow how you determine them to be causal?

- **We thank the reviewer for pointing out that we may have overstated the causal relationship between HBB, iron, red cell biology, and MI. We have modified the language to suggest a relationship whose mechanism is yet to be established (lines 461-465):**
 - o **PGR is powered to reveal directional relationships between genetic variants and phenotypes, including circulating biomarkers. Such relationships were established between *TIMD4* pLoF and circulating lipids, and between *HBB*, circulating lipids and apolipoproteins. In PGR, *HBB* pLoF is associated with lower risk of myocardial infarction. This suggests that *HBB* pLoF could modulate heart attack risk through hemoglobin or its constituent iron, or through lipids and lipoproteins⁸⁸.**

2.10.5 As below, it would be useful to include some comment on participant recruitment methods, and potential sources of bias associated with this.

- **We thank the reviewer for highlighting ways to improve our description of the study. Participant enrollment occurs in hospital settings (emergency wards, outpatient clinics, and in wards) where dedicated and trained Research Medical Officers identified eligible individuals based on pre-defined inclusion criteria, then obtained informed consent. Recruitment took place in both private and public hospitals, with the latter comprising the majority due to their large patient volume. Participants were enrolled from urban and suburban hospitals that also serve nearby rural populations. Uniform eligibility criteria and standardized recruitment procedures were applied across all sites. Case and control participants were enrolled from the same hospital to minimize socio-economic bias. We have now added this information to the revised Methods (lines 506-514):**
 - o **Recruitment.** Cases and controls were recruited in both public and private hospitals across Pakistan. A breakdown of cases and controls is outlined in Supplementary Table 1. Phenotypes were defined based on confirmed clinical diagnoses established by primary physicians, supported by relevant laboratory investigations and standardized diagnostic criteria. Participants were identified at the time of hospital presentation and enrolled following diagnostic confirmation and written and oral informed consent was obtained. Cases were identified according to predefined inclusion and exclusion criteria. Controls were selected from the same hospitals, typically unrelated or distantly related case attendants who met control criteria. To minimize selection bias, enrollment was carried out across both urban and suburban hospitals that also serve surrounding rural populations.

2.11 Methods

2.11.1 509-518: Did you set a specific minimum length for a ROH eg 1Mb? How was a ROH defined? I'm a little unclear after reading this section.

- **We identified regions of ROH using the widely used BCFtools roh algorithm (Narasimhan et al. 2016). The algorithm uses a hidden Markov model that models the probability of transitioning from ROH to non-ROH regions using parameters such as known position-specific recombination rates and population allele frequencies. As such, it does not require a user-defined minimum length threshold.**

2.11.2 610: Your FROH distribution looks skewed (as would be expected) – did you log-transform it prior to logistic regression, and if so why/why not?

- **We thank the reviewer for their close attention to the data. For consistency with previous publications in the field (Clark et al. 2019, Malawsky et al. 2023), we have not transformed the FROH data prior to logistic regression. We do perform rank inverse normal transformation of quantitative traits prior to running linear regression.**

Minor comments ;

2.12 Intro 98: The first line feels a little at odds with the rest of the paper in setting the scene – much of this paper has no relevance to drug discovery?

- **We agree with the reviewer that the paper does not focus exclusively on drug targets. At the same time, findings from human genetics are directly translatable to drug discovery.**

2.13 After reaching the end of the introduction and starting results I was a little unclear on what exactly the aims / objectives of this study are – perhaps a brief sentence to summarise might help to orientate the reader ahead of the analyses being presented in the results

- **We thank the reviewer for this observation. As the reviewer suggested, we have revised the text to summarize the aims/objectives of the study (lines 126-129):**
 - o **This report summarizes the regional genetic diversity of the PGR, the impact of autozygosity on disease risk, the discovery of human knockouts for one-third of coding genes, and examples of recall-by-genotype studies that deepen insights into gene function.**

2.14 103: I think for the non-expert reader it is necessary to expand on why autozygous populations are enriched for recessive / pLoF KO effects when explaining they can “uniquely power genetic discovery” ie spell out that same allele from common ancestor can arrive via each parent / IBD. A bit more detail on autozygosity for the non-expert reader, not necessarily obvious why they are more likely to have 2 copies, perhaps some sort of diagram and explanation of identity-by-descent

- **We thank the reviewer for suggesting this improvement. As the reviewer suggested, we have revised the text to elaborate (lines 114-118):**
 - o **Consanguinity in PGR is driven by a high proportion of individuals who are the offspring of first- or second-cousin unions. These individuals are more likely to be homozygous for (have two alleles of) otherwise rare alleles. Therefore, relative to non-consanguineous cohorts, PGR is powered to discover individuals who are loss-of-function homozygotes, also known as naturally occurring human gene “knockouts”.**

2.15 160: A minor but concerning observation, but the median FROH value of 0.015 doesn't fall within your stated ranges 0.016 – 0.0625 as you say it does.

- **We thank the reviewer for catching this error. It should state that the mean value (0.034) falls within the range. We have corrected this in the text (lines 176-179):**
 - o **The mean F_{ROH} in PGR was within the expected ranges of first ($F_{ROH} = 0.0625$) to second cousin ($F_{ROH} = 0.016$) level of parental relatedness and was consistent with the high percentage of individuals (30.6%) who self-report being born to parents who were first cousins.**

Referee #3 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

- **We thank the reviewer for contributing to the evaluation of our manuscript.**

Referee #4 (Remarks to the Author):

Koch et al present a nice study of 173,303 individuals in Pakistan. They present interesting findings relevant for biomedicine using analyses of autozygosity, LOH variants and human knockouts, and the analyses are generally well done and innovative and the paper is well written.

- **We thank the reviewer for the positive appraisal of our study.**

I have few major concerns:

3.1 For the FROH analysis, Malwasky et al 2023 clearly explain why such analyses can be confounded, and show that restricting analysis to offspring of first cousins is an effective way of reducing confounding due to social/environmental correlates of FROH. Why has no effort to do this been done? The authors should repeat this analysis using such an approach and including covariates for socioenvironmental variables that they have available. They also do not make any mention of multiple testing correction in this analysis and it is not clear if this has been done.

- **Based on the reviewer's suggestion, and similar to Malawsky et al al, we have performed an additional subsetting analysis. We have added the following to the text (lines 214-219):**
 - o **Subsetting to individuals with high autozygosity who were self-reported to have first-cousin parents eliminated the association between autozygosity and educational attainment, indicating that education as a trait may be confounded by social or economic factors (Extended Figure 7). For these individuals, however, the other associations were unchanged, suggesting that the observed effects are driven by autozygosity rather than socioeconomic confounders.**
- **We have added the threshold for multiple testing (liens 182-185):**
 - o **After multiple testing correction (threshold of $P=1.5e-03$), F_{ROH} was significantly associated with reduced odds of having children ($OR=0.78$, $P=4.8e-32$), fewer total children ($\beta=-0.07$, $P=1.2e-10$), fewer years of educational attainment ($\beta=-0.04$, $P=2.0e-38$), and lower odds of having ever used tobacco ($OR=0.89$, $P=3.1e-30$).**

3.2 The authors state that "Causal relationships were established between TIMD4 pLoF and circulating lipids, and between HBB and circulating lipids and apolipoproteins." It's not clear why causal is the right word here and they should edit, they have simply shown associations

- **We thank the reviewer for pointing out that we may have overstated the causal relationship between HBB, iron, red cell biology, and MI. We have modified the language to suggest a relationship whose mechanism is yet to be established (lines 461-465):**
 - o **PGR is powered to reveal directional relationships between genetic variants and phenotypes, including circulating biomarkers. Such relationships were established between *TIMD4* pLoF and circulating lipids, and between *HBB*, circulating lipids and apolipoproteins. In PGR, *HBB* pLoF is associated with lower risk of myocardial infarction. This suggests that *HBB* pLoF could modulate heart attack risk through hemoglobin or its constituent iron, or through lipids and lipoproteins⁸⁸.**

3.3 For this to be a resource, the availability of the dataset for researchers should be made clear, and it is not clear how researchers can access the dataset for other studies. This is a major concern. The authors state "Whole Genome Sequencing data is available for non-profit use at phs001569.v1.p1." but this only leads to the PROMIS dataset which is a tiny fraction of the data presented and analyzed. They also state that "All other academic requests to access relevant Center for Non-Communicable Diseases (CNCD) data or samples should be sent to info@cncdpk.com." How will it be ensured that the data is truly available to interested researchers and fair ethical practices will be applied for those requesting access?

