

1 **Identification of actionable genetic variants in 4,198 Scottish volunteers from the Viking Genes research**  
2 **cohort and implementation of return of results**

3

4 Shona M. Kerr,<sup>1\*</sup> Lucija Klaric,<sup>1\*</sup> Marisa D. Muckian,<sup>2,3</sup> Kiera Johnston,<sup>2</sup> Camilla Drake,<sup>1</sup> Mihail Halachev,<sup>1</sup>  
5 Emma Cowan,<sup>4,5</sup> Lesley Snadden,<sup>4,5</sup> John Dean,<sup>5</sup> Sean L. Zheng,<sup>6,7,8</sup> Prisca K. Thami,<sup>6,7</sup> James S. Ware,<sup>6,7,8</sup>  
6 Gannie Tzoneva,<sup>9</sup> Alan R. Shuldiner,<sup>9</sup> Zosia Miedzybrodzka<sup>4,5</sup> and James F. Wilson<sup>†1,2,10</sup>

7

8 <sup>1</sup>MRC Human Genetics Unit, University of Edinburgh, Institute of Genetics and Cancer, Western General  
9 Hospital, Crewe Road, Edinburgh, EH4 2XU, UK.

10 <sup>2</sup>Centre for Global Health Research, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh EH8  
11 9AG, UK.

12 <sup>3</sup>Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel  
13 Street, London WC1E 7HT, UK.

14 <sup>4</sup>Department of Medical Genetics, Ashgrove House, NHS Grampian, Aberdeen, AB25 2ZA, UK.

15 <sup>5</sup>Medical Genetics Group, University of Aberdeen, Polwarth Building, Aberdeen, AB25 2ZD, UK.

16 <sup>6</sup>National Heart and Lung Institute, Imperial College London, London, UK.

17 <sup>7</sup>MRC Laboratory of Medical Sciences, Imperial College London, London, UK.

18 <sup>8</sup>Royal Brompton & Harefield Hospitals, Guy's and St. Thomas' NHS Foundation Trust, London, UK.

19 <sup>9</sup>Regeneron Genetics Center, Tarrytown, NY, USA.

20 <sup>10</sup>Centre for Genomic and Experimental Medicine, University of Edinburgh, Institute of Genetics and  
21 Cancer, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK.

22

23 \* These authors contributed equally

24 † Correspondence: [jim.wilson@ed.ac.uk](mailto:jim.wilson@ed.ac.uk)

25 **Abstract**

26 The benefits of returning clinically actionable genetic results to participants in research cohorts are  
27 accruing, yet such a genome-first approach is challenging. Here, we describe the return of such results in  
28 two founder populations from Scotland. Between 2005 and 2015, we recruited >4,000 adults with  
29 grandparents from Orkney and Shetland into the Viking Genes research cohort. Return of genetic data  
30 was not offered at baseline, but in 2023 we sent invitations for consent to return of actionable genetic  
31 findings to participants. We generated exome sequence data from 4,198 participants, and used the ACMG  
32 v3.2 list of 81 genes, ClinVar review and pathogenicity status, plus manual curation, to develop a pipeline  
33 to identify potentially actionable variants. We identified 104 individuals (2.5%) carrying 108 actionable  
34 genotypes at 39 variants in 23 genes, and validated these. Working with the NHS clinical genetics service,  
35 which provided genetic counselling and clinical verification of the research results, and after expert clinical  
36 review, we notified 64 consenting participants (or their next of kin) of their actionable genotypes. Ten  
37 actionable variants across seven genes (*BRCA1*, *BRCA2*, *ATP7B*, *TTN*, *KCNH2*, *MUTYH*, *GAA*) have risen 50  
38 to >3,000-fold in frequency through genetic drift in ancestral island localities. Viking Genes is one of the  
39 first UK research cohorts to return actionable findings, providing an ethical and logistical exemplar of  
40 return of results. The genetic structure in the Northern Isles of Scotland, with multiple founder effects,  
41 provides a unique opportunity for a tailored approach to primary and secondary prevention through  
42 genetic screening.

43

44 **Keywords**

45 Return of results, actionable variant, founder effect, exomes, Orkney, Shetland

## 46 Introduction

47 Where exome or genome sequencing is clinically indicated, the opportunity arises to report additional  
48 incidental findings with the aim of diagnosis and pre-emptive clinical intervention. The American College  
49 of Medical Genetics and Genomics (ACMG) has led this approach and annually updates a benchmark list of  
50 genes for report<sup>1</sup>. The list currently has 81 genes, which meet strict criteria: there must be strong evidence  
51 that certain variants within the gene cause disease, and there must be effective medical intervention (for  
52 example screening or medication) which can be offered to the carrier. The great majority of the genes are  
53 associated with either inherited cancer predisposition, cardiovascular risk or metabolic diseases. These  
54 conditions are mostly inherited in an autosomal dominant fashion, but variants in genes causing  
55 autosomal recessive disease which are only actionable if homozygous or compound heterozygous are also  
56 present in the ACMG v3.2 list. Although there has not yet been universal international adoption of this  
57 approach, the ACMG list is the most widely used.

58 The consent obtained from participants in many research cohorts to date, including UK Biobank (UKB)<sup>2</sup>,  
59 does not allow the return of results about actionable variants (AVs) or genotypes. However, some newer  
60 US cohorts, such as *All of Us*, are committed to making their research results, including hereditary disease  
61 risk, accessible to participants<sup>3</sup>. Geisinger's MyCode Community Health Initiative provides a model for  
62 "genome-first" care<sup>4</sup>, and an international policy on return of results acknowledges the potential medical  
63 benefits to the individuals who are participating in genomics research<sup>5</sup>.

64 The populations of the Northern Isles of Scotland – the Orkney and Shetland archipelagos - are the most  
65 genetically isolated of all British and Irish populations, with the highest degree Norse admixture, and an  
66 enrichment of rare and low frequency functional variants<sup>6-9</sup>. Viking Genes ([viking.ed.ac.uk](http://viking.ed.ac.uk)) comprises a set  
67 of Northern and Western Isles population cohort studies, aiming to explore genetic causes of disease – the  
68 Orkney Complex Disease Study (ORCADES), and VIKING I, II and III. ORCADES<sup>9</sup> and VIKING I<sup>10</sup> together  
69 contain a rich data resource of more than 4,000 deeply phenotyped and exome sequenced research  
70 subjects, nearly all with three or four grandparents from Orkney, or Shetland, respectively. In total, Viking  
71 Genes has over 10,000 participants, all of whom consented to analysis of their DNA, and to be contacted  
72 at a later date in connection with future ethically approved studies. We have previously reported founder

73 variants in *KCNH2*, *BRCA1* and *BRCA2* in Shetland and Orkney populations<sup>10-12</sup>, with variants tracing back  
74 to individual island ancestries.

75 In Viking Genes, all cohort members were offered the option of consent to return of actionable genetic  
76 results. Building on a method developed by Kelly *et al*<sup>13</sup> to screen 130,048 adults in the Geisinger MyCode  
77 project, we used the ACMG v3.2 list<sup>1</sup> and variant classification information from the public database  
78 ClinVar<sup>14</sup> to code automated filters to identify potential actionable variants within 4,198 members of the  
79 Viking Genes cohort. Here, we describe the analysis pipeline, summarize actionable results, the process of  
80 returning these, and some challenges associated with implementation of these procedures at scale. We  
81 provide evidence of multiple drifted actionable founder variants that could form the basis for a future  
82 targeted Northern Isles population genetic screening programme.

83

## 84 **Subjects and Methods**

### 85 *Research Volunteer Recruitment and DNA Collection*

86 Recruitment to ORCADES took place from 2005-2011<sup>11</sup>, and to Viking I (Viking Health Study – Shetland)  
87 from 2013-2015<sup>10</sup>. Blood or occasionally saliva samples from participants were collected, processed and  
88 stored using standard operating procedures and managed through a laboratory information management  
89 system at the Edinburgh Clinical Research Facility, University of Edinburgh. Recruitment to VIKING II and  
90 VIKING III took place online from January 2020 to March 2023<sup>15</sup>.

### 91 *Viking Genes Pedigree Information*

92 Records of the births, marriages and deaths in Orkney and Shetland are held at the General Register Office  
93 for Scotland (New Register House, Edinburgh). These records, along with relationship information  
94 obtained from study participants, Orkney and Shetland family history societies, and genealogies available  
95 online, were used to construct pedigrees of study participants using RootsMagic software (S&N Genealogy  
96 Supplies), which was then amended to reflect the genetic kinship between individuals using genotype  
97 data.

### 98 *Genotyping*

99 DNA from all participants was used for genome-wide genotyping on the GSA BeadChip (Illumina) at the  
100 Regeneron Genetics Center. Monomorphic genotypes and genotypes with more than 2% of missingness  
101 and Hardy-Weinberg equilibrium (HWE)  $p < 10^{-6}$  were removed, as well as individuals with more than 3% of  
102 missingness.

### 103 *Exome Sequencing*

104 Quality controlled exome sequence data sets were prepared at the Regeneron Genetics Center, following  
105 the process detailed for UKB<sup>2</sup>, with 94.6% of targeted bases having >20X coverage.

### 106 *Pipeline Development*

107 To find carriers of actionable variants, we first extracted all variants that map to genes in the ACMG v3.2  
108 list from the whole exome sequences (WES)<sup>1</sup>. These variants were annotated with the dbsnp version 151  
109 using the bcftools 1.9 annotate function, using CHROM, POS, ID, REF, ALT columns. The dbsnp 151 data was  
110 downloaded from [https://ftp.ncbi.nlm.nih.gov/snp/organisms/human\\_9606\\_b151\\_GRCh38p7/VCF/](https://ftp.ncbi.nlm.nih.gov/snp/organisms/human_9606_b151_GRCh38p7/VCF/) (00-  
111 All.vcf.gz file) in March 2022. To assess their pathogenicity, these variants were then merged with the  
112 ClinVar data. The variant\_summary.txt.gz file of the Clinvar data was downloaded on 18<sup>th</sup> March 2024,  
113 from [https://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab\\_delimited/](https://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/). Merging by chromosome, position  
114 (genome build GRCh38), reference (REF) and alternate (ALT) alleles was necessary, since rsids can be  
115 multiallelic - containing multiple variants at the same genomic position, including different nucleotide  
116 substitutions, for example tri-allelic SNPs (A>C, A>G) as well as indels, and therefore merging solely by  
117 rsids might result in merging different variants.

118 Some of the variants present in the exomes were not identical to mutations listed in ClinVar, but result in  
119 the same amino-acid change as variants in ClinVar and thus the same consequence, commonly a  
120 frameshift with comparable functional effects. Therefore, we also merged the variants from the WES with  
121 ClinVar data based on the change of the amino-acid that the variant causes, instead of the CHR, POS, REF,  
122 ALT. Furthermore, if the following hypothetical indel occurred, TCATCTA > TCATCTCTA, it could be  
123 named either as c.4insCT or c.5insTC or c.6insCT, causing problems for merging the data using either of  
124 the two approaches. This is illustrated by the real variant rs886044536, *TTN* c.93396\_93400del AGCTT

125 ([MIM: 188840], Results). This is a deletion: GCCAAGCTAAGACT > GCCAAGACT which can be rendered  
126 either as GCCAAG [CTAAG/\*] ACT or as GCC [AAGCT/\*] AAGACT.

