

1 **TITLE: Titin-Truncating variants Predispose to Dilated Cardiomyopathy in Diverse**
2 **Populations**

3
4
5 **Authors:** John DePaolo, MD, PhD,¹ Marc Bornstein, AB,² Renae Judy, MS,¹ Sarah
6 Abramowitz, BA,¹ Shefali S. Verma, PhD,³ Michael G. Levin, MD,^{2,4} Zoltan Arany, MD,
7 PhD,^{2*} Scott M. Damrauer, MD,^{1,2,4,5*}

8
9 **Affiliations:**

10 ¹ Department of Surgery, Perelman School of Medicine, University of Pennsylvania,
11 Philadelphia, PA 19104, USA.

12 ² Cardiovascular Institute, Department of Medicine, Perelman School of Medicine,
13 University of Pennsylvania, PA 19104, USA.

14 ³ Department of Pathology and Laboratory Medicine, Perelman School of Medicine,
15 University of Pennsylvania, PA 19104, USA.

16 ⁴ Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA 19104, USA.

17 ⁵ Department of Genetics, Perelman School of Medicine, University of Pennsylvania,
18 Philadelphia, PA 19104, USA.

19 * Jointly supervised this work

20
21 **Text Word Count:** 2704

22 **References:** 25

23 **Tables:** 1

24 **Figures:** 3

25
26
27 **Address for correspondence:**

28 Scott Damrauer, MD

29 Perelman School of Medicine

30 Division of Vascular Surgery

31 Hospital of the University of Pennsylvania

32 3400 Spruce Street, 14th Floor South Perelman Center

33 Philadelphia, PA 19104

34 Office: 215-615-1698

35 scott.damrauer@penmedicine.upenn.edu

36

37 **Key Points**

38 **Question:** Do high percentage spliced in (hiPSI) titin truncating variants (TTNtvs) confer
39 similar levels of risk for dilated cardiomyopathy (DCM) across diverse populations?

40 **Findings:** In a cohort study that comprised 43,731 individuals of diverse genetic
41 background with electronic health records linked to whole exome sequencing data,
42 hiPSI TTNtvs conferred increased risk of DCM across all individuals irrespective of
43 genetic background as measured by genetic distance from the 1000 Genomes Project
44 European centroid.

45 **Meaning:** The findings of this study suggest that TTNtvs increase risk of DCM among
46 individuals independent of genetic background and that genetic similarity to a reference
47 population should not play a role in screening for genetic causes of dilated
48 cardiomyopathy.

49

50 **Abstract**

51 **Importance:** The effect of high percentage spliced in (hiPSI) *TTN* truncating variants
52 (TTNtvs) on risk of dilated cardiomyopathy (DCM) has historically been studied among
53 population subgroups defined by genetic similarity to European reference populations.
54 This has raised questions about the effect of TTNtvs in diverse populations, especially
55 among individuals genetically similar to African reference populations.

56 **Objective:** To determine the effect of TTNtvs on risk of DCM in diverse population as
57 measured by genetic distance (GD) in principal component (PC) space.

58 **Design:** Cohort study

59 **Setting:** Penn Medicine Biobank (PMBB) is a large, diverse biobank.

60 **Participants:** Participants were recruited from across the Penn Medicine healthcare
61 system and volunteered to have their electronic health records linked to biospecimen
62 data including DNA which has undergone whole exome sequencing.

63 **Main Outcomes and Measures:** Risk of DCM among individuals carrying a hiPSI
64 TTNtv.

65 **Results:** Carrying a hiPSI TTNtv was associated with DCM among PMBB participants
66 across a range of GD deciles from the 1000G European centroid; the effect estimates
67 ranged from odds ratio (OR) = 3.29 (95% confidence interval [CI] 1.26 to 8.56) to OR =
68 9.39 (95% CI 3.82 to 23.13). When individuals were assigned to population subgroups
69 based on genetic similarity to the 1000G reference populations, hiPSI TTNtvs conferred

70 significant risk of DCM among those genetically similar to the 1000G European
71 reference population (OR = 7.55, 95% CI 4.99 to 11.42, $P < 0.001$) and individuals
72 genetically similar to the 1000G African reference population (OR 3.50, 95% CI 1.48 to
73 8.24, $P = 0.004$).

74 **Conclusions and Relevance:** TTNtvs are associated with increased risk of DCM
75 among a diverse cohort. There is no significant difference in effect of TTNtvs on DCM
76 risk across deciles of GD from the 1000G European centroid, suggesting genetic
77 background should not be considered when screening individuals for titin-related DCM.

78

79 Introduction

80 Titin, the protein encoded by the *TTN* gene, is the largest protein in the human
81 body. It is found in the sarcomere where it spans from the Z-disk to the M-band, and is
82 critical for sarcomere assembly, contraction and relaxation in striated cardiac muscle.¹
83 In populations largely composed of individuals genetically similar to the 1000 Genomes
84 Project (1000G) European reference population (EUR), heterozygous *TTN* truncating
85 variants (TTNtvs) that encode for shortened forms of the titin protein have been
86 identified as a common genetic cause of dilated cardiomyopathy (DCM). These TTNtv
87 have been associated with 25% of familial cases of DCM, and 10-20% of sporadic
88 cases.²⁻⁷ Only variants located in exons that are highly likely to be spliced into adult
89 cardiac *TTN* transcripts, known as high percentage spliced in (hiPSI) variants, are
90 pathogenic.³ hiPSI TTNtvs may cause DCM by reducing abundance of full-length TTN
91 protein (haploinsufficiency) and/or through dominant negative effects.^{8,9}