- **We agree with the reviewers that the PGR is a valuable resource. We provide a revised data availability statement with additional transparency and information regarding data access (lines 1082-1091):**
 - o **PGR. Supplementary Tables include summary statistics for autozygosity, summary statistics for biomarker burden tests, and a full exome-wide list of homozygous loss-of-function (KO) variants listed by chromosome, position, reference allele, alternative allele, and homozygous carrier count. Additional summary statistics for all biomarker burden tests, including non-significant associations not reported**

here, have been uploaded to figshare. WGS data are available in dbGaP (accession phs001569.v1.p1). A subset of WES data is available in dbGaP (accession phs000917.v1.p1). Individual-level exomes and phenotype data are stored at CNCD, with access restricted to protect participant data. Academic researchers may apply for access to individual-level data by contacting the CNCD at info@cncdpk.com. Each request will require a detailed research plan and data privacy plan, including an agreement to maintain de-identification of PGR participants.

3.4 What plans do the authors have for community engagement and giving results back to participants?

- **The sequencing data collected in this study was for research purposes. The CNCD is currently building a CLIA-certified sequencing center along with a genetic counselling center for the ethical return of clinically actionable genetic data to participants and their physicians in Pakistan.**

Response to reviewers

We thank the reviewers and editor for their helpful comments and guidance on improving our manuscript. We were pleased that most of our revisions addressed the reviewers' suggestions. For brevity, we list and reply below only to the remaining questions from reviewers.

Editor's comments

Your manuscript, "Translational Insights from Human Knockouts in One-Third of Coding Genes", has now been seen again by 4 referees. You will see from their comments below that while they find your work strengthened, some important points are raised. We are interested in the possibility of publishing your study in Nature, but would like to consider your response to these concerns in the form of a revised manuscript before we make a final decision on publication.

We therefore invite you to revise and resubmit your manuscript, taking into account the points raised by the reviewers. Please highlight all changes in the manuscript text file.

- **Thank you for the opportunity to submit another revision. We address the reviewers' comments in our point-by-point response below and in the attached manuscript revision.**

I want to specifically draw your attention to the concerns raised by reviewer #4 regarding data availability. This is a concern of utmost importance to us and the data availability statement in its current form still does not address all the points that we had raised previously. It still needs to be clearer about which dataset and subset of the data can be found where (it is very confusing in its current form), and please also include a response time to access requests. We further request that you provide us with a brief workflow of what happens when a researcher requests access to the data from you (this can be in the form of a Gantt chart or similar that you include in your cover letter). We feel that ensuring that the research community can make use of the dataset presented here is vital, and we consider this a matter of reproducibility and research integrity that we will also monitor post publication.

- **We have further clarified the data availability statement, include a Gantt chart with our resubmission, and specified a response time for data requests.**

Please also remove the reference to "human knockouts" as per reviewer #4.

- **We welcome the opportunity to move the field away from the term "human knockout." As described in our revised introduction, we propose replacing the term with the more precise description "homozygous loss-of-function (homLoF)" in the context of humans.**

You will also need to make some editorial changes to your paper so that it is as brief as possible and complies with our Guide to Authors. We also strongly suggest that your revised manuscript has tracked changes, which is increasingly requested by referees to aid in their re-review.

- **We edited the manuscript to comply with the Guide to Authors. Edits include trimming for length and ensuring figures meet size requirements.**

Referee #1, review by Vagheesh M. Narasimhan (I sign all my reviews).

1.1 My last concern is that the current population genetic figures (1, 2, and 3) present quite a lot of detail making them difficult to interpret. I would combine the relevant information into a single figure, highlight where the samples are from and use the same color coding on the PCA, admixture (the admixture plot doesn't separate the ancestries either) and location plots. I would also include figure 4a in this plot. and again follow the coloring with the other plots above. Otherwise the reviewers have addressed all my

concerns - this is a remarkable paper and the end of an almost 10 year effort to understand the impact of gene knockouts on 1/3 of the human genome.

- **We thank the reviewer for his positive appraisal of our study**
- **We agree with the reviewer's suggestions for improving presentation of the population genetic data. As recommended by the reviewer:**
 - o **We have combined previous Figure 1 (map of Pakistan) and previous Figure 3 (PCA) into a revised Figure 1.**
 - o **Due to space considerations, we have moved previous Figure 2 (Admixture) to Extended Data Figure 3 and re-numbered subsequent figures.**
 - o **We have used a consistent color scheme to represent self-reported ethnicities and 1KG populations across the PCA, F_{ROH} distribution, and homLoF statistics figures. The admixture plots use a different color scheme since there is not a one-to-one mapping between self-reported ethnicity and admixture components.**

1.2 If at all possible performing sperm typing on the male prdm9 knockout would be a remarkable addition to the evolutionary genomics community, because we would understand how localization of recombination happens in humans in the absence of prdm9. The process is relatively inexpensive but I'm not sure if the individual will consent for additional sampling.

- **We agree with the reviewer that sperm typing of the male PRDM9 homLoF would be informative. We were unable to get consent for this sampling.**

Referee #2 (Remarks to the Author):

Thank you for the opportunity to review this paper – it is clearly a study of monumental scale with potential to further our understanding of autozygosity and new gene effects through population-specific and enriched LoF variants.

• We thank the reviewers for their positive overall assessment of the manuscript and provide answers to their questions below.

We thank the authors for another opportunity to review this excellent manuscript, which has been substantially improved during these revisions.

- **We thank the reviewer for their positive assessment of our manuscript and for helping strengthen our paper. We also appreciate that the reviewer noted that most of our edits improved the manuscript. Below, we will focus only on the remaining questions.**

2.4 Results: A flow chart of which participants are included in which part of each analysis would be very helpful to orientate the reader (this feeds into comments above) – it is essential to orient the reader to what PROMIS-PGR contains in terms of recruitment, genetic data generation (different sources and technologies), omic data similarly and finally what clinical data is collected and contributes to the compendium described.

• We thank the reviewer for the suggestion. To clarify, we have added a structured breakdown of PGR and our study to Supplementary Table 1. Supplementary Table 1 now additionally includes a list of all biomarkers and binary outcomes analyzed in this study, along with the number of individuals for which the data is available. Additionally, for each analysis presented in the paper, we have added to Supplementary Table 1 a breakdown of which individuals were included.

- *Review 2: This is very helpful to orientate the reader. Please include the WES technology used and also list what other sources of genomic data exist other than WES – as the total number of PGR samples that contributed to coding variant counts was 173,303 whereas WES is available is 166,625. This can be added to the “Notes” section of the table.*

- **We thank the reviewer for this suggestion. A total of 166,625 WES samples and 6,678 WGS samples were analyzed in this study. We have added a breakdown of the counts of WES and WGS samples used for each analysis to Supplementary Table 1 and have added more detail to the methods.**

2.5.1 In the population genetics portion of the manuscript, it is not clear why self-reported ethnicity was used for FST as opposed to the just-projected PCs? Was it not possible to project the sub-populations within PGR with sufficient granularity? This could be expanded upon for clarity. Is it possible to clarify for how many individuals this was done?

• We agree with the reviewer that self-reported ethnicity has limitations. We prepared a new Figure 3 to visualize clustering of South Asian populations. Individuals of the same ethnicity tend to be close to each other in PCA. This observation indicates that ethnicity was an appropriate input to the FST analysis, which included all participants.

- *Review 2: We agree with the authors that these changes help strengthen the paper. Can the authors explain what they mean by “ethnicity” because in PGR there is a group clubbed as “Urdu” which is a language rather than ethnic group.*
- **The reviewer is correct that Urdu is a language and we thank them for raising this point. Individuals who identify themselves as Urdu represent a large, mixed group in Pakistan who have largely immigrated from North India during the post-colonial partition of the subcontinent. Individuals who identify as Urdu are ethnically distinct such that they don't recognize themselves as any other ethnic group in Pakistan. For instance, a person who recognizes himself or herself as "Urdu speaking" will not classify himself as "Punjabi", "Sindhi", "Pathan", etc. Importantly, all ethnicity data in PGR is based on the self-reporting of participants. We state this information in the revised main text and revised Methods.**

2.5.2 I'm afraid as a non-expert in pop gen, the conclusions and take-away message from this portion of the manuscript were not apparent. Can this be better fleshed out and more explicit?