127 After merging, we next filtered the output using the ClinVar annotations to retain only pathogenic or likely  
128 pathogenic variants (P/LP), that had status 1 star or higher, that is with assertion of pathogenicity.

129 Moreover, the variant class had to match the disease-causing variant classes for that gene, e.g. only loss-  
130 of-function (truncating) variants for titin (*TTN*). Variants were also subject to manual review before listing  
131 as potentially actionable.

### 132 *Whole Genome Sequencing*

133 Whole genome sequencing (WGS) data is available for 1,300 members of the ORCADES cohort<sup>16</sup> and 500  
134 members of the Viking I cohort<sup>17</sup>. These datasets were used for look-ups, as part of the verification  
135 process of the actionable variant results obtained from exome sequencing.

### 136 *Sanger Sequencing*

137 Pre-designed primer pairs for PCR and Sanger sequencing were selected from a collection that covers over  
138 95% of human coding exons (Primer Designer, ThermoFisher Scientific). If validated primers were not  
139 available, primer pairs were designed using Primer3 software (ThermoFisher Scientific) and checked using  
140 UCSC In Silico PCR<sup>18</sup>. A standard protocol for PCR and Sanger sequencing on an Applied Biosystems 3730xl  
141 Genetic Analyser was applied across all samples. Output files were analysed using Sequencher DNA  
142 sequence analysis software version 5.4.6 (Gene Codes Corporation).

143 Validation of heterozygous short insertions/deletions (indels) can be challenging with Sanger sequencing.  
144 Although still considered the gold standard, the sequencing ladders derived from, for example, the  
145 deleted and non-deleted alleles become out of synchronicity with one another (by the length of the indel  
146 in base-pairs), leading to double sequence after the indel in each direction. Illumina short read  
147 sequencing, as used to generate the exome data, does not suffer from this problem. A further issue is that  
148 certain indels with short repeats within or near them (or in low complexity DNA) can be rendered in  
149 multiple ways in cDNA nomenclature, e.g. *TTN* c.93396\_93400delAGCTT (see above). More common are  
150 single base indels altering the length of homopolymers, which can be annotated as a change at any  
151 position in, for example, the polyA tract. Chromosome positions are affected in the same way, and so

152 when interrogating ClinVar, we performed a secondary merge using the protein sequence nomenclature,  
153 to pick up such instances.

#### 154 *Clinical review of potentially actionable genotypes*

155 All potentially actionable variants and genotypes presented in Table S1 other than those noted as not  
156 determined were discussed with the NHS clinical team in order to agree suitability for return of results to  
157 participants. Local clinical knowledge in NHS Grampian was used as part of this review process.

#### 158 *Electronic Health Record (EHR) Data Linkage*

159 This work used data provided by patients and collected by the NHS as part of their care and support,  
160 reused with permission from Public Health Scotland. NHS routine datasets linked to ORCADES and VIKING  
161 I participants in July 2021, including SMR01, the general/acute hospital inpatient and day case dataset of  
162 episode level data, and SMR00, outpatient appointments and attendances, were accessed using a secure  
163 process, as described previously<sup>11</sup>.

164

## 165 **Results**

### 166 *Exome sequencing and validation*

167 Exome sequences were generated from 4,198 participants from the Viking Genes cohort using industry-  
168 standard protocols as described in the Methods section. The exome sequence calls are of high quality, but  
169 it is important to verify them in the research laboratory, before considering the process of notifying a  
170 potential carrier. We therefore attempted to validate 149 potential actionable variant genotypes: 77 by  
171 Sanger sequencing, 48 by WGS and 24 by both sequence analysis methods. The total of 149 includes a  
172 number of heterozygous carriers of recessive variants in genes on the ACMGv3.2 list. 147 out of the 149  
173 assessed were validated (98.7%). Notably, both of the variants that failed (one by Sanger sequencing  
174 alone, and one by both Sanger sequencing and WGS) were only 2 kb apart in the same gene, *MSH2* [MIM:  
175 609309]. Manual review of the exome read stacks suggests that these variants are both false positives,  
176 with low read numbers and increased noise: they were not included in the total counts of actionable  
177 variants found.

178 One variant, in *DES* [MIM: 125660], c.973C>T, was found to be most likely a somatic mosaic in the sample  
179 of venous blood and so was not included in our counts and not returned. In Sanger sequencing from  
180 forward and reverse strands, only a very weak T was observed, and the read counts in WGS were 19 C and  
181 1T. In 85X exome data, the total read counts (sum of forward and reverse) are 72C, 13T, for a variant allele  
182 distribution of 15%.

### 183 *Pipeline for identification of potential actionable variants*

184 Exome sequence data from 2,090 ORCADES participants (820 male and 1,270 female) and 2,108 VIKING I  
185 participants (843 male and 1,265 female) passed all sequence and genotype quality control thresholds  
186 (Methods). The data were run through our pipeline (Methods), a key component of which is the ClinVar  
187 resource ([www.ncbi.nlm.nih.gov/clinvar/](http://www.ncbi.nlm.nih.gov/clinvar/)). ClinVar aggregates information about genomic variation and its  
188 relationship to human health and allocates a clinical significance category to each variant<sup>14</sup>.

189 In total, 59 potential actionable variants were observed, consisting of 27 singletons (Supplemental Table  
190 S1) and 32 non-singleton variants (Tables 1 and 2). Half of the singleton variants (13/27) were observed in  
191 participants with one non-islander parent, and a further 4/27 were in volunteers with one non-islander  
192 grandparent, for a total of 63% of singletons with at least one quarter non-Northern Isles DNA. The  
193 remaining 10 singletons presumably arose by mutation in the last few generations and are thus very rare,  
194 or were brought into the islands by non-parental events. The variants were found in 26 different genes.  
195 For three of these genes, *BTD* [MIM: 609019], *CASQ2* [MIM: 114251] and *RPE65* [MIM: 180069], each  
196 carrier only had a single copy of a variant that acts recessively, thus not an actionable genotype. In this  
197 paper we are careful to distinguish between actionable variants (from the ACMG list) and actionable  
198 genotypes, the combinations of such variants which are actionable, for example compound heterozygotes  
199 or homozygotes for autosomal recessive conditions. Altogether, using the 81-gene ACMG 3.2 list, we  
200 observed actionable genotypes for 104 individuals (2.5% of the population) carrying 108 actionable  
201 genotypes at 39 of the variants, across 23 genes. No X-linked actionable variants were found.

202 The level of support for each review status in ClinVar decreases from four star downwards  
203 ([https://www.ncbi.nlm.nih.gov/clinvar/docs/review\\_status](https://www.ncbi.nlm.nih.gov/clinvar/docs/review_status)). The great majority of the actionable variants



204 we found were 3 (reviewed by expert panel) or 2 (criteria provided, multiple submitters, no conflicts) star  
205 (Table S1). Three of the non-singleton autosomal dominant actionable variants (Table 1) had a review  
206 status in ClinVar of 1 star, all of which had a single submitter. The paucity of submissions often indicates  
207 rarity of the variant, which in turn makes assessment of pathogenicity and hence actionability more  
208 challenging.

#### 209 *Actionable genotype co-occurrences*

210 We identified four individuals with actionable genotypes in two separate genes (Table 3). We expect some  
211 such co-occurrences because of independent assortment, and we should see these according to the  
212 product of the allele frequencies of the variants in question. If one or both have drifted upwards in  
213 frequency, more will be seen. We found that two of these four Viking Genes individuals had AVs in two  
214 different cancer susceptibility genes, while two had actionable genotypes in a cancer susceptibility gene  
215 and a metabolic gene.

#### 216 *Drifted actionable variants in Northern Isles populations*

217 The isolated Northern Isles populations carry different sets of rare variants compared to cosmopolitan  
218 populations. In total, 10 actionable variants across 7 genes (*BRCA1* [MIM: 113705], *BRCA2* [MIM:600185],  
219 *ATP7B* [MIM: 606882], *TTN* [MIM: 188840], *KCNH2* [MIM: 152427], *MUTYH* [MIM: 604933], *GAA* [MIM:  
220 606800]) have risen in frequency by at least 50-fold through the action of genetic drift in the Northern  
221 Isles (Table 4). These include results published previously on a *BRCA1* breast and ovarian cancer  
222 predisposition variant that is 470-fold more common in Orkney than in the UK<sup>11</sup>, a *KCNH2* Long QT  
223 Syndrome variant that is ~90-fold more common in Shetland<sup>10</sup>, and a pathogenic *BRCA2* variant ~155-fold  
224 more common in Shetland<sup>12</sup>.

225 Most actionable variants in the ACMG list have a dominant mode of inheritance and so are actionable if  
226 heterozygous, but some are only considered actionable if present in two copies. Table 4 includes five  
227 drifted variants in Northern Isles populations that act recessively (also listed in Table 2). These are in  
228 *ATP7B* (three drifted variants segregating in our cohorts), *MUTYH* and *GAA*. Wilson Disease [MIM:  
229 277900] is a disorder of copper metabolism, caused by variants in the copper transporting ATPase gene  
230 *ATP7B*. The marked discrepancy between the “genetic prevalence” in populations and the number of

231 clinically diagnosed cases of Wilson Disease is probably due both to reduced penetrance of many *ATP7B*  
232 mutations, and failure to diagnose patients with this highly treatable disorder<sup>19</sup>. We identified three  
233 founder effects on recessive *ATP7B* Wilson disease variants, two in Shetland, which gives rise to  
234 substantial risk of homozygosity or compound heterozygosity. The combined allele frequency in  
235 Shetlanders is 0.8%. Together with five other alleles segregating at lower frequencies, the sum  
236 Pathogenic/Likely Pathogenic (P/LP) allele frequency is 1.12%; thus about 1/45 Shetlanders are carriers of  
237 P/LP *ATP7B* variants. Under random mating, this predicts about 1/8000 will be homozygotes or compound  
238 heterozygotes, at risk of being affected. Indeed, one homozygote for the drifted variant c.3007G>A was  
239 found.

240 There is a founder effect on a recessive *MUTYH* variant (c.1130C>T) in Orkney (0.86% carriers, 70x more  
241 common than UKB), causing *MYH*-associated polyposis [MIM: 608456] (Table 4). While not increased in  
242 frequency compared to other populations, there are two further *MUTYH* variants segregating in the  
243 Orcadian population, bringing the combined carrier frequency to 2.9%. This suggests that 1/35 Orcadians  
244 are carriers and, assuming random mating among Orcadians, an expectation that ~1/5000 will be at risk of  
245 *MYH*-associated polyposis. Consistent with this, we did observe a compound heterozygote in the Orkney  
246 data.