92 We previously demonstrated the association of hiPSI TTNtvs with DCM
93 exclusively in EUR individuals, but were unable to detect an association between hiPSI
94 TTNtvs and DCM among Penn Medicine Biobank (PMBB) participants genetically
95 similar to the 1000G African reference population (AFR; odds ratio [OR] 1.8, 95% CI 0.2
96 to 13.7, $P=0.57$), or among AFR participants of the Jackson Heart Study.⁷ Recently, a
97 cross-sectional analysis of individuals with DCM demonstrated a statistically significant
98 but attenuated effect of predicted loss-of-function (pLOF) variants in *TTN* among AFR
99 individuals compared to EUR individuals, potentially underscoring the limitations of
100 applying genetic understanding of disease derived from a single ancestry group to
101 diverse populations.¹⁰

102 Common and rare genetic risk factors for disease have historically been
103 analyzed among different populations separated by ancestry.¹¹⁻¹⁴ This reflects the belief
104 that the artificial grouping of individuals into genetically common cohorts allows
105 improved assignment of genetic risk. However, greater understanding of genetic
106 similarities between groups of individuals combined with the results of genetic admixture
107 within different populations suggests that instead of strict dichotomization into groups,
108 genetic ancestry is better thought of as a continuum of relative similarity.¹²

109 Here we describe the assessment of DCM risk conferred by hiPSI TTNtvs in a
110 large, diverse biobank. Instead of using genetic similarity to group individuals together,
111 each participant's genetic distance (GD) from the 1000G European centroid was
112 calculated and risk of DCM was assessed across a range of genetic distances. As a
113 sensitivity analysis, participants were assigned an "population group" according to their
114 genetic similarity to one of the 1000G continental-level reference populations and risk of
115 DCM was stratified by "genetic similar" group. Due to previous research identifying
116 TTNtvs risk factors for atrial fibrillation (Afib),¹⁵⁻¹⁷ we performed a similar investigation
117 into the link between TTNtvs and Afib.

118

119 **Methods**

120 *Study Population*

121 The Penn Medicine BioBank is a genomic and precision medicine cohort
122 comprising participants who receive care in the Penn Medicine health system and who
123 consent to linkage of electronic health records with biospecimens, including 43,731 with
124 DNA which has undergone whole exome sequencing (pmbb.med.upenn.edu). As

125 previously described,⁷ DCM was defined either as ≥ 2 outpatient or ≥ 1 inpatient
126 encounters with: 1) the *International Classification of Diseases, 10th Revision* (ICD10)
127 diagnosis code of I42.0; or 2) ICD10 codes I42.8 or I42.9 or the *International*
128 *Classification of Diseases, Ninth Revision* (ICD9) codes 425.4, 425.8, or 425.9, and
129 mention of “dilated cardiomyopathy” or “DCM” in free text encounter notes. Ischemic
130 cardiomyopathy (ICM) was defined as ≥ 2 outpatient or ≥ 1 inpatient encounters with
131 ICD10 codes I24 or I25, or ICD9 codes 411 or 414. Afib was defined as ≥ 2 outpatient or
132 ≥ 1 inpatient encounters with ICD10 codes I48, I48.1, I48.2, or I48.9, or ICD9 codes
133 427.2 or 427.21. A total of 9,020 individuals in PMBB had transthoracic
134 echocardiography (TTE) data available that included left ventricular ejection fraction
135 (LVEF).

136 *Genetic Data*

137 Whole exome sequencing was performed as previously described by Regeneron
138 Genomics Center.¹⁸ Individual patient DNA samples were processed and sequenced on
139 the Illumina NovaSeq 6000 (Albany, NY, USA). WeCall variant caller (v2.0.0) was
140 employed for sequence alignment (GRCh38), variant identification, and genotype
141 assignment. Quality control exclusions included sex errors, high rates of heterozygosity
142 (D-statistic > 0.4), low sequence coverage, and genetically identified sample duplicates.
143 Single nucleotide variants (SNVs) were filtered for a read depth ≥ 7 and were retained if
144 they either had at least one heterozygous variant genotype with an allele balance ratio
145 ≥ 0.15 , or a homozygous variant genotype.¹⁹ Insertion-deletion variants (INDELs) were
146 filtered for a read depth ≥ 10 and either a heterozygous variant genotype with an allele
147 balance ≥ 0.20 , or a homozygous variant genotype.

148 *TTN Truncating Variant Calls*

149 TTNtvs were filtered for minor allele frequency < 0.001 and selected based on
150 predicted loss of function, truncating variants using ANNOVAR.²⁰ Splice site variants
151 were screened for those affecting canonical donor or acceptor splice sites (two bases
152 flanking either exon). Variants were considered hiPSI if percent spliced in was greater
153 than 90% (PSI>0.9).³

154 *Genetic Distance*

155 To consider differences in genetic ancestry along a continuous scale, genetic
156 distance (GD) from the 1000G European centroid was calculated as previously
157 described.¹⁴ Briefly, the 43,731 individuals in the PMBB were projected onto the 1000G²¹
158 principal component (PC) space. The GD was calculated by taking the Euclidean
159 distance for each individual from the 1000G European centroid using the equation

$$160 \quad d_i = \sqrt{\sum_{j=1}^{J=10} (pc_{ij} - pc_{Ej})^2}$$

where pc_{ij} is the j th PC of individual i , pc_{Ej} is the 1000G

161 European mean of the j th PC, and J is set to 10. GD was used to bin each individual in
162 PMBB into deciles from the closest 10% to the European centroid to the furthest 10%
163 from the European centroid.