• We thank the reviewer for suggesting ways to make this section more insightful. To address the reviewer's point, we re-ordered the population analysis, which now presents Fst analysis, admixture, and PCA, for a clearer, stepwise presentation of South Asian ancestries and their genetic make-up. We also provided a PCA of the individuals in our study alongside South Asian 1000 Genomes (1KG-SAS) that we separated into its constituent ethnicities, as recommended by the reviewer. As the reviewer predicted, the new analysis projected PGR ethnicities along a geographic gradient of proximity to Europeans (PC1, Figure 3c,d). At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KG-SAS, and at the other end were the western populations of Pakistan, such as the Pathan and Balochi of PGR. The revised text includes the following information (lines 159-170):

o Principal Component Analysis (PCA) quantified the heterogeneity of PGR subpopulations relative to each other and to non-PGR populations from the 1KG (Figure 3). African, East Asian, European, and South Asian ancestries separated into clusters, with South Asians in PGR following a gradient along PC2 from European ancestries to East Asians (Figure 3a,b). Consistent with the admixture analyses, Pathan and Balochi ethnicities clustered more closely with Europeans than did other PGR ethnicities, while a subset of Balochis clustered more closely with Africans. Among individuals without recent African admixture, PC1 distinguished PGR ethnicities along a gradient of proximity to Europeans, consistent with previous observations (Figure 3c,d)^{19,20}. At one end of this spectrum were South Indian populations, such as the Telugu and Tamil of 1KGSAS, and at the other end were populations of western Pakistan, such as the Pathan and Balochi of PGR. PC2 and PC3 identified two additional clusters, one of which was predominantly Balochi and the other predominantly Punjabi, consistent with previous observations of substructure among the latter population²¹.

- *Review 2: We agree this adds some clarity to results for the non-expert reader. Can authors please differentiate between their use of ethnicity and genetically determined ancestry as these are not the same and, in many cases, it appears they are referencing the latter, certainly in Fig 2 and 3 which are PCA-reliant.*
- **The reviewer is correct that the admixture and PCA plots each provide a view of genetically determined ancestry. Within these plots, the grouping (admixture plots) and labelling (PCA plots) are done by self-reported ethnicity to show the genetic variation that exists both within and across these self-reported ethnic groups.**

2.6 Supplemental Table 4 is an excellent presentation of these data. Apologies if I have not been able to spot this, but is there a histogram that plots FROH (on Y axis) against age (on X)? There appears to be an age-based stratification of the effect of FROH on hypertension, CKD, and T2D and perhaps others. Wherein, if you were to meta-analyse the effects of FROH across these age-groups you would not find agreement in effect. Indeed this is captured in the p value for heterogeneity across groups. It could be, that individuals recruited at older ages have lower genome-wide/exome-wide runs of homozygosity, hence the question about distribution of FROH across age groups. There are also ~ 32,200 individuals recruited < 40 years whereas over 130k are over the age of 40 at recruitment.

The impact / heterogeneity across age groups makes it hard digest the results from the FROH analysis without a degree of scepticism. For example, if we were to accept the overall finding that higher FROH is protective from T2D, MI, what is the biological basis of this?

- **We thank the reviewer for these comments and suggestions. As proposed, we have now added additional considerations about the possibility of these biases to the discussion (lines 431-433):**
 - o *“Future longitudinal studies may elucidate these age-dependent associations and exclude the possibility of ascertainment or survivorship bias.”*
- **Regarding the biological basis of our findings, we offer an evolutionary hypothesis in the discussion. Future studies will focus on understanding possible biological mechanisms.**

2.6.6 190-200: Again these are extremely interesting results. Why was age 40 chosen as a cutoff? This is a typical young-onset threshold for T2D, but is perhaps a little early for MI. It would be more interesting to see these effect sizes plotted out by 10-year age bin (ie 18-30, 30-40, etc) to see how the association varies across the full breadth of ages (accepting power will be limited here), as opposed to a binary </>40 bin which leaves the reader wondering where the effect arises from. These divergent effects across the age range are very surprising and a bit more granularity would be most interesting to the reader.

• We agree with the reviewer that the cutoff of 40 was arbitrary. We have now stratified our analyses into multiple age bins and describe the results in the revised text (lines 219-224): o To determine if the associations were age-dependent, we binned across age groups (Extended Figure 8). Significant heterogeneity of effect was observed for educational attainment ($P=3.4e-08$), number of children ($P=3.2e-05$), having had children ($P=6.2e-06$), diabetes ($P=9.0e-06$), and hypertension ($P=1.9e-08$), with autozygosity trending towards increased risk of hypertension and diabetes in individuals <40 years and protection from hypertension and diabetes in individuals 50 years old (Extended Figure 8).

This approach is much more informative to the reader and strengthens the manuscript

- *Review2: agree. Thank you for making this table. See previous comment regarding acknowledging the lack of longitudinal data, recruitment of as many younger individuals <40 compared to >40 and >50, and heterogeneity of effects across age groups as a critical limitation of this analysis and generalisability of this finding.*
- **Please see our response above.**

2.6.7 Is there some sort of artificial clumping driving this association? Is FROH associated with age in some way, i.e., is the population becoming more autozygous over time, more FROH in young people, who do not yet have T2D/MI etc?

- We see less autozygosity with age in PGR ($P=8.4e-286$), albeit with a small effect size ($\beta=0.00026$ FROH per year of age). In Figure 1 of Malawsky et al., the authors report a similar magnitude of effect between age and autozygosity in British Pakistanis, but not in British Bangladeshis, arguing that survivorship bias or ascertainment bias are unlikely to explain the association between age and autozygosity. Additionally, we have included both age and age*age as covariates in all association analyses.

- *Review 2: We agree that through the lens the authors describe, survivorship bias does not necessarily explain these findings. But the age-based heterogeneity of effects is hard to ignore. There appears to be an age-independent but somehow intertwined artefact. Do those who survive to age 50, 60 and higher have runs of homozygosity around different loci that are protective compared to those who are younger and indeed theoretically, since you cannot measure this –in those who did not survive to participate in the study?*
- **Please see our response above.**

2.6.11 Finally, the discussion section is lacking depth as to why these findings diverge so entirely with previous studies. There is not the uncertainty on the effects of autozygosity as the authors suggest. There is in fact a consensus on the increased risk of ill-health. So the protective effects observed do require considered sensitivity testing and perhaps even external validation.

- We agree with the reviewer that the lower diabetes and MI risk in our study contrast with conventional expectations. We do however observe many associations consistent with previous reports (Clark 2019): higher rates of childlessness, fewer children, lower educational attainment, lower tobacco usage, shorter stature, lower weight, and lower cholesterol. PGR is the largest consanguineous cohort in the world and can detect associations between autozygosity and traits that are not present in other cohorts in sufficient numbers, such as myocardial infarction. For these reasons, external validation will require the expansion of other highly autozygous cohorts to similar size and case/control composition. We have performed sensitivity analyses within PGR, such as subsetting by ethnicity.

- *Review 2: External validation of a finding as controversial as higher FROH being associated with lower risk of important health outcomes like MI and T2D etc. Is required if this finding is to be asserted. Why not utilise resources like G&H where there seem to be >20k individuals of pakistani ancestry with WES data available based on this preprint: <https://www.medrxiv.org/content/10.1101/2025.09.24.25336527v1.supplementary-material>*
- **We agree with the reviewer that external validation, wherever possible, would be beneficial. We cite Malawsky et al. (2023) which finds an association between F_{ROH} and increased risk of T2D using the G&H data described by the reviewer. For many outcomes, such as MI, PGR is uniquely powered due not only to its large number of highly consanguineous individuals, but its case/control breakdown. Additional large, highly consanguineous cohorts with similar phenotypic distributions are not available.**

2.7.2 225: These are extremely interesting and well-presented results. To orientate the reader, it would be useful to have a short section explaining how many GENE*phenotype associations were tested, and what the associated Bonferroni-corrected p value would therefore be (for example, the “lower random glucose ($\beta=-0.05$, $P=2.3e-02$)” seems unlikely to be anything other than nominally significant, but it is not clear to the reader which other association would survive multiple testing correction.

- We thank the reviewer for the kind commentary. Regarding multiple testing, we perform an initial PheWAS where we report all gene-biomarker associations that are significant at an exome-wide Bonferroni threshold of $P<9.9e-08$ in Supplementary Table 5. Elsewhere, when following up on specific genes we report additional, nominal associations. We have specified in the text that these associations are nominally significant at this threshold.

- *Review 2: Thank you for specifying where associations are purely nominal. It would be helpful for the reader to understand how “specific genes” in this context were identified – on the basis of specifically interesting phenotypic associations, or on the basis of specifically interesting genes, if*

so how were these defined, and if not, so how were they selected?