247 Finally, there are also founder effects in Shetland on two dominant loss-of-function titin (*TTN*) variants,  
248 increasing the risk of dilated cardiomyopathy [MIM: 604145] (Table 4). One of these, c.93396\_93400del, is  
249 >3,000-fold more common in Shetland than in UKB, where there is only a single instance in 917,682  
250 alleles. Together with a third returnable *TTN* variant that does not meet our thresholds to be defined as  
251 drifted, the combined allele frequency is 0.5%, meaning 1/100 Shetlanders are carriers. In total, we  
252 observed 24 carriers of *TTN* loss-of-function (LoF) variants (23 of which are accounted for by these three  
253 alleles).

#### 254 *An ultra-rare drifted titin variant*

255 The penetrance of the *TTN* alleles included in ClinVar as pathogenic or likely pathogenic is variable and age  
256 related. However, the deletion of 5 nucleotides c.93396\_93400del causes a frameshift, which creates a  
257 premature translational stop signal (p.Trp31134\*), predicted to result in an absent or disrupted protein

258 product. This variant is in the A-band of the *TTN* gene. Truncating variants in the A-band of *TTN* are  
259 significantly overrepresented in patients with dilated cardiomyopathy (DCM), and are considered to be  
260 likely pathogenic for the disease<sup>20</sup>. *TTN* truncating variants (TTNtv) located in exons highly expressed in  
261 the heart (proportion spliced in [PSI]>0.9, or hiPSI), such as this one, increase the odds of DCM by 11-19-  
262 fold<sup>21</sup>. This variant has therefore been classified as Likely Pathogenic in ClinVar. We examined data from  
263 347 individuals ascertained with DCM and found to carry hiPSI TTNtv in DCM cohort studies from Europe,  
264 USA and Australia, and also interrogated data from TTNtv carriers in the UKB<sup>22</sup>, the Geisinger MyCode  
265 Biobank and the PennMedicine Biobank (168 affected with DCM, 2,409 unaffected)<sup>21</sup>. In total we observed  
266 1,200 distinct hiPSI TTNtv in 2,094 individuals. The c.93396\_93400del variant was observed twice (1 in  
267 UKB and 1 in MyCode), both times in individuals without a diagnosis of DCM. The variant was not  
268 observed in the 100,000 Genomes Project<sup>23</sup> (670 DCM cases, 62,330 non-DCM participants) or All of Us<sup>24</sup>  
269 (245,460 total participants, of whom 125,860 were reported as European heritage), and therefore can be  
270 considered ultra-rare.

271 Analysis of linked data from the EHR for cardiomyopathy (ICD-10 code I42), atrial fibrillation (ICD-10 I48)  
272 or heart failure (ICD-10 I50) shows that the 16 carriers of the drifted *TTN* variant c.93396\_93400del we  
273 identified in VIKING I are nearly four times more likely ( $P=0.02$ , Fisher's exact test) to have an inpatient or  
274 day patient record for cardiomyopathy (1/16), atrial fibrillation (2/16) or heart failure (2/16) than non-  
275 carriers (4/16 carriers have entries for at least one of these conditions, i.e. a nominal combined  
276 penetrance of 25%, versus a prevalence of 6.8% in 4,133 individuals without *TTN* LoF variants). Analysis of  
277 ECG data collected in the VIKING 1 recruitment clinic from 15 of the carriers also shows an over-  
278 representation of features such as first degree atrioventricular block (3/15) or right bundle branch block  
279 (4/15), affecting a total of 7 carriers. In one Shetland *TTN* c.93396\_93400del family investigated clinically,  
280 the condition has typically been detectable on echocardiography around age 30, but symptoms have not  
281 developed until later. One patient required a cardiac resynchronisation therapy-defibrillator, while  
282 another is asymptomatic despite an abnormal echocardiogram in their late 50s. Together with the 7  
283 entries in ClinVar submitted to date, we now report this further direct evidence for pathogenicity of this  
284 ultra-rare drifted *TTN* variant.

285 *Impact of sub-population gene pools*

286 The drifted variants are not shared between the two archipelagos, except *TTN* p.Gln22507Ter, which is  
287 observed in Shetland among individuals with recent Orcadian ancestry; and *MUTYH* p.Pro377Leu, for  
288 which the Shetland carriers have no recent Orcadian ancestry. This latter variant may have been  
289 introduced to the Northern Isles more than once, or there may be unrecorded non-parental events  
290 connecting the Orcadian and Shetlandic kindreds.

291 The great majority of these actionable variants can be traced back to individual isles within Orkney or  
292 Shetland, or in some cases individual parishes: 13-15 ancient geographic, legal, and religious subdivisions  
293 of the Mainlands of Orkney and Shetland (Figure 1). Among published variants, 95% of carriers of the  
294 drifted *BRCA1* variant traced their genealogies to Westray, in the North Isles of Orkney<sup>11</sup>, while 78% of the  
295 drifted *BRCA2* variant carriers trace back to Whalsay, an isle in Shetland<sup>12</sup>. Likewise, drifted *KCNH2* variant  
296 carriers from Viking Genes and NHS Clinical Genetics<sup>10</sup> trace back either to the parish of Aithsting in the  
297 West Mainland of Shetland or Unst in the North Isles of Shetland (Figure 1 and Table 4).

298 Among the drifted variants reported here, 80% of the carriers of the most drifted *ATP7B* variant have  
299 genealogies going back to Burra Isle, Shetland and there is also a strong founder effect on the most highly  
300 drifted *TTN* variant c.93396\_93400del, this time in Yell, in the North Isles of Shetland, with 100% of  
301 carriers tracing their family trees back to there. Further fine scale, within-archipelago population  
302 structuring of clinically important variants is observed in Westray, in the North Isles of Orkney, with  
303 another *ATP7B* variant having drifted to high frequencies there, and 100% of carriers tracing pedigrees  
304 there, similarly to a *MUTYH* variant, where 60% of carriers have Westray genealogies; indeed, we  
305 observed a compound heterozygote with ancestry from that isle.

306 *Genetic down-drift*

307 It is more difficult to quantify which variants have drifted down in frequency, but the near absence of  
308 familial hypercholesterolaemia (FH) [MIM: 143890 and MIM:144010] variants is notable, with only one  
309 carrier across >4,000 subjects. This *LDLR* [MIM: 606945] variant was observed in Shetland and is the  
310 known p.C163Y West of Scotland founder variant<sup>25</sup>; it may have been brought to the isles by recent gene  
311 flow from Mainland Scotland; no FH variants were observed in Orkney. No P/LP *APOB* [MIM:107730] or

312 *PCSK9* [MIM: 607786] variants were observed. Significantly fewer volunteers from the Northern Isles  
313 carried FH variants than observed in European ancestry participants from the UKB ( $P < 0.0001$ , Fisher's  
314 exact test). The carrier frequency in the Northern Isles (1/4198) is more than 10-fold lower than that in  
315 UKB European ancestry subjects (1/288)<sup>26</sup>.

### 316 *The Return of Results (RoR) process*

317 With reference to the Association of Clinical Genomic Science (ACGS) guidelines

318 (<https://www.acgs.uk.com/media/12533/uk-practice-guidelines-for-variant-classification-v12-2024.pdf>)

319 and local clinical knowledge, four variants in a total of 10 genotype carriers were not deemed returnable  
320 by the NHS clinical team (Table 5 and Table S1). Our next goal was to implement a process of return of  
321 results (RoR) to carrier participants, summarised in Figure 2. Key principles include: the use of informed  
322 consent; respecting the right not to know; checking the research results as far as possible before sharing a  
323 potentially actionable variant ID; decisions on whether to return a variant result taken by NHS clinical  
324 experts; counselling provided through the NHS, and verification of all the research results in an accredited  
325 clinical laboratory. The consent form and RoR participant information sheets sent to cohort members are  
326 provided as supplemental information. Frequently asked questions about return of results are available  
327 for participants on the study website at [https://viking.ed.ac.uk/for-viking-genes-volunteers/faqs/return-](https://viking.ed.ac.uk/for-viking-genes-volunteers/faqs/return-of-results)  
328 [of-results](https://viking.ed.ac.uk/for-viking-genes-volunteers/faqs/return-of-results)

329 We obtained a favourable opinion from the Research Ethics Committee to return results of genotypes that  
330 have the potential to improve the health of the individuals carrying them. We subsequently obtained  
331 further permission to contact male carriers of pathogenic *BRCA1* variants. This was justified by the fact  
332 that although there is no clinical action to be taken with respect to the letter recipient, the information  
333 could be of value to their female relatives, who may not be Viking Genes research participants. There are  
334 13 male carriers of the *BRCA1* p.V1736A variant in the ORCADES cohort<sup>11</sup> and none in VIKING I, so addition  
335 of these participants increases the total number of actionable genotypes from 108 to 121. Since there are  
336 four individuals with two actionable genotypes each (Table 3), this corresponds to a total of 117 total  
337 actionable genotype carriers (Table 5).

338 We further obtained ethical permission to offer actionable results to the next of kin of deceased  
339 participants, if we found any that might have been relevant to the health of the participant who died, and  
340 therefore potentially their family, during our research (Figure 2). To date, six next of kin have provided  
341 their consent to receive this information, with one actionable genotype disclosed.

342 The process we have implemented is in good accordance with a checklist of components for return of  
343 results developed by an interdisciplinary panel of European experts<sup>27</sup>. In Viking Genes, the return of  
344 medically actionable genetic results to participants is optional, i.e. an opt-in approach. This has legal,  
345 philosophical, and ethical foundations, to allow patient autonomy and the right not to know. This is  
346 achieved at three separate points in the process of implementation (Figure 2). Letters containing results  
347 are only sent with the consent of the recipient. Upon receipt, the participant can decline any further  
348 investigation by not contacting the NHS about the research finding. Finally, after genetic counselling, the  
349 patient can decide not to have the research result confirmed in a clinical setting. It should also be noted  
350 that a participant can get in touch with Viking Genes at any time to request to change the consent they  
351 gave about RoR, in either direction, or to withdraw from the study. The research team communicates with  
352 Viking Genes volunteers by regular newsletters and social media, and invites them to complete occasional  
353 research surveys. A subset of participants in the study forms the public - participant involvement (PPI)  
354 group. We therefore aim to ensure both that understanding of the opportunity to receive results is high,  
355 and the processes we are implementing have the support of the study participants and the populations  
356 they are drawn from.

#### 357 *Initial outcomes from issuing RoR letters*

358 The text of the letters alerting the volunteer to the research result is written in lay language, agreed by  
359 the research and NHS teams, and the template document was given a favourable opinion by the Research  
360 Ethics Committee. The NHS clinic team is not made aware of the identity of the research participants, but  
361 approves return of the variants, which are also noted in the letters. Letters were not sent to all carrier  
362 participants at one time. Instead, the 64 RoR letters issued (Table 5) were posted over the course of  
363 almost exactly one year. This was designed to prevent the NHS clinics from being overwhelmed, but has  
364 meant that outcomes are still being accumulated. For example, it is too early to tally the number of family

365 members contacted by cascade testing from each index case. In the NHS, each index case on average  
366 leads to three subsequent cascade tests, half of which are positive<sup>28</sup>. In the case of the Shetland *BRCA2*  
367 founder variant c.517-2A>G, a total of 56 people were eligible for cascade letters from just 5 consented  
368 carriers in Viking Genes, suggesting a higher level of knowledge of family structure among our  
369 participants. We observe that RoR letter recipients take a variable length of time to decide whether to  
370 contact the NHS, and not all choose to do so.