164 *Statistical Analysis*

165 Logistic regression was employed to evaluate the association of hiPSI TTNtvs
166 with prevalent DCM adjusting for age and sex. Linear regression was used to assess
167 the effect of hiPSI TTNtvs on minimum left ventricular ejection fraction (LVEF) adjusting
168 for age and sex. Statistical analyses were performed using R (version 4.2.0).

169

170

171 **Results**

172 *Study Population*

173 There were 43,371 participants in the analytic cohort (**Table 1**). The median age
174 at analysis was 57 years (IQR: 45-69 years) and 21,907 (50%) were female. One
175 percent (436 individuals) carried a hiPSI TTNtv, 1,112 individuals (2.5%) had a DCM
176 diagnosis, and 4,920 individuals (11%) had an Afib diagnosis. There were 365 unique
177 TTNtvs identified including 174 stop-gains, 44 frameshifting insertion-deletions, and 147
178 essential splice site variants.

179 *Calculating genetic distance*

180 GD as the calculated Euclidean distance provides a method to assess the risk of
181 DCM conferred by TTNtvs across a genetically diverse population without relying on
182 artificial grouping based on genetic similarity to a reference panel. Euclidean distance is
183 the distance between two points in n^{th} -dimensional space. The vast majority of
184 understanding of the genetic risk of DCM, including that conferred by TTNtvs, was
185 derived from investigation of cohorts primarily comprised of individuals genetically
186 similar to the 1000G European reference population. Integrating the genetic distance
187 from the 1000G European centroid into the analysis of DCM risk among TTNtv carriers
188 accounts for population differences between individuals and the cohort from which
189 genetic understanding of disease has been derived. Therefore, GD provides two useful
190 tools in the current analysis: 1) a method to treat genetic similarity as a continuum
191 rather than an artificially dichotomized value; and 2) a way to test if decreasing genetic
192 similarity indicates a change in TTNtv-derived DCM risk that might support previous
193 findings of differing effects among genetically dissimilar groups.

194 GD was calculated for each individual in PMBB and genetic diversity by GD was
195 compared to categorical assignment to population subgroups similar to the 1000G
196 reference populations based on PC analysis (**Figure 1A-D**). To demonstrate the extent
197 of genetic admixture and the arbitrary nature of grouping individuals genetically similar
198 to reference populations, PMBB individuals were compared to 1000G reference panel
199 individuals based on PC analysis (**Figure 1E-G**). Individuals were then binned by GD
200 decile; each decile included a range of individuals genetically similar to a 1000G
201 reference population (**Supplemental Table 1**) and PC analysis by GD decile
202 redemonstrated the extent of genetic admixture (**Supplemental Figure 1**). Together
203 these analyses characterize the genetic heterogeneity in the PMBB population.

204

205 *Effect of TTNtvS on risk of DCM across genetic distances*

206 In the population overall, there is a 5.6-fold increased risk of DCM conferred by
207 carrying a TTNtv when adjusting for age, sex, and the first five genetic PCs
208 [**Supplemental Table 2**]. To determine the effect of TTNtvS on DCM risk at different
209 GDs from the European centroid, logistic regression analysis was performed among
210 PMBB individuals binned by GD decile adjusting for age and sex within each bin.
211 Among those individuals closest to the 1000G European centroid by GD, hiPSI TTNtvS
212 were associated with DCM (Decile 1 OR = 5.82, 95% CI 2.42 to 14.01, $P < 0.001$)
213 [**Figure 2A**]; there was no significant effect decomposition as GD decile increased
214 (Decile 10 OR = 9.39, 95% CI 3.82 to 23.13, $P < 0.001$). Meta-analysis of binned
215 results demonstrated an overall OR = 5.42 (95% CI 4.01 to 7.32, $P < 0.001$); there
216 was no evidence of significant heterogeneity ($I^2 = 0\%$, $\chi^2 = 4.59$, $P = 0.87$). This was

217 supported by a second analysis among all participants in which either GD from the
218 1000G European centroid alone nor the interaction between GD and hiPSI TTNtvs had
219 a statistically significant effect on the risk of DCM (**Supplemental Table 2**). Taken
220 together these results suggest that GD does not alter DCM risk. Similarly, there was no
221 evident trend of greater or lesser effect on minimum LVEF as GD increased (**Figure**
222 **2B**), though there was a trend towards increased heterogeneity ($I^2=42\%$, $\chi^2=15.48$,
223 $P=0.08$). We conclude that hiPSI TTNtvs associate with increased DCM risk and
224 reduced minimum LVEF across diverse populations.

225

226 *Comparison to categorical assignment of genetic diversity*

227 We previously demonstrated a significant effect on DCM risk conferred by
228 carrying TTNtvs among individuals genetically similar to the 1000G EUR reference
229 population, however could not detect an effect among individuals genetically similar to
230 the 1000G AFR reference population.⁷ In this updated study, encompassing a
231 substantially larger number of participants, we evaluated the effect of TTNtvs among
232 individuals artificially separated into groups to compare both to our previous published
233 results and to the results above. Carrying a hiPSI TTNtv was associated with DCM
234 among individuals genetically similar to the 1000G EUR reference population, whether
235 including (**Supplemental Figure 2**) or excluding individuals with ischemic
236 cardiomyopathy [ICM] (odds ratio [OR] 7.55, 95% CI 4.99 to 11.42, $P<0.001$) [**Figure**
237 **3A**]. Similar, but attenuated, results were observed among individuals genetically similar
238 to the 1000G AFR reference population (OR 3.50, 95% CI 1.48 to 8.24, $P=0.004$).
239 Linear regression analysis demonstrated a consistent effect on minimum LVEF

240 conferred by carrying a TTNtv without attenuation in the AFR population group **[Figure**
241 **3B]**. We conclude that when individuals are artificially grouped by genetic similarity to a
242 reference population, hiPSI TTNtvs continue to associate with a diagnosis of DCM and
243 negatively associate with decreased minimum LVEF across diverse groups.