My apologies for repeating the point, but in a PheWAS where Bonferroni correction requires a $p < 9.9e-8$ discussing as a major result, an association with a $P = 0.023$ seems far from statistically meaningful. If 1/3 of coding genes contain knockouts, as is stated in the title, then by chance one would expect many thousands of associations in the $p = 0.023$ threshold. Was there some other a priori definition of genes of interest which would confer a more relaxed P value threshold? I think some explicit justification of why results at relaxed P value thresholds are highlighted, given the scale of PheWAS association testing, is needed, or the authors should focus discussion on those associations with more convincing P values.

- **We thank the reviewer for these comments. We report all gene-biomarker associations that are significant after multiple test correction ($p < 9.9e-8$). For select genes of clinical importance that are discovered through this approach, such as *HBB*, we show all nominal associations to provide a full picture of the gene, specifying that these additional associations are nominal ($p < 0.05$).**

2.7.5 229: It would be useful to compare to G&H recent effort characterising pLoF human KOs, recent preprint: <https://doi.org/10.1101/2025.06.05.25329068>. For example the Genes and Health authors report 2991 homozygous carriers of pLoF in a sample size of ~44k individuals and this could be compared to findings from PROMIS-PGR for a better understanding of the landscape of genomic studies in SAS populations and sub-populations.

• Based on the reviewer's feedback, we have included a comparison to South Asians in gnomAD using the same set of pLoF filters that were applied to PGR. gnomAD South Asians is comparable to Genes and Health in size and number of KOs: gnomAD has 45,546 South Asians, from which KOs are present in 3,022 genes. We have not analyzed the Genes & Health cohort as this preprint is not yet peer-reviewed and could be subject to change.

- *Review 2: We agree this is a pre-print and may be subject to change, but would also contest that referencing and learning from pre-printed manuscripts is increasingly the norm in the modern scientific era; particularly when making claims of particular novelty.*
- **We recognize that Genes & Health is a valuable resource and appreciate the reviewer's suggestion to include a direct comparison of the homLoF genes present in PGR and Genes & Health. We do not have access to variant-level data in G&H that would allow us to apply the additional loss-of-function variant filters applied to PGR—including correction of errors in LOFTEE—which we describe in the methods, supplementary text, and Supplementary Figure 2, and would allow for a relevant head-to-head comparison of homLoF genes between these two cohorts.**

2.7.6 217: Some of these associations with *HBB*, including iron and lipid traits, in addition to bilirubin, are reported in the Kim et al paper above and should be discussed in this context. - Please see above.

- *Review 2: It is unlikely that an independent cohort reporting a consistent result with PGR is in any way problematic. If anything, a future reader of this paper would benefit from knowing that these findings are consistent and coherent with available literature from the limited other resources of south Asian data that exist.*
- **We thank the reviewer for the suggestion. We have updated the manuscript and added this citation (lines 234-235):**
 - ***"HBB pLoF variants are associated with reduced hematocrit, increased protection against malaria parasite, increased risk of β -thalassemia, bilirubin levels, and various lipid traits³⁰⁻³²."***

2.9.3 326: Similarly, some more phenotypic information on the human KOs for SLC2A4— if available — would be very interesting to the reader. Given they are first reported human Kos for this gene, there is likely to be enormous interest in knowing more about glycaemic measures in these individuals. The T2D knowledge portal reports findings with thyroid function at the gene-level, rare variants with ApoA1 and blood pressure. The authors have a unique opportunity to explore and characterise these rare individuals.

• We agree with the reviewer that this additional information would provide insights. We have added available glucose and HbA1c levels for SLC2A4 and the KOs in Supplementary Table 10. We were unable to recontact the SLC2A4 knockout for additional characterization.

- *Review 2: this is unfortunate but understandable. It would be good to include a statement saying so in text. Or simply, recontact was not possible for deeper phenotyping.*
- **We agree with the reviewer and added a statement to the text (lines 348-349):**
 - ***“As predicted by preclinical models⁵⁷, this participant, who was not available for RBG, has type 2 diabetes.”***

2.9.4 The findings in DENND1B KOs are interesting, but at odds with published summary stats from UKBiobank. How do authors reconcile this?

• We are similarly puzzled that the two new DENND1B KOs are not obese, in contrast to the report by Wallis et al. The difference may be due to variable penetrance, ancestral differences in weight gain, statistical limitations due to the rarity of DENND1B pLoF, or a combination of these factors.

• Wallis et al. does not show a formal burden analysis of the association between damaging or high confidence DENND1B loss-of-function variants and BMI with a corresponding effect size. Previously published analyses showed only nominally significant enrichment of combined damaging and protein truncating variants and a magnitude-free directionality of BMI (no reported effect size). Additionally, the paper provided supporting evidence of a single pediatric obesity case homozygous for a damaging missense variant (R501C)—the parents of this child were both heterozygous carriers and of normal BMI.

• In a formal burden analysis in PGR, we see no significant associations between DENND1B damaging/high confidence loss-of-function variants and BMI. Burden analyses in AZPheWAS also show no significant association between DENND1B rare damaging/loss of function variants and BMI. In addition, PGR contains two adult, high confidence loss-of-function homozygotes with normal BMI. Collectively, this suggests that DENND1B loss-of-function has little effect on BMI in adult humans.

- *Review 2: Thank you. As throughout the manuscript, I think external validation of your findings is reassuring and should be leveraged. Could authors please add the agreeing information from AZPheWAS portal to text to head off any concerns linked to disagreement with the Wallis paper.*
- **We thank the reviewer for this suggestion. We have added the following to the text (lines 352 – 353):**
 - ***“In the UK Biobank, predicted damaging variants in DENND1B are nominally associated with higher BMI in European ancestries²⁶”***

2.9.6 Given the novel finding, which is at odds with the AstraZeneca phewas portal findings for LoF DNND1B, it would be valuable to undertake replication with other data sets, UK Biobank, Genes & Health to confirm this findings.

• With apologies to the reviewer if we misunderstood their comment, we do not see significant associations between DENND1B LoF and BMI in the AZPheWAS portal.

- *Review 2: With apologies to the authors, the results we were drawing their attention to were for the AZPheWAS “flexdmg” mask. In any case including AZPheWAS portal in Supp Table 5 helps orientate the reader to previously described results. This is not cross-referenced in the results paragraph on DENND1B – could you please do so and reassure readers about consistency?*
- **We thank the reviewer for this comment. We have checked the manuscript throughout for consistency in referencing, including Supplemental Table 5.**

2.9.7 338: Associations with RXFP1 have been demonstrated in UK Biobank, data available on AZphewas portal. The novelty here is unclear – but would benefit from highlighting and meta-analyses with UKB.

- With apologies to the reviewer if we misunderstood their comment, we do not see significant associations in RXFP1 in the AZPheWAS portal.
- We believe that our report of RXFP1 KO is important because the healthy cardiovascular, metabolic, and reproductive profiles of these individuals strongly invalidate the longstanding pre-clinical observations that formed the foundation of past clinical trials. Decades of publications using rodent models suggested that RXFP1 is critical for nursing, fibrosis, and cardiac health, and metabolism. These rodent studies motivated many clinical trials testing RXFP1 activators in humans, including two large Phase 3 trials, to treat a broad spectrum of diseases, including impaired cervical ripening during childbirth, scleroderma, and acute heart failure. These trials have been unsuccessful.

- *Review 2: On the AZPheWAS portal RXFP1 rare variants are associated with scleroderma|systemic sclerosis ($p = 0.04$), heart failure|pulmonary oedema ($p = 5e-3$), and arterial stricture ($p = 1.2e-4$). We agree in principle with the author's contention that a healthy RXFP1 KO is of substantial interest, but the point remains that the p values the authors consider significant must be consistent across analyses – for example if they wish to highlight a p value of $p = 0.02$ in their results (as discussed in 2.7.2 above), then we would argue they should also draw the reader's attention to results for other findings elsewhere at a similar threshold and cannot dismiss these as not significant.*
- **We agree with the reviewer, including the point that the RXFP1 homLoF is of substantial interest. Space considerations preclude adding supplementary association statistics.**

2.9.9 All these findings would be more impactful, if where possible, external validation or replication was undertaken. Coherent and directionally stable associations with pTvs in UK Biobank, Genes & Health which is also likely enriched for LoF variants (albeit in a smaller sample size), PRECISE etc.