371 In all three of the previously reported actionable founder variants<sup>10-12</sup>, we identified carriers in our  
372 research populations who could not have been ascertained from oral history-based cascade testing.  
373 Likewise, many of the index cases described here were found to be in family groups not previously  
374 recorded by the clinical geneticists. Out-patient NHS EHR data showed that only 9% of our carriers had  
375 interacted with clinical genetics, compared with 3.6% of the non-carriers, a 2.5-fold increase ( $P < 0.002$ ,  $\chi^2$ ).  
376 This provides further evidence that a substantial number of the research participants who received a  
377 return of results letter were previously unaware of the genomic medicine finding, and have the  
378 opportunity for preventative medical intervention as a result of the Viking Genes return of results  
379 programme.

380

## 381 **Discussion**

### 382 *Survey of Actionable Variants*

383 We observed actionable genotypes for 39 autosomal actionable variants, 34 dominant and 5 recessive,  
384 and no X-linked variants. Ten founder variants showed strong enrichment, having risen >50-fold in  
385 frequency in Orkney or Shetland, while otherwise commonly seen variants, notably those causing familial  
386 hypercholesterolaemia, were at least ten-fold lower in frequency. Importantly, both *TTN* and *ATP7B*  
387 contain multiple drifted variants, which sum together to even higher frequencies of (likely) pathogenic  
388 variation.

389 The ten actionable variants which have drifted up to much higher frequencies in Orkney or Shetland than  
390 in the general UK population mostly have a clear origin in a particular parish or isle within the Northern  
391 Isles, 200 or more years ago, highlighting the continuing clinical importance of these historic sub-

392 population gene pools. As is the case with other genetic isolate populations<sup>29</sup>, the combination of multiple  
393 founder variants can result in a significant summed population frequency, sometimes in conjunction with  
394 non-drifted variants. At the same time, other variants have drifted towards loss, balancing the genetic  
395 burden of disease risk.

#### 396 *Comparisons with other populations surveyed for actionable variants*

397 Overall, these two UK isolate populations have a similar percentage of actionable genotypes (2.5%),  
398 compared to cosmopolitan populations worldwide. However, the number carrying one or more actionable  
399 genotypes depends on the gene list, rules applied (including bespoke curation) and date of analysis. The  
400 ACMG list of recommendations of genes and phenotypes is updated annually, whereas ClinVar is designed  
401 to enable the ongoing evolution and development of knowledge regarding variations and associated  
402 phenotypes, and its website is updated weekly. The status of a variant, particularly if observed rarely  
403 worldwide, may therefore change over time as more data are submitted, interpreted and reported in  
404 ClinVar. This makes it challenging to compare our Viking Genes actionable variant carrier percentage with  
405 populations reported in other publications. In the first 50,000 exome sequences from the UKB population  
406 cohort, an actionable genotype figure of 2.0% was reported<sup>2</sup>. However, that analysis used v2.0 of the  
407 ACMG list<sup>30</sup>, containing 59 genes, rather than the 81 in v3.2 that were surveyed here. Similarly, a figure of  
408 2.8% can be calculated from the DiscovEHR study<sup>31</sup> and 2.3% in Iceland<sup>32</sup>, using the ACMG list v2.0.  
409 Another way to compare with studies in other populations is to use the ACMG 2.0 (59 gene) list to assess  
410 our data in the Northern Isles population, in which case the percentage with actionable genotypes drops  
411 to 1.4%.

412 The inclusion of *HFE* [MIM: 613609] p.Cys282Tyr homozygotes, in later versions of the list, would on its  
413 own increase the percentage of actionable genotype carriers in the UKB from 2.0% to 2.6%. This is based  
414 on the *HFE* variant having a reported homozygous frequency of 0.6%, or 1 in 156, in the UKB<sup>33</sup>. We  
415 therefore show here that the Northern Isles populations have an AV percentage that is in a similar range  
416 to the rest of the population of the UK as sampled in the UKB, to the DiscovEHR study of people in  
417 Pennsylvania, USA, and to the Icelandic population (4.0% in the most recent analysis, including manual  
418 curation)<sup>32</sup>, but much lower than that reported for the Old Order Amish of Pennsylvania<sup>34</sup>.



419 We identified a sufficient number of carriers to provide specific evidence for pathogenicity of one drifted  
420 variant in *TTN*, c.93396\_93400del, p.Trp31134\*. A total of 44 genes were recently asserted to be  
421 implicated in non-syndromic Dilated Cardiomyopathy (DCM) by the ClinGen DCM Gene Curation Expert  
422 Panel<sup>35</sup>. *TTN* is one of 11 genes with definite evidence, and featured in a gene-based analysis of  
423 penetrance and clinical phenotype in 18,665 UKB participants<sup>36</sup>. A sub-analysis for genes including *TTN*  
424 revealed DCM or early DCM features in 45.4% of cases<sup>36</sup>. Moreover, significant excess mortality was  
425 observed among carriers of *TTN* truncating variants in Dutch founder pedigrees, driven by subjects  $\geq 60$   
426 years<sup>37</sup>.

#### 427 *Return of Results to Research Volunteers - Benefits and Challenges*

428 The return of results processes that we successfully implemented for Viking Genes participants as  
429 described here were somewhat more straightforward than they would be in cosmopolitan populations,  
430 due to the reduced number of different actionable variants in the ancestral populations of the Northern  
431 Isles of Scotland, and the high support for and trust in scientific research in these communities.  
432 Nonetheless, the processes and experiences of returning actionable genetic results in Viking Genes are  
433 directly relevant as an exemplar (especially in the UK) for other research cohorts interested in the  
434 feasibility and logistics of the return of actionable results to their participants. Our recent recruitment to a  
435 new cohort study, VIKING II, was eligible to people of Northern Isles ancestry regardless of domicile<sup>15</sup>. At  
436 the outset of recruitment, VIKING II offered the option of consent to return of selected clinically  
437 actionable results. This option was chosen by an overwhelming majority (98%), 6,054 of the 6,178  
438 participants who consented and completed the study questionnaire. This is in good agreement with  
439 several other studies that also show strong support for receiving actionable findings, both in theory and in  
440 practice<sup>38,39</sup>. These views may mean that the return of actionable genetic findings from research will in  
441 future become the default practice<sup>39</sup>. However, the costs, harms and benefits of returning actionable  
442 results are still at an early stage of being fully characterized.

443 For some diseases in the ACMG list, there can often be a long delay before accurate diagnosis, due to the  
444 overlap of symptoms with other more prevalent conditions. Examples we have found in Viking Genes  
445 include Pompe disease [MIM: 232300], caused by a deficiency of the enzyme acid alpha-glucosidase

446 (GAA), which can lead to nerve and muscle problems, Wilson disease [MIM: 277900], caused by copper  
447 overload, and the iron overload condition hereditary haemochromatosis [MIM: 235200] (*HFE*). Early  
448 diagnosis and thus early intervention can have an impact on future quality of life. Furthermore, in addition  
449 to the medical intervention available for each actionable result, a clearer diagnosis/etiology following the  
450 return of results can also be of value, for those already affected, particularly for conditions like  
451 cardiomyopathy which have complex symptoms and presentation. Increased awareness of these  
452 otherwise rare conditions among healthcare workers in the Northern Isles could lead to decreased times  
453 to diagnosis, with concomitant benefits to patients.

454 Cross-sectional approaches, for example to understand the penetrance of rare-variants in  
455 cardiomyopathy-associated genes, can provide valuable data<sup>40</sup>, and genome-first evaluation at scale can  
456 enable new diagnoses missing from clinical healthcare<sup>41</sup>. However, the scientific literature is only just  
457 beginning to provide insight into how the ACMG guidelines have been translated into precision health  
458 outcomes for the recipients of the findings: for some genes benefits clearly outweigh harms, while for  
459 others there is little evidence as yet<sup>42-44</sup>. This is likely to gather pace in the near future, as more large scale  
460 initiatives such as All of Us<sup>3</sup>, the Healthy Oregon Project<sup>45</sup> and the Colorado Center for Personalized  
461 Medicine<sup>46</sup> implement the return of actionable results. Furthermore, disease manifestation, healthcare  
462 outcomes and costs of disclosure have recently been described for actionable findings in a UK setting from  
463 the Genomics England 100,000 Genomes Project<sup>47</sup>. Genomics England returned results for actionable  
464 genotypes in 13 genes within the 81 genes on the ACMG3.2 list, for hereditary cancer syndromes and  
465 familial hypercholesterolaemia, plus carrier status for cystic fibrosis ([MIM: 219700], not on the ACMG  
466 list). Due to a lack of consensus on benefit versus harm and the challenges of reduced penetrance,  
467 Genomics England chose not to inform carriers of actionable genotypes increasing the risk of Long QT  
468 syndrome, cardiomyopathies, aortopathies, further inherited cancers, metabolic diseases and the most  
469 frequent genetic disease, hereditary haemochromatosis ([www.genomicsengland.co.uk/initiatives/100000-](http://www.genomicsengland.co.uk/initiatives/100000-genomes-project/additional-findings)  
470 [genomes-project/additional-findings](http://www.genomicsengland.co.uk/initiatives/100000-genomes-project/additional-findings)).

471 Jensson *et al*<sup>32</sup> found shorter median survival among persons carrying actionable genotypes than among  
472 non-carriers, in Iceland. Specifically, they report that carrying an actionable genotype in a cancer gene on

473 the ACMG list was associated with life span that was three years shorter than that among noncarriers,  
474 with causes of death among carriers attributed primarily to cancer-related conditions. Furthermore, while  
475 many studies have demonstrated the diagnostic and therapeutic value of exome or whole genome  
476 sequencing in critically ill paediatric cases, a recent study shows that the diagnostic utility of exome  
477 sequencing in critically ill adults is similar, largely due to uncovering variants in medically actionable  
478 genes<sup>48</sup>.

479 Our data highlight the clinical relevance of local genetic sub-populations, as most of the founder variants  
480 are also associated with one or other parish or isle, with deep genealogies linking most carriers to  
481 ancestors from there (notwithstanding ancestral non-parental events which obscure the connection  
482 between genetics and geography). Within each archipelago, it is natural that historic barriers to marriage  
483 in the form of storm-prone, strongly tidal ocean and sea channels and sounds, have given rise to  
484 subpopulation gene pools or micro-isolates<sup>7,49</sup>. Within the main islands, the parishes are sometimes  
485 separated by low hills or sea firths, and sometimes not, but marriage records show they were  
486 demographically important over the centuries.