244

245 *Effect of titin truncating variants on risk of atrial fibrillation*

246 To test the effect of TTNtvs on risk of Afib, another cardiovascular diagnosis
247 previously shown to associate with TTNtvs, we used logistic regression analysis by GD
248 and by genetically similar group. Among those individuals closest to the 1000G
249 European centroid by GD, hiPSI TTNtvs were associated with Afib (Decile 1 OR = 2.68,
250 95% CI 1.40 to 5.16, $P=0.003$) **[Supplemental Figure 3A]**. While there was increased
251 variability across deciles (deciles 2,4, and 5 had statistically insignificant associations
252 between hiPSI TTNtvs and Afib, $I^2=40\%$, $\chi^2=14.96$, $P=0.09$), there was no significant
253 trend in effect as GD from the 1000G European centroid increased (Decile 10 OR =
254 4.36, 95% CI 2.21 to 8.61, $P<0.001$). When PMBB individuals were grouped by genetic
255 similarity to a reference population, the effect estimate among individuals genetically
256 similar to the 1000G EUR reference population (OR = 2.11, 95% CI 1.62 to 2.77,
257 $P<0.001$) was similar to the effect estimate among individuals genetically similar to the
258 1000G AFR reference population (OR = 1.98, 95% CI 1.02 to 3.85, $P=0.04$)
259 **[Supplemental Figure 3B]**. We conclude that TTNtvs associate with increased risk of
260 Afib across diverse populations.

261

262

263 **Discussion**

264 Using continuous genetic distance as an alternative method to investigate the
265 effect of genetic similarity on rare variant cardiovascular disease risk, we demonstrated
266 that hiPSI TTNtvs associated with increased risk of DCM and reduced minimum LVEF
267 across diverse populations. These results were supported when individuals were
268 artificially dichotomized by genetic similarity to reference populations, although the
269 effect estimate among individuals genetically similar to the AFR reference population
270 may be attenuated. We also demonstrated a consistent though less robust effect of
271 hiPSI TTNtvs on risk of Afib across diverse populations.

272 Although genetic epidemiology studies have historically relied on arbitrary cutoffs
273 to differentiate and label populations of similar ancestry, genetic ancestry is more
274 appropriately considered as a continuum due to the degree of admixture even among
275 similar populations.²²⁻²⁴ The question of TTNtv effect among individuals genetically
276 similar to the 1000G AFR reference population provided an opportunity to reassess the
277 role ancestry plays in how we approach individual risk in cardiovascular disease. The
278 homogeneity of effect of TTNtvs across deciles of GD from the 1000G European
279 centroid removes barriers constructed by presenting results separated by artificially
280 constructed population subgroups, such as within group variation and small sample
281 size. Future evaluation of monogenic risk of cardiovascular disease may benefit from
282 utilizing a similar approach, which may be more broadly applicable than analyses using
283 genetically similar group dichotomization alone.

284 These results refine our previous findings that hiPSI TTNtvs are associated with
285 DCM among individuals genetically similar to EUR reference populations in PMBB.

286 Previously, we reported an OR of 18.7 (95% CI 9.1 to 39.4) for DCM among EUR
287 individuals with hiPSI TTNtvs.⁷ The range of effect estimates in our analysis of risk by
288 GD from the 1000G European centroid was OR 3.3 (95% CI 1.3 to 8.6) to OR 9.4 (95%
289 CI 3.8 to 23.1), and the effect estimate for individuals genetically similar to EUR
290 reference panels was contained within that range (OR 7.6, 95% CI 5.0 to 11.4,
291 $P<0.001$). The attenuation of effect we observed between our initial publication and the
292 present study may reflect a winner's curse in our initial study, where the initially-reported
293 effect size tends to be over-estimated in genetic association studies.²⁵ However, the
294 makeup of PMBB has also changed substantially in the past five years such that it is
295 less enriched for cardiovascular disease (6.2% prevalence of DCM previously
296 compared to 2.5% now) due to increasingly broad regional enrollment from centers of
297 primary care which may have had a more meaningful impact on our current results.

298 Our findings also clarify the effect of hiPSI TTNtvs in individuals genetically
299 similar to the AFR reference population. We previously identified no statistically
300 significant effect among PMBB participants genetically similar to the AFR reference
301 population.⁷ This corroborated other evidence suggesting the possibility of differing
302 effects across diverse populations.³ Notably, our previous PMBB cohort included only
303 2,123 participants genetically similar to the AFR reference population, 20 of whom
304 carried a hiPSI TTNtv, representing a power of 55% to identify an OR of 3.3, suggesting
305 we were previously underpowered to identify an effect similar to what we presently
306 report. We now identify a significant effect of TTNtvs (OR = 3.5, 95% CI 1.5 to 8.2,
307 $P=0.004$) in this population. Moreover, when we consider the extent of admixture that
308 exists in this population, the dispersion of PMBB participants genetically similar to the

309 1000G AFR reference population across a range of GD from the 1000G European
310 centroid, and the similar effect estimates across those distances, these findings are
311 convincing that hiPSI TTNtvs confer increased DCM risk among individuals within these
312 GD deciles corresponding to those previously labeled genetically similar to the 1000G
313 AFR reference population.

314

315 **Limitations**

316 This work has some limitations. Reliance on electronic health records to establish
317 a diagnosis of DCM may have misclassified a subset of individuals. Similarly, TTEs
318 were not interpreted by a uniform reader. Our cohort is also largely composed of EUR
319 and AFR individuals, limiting generalizability of the GD continuum concept to other
320 genetically similar groups.