- We have taken care to highlight examples of novel high confidence loss-of-function homozygotes that have not been previously reported or characterized by other cohorts. The absence of KOs in other cohorts precludes external validation of the phenotypes associated (or not associated) with these KOs.

Review2:: The logic of not replicating or meta-analysing novel findings where possible is not apparent to me. Particularly where p values are not entirely convincing.

- *Review 2: If attempts have been made to identify LoF KOs in other cohorts and none were found this should be explicitly stated in text and will enhance the novelty of PGR.*
- **We agree that meta-analysis of novel findings is valuable wherever possible. In the text, we clarified that homLoF individuals for these genes have not previously been described.**

2.9.10 Similar to the previous section, it would be helpful to know how many GENE*trait associations (*model type (additive / recessive)) etc were performed across this section, and an associated Bonferroni-corrected p value presented to the reader. Presumably not just the genes included in the vignettes were included in analysis, but it is difficult to get a sense of what a meaningful p value is without knowing the number of tests performed.

- The associations in this section are generally hypothesis-based, motivated by the observations of KOs in genes for which KOs have not been previously identified. Where appropriate, we have added additional descriptions of statistical tests.

- *Review 2: Our apologies, we struggled to find these in the text, please could the authors highlight which additional descriptions of tests have been included?*
- **Apologies to the reviewer. We intended to state that we added additional information, such as Bonferroni thresholds, to the other sections where they are relevant.**

2.10.3 411: It is unconvincing that the thrifty genotype hypothesis explains why these (age-stratified) results are at odds with the Genes & Health results in Malawsky et al – I think more fleshing out of the proposed mechanism is needed here to convince the reader.

- Given the retrospective nature of our observations, we offer only a hypothesis rather than a mechanism supported by data. In the revised text, we have clarified that the thrifty gene hypothesis and the possibility of past resource-limited environments may explain only a subset of the protective phenotypes, such as lower LDL-C, triglycerides, BMI, and MI (lines 456-460):

- o The mildly protective effects of autozygosity on some cardiometabolic biomarkers (LDL-C, triglycerides, BMI) and outcomes (MI) may represent reduced fitness of traits that were protective in past resource-limited environments, consistent with the “thrifty genotype” and related hypotheses that aim to explain high rates of obesity, diabetes, hypertension, and cardiovascular disease as a byproduct of adaptation to a divergent ancestral environment⁸³⁻⁸⁷.

- *Review 2: Given the age-based heterogeneity it would be useful to readers to understand how that aspect squares off with the mechanism proposed here.*
- **Please see our response above.**

Referee #3 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports.

Referee #4 (Remarks to the Author):

The authors have addressed most of my concerns. I still have two major concerns:

- human knockouts seems to be a ill-suited term to refer to humans and can lead to all kinds of stigma and miscommunication. I suggest the authors take the time to think of a better way to refer to said individuals and as such alter the title and use in the manuscript. The terminology we use is key to not go down a eugenics path again.

- **We replaced the abbreviation “KO” with “homLoF” (homozygous loss-of-function) when referring to individuals homozygous for naturally-occurring loss-of-function variants. We describe the reasoning for this change of phrasing in the introduction (lines 116-119):**
 - o ***“To date, these human loss-of-function homozygotes have been called human gene “knockouts”. Since “knockout” originally denoted purposeful deletion of a gene in model organisms, such as mice, we replace the term, in the context of humans, with homozygous loss-of-function (homLoF), which precisely describes naturally occurring genotypes.”***

The access to the dataset is still not clear, and generates a lot of concern if academics will truly be able to access the data for academic projects. If the data has to be hosted at CNCD, how will the authors ensure a fair and transparent process post-publication?

- **We have revised the data availability statement. It now states (lines 1054-1065):**
 - o ***“PGR. Supplementary Tables include summary statistics for autozygosity, summary statistics for biomarker burden tests, and a full exome-wide list of homozygous loss-of-function (homLoF) variants listed by chromosome, position, reference allele, alternative allele, and homozygous carrier count. Additional summary statistics for all biomarker burden tests, including non-significant associations, are uploaded to figshare. A subset (n=7,298) of the WES samples (accession phs000917.v1.p1) used in this study and all (n=6,678) of the WGS samples (accession phs001569.v1.p1) used in this study were previously made***

available to researchers in dbGaP^{2,93}. All phenotype data and individual-level exomes and genomes, including those previously deposited in dbGaP, are stored at CNCD, with access protecting participant data. Academic researchers may apply for access to all individual-level data from this study by contacting the CNCD at info@cncdpk.com, with an initial review response time of 45 business days. Each request will require a detailed research plan and data privacy plan, including an agreement to maintain de-identification of PGR participants."

Response to Editor and Reviewers:

The autozygosity association results *must* be further de-emphasized for the reasons that reviewer #2 has been laying out since the original review (underpowered and unreplicated analysis). Beside much more cautious description of these results, we also ask that you move them to further back in the results section; they should not be presented as the main feature of the work. This also means that Figure 3 should be turned into an Extended Data Figure.

- **At the request of this Reviewer and the Editor, we have de-emphasized our autozygosity findings, moved these findings to a later section in Results, removed them from the Abstract, converted Figure 3 into an Extended Data Figure, emphasized the future need to replicate our findings, removed nominal associations, and removed from the Discussion a description of the “thrifty gene” evolutionary hypothesis.**

Further, we agree with reviewer #4 that “homozygous loss-of-function humans” is only marginally better than “human knockouts”. It’s the genes that have loss-of-function variants, not the humans who have ‘lost function’, and the writing of the manuscript needs to reflect this.

- **We appreciate the concerns of the reviewer and Editor. We have:**
 - o **Revised the title.**
 - o **Replaced the terms “human knockouts” and “homozygous loss-of-function humans” and now refer precisely to homozygous loss-of-function variants.**
 - o **Edited the text for clarity and sensitivity.**

For the manuscript title, it might be better to highlight the resource than the ‘insights’ (which is vague anyway).

- **We revised the title to focus on the resource and match the convention of other similar biobank studies: “Analysis of 173,303 exomes and genomes in the Pakistan Genome Resource”**

We strongly recommend that the entire manuscript undergoes a sensitivity reading before resubmission.

- **Completed.**

Parts of the manuscript appear somewhat awkwardly written and it would be helpful for the text to be further reviewed for language and clarity.

- **Completed.**

Please add references to the abstract.

- **Completed.**

Please reduce the Abstract to ~230 words or fewer.

- **Completed.**

For a regular research article we recommend ~60 main text references.

- **After a careful rereading and splitting the references into separate main and methods-only sections, we now have 82 essential main text references.**

Please split the main text and methods references into two separate, but continuously numbered, lists that follow the discussion and the methods section, respectively.

- **Completed.**

You have more than one account on EJP. Please contact Nature Manuscripts (naturemanuscripts@nature.com) to confirm all accounts which belong to you, and your primary email address.

- **Completed.**

We have noticed that not all corresponding authors of the manuscript are ORCID linked. As part of our efforts to improve transparency in authorship, we now request that all authors identified as 'corresponding author' create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS) prior to acceptance. For more information please visit <http://www.springernature.com/orcid>. If you have an alternative account which is ORCID linked, please contact johannes.black@springernature.com

- **Completed.**

7. Please provide a supplementary information guide as a separate word document.

- **Completed.**

8. There are potential third party rights issues in the figures. Please check the sources of all illustrations and clarify whether permissions are needed to adapt or reproduce them. Please make sure to include the relevant details in the third party rights table when you resubmit (more information below). If Biorender or a similar software has been used, please also ensure to provide relevant licenses. In particular please check figure/s: main figure 1a

- **Completed. OpenStreetMap (OSM) data is distributed under the Open Database License (ODbL) which allows one to freely share, create, or adapt, so long as proper attribution is provided. Static and print maps created using Mapbox require Mapbox logo and text attribution. We have provided all necessary attribution in Fig 1A.**

10. Please provide an author contributions statement in the main article file.

- **Completed.**

11. Since your author list includes consortia, please consult the attached guidelines for formatting instructions and make any necessary changes.

- **Completed.**

12. Please remove the supplementary information from the article file and resupply it as separate file.

- **Completed. Supplementary information is now provided as a separate file which includes supplemental text, GANTT chart, supplemental figures, supplemental table legends, and a table of contents.**

13. Please ensure that the text size in all figures is at least 5 pt Arial, keeping in mind that the maximum figure size is 18cm width by 17cm height. It looks like Figure 1 may be too large in its current form and you may have to move panels to the ED section (please add these panels to existing EDs and/or rearrange ED figures as needed so as not to increase the number of ED items).