#### 487 *Population Screening Opportunities*

488 Population screening in a clinical context is currently available in some populations that show founder  
489 effects, such as an NHS programme offering *BRCA*-testing to those with Jewish ancestry  
490 (<https://www.nhsjewishbrcaprogramme.org.uk/>). Charitable organisations are also increasingly making  
491 specific genetic tests available to populations of higher risk. Examples include improving the prevention  
492 and diagnosis of Jewish genetic disorders in the UK (<https://www.jnetics.org/>) and the provision of postal  
493 tests for genetic haemochromatosis (<https://www.haemochromatosis.org.uk/>).

494 Our estimates of the frequency of the *BRCA1* and *BRCA2* actionable variants in Westray and Whalsay,  
495 respectively, (1/19 and 1/43)<sup>12</sup> show that the local gene pool frequencies can be much higher than the  
496 archipelago-wide frequencies. This in turn suggests that individuals with ancestry from islands such as Yell,  
497 Burra and Westray are at higher risk of being carriers than the Shetland- or Orkney-wide data suggest. The  
498 ability to target testing to those with an origin in a particular parish may improve the cost-effectiveness of  
499 future screening. At the same time, with increasing movement, younger generations and inhabitants of

500 the towns of Kirkwall and Lerwick (and of course the diasporas outside the Northern Isles) tend to have  
501 more mixed ancestry, so the importance of these micro-isolates is waning.

502 We report four actionable genotype co-occurrences. Many more instances in other populations are likely  
503 to be uncovered in the future, as genetic analyses move from limited gene panels towards WES and WGS.

504 A recent review, for example, describes Multi-locus Inherited Neoplasia Allele Syndrome (MINAS), which  
505 refers to individuals with germline pathogenic variants in two or more cancer susceptibility genes<sup>50</sup>.

506 We have not investigated LoF variants that may be present in our data, but which are not reported in  
507 ClinVar. There may be potentially important ultra-rare variants in this class, which warrants future  
508 research. Furthermore, clinically relevant variants that lie outside of exomes, and structural variants, may  
509 be uncovered in subsequent research. However, except for the most drifted variants, the assessment of  
510 pathogenicity will always be limited by the modest numbers of carriers likely to be found in our cohorts.

511 We also note that the high levels of engagement and knowledge of kinship led to higher uptake of cascade  
512 testing, and thus a higher-than-expected demand on clinical services, compared to previous cascade  
513 testing statistics reported in the NHS<sup>28</sup>.

514 Studies of population genetic structure using unselected genome-wide markers<sup>7,51</sup> are not always  
515 considered relevant for clinical genetics. However, a number of reports now demonstrate the  
516 consequences of such population structure for clinical genetics at the population level. A case in point is  
517 the Irish Travellers, shown to be a distinct population isolate of Irish origin<sup>52</sup> with their own suite of  
518 Mendelian conditions at high frequencies<sup>53</sup>. Similarly, the genetic isolation of Orkney and Shetland has  
519 been shown to translate into clinically important differences at the population level, from the earliest  
520 studies of Northern Isles genetics, showing a founder effect on the *SOD1* [MIM: 147450] p.D90A allele<sup>54</sup>,  
521 to novel Mendelian conditions<sup>55-57</sup>, and distinct blood group/isozyme<sup>58</sup> and uniparental<sup>6</sup> markers. More  
522 recently, numerous founder effects, both on Mendelian recessive alleles<sup>8</sup>, and on dominant actionable  
523 variants<sup>10-12</sup> have been described.

524 Orcadians and Shetlanders thus belong to a group of European-heritage populations, including Ashkenazi  
525 Jews and Irish Travellers, where high frequencies of important clinical variants have arisen from genetic

526 drift, with concomitant simplification of the gene pools providing valuable opportunities for population-  
527 wide, cost-effective targeted genetic screening.

528

#### 529 **Data and Code Availability Statement**

530 There is neither Research Ethics Committee approval, nor consent from Viking Genes participants, to  
531 permit open release of the individual level research data underlying this study. The datasets generated  
532 and analysed during the current study are therefore not publicly available. Instead, the research data  
533 and/or DNA samples are available from [accessQTL@ed.ac.uk](mailto:accessQTL@ed.ac.uk) on reasonable request, following approval by  
534 the Data Access Committee and in line with the consent given by participants. The ClinVar-exome pipeline  
535 code is available on GitHub ([https://github.com/viking-genes/clinvar\\_pipeline](https://github.com/viking-genes/clinvar_pipeline)). Frequencies for the ten  
536 drifted variants reported here were derived from the UK Biobank Whole Genome Sequencing (WGS)  
537 project and were obtained from the publicly available UK Biobank Allele Frequency Browser  
538 (<https://afb.ukbiobank.ac.uk/>), which was generated by the WGS consortium under the UK Biobank  
539 Resource (project ID 52293).

#### 540 **Acknowledgments**

541 This work was funded by the Medical Research Council University Unit award to the MRC Human Genetics  
542 Unit, University of Edinburgh, MC\_UU\_00007/10 and a Wellcome Trust Institutional Translational  
543 Partnership Award (University of Edinburgh 222060/Z/20/Z -PIII031). LK was supported by an RCUK  
544 Innovation Fellowship from the National Productivity Investment Fund (MR/R026408/1). ORCADES was  
545 supported by the Chief Scientist Office of the Scottish Government (CZB/4/276 and CZB/4/710), a Royal  
546 Society URF to JFW, and Arthritis Research UK. JSW was supported by the Medical Research Council (UK),  
547 Sir Jules Thorn Charitable Trust (21JTA), British Heart Foundation (RE/18/4/34215), and the NIHR Imperial  
548 College Biomedical Research Centre. This research has been conducted using the UK Biobank Resource  
549 under Application Numbers 47602 and 19655.  
550 Emily Weiss and Reka Nagy assembled the Orkney pedigree, and Barbara Gray the Shetland pedigree,  
551 using records at the General Register Office, the Shetland Family History Society, and study information,  
552 building on earlier pedigree work by Ruth McQuillan and Jim Wilson. Sanger sequencing was performed

553 by the technical services team at the MRC Human Genetics Unit. We thank the NHS Grampian clinical  
554 genetics team and the Viking Genes research team (in particular David Buchanan, Rachel Edwards and  
555 Craig Sinclair) for their contributions to the implementation of the return of results process. Finally, we  
556 thank the people of the Northern and Western Isles for their involvement in and ongoing support for our  
557 research.

#### 558 **Author Contributions**

559 SMK managed the project and drafted the manuscript. LK created and implemented the ClinVar-exome  
560 pipeline, with input and support from KJ, and analysed the exome datasets. MDM updated the pipeline.  
561 CD validated exome data by PCR and Sanger sequencing of genomic DNA. MH validated exome data by  
562 analysis of WGS data. SMK, ZM and JFW planned the RoR methodology. ZM and JD provided information  
563 on the clinical epidemiology of the Northern Isles and clinically reviewed the variants to be reported. JSW,  
564 SLZ and PKT interpreted phenotype data in cardiomyopathy variant carriers. GT and ARS conceived and  
565 managed the Viking Genes exome sequencing. ZM led the NHS RoR clinical team, which included genetic  
566 counsellors EC and LS, who discussed actionable results with participants. JFW is the Chief Investigator of  
567 Viking Genes, was awarded funding to implement the work, analysed the data and helped draft the  
568 manuscript. All authors provided input and feedback on drafts of the manuscript.

#### 569 **Ethical Approval**

570 Eligible participants were recruited to Viking Genes, NHS Lothian South East Scotland Research Ethics  
571 Committee reference 19/SS/0104. Research participants gave informed consent for research procedures  
572 including DNA sequencing (mandatory) and return of actionable results (optional).

573

574 For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to  
575 any Author Accepted Manuscript version arising from this submission.

576

#### 577 **Supplemental Files**

578 Supplemental Table S1 All actionable variants reported by our ClinVar pipeline.

579 Supplemental Information – RoR Participant Information Sheet (PIS), RoR Consent Form and RoR PIS for  
580 Next of Kin.

581

## 582 Declaration of Interests

583 AS and GT are employees and/or stockholders of Regeneron Genetics Center or Regeneron  
584 Pharmaceuticals. LK is an employee of BioAge Labs and holds share options. JSW has received research  
585 support from Bristol Myers Squibb, and has acted as a consultant for MyoKardia, Pfizer, Foresite Labs,  
586 Health Lumen, and Tenaya Therapeutics.

587

## 588 References

- 589 1. Miller, D.T., Lee, K., Abul-Husn, N.S., Amendola, L.M., Brothers, K., Chung, W.K., Gollob, M.H.,  
590 Gordon, A.S., Harrison, S.M., Hershberger, R.E., et al. (2023). ACMG SF v3.2 list for reporting of  
591 secondary findings in clinical exome and genome sequencing: A policy statement of the American  
592 College of Medical Genetics and Genomics (ACMG). *Genet Med* 25, 100866.  
593 [10.1016/j.gim.2023.100866](https://doi.org/10.1016/j.gim.2023.100866).
- 594 2. Van Hout, C.V., Tachmazidou, I., Backman, J.D., Hoffman, J.D., Liu, D., Pandey, A.K., Gonzaga-  
595 Jauregui, C., Khalid, S., Ye, B., Banerjee, N., et al. (2020). Exome sequencing and characterization  
596 of 49,960 individuals in the UK Biobank. *Nature* 586, 749-756. [10.1038/s41586-020-2853-0](https://doi.org/10.1038/s41586-020-2853-0).
- 597 3. Venner, E., Muzny, D., Smith, J.D., Walker, K., Neben, C.L., Lockwood, C.M., Empey, P.E., Metcalf,  
598 G.A., Kachulis, C., All of Us Research Program Regulatory Working, G., et al. (2022). Whole-  
599 genome sequencing as an investigational device for return of hereditary disease risk and  
600 pharmacogenomic results as part of the All of Us Research Program. *Genome Med* 14, 34.  
601 [10.1186/s13073-022-01031-z](https://doi.org/10.1186/s13073-022-01031-z).
- 602 4. Schwartz, M.L.B., McCormick, C.Z., Lazzeri, A.L., Lindbuchler, D.M., Hallquist, M.L.G., Manickam,  
603 K., Buchanan, A.H., Rahm, A.K., Giovanni, M.A., Frisbie, L., et al. (2018). A Model for Genome-First  
604 Care: Returning Secondary Genomic Findings to Participants and Their Healthcare Providers in a  
605 Large Research Cohort. *Am J Hum Genet* 103, 328-337. [10.1016/j.ajhg.2018.07.009](https://doi.org/10.1016/j.ajhg.2018.07.009).
- 606 5. Lewis, A.C.F., Knoppers, B.M., and Green, R.C. (2021). An international policy on returning  
607 genomic research results. *Genome Med* 13, 115. [10.1186/s13073-021-00928-5](https://doi.org/10.1186/s13073-021-00928-5).
- 608 6. Wilson, J.F., Weiss, D.A., Richards, M., Thomas, M.G., Bradman, N., and Goldstein, D.B. (2001).  
609 Genetic evidence for different male and female roles during cultural transitions in the British Isles.  
610 *Proc Natl Acad Sci U S A* 98, 5078-5083. [10.1073/pnas.071036898](https://doi.org/10.1073/pnas.071036898).
- 611 7. Gilbert, E., O'Reilly, S., Merrigan, M., McGettigan, D., Vitart, V., Joshi, P.K., Clark, D.W., Campbell,  
612 H., Hayward, C., Ring, S.M., et al. (2019). The genetic landscape of Scotland and the Isles. *Proc Natl*  
613 *Acad Sci U S A* 116, 19064-19070. [10.1073/pnas.1904761116](https://doi.org/10.1073/pnas.1904761116).
- 614 8. Halachev, M., Gountouna, V.E., Meynert, A., Tzoneva, G., Shuldiner, A.R., Semple, C.A., and  
615 Wilson, J.F. (2024). Regionally enriched rare deleterious exonic variants in the UK and Ireland. *Nat*  
616 *Commun* 15, 8454. [10.1038/s41467-024-51604-2](https://doi.org/10.1038/s41467-024-51604-2).
- 617 9. McQuillan, R., Leutenegger, A.L., Abdel-Rahman, R., Franklin, C.S., Pericic, M., Barac-Lauc, L.,  
618 Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., et al. (2008). Runs of homozygosity in  
619 European populations. *Am J Hum Genet* 83, 359-372. [10.1016/j.ajhg.2008.08.007](https://doi.org/10.1016/j.ajhg.2008.08.007).
- 620 10. Kerr, S.M., Klaric, L., Halachev, M., Hayward, C., Boutin, T.S., Meynert, A.M., Semple, C.A.,  
621 Tuiskula, A.M., Swan, H., Santoyo-Lopez, J., et al. (2019). An actionable KCNH2 Long QT Syndrome