321

322 **Conclusion**

323 In summary, our updated analysis demonstrates a strong association between
324 hiPSI TTNtvs and DCM across a range of genetic backgrounds, including subjects
325 genetically similar to African reference population. These data suggest that
326 recommendations for genetic testing and counseling should not differ between diverse
327 populations.

328

329 **Acknowledgements, Conflicts of Interest and Disclosures**

330 We acknowledge the Penn Medicine BioBank (PMBB) for providing data and
331 thank the patient-participants of Penn Medicine who consented to participate in this

332 research program. We would also like to thank the Penn Medicine BioBank team and
333 Regeneron Genetics Center for providing genetic variant data for analysis. The PMBB is
334 approved under IRB protocol# 813913 and supported by Perelman School of Medicine
335 at University of Pennsylvania, a gift from the Smilow family, and the National Center for
336 Advancing Translational Sciences of the National Institutes of Health under CTSA
337 award number UL1TR001878.

338 M.G.L. receives research support to his institution from MyOme outside of this
339 work. S.M.D receives research support to his institution from RenalytixAI and in kind
340 support from Novo Nordisk, both outside of this work.

341

342 **Sources of Funding**

343 J.D. is supported by the American Heart Association (23POST1011251). M.G.L.
344 received support from the Institute for Translational Medicine and Therapeutics of the
345 Perelman School of Medicine at the University of Pennsylvania, the NIH/NHLBI National
346 Research Service Award postdoctoral fellowship (T32HL007843), the Measey
347 Foundation, and the Doris Duke Foundation (2023-0224). Z.A. is supported by the
348 NIH/NHLBI (R01-HL152446) and the Department of Defense (W81XWH18-1-0503).

349

350

351 **References**

- 352
- 353 1. Freiburg A, Trombitas K, Hell W, Cazorla O, Fougerousse F, Centner T,
354 Kolmerer B, Witt C, Beckmann JS, Gregorio CC, et al. Series of exon-skipping
355 events in the elastic spring region of titin as the structural basis for myofibrillar
356 elastic diversity. *Circ Res*. 2000;86:1114-1121. doi: 10.1161/01.res.86.11.1114
- 357 2. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D,
358 Conner L, DePalma SR, McDonough B, Sparks E, et al. Truncations of titin
359 causing dilated cardiomyopathy. *N Engl J Med*. 2012;366:619-628. doi:
360 10.1056/NEJMoa1110186
- 361 3. Roberts AM, Ware JS, Herman DS, Schafer S, Baksi J, Bick AG, Buchan RJ,
362 Walsh R, John S, Wilkinson S, et al. Integrated allelic, transcriptional, and
363 phenomic dissection of the cardiac effects of titin truncations in health and
364 disease. *Sci Transl Med*. 2015;7:270ra276. doi: 10.1126/scitranslmed.3010134
- 365 4. Akinrinade O, Alastalo TP, Koskenvuo JW. Relevance of truncating titin
366 mutations in dilated cardiomyopathy. *Clin Genet*. 2016;90:49-54. doi:
367 10.1111/cge.12741
- 368 5. Haas J, Frese KS, Peil B, Kloos W, Keller A, Nietsch R, Feng Z, Muller S,
369 Kayvanpour E, Vogel B, et al. Atlas of the clinical genetics of human dilated
370 cardiomyopathy. *Eur Heart J*. 2015;36:1123-1135a. doi:
371 10.1093/eurheartj/ehu301
- 372 6. Akinrinade O, Ollila L, Vattulainen S, Tallila J, Gentile M, Salmenpera P, Koillinen
373 H, Kaartinen M, Nieminen MS, Myllykangas S, et al. Genetics and genotype-

- 374 phenotype correlations in Finnish patients with dilated cardiomyopathy. *Eur Heart*
375 *J.* 2015;36:2327-2337. doi: 10.1093/eurheartj/ehv253
- 376 7. Haggerty CM, Damrauer SM, Levin MG, Birtwell D, Carey DJ, Golden AM,
377 Hartzel DN, Hu Y, Judy R, Kelly MA, et al. Genomics-First Evaluation of Heart
378 Disease Associated With Titin-Truncating Variants. *Circulation.* 2019;140:42-54.
379 doi: 10.1161/CIRCULATIONAHA.119.039573
- 380 8. McAfee Q, Chen CY, Yang Y, Caporizzo MA, Morley M, Babu A, Jeong S,
381 Brandimarto J, Bedi KC, Jr., Flam E, et al. Truncated titin proteins in dilated
382 cardiomyopathy. *Sci Transl Med.* 2021;13:eabd7287. doi:
383 10.1126/scitranslmed.abd7287
- 384 9. McAfee Q, Caporizzo MA, Uchida K, Bedi KC, Jr., Margulies KB, Arany Z,
385 Prosser BL. Truncated titin protein in dilated cardiomyopathy incorporates into
386 the sarcomere and transmits force. *J Clin Invest.* 2023. doi: 10.1172/JCI170196
- 387 10. Jordan E, Kinnamon DD, Haas GJ, Hofmeyer M, Kransdorf E, Ewald GA, Morris
388 AA, Owens A, Lowes B, Stoller D, et al. Genetic Architecture of Dilated
389 Cardiomyopathy in Individuals of African and European Ancestry. *JAMA.*
390 2023;330:432-441. doi: 10.1001/jama.2023.11970
- 391 11. Astore C, Sharma S, Nagpal S, Consortium NIG, Cutler DJ, Rioux JD, Cho JH,
392 McGovern DPB, Brant SR, Kugathasan S, et al. The role of admixture in the rare
393 variant contribution to inflammatory bowel disease. *Genome Med.* 2023;15:97.
394 doi: 10.1186/s13073-023-01244-w
- 395 12. Lewis ACF, Molina SJ, Appelbaum PS, Dauda B, Di Rienzo A, Fuentes A,
396 Fullerton SM, Garrison NA, Ghosh N, Hammonds EM, et al. Getting genetic