- **Completed. The figure was resized.**

14. Please include the GANTT chart detailing the data access workflow in the SI file as a Supplementary Note and reference it in the data availability statement.

- **Completed.**

15. In the data availability statement, you mention “Additional summary statistics for all biomarker burden tests, including non-significant associations, are uploaded to figshare.” Please include the figshare doi and make sure that the data are live on final resubmission.

- **Completed.**

16. The competing interests statement must refer to all co-authors.

- **Completed.**

17. Please clarify if any custom code was developed for this study. If so, it should be made available via a suitable repository such as github.

- **No custom code was developed for this study.**

18. For any Supplementary Figures, please check and confirm that:

- * If data is presented as bar charts, individual data points are shown using overlaid dot plots.
- * The n number (i.e. the sample size used to derive statistics) is provided and defined as a precise value (not a range), using the wording “n=X samples/cells/independent experiments” etc. where applicable.
- * Any chart axis, error bars, scale bars, symbols and colour scales are defined.
- * Any statistical tests used for data analysis are specified and exact p-values are provided either on the figures themselves, in the legend or in the Source Data file.
- * Wherever representative data such as micrographs are shown, the legend indicates how many times the experiment was repeated with the same results.

- **Completed.**

TRANSPARENT PEER REVIEW: Nature uses a transparent peer review system for original research manuscripts. We encourage increased transparency in peer review by publishing the reviewer comments, author rebuttal letters and editorial decision letters. Such peer review material is made available as a supplementary peer review file.

Please note: we allow redactions to authors’ rebuttal and reviewer comments in the interest of confidentiality. If you are concerned about the release of confidential data, please let us know specifically what information you would like to have removed. Please note that we cannot incorporate redactions for any other reasons. Reviewer names will be published in the peer review files if the reviewer signed the comments to authors, or if reviewers explicitly agree to release their name. For more information, please refer to our [FAQ page](#).

- **No redactions requested.**

ORCID--IMPORTANT: All authors identified as ‘corresponding author’ on the manuscript must have an ORCID associated with their Nature account before submitting the final version of the manuscript. While non-corresponding authors do not have to link their ORCIDs, they are encouraged to do so. Please note that it will NOT be possible to add/modify ORCIDs at the proof stage. Thus, if they wish to have their ORCID added to the paper they must follow the above procedure prior to acceptance. If you have any issues attaching an ORCID identifier to your Nature account, please contact the [Platform Support Helpdesk](#).

- **Completed.**

STATISTICS: When revising your manuscript, you should ensure that any statistical analysis used is sound and that it conforms to our [guidelines](#). A collection of articles explaining the basics of statistical analysis and advice on how to best present it can be found [here](#).

- **Completed.**

REPRODUCIBILITY: To ensure that the quality and transparency of methods and statistical reporting (as discussed [here](#)) are sound before the paper is published, we have reviewed your Reporting summary editorially. I have attached two documents: one listing specific issues related to your manuscript and one containing an annotated version of the Reporting summary. Please ensure that, as well as the more general points below, the points highlighted in the attached documents are addressed in full, both on these forms and within the manuscript. Both forms should be uploaded as a "Related Manuscript" file type. The Reporting summary will be published with your paper.

- **Completed. We have provided the information requested in the Reporting summary and manuscript annotations from the editors.**

TITLE: Titles cannot exceed 75 characters (including spaces); they must not contain punctuation.

- **Completed.**

SUMMARY PARAGRAPH: Papers start with a fully referenced, bold paragraph, ideally of about 200 words, aimed at readers in other disciplines. Numbers, abbreviations, acronyms or measurements should be avoided unless essential. The summary paragraph consists of 2 to 3 sentences of basic-level introduction to the field; a brief account of the background and rationale of the work; a statement of the main conclusions (introduced by the phrase 'Here we show' or its equivalent); and a conclusion of 2 to 3 sentences putting the main findings into general context so it is clear how the results described in the paper have moved the field forward. A downloadable, annotated example is available [here](#).

- **We followed these guidelines.**

MAIN TEXT: If further introductory material is necessary, the main text can begin with up to 500 words of introduction expanding on the background to the work (some overlap with the summary is acceptable), before proceeding to a concise, focused account of the findings, and ending with 1 or 2 short paragraphs of discussion. Sections are separated with subheadings (up to 40 characters including spaces) to aid navigation.

- **We followed these guidelines.**

REFERENCES: As a guideline, most papers should include no more than 50 main text references; all additional references can be cited in (and listed after) the Methods section, as detailed below.

- **We followed these guidelines.**

FIGURE LEGENDS: These should be listed sequentially after the main text references and not in the figure files. Each legend should begin with a brief title for the whole figure and continue with a short description of each panel and the symbols used. Legends should not exceed 300 words each. Each figure legend should contain, for each panel where relevant, the following information:

- * the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- * a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc);
- * a statement of how many times the experiment shown was replicated;
- * definitions of statistical methods and measures:
- * very common tests (e.g. t-test, simple Chi-square tests, Wilcoxon and Mann-Whitney tests) can be identified by name only, but more complex techniques should be described in the Methods;
- * whether tests are one-sided or two-sided;
- * whether there are adjustments for multiple comparisons;
- * the statistical test results (e.g., P values);
- * the definition of 'center values' as median or average;

* the definition of error bars as s.d. or s.e.m.

Descriptions that are too long for the figure legend should be included in the Methods section.

- **We followed these guidelines.**

METHODS: The Methods section, which provides the full, step-by-step instructions that would allow other researchers to replicate the results, is included after the main text figure legends. The Methods section will not appear in print but will appear online in the full-text HTML and PDF versions. The Methods section should be written as concisely as possible but should contain all elements necessary to allow interpretation and reproduction of the results. If there are additional references (in the Methods section, Supplementary Information, etc), their numbering should continue from the last entry in the main text reference list, and they should be listed following the Methods section. Specialized methods that require chemical structures, figures, or tables cannot be accommodated in the Methods section of the main text file. If such information is part of the Methods, the entire Methods section must instead be included within a Supplementary Information text file.

- **We followed these guidelines.**

ETHICS STATEMENT: For research involving human research participants, the Methods section must include an ethics statement. This statement should provide the name of the committee that approved the study; confirm that the research was performed in accordance with all relevant guidelines and regulations; and confirm that informed consent was obtained from all participants. If the study was granted an exemption from requiring ethics approval, details of the committee granting the exemption must be included.

- **Completed.**

Research involving human embryos or gametes, or human stem cells in contexts requiring ethical oversight, also must include an ethics statement in the Methods section. This statement should provide the name of the committee that approved the study and confirm that the research was performed in accordance with all relevant guidelines and regulations. We encourage authors to follow the principles laid out in the [2021 ISSCR Guidelines for Stem Cell Research and Clinical Translation](#). Where necessary, the ethics statement should also describe the conditions of donation of materials, such as human embryos or gametes, and confirm that informed consent was obtained from all donors of cells or tissues.

- **Not applicable to our study.**

MAIN TEXT STATEMENTS: Several statements (which will not appear in print but will appear online in the full-text HTML and PDF) are required after the Methods (and additional references, if present). First, there should be an Acknowledgements section, listing grant/financial support. Next, we require a detailed Author Contribution statement; the specific contributions of each author, particularly in terms of which authors performed which specific experiments, must be listed. This is followed by a Competing Interest statement. Financial and non-financial interests should be noted here, as well as any patents; patent information should include at a minimum patent number, what is covered by the patent, and who submitted the patent application. Finally, an Additional Information statement should include information regarding reprints and permissions and name the author(s) to whom correspondence and requests for materials should be addressed. Formatting details and an example are available [here](#).

- **We followed these guidelines.**

DATA AND CODE AVAILABILITY STATEMENTS: Any manuscript reporting original research must include a Data Availability statement that makes transparent to the reader the conditions of access to the “minimum dataset” that is necessary to interpret, verify and extend the research in the article. This minimum dataset may be provided through deposition in public community/discipline-specific repositories, custom proprietary repositories (for certain types of datasets), or general repositories like Figshare, Zenodo and Dryad. We strongly discourage providing large datasets in Supplementary Information; the

preferred approach is to make data available in repositories. More information on Nature Portfolio's reporting standards and guidance on preparing your Data Availability statement can be found [here](#).