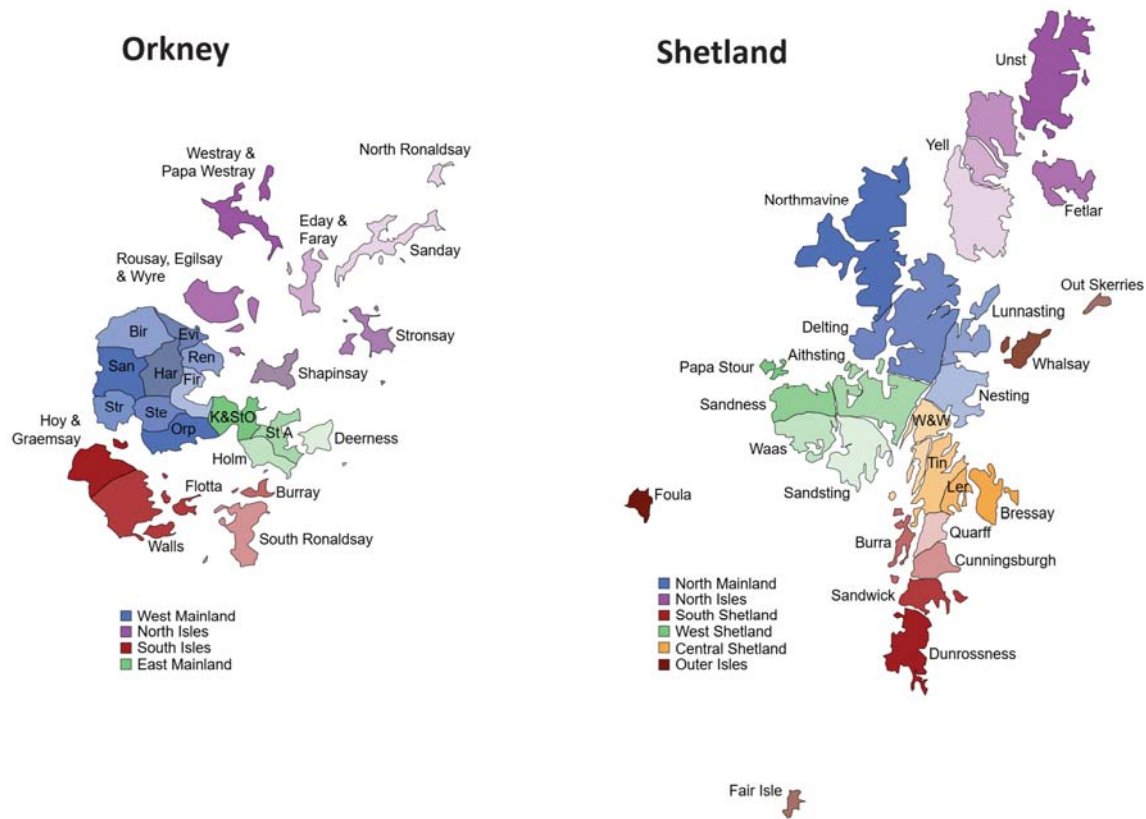
- 622 variant detected by sequence and haplotype analysis in a population research cohort. *Sci Rep* 9,  
623 10964. [10.1038/s41598-019-47436-6](https://doi.org/10.1038/s41598-019-47436-6).
- 624 11. Kerr, S.M., Cowan, E., Klaric, L., Bell, C., O'Sullivan, D., Buchanan, D., Grzymiski, J.J., van Hout, C.V.,  
625 Tzoneva, G., Shuldiner, A.R., et al. (2023). Clinical case study meets population cohort:  
626 identification of a BRCA1 pathogenic founder variant in Orcadians. *Eur J Hum Genet* 31, 588-595.  
627 [10.1038/s41431-023-01297-w](https://doi.org/10.1038/s41431-023-01297-w).
- 628 12. Kerr, S.M., Klaric, L., Muckian, M.D., Cowan, E., Snadden, L., Tzoneva, G., Shuldiner, A.R.,  
629 Miedzybrodzka, Z., and Wilson, J.F. (2024). Two founder variants account for over 90% of  
630 pathogenic BRCA alleles in the Orkney and Shetland Isles in Scotland. *European Journal of Human*  
631 *Genetics*. [10.1038/s41431-024-01704-w](https://doi.org/10.1038/s41431-024-01704-w).
- 632 13. Kelly, M.A., Leader, J.B., Wain, K.E., Bodian, D., Oetjens, M.T., Ledbetter, D.H., Martin, C.L., and  
633 Strande, N.T. (2021). Leveraging population-based exome screening to impact clinical care: The  
634 evolution of variant assessment in the Geisinger MyCode research project. *Am J Med Genet C*  
635 *Semin Med Genet* 187, 83-94. [10.1002/ajmg.c.31887](https://doi.org/10.1002/ajmg.c.31887).
- 636 14. Landrum, M.J., Chitipiralla, S., Brown, G.R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur,  
637 K., Liu, C., et al. (2020). ClinVar: improvements to accessing data. *Nucleic Acids Res* 48, D835-  
638 D844. [10.1093/nar/gkz972](https://doi.org/10.1093/nar/gkz972).
- 639 15. Kerr, S., Edwards, R., Buchanan, D., Dean, J., Miedzybrodzka, Z., and Wilson, J. (2023). VIKING II, a  
640 Worldwide Observational Cohort of Volunteers with Northern Isles Ancestry. *International Journal*  
641 *of Population Data Science* 8. [10.23889/ijpds.v8i1.2121](https://doi.org/10.23889/ijpds.v8i1.2121).
- 642 16. Gilly, A., Klaric, L., Park, Y.C., Png, G., Barysenka, A., Marsh, J.A., Tsafantakis, E., Karaleftheri, M.,  
643 Dedoussis, G., Wilson, J.F., and Zeggini, E. (2022). Gene-based whole genome sequencing meta-  
644 analysis of 250 circulating proteins in three isolated European populations. *Mol Metab* 61,  
645 101509. [10.1016/j.molmet.2022.101509](https://doi.org/10.1016/j.molmet.2022.101509).
- 646 17. Halachev, M., Meynert, A., Taylor, M.S., Vitart, V., Kerr, S.M., Klaric, L., Consortium, S.G.P., Aitman,  
647 T.J., Haley, C.S., Prendergast, J.G., et al. (2019). Increased ultra-rare variant load in an isolated  
648 Scottish population impacts exonic and regulatory regions. *PLoS Genet* 15, e1008480.  
649 [10.1371/journal.pgen.1008480](https://doi.org/10.1371/journal.pgen.1008480).
- 650 18. Raney, B.J., Barber, G.P., Benet-Pages, A., Casper, J., Clawson, H., Cline, M.S., Diekhans, M.,  
651 Fischer, C., Navarro Gonzalez, J., Hickey, G., et al. (2024). The UCSC Genome Browser database:  
652 2024 update. *Nucleic Acids Res* 52, D1082-D1088. [10.1093/nar/gkad987](https://doi.org/10.1093/nar/gkad987).
- 653 19. Wallace, D.F., and Dooley, J.S. (2020). ATP7B variant penetrance explains differences between  
654 genetic and clinical prevalence estimates for Wilson disease. *Hum Genet* 139, 1065-1075.  
655 [10.1007/s00439-020-02161-3](https://doi.org/10.1007/s00439-020-02161-3).
- 656 20. Roberts, A.M., Ware, J.S., Herman, D.S., Schafer, S., Baksi, J., Bick, A.G., Buchan, R.J., Walsh, R.,  
657 John, S., Wilkinson, S., et al. (2015). Integrated allelic, transcriptional, and phenomic dissection of  
658 the cardiac effects of titin truncations in health and disease. *Sci Transl Med* 7, 270ra276.  
659 [10.1126/scitranslmed.3010134](https://doi.org/10.1126/scitranslmed.3010134).
- 660 21. Haggerty, C.M., Damrauer, S.M., Levin, M.G., Birtwell, D., Carey, D.J., Golden, A.M., Hartzel, D.N.,  
661 Hu, Y., Judy, R., Kelly, M.A., et al. (2019). Genomics-First Evaluation of Heart Disease Associated  
662 With Titin-Truncating Variants. *Circulation* 140, 42-54. [10.1161/CIRCULATIONAHA.119.039573](https://doi.org/10.1161/CIRCULATIONAHA.119.039573).
- 663 22. Backman, J.D., Li, A.H., Marcketta, A., Sun, D., Mbatchou, J., Kessler, M.D., Benner, C., Liu, D.,  
664 Locke, A.E., Balasubramanian, S., et al. (2021). Exome sequencing and analysis of 454,787 UK  
665 Biobank participants. *Nature* 599, 628-634. [10.1038/s41586-021-04103-z](https://doi.org/10.1038/s41586-021-04103-z).
- 666 23. Investigators, G.P.P., Smedley, D., Smith, K.R., Martin, A., Thomas, E.A., McDonagh, E.M., Cipriani,  
667 V., Ellingford, J.M., Arno, G., Tucci, A., et al. (2021). 100,000 Genomes Pilot on Rare-Disease  
668 Diagnosis in Health Care - Preliminary Report. *N Engl J Med* 385, 1868-1880.  
669 [10.1056/NEJMoa2035790](https://doi.org/10.1056/NEJMoa2035790).
- 670 24. Bick, A.G., Metcalf, G.A., Mayo, K.R., Lichtenstein, L., Rura, S., Carroll, R.J., Musick, A., Linder, J.E.,  
671 Jordan, I.K., Nagar, S.D., et al. (2024). Genomic data in the All of Us Research Program. *Nature*  
672 627, 340-346. [10.1038/s41586-023-06957-x](https://doi.org/10.1038/s41586-023-06957-x).
- 673 25. Lee, W.K., Haddad, L., Macleod, M.J., Dorrance, A.M., Wilson, D.J., Gaffney, D., Dominiczak, M.H.,  
674 Packard, C.J., Day, I.N., Humphries, S.E., and Dominiczak, A.F. (1998). Identification of a common