- 397 ancestry right for science and society. *Science*. 2022;376:250-252. doi:
398 10.1126/science.abm7530
- 399 13. Busby GB, Kulm S, Bolli A, Kintzle J, Domenico PD, Botta G. Ancestry-specific
400 polygenic risk scores are risk enhancers for clinical cardiovascular disease
401 assessments. *Nat Commun*. 2023;14:7105. doi: 10.1038/s41467-023-42897-w
- 402 14. Ding Y, Hou K, Xu Z, Pimplaskar A, Petter E, Boulier K, Prive F, Vilhjalmsson BJ,
403 Olde Loohuis LM, Pasaniuc B. Polygenic scoring accuracy varies across the
404 genetic ancestry continuum. *Nature*. 2023;618:774-781. doi: 10.1038/s41586-
405 023-06079-4
- 406 15. Ahlberg G, Refsgaard L, Lundegaard PR, Andreassen L, Ranthe MF, Linscheid N,
407 Nielsen JB, Melbye M, Haunso S, Sajadieh A, et al. Rare truncating variants in
408 the sarcomeric protein titin associate with familial and early-onset atrial
409 fibrillation. *Nat Commun*. 2018;9:4316. doi: 10.1038/s41467-018-06618-y
- 410 16. Corden B, Jarman J, Whiffin N, Tayal U, Buchan R, Sehmi J, Harper A,
411 Midwinter W, Lascelles K, Markides V, et al. Association of Titin-Truncating
412 Genetic Variants With Life-threatening Cardiac Arrhythmias in Patients With
413 Dilated Cardiomyopathy and Implanted Defibrillators. *JAMA Netw Open*.
414 2019;2:e196520. doi: 10.1001/jamanetworkopen.2019.6520
- 415 17. Akhtar MM, Lorenzini M, Cicerchia M, Ochoa JP, Hey TM, Sabater Molina M,
416 Restrepo-Cordoba MA, Dal Ferro M, Stolfo D, Johnson R, et al. Clinical
417 Phenotypes and Prognosis of Dilated Cardiomyopathy Caused by Truncating
418 Variants in the TTN Gene. *Circ Heart Fail*. 2020;13:e006832. doi:
419 10.1161/CIRCHEARTFAILURE.119.006832

- 420 18. Dewey FE, Gusarova V, O'Dushlaine C, Gottesman O, Trejos J, Hunt C, Van
421 Hout CV, Habegger L, Buckler D, Lai KM, et al. Inactivating Variants in
422 ANGPTL4 and Risk of Coronary Artery Disease. *N Engl J Med*. 2016;374:1123-
423 1133. doi: 10.1056/NEJMoa1510926
- 424 19. Verma A, Damrauer SM, Naseer N, Weaver J, Kripke CM, Guare L, Sirugo G,
425 Kember RL, Drivas TG, Dudek SM, et al. The Penn Medicine BioBank: Towards
426 a Genomics-Enabled Learning Healthcare System to Accelerate Precision
427 Medicine in a Diverse Population. *J Pers Med*. 2022;12. doi:
428 10.3390/jpm12121974
- 429 20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic
430 variants from high-throughput sequencing data. *Nucleic Acids Res*.
431 2010;38:e164. doi: 10.1093/nar/gkq603
- 432 21. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample
433 Resource (IGSR) collection of open human genomic variation resources. *Nucleic*
434 *Acids Res*. 2020;48:D941-D947. doi: 10.1093/nar/gkz836
- 435 22. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, Vergara
436 C, Torgerson DG, Pino-Yanes M, Shringarpure SS, et al. A continuum of
437 admixture in the Western Hemisphere revealed by the African Diaspora genome.
438 *Nat Commun*. 2016;7:12522. doi: 10.1038/ncomms12522
- 439 23. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, Hirbo
440 JB, Awomoyi AA, Bodo JM, Doumbo O, et al. The genetic structure and history of
441 Africans and African Americans. *Science*. 2009;324:1035-1044. doi:
442 10.1126/science.1172257

- 443 24. Gouveia MH, Borda V, Leal TP, Moreira RG, Bergen AW, Kehdy FSG, Alvim I,
444 Aquino MM, Araujo GS, Araujo NM, et al. Origins, Admixture Dynamics, and
445 Homogenization of the African Gene Pool in the Americas. *Mol Biol Evol.*
446 2020;37:1647-1656. doi: 10.1093/molbev/msaa033
- 447 25. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of
448 genetic association studies supports a contribution of common variants to
449 susceptibility to common disease. *Nat Genet.* 2003;33:177-182. doi:
450 10.1038/ng1071
- 451
- 452