For all studies using custom code or mathematical algorithms that are deemed central to the conclusions, a Code Availability statement must be included, indicating whether and how the code or algorithm can be accessed, including any restrictions to access. The Code Availability statement is listed as a separate section after the Data Availability statement but before any additional references. Code should be deposited in a DOI-minting repository such as Zenodo, Gigantum or Code Ocean and cited in the reference list. Authors are encouraged to manage subsequent code versions and to use a license approved by the open source initiative. Additional details can be found [here](#).

- **We followed these guidelines.**

DISPLAY ITEMS: We suggest that you take stock of all data that have been generated throughout the review process and ensure that only the data most central to the conclusions are presented in the main text figures. Any figures included within the main text file during the review process must be removed from the final main text file and uploaded as separate, individual files; they will be integrated into the main paper in print and online. An overview of the key features of this presentation may be found [here](#).

Figures should be comprehensible to readers in other disciplines and assist in understanding of the paper. Main text figures (but **not** Extended Data) must be provided in production-quality versions in an editable format (i.e., .ai, .cmx, .cdr, .doc, .eps, .pdf, .ppt, .ps, .psd, .svg and .xls); we cannot accept figures in .cvs, .gif, .jpg, .png and .tif formats. We highly encourage you to consult our [artwork guidelines](#). They should be as small and simple as is compatible with clarity. All panels of a figure should be logically connected and assembled on a single page in a rectangular shape; any essential alignments (parts horizontal, vertical, spacings, etc) should be indicated. Each panel of a multipart figure should be sized so that the whole figure can be proportionally reduced and reproduced on the printed page at the smallest size at which essential details are visible. Nature's standard figure sizes are either 9 or 18 cm wide; the maximum permitted height is 17 cm. Panels should be arranged to fit these widths while minimizing excess space around the panels. Tables should be prepared using the Table menu in Word. As we must be able to edit the figures so that they conform to our house style, the submission of files that are incorrectly formatted, flattened, or of insufficient resolution may delay final acceptance of your manuscript.

- **We followed these guidelines.**

THIRD PARTY RIGHTS: You must provide proof that you have secured permission to use any third party materials that appear in any part of your manuscript, including Extended Data and Supplementary Information. Please fill out a [Third Party Rights Table](#), and upload this with the final version of your manuscript. Third party materials include any figures, tables, images, videos or text boxes that are reproductions or adaptations of items that have previously been published elsewhere and/or are owned by a third party. This includes pictures taken by professional photographers, maps and images downloaded from the internet. You will need to obtain the right to use each of these items before your paper can be accepted for publication. You will also need to give proper attribution to the copyright holders in your paper. Please ensure you upload any necessary grants of rights alongside the final version of your manuscript. More information is available on our [Rights and permissions page](#). Failure to obtain the appropriate rights and to supply a completed third party rights table will delay the publication of your article. The editorial assistant (cc'd) can help with any questions.

- **Completed. OpenStreetMap (OSM) data is distributed under the Open Database License (ODbL) which allows one to freely share, create, or adapt, so long as proper attribution is provided. Static and print maps created using Mapbox require Mapbox logo and text attribution. We have provided all necessary attribution in Fig 1A.**

IMAGE INTEGRITY: We strongly advise that you go carefully through all the data (including Extended Data and Supplementary Information) to ensure there are no accidental image/data duplications, other image manipulations or data errors. Such issues generally require correction after publication. Any image

provided for publication, either in print or online (including Extended Data and Supplemental Information), may be subject to a quality control process to check for image integrity and manipulation. A discussion of our standards regarding how images should be prepared and presented can be found [here](#).

- **Completed.**

EXTENDED DATA: Extended Data do not appear in print but are included online within the full-text HTML and integrated in the downloadable PDF. Extended Data are an integral part of the paper, and only data that directly contribute to the main message should be included. All Extended Data must be referred to in the main text, and their legends should be listed sequentially at the end of the main text file, not in the Extended Data files. Extended Data should be assembled into a maximum of 10 A4 size, multi-panelled display items. They must be supplied as individual files in .jpg, .tif or .eps format **only**. They should be of the same quality as the main figures, but there are important differences in their formatting. More specific instructions are provided [here](#). If you need to describe a complex process, we encourage you to add a schematic of the main finding as part of the Extended Data to aid readers unfamiliar with the immediate discipline.

- **Completed.**

SUPPLEMENTARY INFORMATION: Supplementary Information (SI) is online-only, peer-reviewed material that is essential background to the study (e.g., large data sets, more complex methods, and calculations), but which is too large or impractical, or of interest only to a few specialists, to justify inclusion in the print version of the paper (see [here](#) for further details). While SI should not typically contain data figures (any figures additional to those appearing in the main text should be formatted as Extended Data), we require that the raw, uncropped data for gels be presented as an SI figure (see below). Tables may be included in SI, but only if they are unsuitable for formatting as Extended Data (e.g., tables containing large data sets that cannot fit a single page or raw data tables that are best suited to Excel files). If a manuscript has SI, each discrete SI item (e.g., videos, tables) must be referred to at an appropriate point in the main text file. You must also provide a Word file entitled "SI Guide", containing a cover page with manuscript title and author information; a table of contents (preferably with page numbers); and then any SI text, notes, figures, and titles and legends for any separate SI files; for additional information see [here](#).

- **Completed.**

CELL LINE IDENTIFICATION: To help curb the inadvertent use of cross-contaminated or misidentified cell lines, we ask that you check your reagents against the [list of commonly misidentified cell lines](#) maintained by the International Cell Line Authentication Committee, which is also accessible through the NCBI BioSample [database](#). If you have used a cell line that is on this list, you must provide a scientific justification and state the identity issue in the Methods. The editors reserve the right to demand that the data be removed from the paper if the justification is deemed unsatisfactory. In addition, authors must identify the source of cell lines (with catalog number if obtained from a vendor or cell bank) and report whether the cell lines have been authenticated, including the method used, the results, and the date authentication testing was last performed for that cell line. You should be able to provide the test results upon request. Mycoplasma contamination testing status must also be reported. These requirements will be particularly scrutinized for cancer studies, where the issue of cell line misidentification has been well documented. Resources on cell line authentication are available [here](#).

- **Completed.**

SOURCE DATA (GRAPHS): To increase transparency, we strongly encourage you to provide, in spreadsheet form, the data underlying the graphical representations used in figures. For all experiments presenting data from animal models, this is a requirement and is not optional. This is in addition to our well-established data-deposition policy for specific types of experiments and large datasets. Online readers of the manuscript will be able to access the graphical source data directly from the figure legend. Spreadsheets must be submitted in .xls, .xlsx or .csv formats. One file per figure is permitted. If there is a

multi-panelled figure, the source data for each panel should be clearly labeled in the file; alternatively the source data for a figure can be included in multiple, clearly labeled sheets within an Excel file. File sizes of up to 30 MB are permitted, but it is expected that the vast majority of graphical source data files will be considerably smaller than this. When submitting these files with your manuscript, you should select the "Source Data" file type and use the title field in the file description tab to indicate the figure(s) to which the source data pertain. Source data should not be provided as Extended Data.

- No animal model data is presented in our study.

RAW DATA (GELS): You must provide the original source images for all data obtained by electrophoretic separation (e.g., EMSA, northern/Southern/western blots, etc). The raw images must be assembled into a single .pdf or .tif file (multiple gels on a single page is encouraged). The file should be uploaded as Supplementary Figure 1. The full scanned images must be in uncropped form and contain labeled size/molecular weight markers and loading controls. There should be an accurate indication of how the gels were cropped for the final figure. The figure legends and raw data files should indicate whether controls (such as beta-actin) were run on the same gel as loading controls, or on separate gels as sample processing controls (see [here](#) for guidance). While the data can be displayed in a relatively informal style, there must be a correspondence between each source data image and a specific main text or Extended Data figure. The main text or Extended Data figure legends should refer to the uncropped scans explicitly (e.g., "For gel source data, see Supplementary Figure 1."). For examples, see [here](#) or [here](#).

- Provided.

DATA DEPOSITION: The following specific points may be relevant to your paper, so please ensure that you provide the following information:

* Sequences for any RNAi/small RNA constructs must be included.

* Accession numbers for gene expression data or RNA sequencing data must be listed.

* Papers reporting protein structures must conform to our standards listed in the [Guide to Authors](#). The Data Availability statement must state that the X-ray crystallographic coordinates and structure factor files (or comparable NMR or cryoEM data) have been deposited in the appropriate, named, public database, along with all relevant accession number(s). You must use the standard Nature [templates for structural data](#); there are separate links to tables for X-ray crystallographic, NMR and cryoEM structures. These tables must be presented as Extended Data; if the number of entries causes the table to exceed a page, it must be divided into two Extended Data items. The contour level of any electron density maps presented, as well as the type of map (i.e., Fo-Fc or 2Fo-Fc), should be explicitly stated in the figure legend.