- 675 low density lipoprotein receptor mutation (C163Y) in the west of Scotland. *J Med Genet* 35, 573-  
676 578. [10.1136/jmg.35.7.573](https://doi.org/10.1136/jmg.35.7.573).
- 677 26. Gratton, J., Humphries, S.E., and Futema, M. (2023). Prevalence of FH-Causing Variants and Impact  
678 on LDL-C Concentration in European, South Asian, and African Ancestry Groups of the UK Biobank-  
679 Brief Report. *Arterioscler Thromb Vasc Biol* 43, 1737-1742. [10.1161/ATVBAHA.123.319438](https://doi.org/10.1161/ATVBAHA.123.319438).
- 680 27. Vears, D.F., Hallowell, N., Bentzen, H.B., Ellul, B., Nost, T.H., Kerasidou, A., Kerr, S.M., Th  
681 Mayrhofer, M., Mezinska, S., Ormondroyd, E., et al. (2023). A practical checklist for return of  
682 results from genomic research in the European context. *Eur J Hum Genet* 31, 687-695.  
683 [10.1038/s41431-023-01328-6](https://doi.org/10.1038/s41431-023-01328-6).
- 684 28. Woodward, E.R., Green, K., Burghel, G.J., Bulman, M., Clancy, T., Lalloo, F., Schlecht, H., Wallace,  
685 A.J., and Evans, D.G. (2022). 30 year experience of index case identification and outcomes of  
686 cascade testing in high-risk breast and colorectal cancer predisposition genes. *Eur J Hum Genet*  
687 30, 413-419. [10.1038/s41431-021-01011-8](https://doi.org/10.1038/s41431-021-01011-8).
- 688 29. Hamel, N., Kotar, K., and Foulkes, W.D. (2003). Founder mutations in BRCA1/2 are not frequent in  
689 Canadian Ashkenazi Jewish men with prostate cancer. *BMC Med Genet* 4, 7. [10.1186/1471-2350-  
690 4-7](https://doi.org/10.1186/1471-2350-4-7).
- 691 30. Kalia, S.S., Adelman, K., Bale, S.J., Chung, W.K., Eng, C., Evans, J.P., Herman, G.E., Hufnagel, S.B.,  
692 Klein, T.E., Korf, B.R., et al. (2017). Recommendations for reporting of secondary findings in clinical  
693 exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American  
694 College of Medical Genetics and Genomics. *Genet Med* 19, 249-255. [10.1038/gim.2016.190](https://doi.org/10.1038/gim.2016.190).
- 695 31. Dewey, F.E., Murray, M.F., Overton, J.D., Habegger, L., Leader, J.B., Fetterolf, S.N., O'Dushlaine, C.,  
696 Van Hout, C.V., Staples, J., Gonzaga-Jauregui, C., et al. (2016). Distribution and clinical impact of  
697 functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* 354.  
698 [10.1126/science.aaf6814](https://doi.org/10.1126/science.aaf6814).
- 699 32. Jensson, B.O., Arnadottir, G.A., Katrinardottir, H., Fridriksdottir, R., Helgason, H., Oddsson, A.,  
700 Sveinbjornsson, G., Eggertsson, H.P., Halldorsson, G.H., Atlason, B.A., et al. (2023). Actionable  
701 Genotypes and Their Association with Life Span in Iceland. *N Engl J Med* 389, 1741-1752.  
702 [10.1056/NEJMoa2300792](https://doi.org/10.1056/NEJMoa2300792).
- 703 33. Pilling, L.C., Tamosauskaite, J., Jones, G., Wood, A.R., Jones, L., Kuo, C.-L., Kuchel, G.A., Ferrucci, L.,  
704 and Melzer, D. (2019). Common conditions associated with hereditary haemochromatosis genetic  
705 variants: cohort study in UK Biobank. *BMJ* 364, k5222. [10.1136/bmj.k5222](https://doi.org/10.1136/bmj.k5222).
- 706 34. Lynch, M.T., Maloney, K.A., Pollin, T.I., Streeten, E.A., Xu, H., Regeneron Genetics, C., Shuldiner,  
707 A.R., Van Hout, C.V., Gonzaga-Jauregui, C., and Mitchell, B.D. (2021). The burden of pathogenic  
708 variants in clinically actionable genes in a founder population. *Am J Med Genet A* 185, 3476-3484.  
709 [10.1002/ajmg.a.62472](https://doi.org/10.1002/ajmg.a.62472).
- 710 35. Jordan, E., Peterson, L., Ai, T., Asatryan, B., Bronicki, L., Brown, E., Celeghin, R., Edwards, M., Fan,  
711 J., Ingles, J., et al. (2021). Evidence-Based Assessment of Genes in Dilated Cardiomyopathy.  
712 *Circulation* 144, 7-19. [10.1161/CIRCULATIONAHA.120.053033](https://doi.org/10.1161/CIRCULATIONAHA.120.053033).
- 713 36. Shah, R.A., Asatryan, B., Sharaf Dabbagh, G., Aung, N., Khanji, M.Y., Lopes, L.R., van Duijvenboden,  
714 S., Holmes, A., Muser, D., Landstrom, A.P., et al. (2022). Frequency, Penetrance, and Variable  
715 Expressivity of Dilated Cardiomyopathy-Associated Putative Pathogenic Gene Variants in UK  
716 Biobank Participants. *Circulation* 146, 110-124. [10.1161/CIRCULATIONAHA.121.058143](https://doi.org/10.1161/CIRCULATIONAHA.121.058143).
- 717 37. Jansen, M., Baas, A.F., van Spaendonck-Zwarts, K.Y., Ummels, A.S., van den Wijngaard, A.,  
718 Jongbloed, J.D.H., van Slegtenhorst, M.A., Lekanne Deprez, R.H., Wessels, M.W., Michels, M., et al.  
719 (2019). Mortality Risk Associated With Truncating Founder Mutations in Titin. *Circ Genom Precis  
720 Med* 12, e002436. [10.1161/circgen.118.002436](https://doi.org/10.1161/circgen.118.002436).
- 721 38. Middleton, A., Morley, K.I., Bragin, E., Firth, H.V., Hurles, M.E., Wright, C.F., Parker, M., and study,  
722 D.D.D. (2016). Attitudes of nearly 7000 health professionals, genomic researchers and publics  
723 toward the return of incidental results from sequencing research. *Eur J Hum Genet* 24, 21-29.  
724 [10.1038/ejhg.2015.58](https://doi.org/10.1038/ejhg.2015.58).
- 725 39. Schupmann, W., Miner, S.A., Sullivan, H.K., Glover, J.R., Hall, J.E., Schurman, S.H., and Berkman,  
726 B.E. (2021). Exploring the motivations of research participants who chose not to learn medically

- 727 actionable secondary genetic findings about themselves. *Genet Med* 23, 2281-2288.  
728 10.1038/s41436-021-01271-1.
- 729 40. McGurk, K.A., Zhang, X., Theotokis, P., Thomson, K., Harper, A., Buchan, R.J., Mazaika, E.,  
730 Ormondroyd, E., Wright, W.T., Macaya, D., et al. (2023). The penetrance of rare variants in  
731 cardiomyopathy-associated genes: A cross-sectional approach to estimating penetrance for  
732 secondary findings. *Am J Hum Genet* 110, 1482-1495. 10.1016/j.ajhg.2023.08.003.
- 733 41. Forrest, I.S., Duffy, A., Park, J.K., Vy, H.M.T., Pasquale, L.R., Nadkarni, G.N., Cho, J.H., and Do, R.  
734 (2024). Genome-first evaluation with exome sequence and clinical data uncovers underdiagnosed  
735 genetic disorders in a large healthcare system. *Cell Rep Med*, 101518.  
736 10.1016/j.xcrm.2024.101518.
- 737 42. Sapp, J.C., Facio, F.M., Cooper, D., Lewis, K.L., Modlin, E., van der Wees, P., and Biesecker, L.G.  
738 (2021). A systematic literature review of disclosure practices and reported outcomes for medically  
739 actionable genomic secondary findings. *Genet Med* 23, 2260-2269. 10.1038/s41436-021-01295-7.
- 740 43. Blout Zawatsky, C.L., Shah, N., Machini, K., Perez, E., Christensen, K.D., Zouk, H., Steeves, M., Koch,  
741 C., Uveges, M., Shea, J., et al. (2021). Returning actionable genomic results in a research biobank:  
742 Analytic validity, clinical implementation, and resource utilization. *Am J Hum Genet* 108, 2224-  
743 2237. 10.1016/j.ajhg.2021.10.005.
- 744 44. Clayton, E.W., Smith, M.E., Anderson, K.C., Chung, W.K., Connolly, J.J., Fullerton, S.M., McGowan,  
745 M.L., Peterson, J.F., Prows, C.A., Sabatello, M., and Holm, I.A. (2023). Studying the impact of  
746 translational genomic research: Lessons from eMERGE. *Am J Hum Genet* 110, 1021-1033.  
747 10.1016/j.ajhg.2023.05.011.
- 748 45. O'Brien, T.D., Potter, A.B., Driscoll, C.C., Goh, G., Letaw, J.H., McCabe, S., Thanner, J., Kulkarni, A.,  
749 Wong, R., Medica, S., et al. (2023). Population screening shows risk of inherited cancer and  
750 familial hypercholesterolemia in Oregon. *Am J Hum Genet* 110, 1249-1265.  
751 10.1016/j.ajhg.2023.06.014.
- 752 46. Wiley, L.K., Shortt, J.A., Roberts, E.R., Lowery, J., Kudron, E., Lin, M., Mayer, D., Wilson, M.,  
753 Brunetti, T.M., Chavan, S., et al. (2024). Building a vertically integrated genomic learning health  
754 system: The biobank at the Colorado Center for Personalized Medicine. *Am J Hum Genet* 111, 11-  
755 23. 10.1016/j.ajhg.2023.12.001.
- 756 47. Nolan, J., Buchanan, J., Taylor, J., Almeida, J., Bedenham, T., Blair, E., Broadgate, S., Butler, S.,  
757 Cazeaux, A., Craft, J., et al. (2023). Secondary (additional) findings from the 100,000 Genomes  
758 Project: disease manifestation, healthcare outcomes and costs of disclosure. *Genet Med*, 101051.  
759 10.1016/j.gim.2023.101051.
- 760 48. Gold, J., Kripke, C.M., Center, R.G., BioBank, P.M., and Drivas, T.G. (2024). Universal Exome  
761 Sequencing in Critically Ill Adults: A Diagnostic Yield of 25% and Race-Based Disparities in Access to  
762 Genetic Testing. *medRxiv*, 2024.2003.2011.24304088. 10.1101/2024.03.11.24304088.
- 763 49. O'Dushlaine, C., McQuillan, R., Weale, M.E., Crouch, D.J., Johansson, A., Aulchenko, Y., Franklin,  
764 C.S., Polasek, O., Fuchsberger, C., Corvin, A., et al. (2010). Genes predict village of origin in rural  
765 Europe. *Eur J Hum Genet* 18, 1269-1270. 10.1038/ejhg.2010.92.
- 766 50. McGuigan, A., Whitworth, J., Andreou, A., Hearn, T., Genomics England Research, C., Tischkowitz,  
767 M., and Maher, E.R. (2022). Multilocus Inherited Neoplasia Allele Syndrome (MINAS): an update.  
768 *Eur J Hum Genet* 30, 265-270. 10.1038/s41431-021-01013-6.
- 769 51. Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., Hutnik, K., Royrvik, E.C.,  
770 Cunliffe, B., Wellcome Trust Case Control, C., et al. (2015). The fine-scale genetic structure of the  
771 British population. *Nature* 519, 309-314. 10.1038/nature14230.
- 772 52. Gilbert, E., Carmi, S., Ennis, S., Wilson, J.F., and Cavalleri, G.L. (2017). Genomic insights into the  
773 population structure and history of the Irish Travellers. *Sci Rep* 7, 42187. 10.1038/srep42187.
- 774 53. Lynch, S.A., Crushell, E., Lambert, D.M., Byrne, N., Gorman, K., King, M.D., Green, A., O'Sullivan, S.,  
775 Browne, F., Hughes, J., et al. (2018). Catalogue of inherited disorders found among the Irish  
776 Traveller population. *J Med Genet* 55, 233-239. 10.1136/jmedgenet-2017-104974.
- 777 54. Welch, S.G., and Mears, G.W. (1972). Genetic variants of human indophenol oxidase in the  
778 Westray Island of the Orkneys. *Hum Hered* 22, 38-41. 10.1159/000152465.
- 779 55. Rennie, A.G. (1983). The Whalsay cataract. *Trans Ophthalmol Soc U K* (1962) 103 ( Pt 5), 530-531.