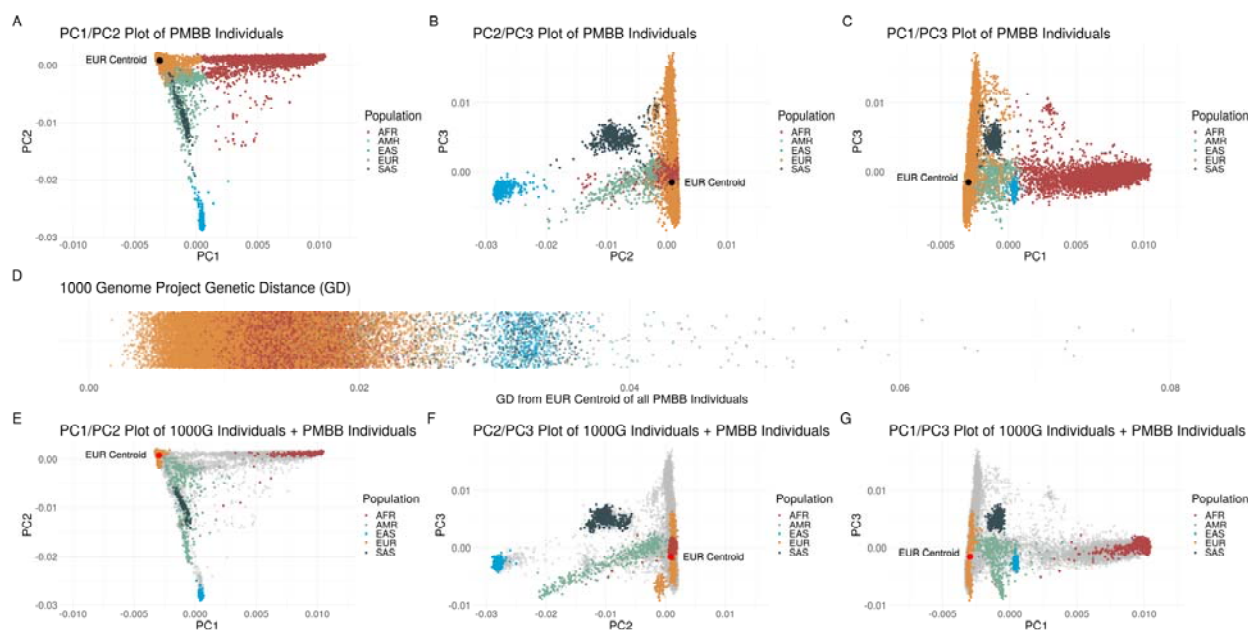
453 **Tables and Figure Legends**

454 **Table 1: Clinical characteristics of individuals in the Penn Medicine Biobank with**
 455 **and without high percentage spliced-in *TTN* truncating variants.**

Demographic	hiPSI <i>TTN</i>tv (436)	non-hiPSI <i>TTN</i>tv (43,295)	P-value Difference
Age, median (IQR)	56 (45-67)	57 (45-69)	0.65
Female Sex, N (%)	184 (42%)	21,723 (50%)	<0.001
"Genetically Similar" Group, N (%)			
EUR	330 (75%)	29,626 (68%)	0.002
AFR	88 (20%)	11,048 (26%)	0.008
AMR	2 (0.4%)	564 (1.3%)	0.12
EAS	4 (0.9%)	466 (1.1%)	0.74
SAS	5 (1.1%)	554 (1.3%)	0.79
Unknown	7 (1.6%)	1,037 (2.4%)	0.28
Dilated Cardiomyopathy, N (%)	52 (12%)	1,060 (2.4%)	<0.001
Ischemic Cardiomyopathy, N (%)	64 (15%)	5,305 (12%)	0.13
Atrial Fibrillation, N (%)	89 (20%)	4,831 (11%)	<0.001

456

457



458

459 **Figure 1: Principal component based clusters of individuals in the Penn Medicine**
460 **Biobank compared to genetic distance from the 1000 Genomes Project European**
461 **centroid and compared to continental-level clusters of 1000 Genomes Project**
462 **individuals demonstrating the degree of overlap of different genetically similar**
463 **groups. A-C:** Discrete labelling of the position of each individual within the Penn
464 Medicine Biobank (PMBB) colored by genetically similar group based on 1000
465 Genomes Project (1000G) continental-level reference populations in a plot of (A)
466 principal component (PC) 1 versus PC2, (B) PC2 versus PC3, and (C) PC1 versus PC3
467 with 1000G European centroid (black dot) identified in each plot. (D) the genetic
468 distance from the 1000G European centroid of each individual within PMBB again
469 colored by genetically similar group based on 1000 Genomes Project (1000G)
470 continental-level reference populations. **E-G:** Discrete labelling of the position of each
471 1000G reference panel individuals colored by continental region with PMBB individuals

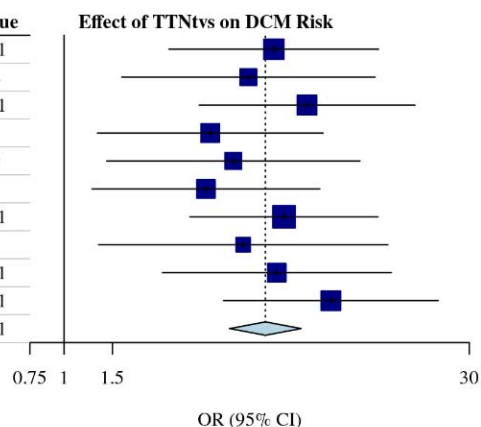
472 represented by grey dots in a plot of **(E)** PC1 versus PC2, **(F)** PC2 versus PC3, and **(G)**