* For every new chemical compound, a complete description of the synthesis and the physical characterization (i.e., NMR, MS, etc) must be included in the Supplementary Information (see [here](#)).

* Papers containing new or revised formal taxonomic nomenclature for animals, whether living or extinct, are accepted conditional on the provision of LSIDs (Life Science Identifiers) by means of registration of such nomenclature with ZooBank, the online registration system for the International Code of Zoological Nomenclature (ICZN). ZooBank LSIDs can be resolved and the associated information viewed through any standard web browser by appending the LSID to the prefix "<http://zoobank.org/>".

* We strongly encourage deposition of 3D morphological data in a suitable repository such as MorphoBank, MorphoSource or similar; the relevant accession numbers should be listed in the Data Availability statement.

* For animal experiments, you must confirm that all experiments were performed in accordance with relevant guidelines and regulations. There should be a statement identifying the institutional and/or licensing committee approving the experiments, including any relevant details. Sex and other

characteristics of animals that may influence the results must be described. Details of housing and husbandry must be included if they are likely to influence experimental results. Further details can be found [here](#).

* Human genotype data (e.g., SNP array data) should be deposited into a public database (dbGAP or EGA) with a controlled access policy.

- **A Data Access policy statement has been provided for human data.**

CONSORTIA: For papers containing one or more consortia, all members of the consortium who contributed to the paper must be listed in the paper (i.e., print/online PDF). If necessary, individual authors can be listed in both the main author list and as a member of a consortium listed at the end of the paper. When submitting your revised manuscript via the online submission system, the consortium name should be entered as an author, together with the contact details of a nominated consortium representative. See [here](#) for our authorship policy and [here](#) for further consortia formatting guidelines, which should be adhered to prior to acceptance.

We will not send your revised paper for further review. If the revised paper is in our format (as detailed above), in accessible style and of appropriate length, we shall begin the acceptance process.

In order to accept your paper, we require the following electronic files:

* A cover letter describing your response to any editorial comments and detailing any format changes during revision, particularly if the overall length is affected.

- **Provided.**

* A point-by-point response (preferably in Word) to any remaining issues raised by our referees.

- **Provided.**

* The final version of your text as a Word document. Word Equation Editor/MathType should be used only for formulae that cannot be produced using normal text or symbol font. If this is not possible, the manuscript can be supplied as a single plain vanilla TeX or LaTeX file that includes all references and abbreviations, with no special formatting, as well as a PDF version that is uploaded as a 'related manuscript file'.

- **Provided.**

* Production-quality versions of all figures (see above).

- **Provided.**

* The final version of the Extended Data.

- **Provided.**

* The final version of any Supplementary Information, presented as one file (ideally a PDF) if feasible, as well as a separate SI Guide.

- **Provided.**

* Source Data, if appropriate.

- **Data Access policy provided.**

* For optimal quality videos we encourage H.264 encoding and a standard aspect ratio of 16:9 (4:3 is second best), without compression.

- **Not applicable.**

* Completed and signed copies of the following **four (or five) forms**, uploaded as a "Related Manuscript File" file type:

1) [Biology editorial checklist](#)

- **Provided.**

2) [Manuscript checklist](#)

- **Provided.**

3) [Reporting summary](#)

- **Provided.**

4) [Third-party rights table](#)

- **Provided.**

5) [Code and software submission checklist](#) (if applicable)

- **Not applicable.**

Referees' comments:

Below, we address the remaining concerns of the reviewers.

Referee #2 (Remarks to the Author):

2.6.11 Regarding the biological basis of our findings, we offer an evolutionary hypothesis in the discussion. Future studies will focus on understanding possible biological mechanisms.

Review 3: It is helpful that these limitations have been added, but I would still argue the overall framing of certainty around this finding which is (1) in conflict with the existing literature and (2) not replicated is too strong.

- **At the request of this Reviewer and the Editor, we have de-emphasized our autozygosity findings, moved these findings to a later section in Results, removed them from the Abstract, converted Figure 3 into an Extended Data Figure, emphasized the future need to replicate our findings, removed nominal associations, and removed from the Discussion a description of the “thrifty gene” evolutionary hypothesis.**

2.7.2 We agree with the reviewer that external validation, wherever possible, would be beneficial. We cite Malawsky et al. (2023) which finds an association between FROH and increased risk of T2D using the G&H data described by the reviewer. For many outcomes, such as MI, PGR is uniquely powered due not only to its large number of highly consanguineous individuals, but its case/control breakdown. Additional large, highly consanguineous cohorts with similar phenotypic distributions are not available.

Review 3: Thank you for your response. Do the authors have, for example, power calculations from other available cohorts to support their assertion that PROMIS-PGR is uniquely powered to detect these effects, even if only at a nominally or borderline significant (with consistent direction of effect) level? I am not suggesting this should be in the main text, but it seems difficult to believe an effect such as that observed for MI ~ FROH that is so well-powered for detection in a

cohort of ~150k individuals would be entirely absent in a highly consanguineous cohort of ~55k individuals (~20k Pakistani) like Genes & Health. Indeed this association was tested and not significant in Malawsky et al.

- **Contrary to the Reviewer's assertion, we are powered to detect these associations, as is evident by the highly significant p-values of the associations. Our study is not only larger than Genes & Health in number of samples but also includes a large MI case-control cohort. Nonetheless, we have de-emphasized our findings to satisfy the Reviewer and to follow the instructions of the Editor.**

2.7.2 We thank the reviewer for these comments. We report all gene-biomarker associations that are significant after multiple test correction ($p < 9.9e-8$). For select genes of clinical importance that are discovered through this approach, such as HBB, we show all nominal associations to provide a full picture of the gene, specifying that these additional associations are nominal ($p < 0.05$).

Review 3: Apologies for reiterating this point, but it is not clear how these “select” genes were defined or selected, whether it was a priori or not, and if not, what the relevance of nominally significant p values of ~0.05 is in the context of a PheWAS – the reader will be much less convinced if they suspect these associations were selectively picked out post-hoc. It is also not entirely clear how the $p < 9.9e-8$ is arrived at, as above, the authors describe this as exome-wide significant, but it is unclear if this is corrected for the multiple phenotypes tested? Shouldn't the value of interest be phenome-wide significant?

This is of course a relatively minor point on which we disagree, in the broader context of a well-run study and fabulous manuscript, and I will be more than content to leave it to the editor's expertise to determine whether this comment should be discounted or needs further consideration.

- **We appreciate the reviewer's concern about the potential for post hoc selection. The genes discussed were not highlighted on the basis of the primary exome-wide analysis. Our intention was not to use the nominal associations as stand-alone evidence, but instead to present them as supportive, hypothesis-generating observations. With respect to the $p < 9.9e-8$ threshold, this threshold was derived by considering all tests across genes/phenotypes as detailed in the methods: “A Bonferroni multiple testing correction was calculated based on the total number of unique tests across all genes and phenotypes, only counting once any gene for which the mask1 and mask2 variant sets were identical. A total of 506,644 tests were performed, giving a Bonferroni multiple test correction threshold of $P < 9.9e-08$.” We believe the manuscript already presents these results with appropriate caution and we thank reviewer for raising this point.**

Referee #4 (Remarks to the Author):

I think article titles and labels used in the article should be considered carefully especially when addressing humans, but obviously in all cases.

homozygous loss-of-function humans is a very strange way to refer to humans, and can easily be misconstrued in all kinds of political and social narratives. This is a term used for certain genetic variants and to apply it to humans like this is not reasonable, as it explicitly makes it seem like these humans have lost functions (extremely eugenic terms) when their own analyses show that certain variant profiles are not always correlated to traits.

I think it is important that the authors think of a more thoughtful title for the study, and also a better label throughout the article. I see this as a major concern to fix, they should likely consult with a bioethicist or historian of science (for e.g Anna Lewis, Tayyaba Jiwani).

- **We appreciate the reviewer's concerns. We have:**
 - o **Revised the title.**

- Replaced the terms “human knockouts” and “homozygous loss-of-function humans” and now refer precisely to homozygous loss-of-function variants.
 - Edited the text for clarity and sensitivity.
- Our study is not about eugenics in any way. Our study, like many other similar studies of biobanks published in this and other journals, describes individuals that carry naturally occurring biallelic predicted loss-of-function variants in specific genes.