473 PC1 versus PC3 with 1000G European centroid (red dot) identified in each plot.

474

A) Effect of TTNts on Risk of DCM

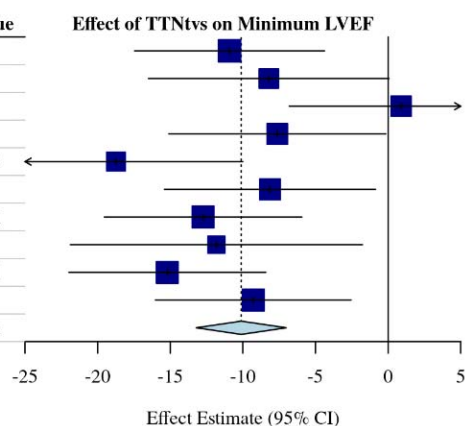
Decile Cohort	OR (95% CI)	P-value
Decile 1 (Closest to EUR Centroid)	5.82 [2.42; 14.01]	< .001
Decile 2	4.70 [1.62; 13.60]	.004
Decile 3	7.69 [3.11; 19.01]	< .001
Decile 4	3.41 [1.32; 8.79]	.01
Decile 5	4.14 [1.43; 11.99]	.009
Decile 6	3.29 [1.26; 8.56]	.01
Decile 7	6.34 [2.88; 13.95]	< .001
Decile 8	4.50 [1.34; 15.16]	.02
Decile 9	5.97 [2.28; 15.59]	< .001
Decile 10 (Furthest from to EUR Centroid)	9.39 [3.82; 23.13]	< .001
Meta-analysis	5.42 [4.01; 7.32]	< .001



Heterogeneity: $\chi^2 = 4.59$ ($P = .87$), $I^2 = 0\%$

B) Effect of TTNts on Minimum LVEF

Decile Population	Beta (95% CI)	P-value
Decile 1 (Closest to EUR Centroid)	-10.92 [-17.45; -4.38]	.001
Decile 2	-8.22 [-16.50; 0.07]	.05
Decile 3	0.89 [-6.80; 8.59]	.82
Decile 4	-7.62 [-15.10; -0.15]	.05
Decile 5	-18.74 [-27.48; -9.99]	< .001
Decile 6	-8.13 [-15.40; -0.86]	.03
Decile 7	-12.74 [-19.53; -5.94]	< .001
Decile 8	-11.83 [-21.88; -1.77]	.02
Decile 9	-15.20 [-21.98; -8.42]	< .001
Decile 10 (Furthest from EUR Centroid)	-9.29 [-16.03; -2.55]	.007
Meta-analysis	-10.11 [-13.22; -7.01]	< .001



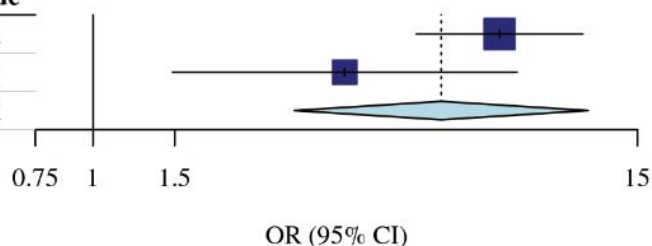
Heterogeneity: $\chi^2 = 15.48$ ($P = .08$), $I^2 = 42\%$

475

476 **Figure 2: Effect of TTN truncating variants on risk of dilated cardiomyopathy and**
 477 **on minimum left ventricular ejection fraction, stratified by genetic distance from**
 478 **the 1000 Genomes Project European centroid. (A)** Logistic regression analysis of the
 479 association between hiPSI TTNts on risk of DCM stratified by genetic distance (GD)
 480 from the 1000 Genomes Project (1000G) European centroid. **(B)** Linear regression
 481 analysis of the association between hiPSI TTNts and minimum left ventricular ejection
 482 fraction stratified by GD from the 1000G European centroid. OR = odds ratio; CI =
 483 confidence interval; EUR = European; LVEF = left ventricular ejection fraction.

A) Effect of TTNtvs on Risk of DCM Dichotomized by Genetically Similar Group

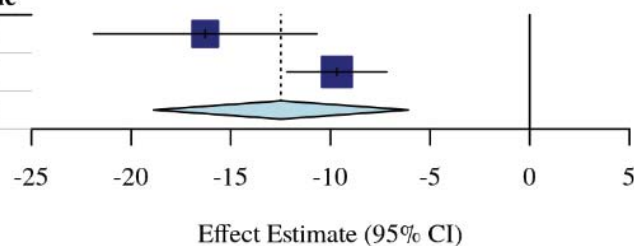
Population	OR (95% CI)	P-value
EUR	7.55 [4.99; 11.42]	< .001
AFR	3.50 [1.48; 8.24]	.004
Meta-analysis	5.65 [2.72; 11.74]	< .001



Heterogeneity: $\chi^2_1 = 2.52$ ($P = .11$), $I^2 = 60\%$

B) Effect of TTNtvs on Minimum LVEF Dichotomized by Genetically Similar Group

Population	Beta (95% CI)	P-value
AFR	-16.28 [-21.88; -10.68]	< .001
EUR	-9.68 [-12.18; -7.18]	< .001
Meta-analysis	-12.48 [-18.88; -6.09]	< .001



Heterogeneity: $\chi^2_1 = 4.45$ ($P = .03$), $I^2 = 78\%$

484

485

486

487

488

489

490

491

492

493

494

495

Figure 3: Effect of high percentage spliced in *TTN* truncating variant on risk of dilated cardiomyopathy diagnosis and minimum left ventricular ejection fraction reduction stratified by genetically similar group excluding those with ischemic cardiomyopathy in the Penn Medicine Biobank. (A) Logistic regression analysis of the association between hiPSI TTNtvs and DCM diagnosis, and (B) linear regression analysis of the association between hiPSI TTNtvs and minimum left ventricular ejection fraction among individuals genetically similar to the 1000 Genomes Project European and African reference population, and meta-analyzed. OR = odds ratio; CI = confidence interval; EUR = individuals genetically similar to the European reference population; AFR = individuals genetically similar to the African reference population.