- 780 56. Muir, V.M. (1978). Tylosis in the Orkney islands. *J Biosoc Sci* *10*, 1-6.  
781 10.1017/s0021932000011408.
- 782 57. Rugg, E.L., Common, J.E., Wilgoss, A., Stevens, H.P., Buchan, J., Leigh, I.M., and Kelsell, D.P. (2002).  
783 Diagnosis and confirmation of epidermolytic palmoplantar keratoderma by the identification of  
784 mutations in keratin 9 using denaturing high-performance liquid chromatography. *Br J Dermatol*  
785 *146*, 952-957. 10.1046/j.1365-2133.2002.04764.x.
- 786 58. Roberts, D.F. (1986). Who are the Orcadians? *Anthropologischer Anzeiger* *44*, 93-104.
- 787
- 788



789

790 **Figure 1** Isles and parishes of Orkney and Shetland. Each archipelago consists of a larger island called the

791 Mainland, with surrounding smaller isles. Both island groups can be divided into 4-6 larger areas, such as

792 West Shetland (known locally as the West side) or the West Mainland of Orkney, which in turn are

793 organised into 25 parishes or isles. Founder effects are usually concentrated in one parish/isle or in one of

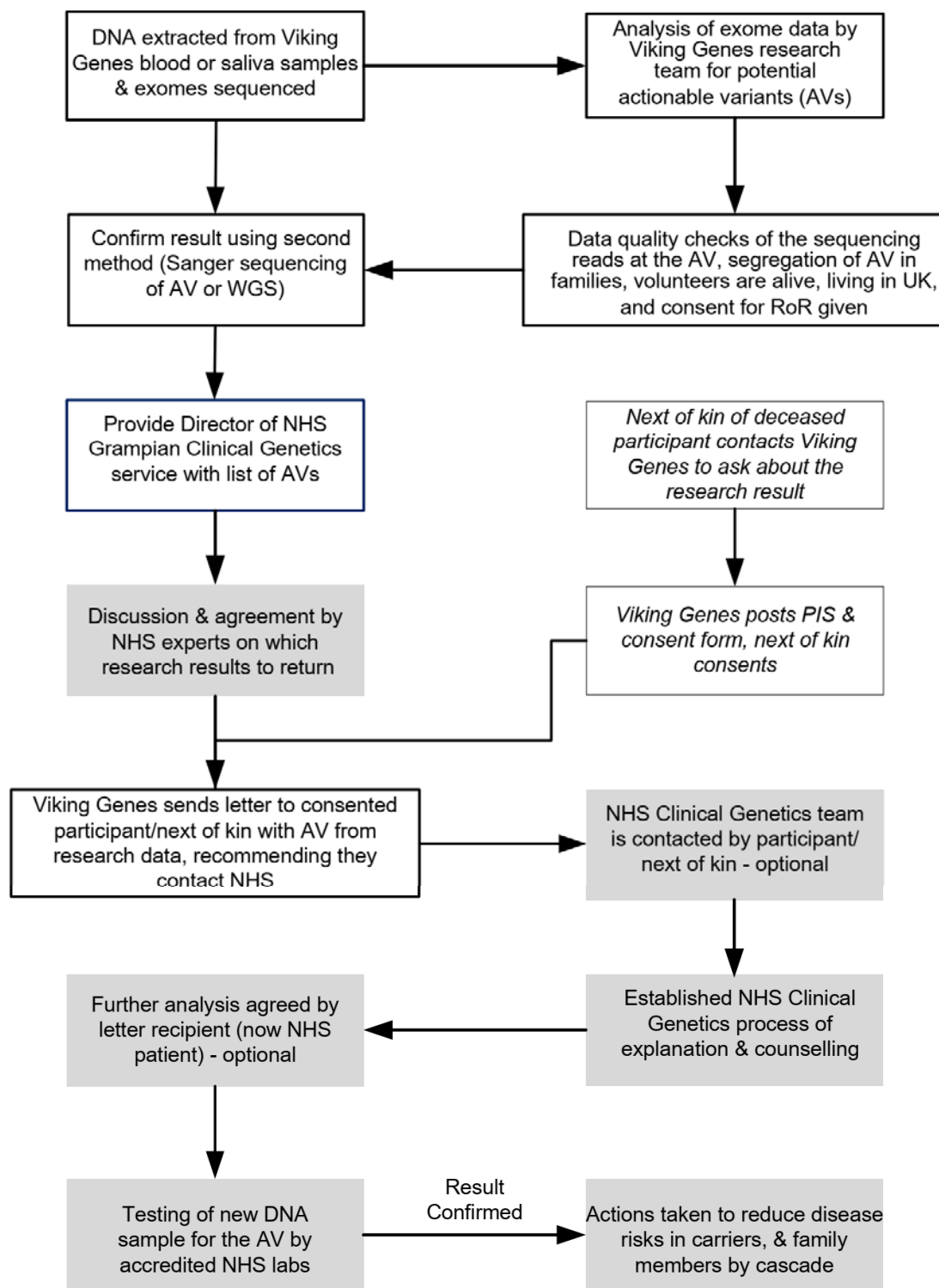
794 the larger contiguous areas. W&W, Whiteness & Weisdale; Tin, Tingwall; Ler, Lerwick; St A, St. Andrews;

795 K&StO, Kirkwall & St. Ola; Orp, Orphir; Ste, Stenness; Str, Stromness; San, Sandwick (Orkney); Bir, Birsay;

796 Evi, Evie; Ren, Rendall; Fir, Firth; Har, Harray.

797

798



799

800 **Figure 2. Overview of Viking Genes return of results.** Processes implemented by the Viking Genes

801 research team are boxed with no shading; processes implemented by the NHS clinical team are shaded.

802 PIS, participant information sheet (see Supplemental information).

category	disease	gene	c.	p.	clin sig	actionable	consequence	LoF	star
cancer	breast-ovarian cancer, familial	<i>BRCA1</i>	c.5207T>C	p.Val1736Ala	pathogenic	All P and LP	missense	0	3
		<i>BRCA2</i>	c.517-2A>G	.	pathogenic	All P and LP	splice acceptor	1	3
	Lynch syndrome	<i>MSH6</i>	c.3261dup	p.Phe1088fs	pathogenic	All P and LP	frameshift	1	3
	attenuated familial adenomatous polyposis	<i>APC</i>	c.154C>T	p.Gln52Ter	pathogenic	All P and LP	stop gain	1	2
	multiple endocrine neoplasia	<i>RET</i>	c.2410G>A	p.Val804Met	P/LP	All P and LP	missense	0	2
metabolic	hereditary transthyretin amyloidosis	<i>TTR</i>	c.262A>T	p.Ile88Leu	P/LP	All P and LP	missense	0	2
cardiovascular	long QT syndrome	<i>KCNH2</i>	c.1750G>A	p.Gly584Ser	P/LP	All P and LP	missense	0	2
		<i>SCN5A</i>	c.4716C>T	p.Gly1572=	P/LP	All P and LP	splice donor	1	2
	dilated cardiomyopathy	<i>TTN</i>	c.1558dupA*	p.Thr520fs	LP	only LOF	frameshift	1	1
		<i>TTN</i>	c.67519C>T	p.Gln22507Ter	LP	only LOF	stop gain	1	1
		<i>TTN</i>	c.88837A>T	p.Lys29613Ter	P/LP	only LOF	stop gain	1	2
		<i>TTN</i>	c.93396_93400del	p.Ala31133_Trp31134insTer	P/LP	only LOF	frameshift	1	2
vascular Ehlers-Danlos syndrome	<i>COL3A1</i>	c.944G>C*	p.Gly315Ala	LP	All P and LP	missense	0	1	

**Table 1 Summary of non-singleton autosomal dominant actionable variants in the cohort of 4,198 participants.** The diseases are grouped into the categories of cancer, metabolic and cardiovascular. c., coding sequence change; p., protein sequence change; clin sig, clinical significance according to ClinVar; LP, likely pathogenic; P/LP, pathogenic/likely pathogenic; actionable, variant categories considered to be actionable; LOF, loss of function; consequence, the molecular consequence of the variant; star, ClinVar star status. *TTN* c.1558dupA does not appear in ClinVar, so the star status and clinical significance are given for a variant causing the identical amino-acid change. \*These variants were not considered actionable by the NHS Clinical Genetics team and were not returned.

category	disease	gene	c.	p.	clin sig	actionable	consequence	LoF	star	
cancer	MYH-associated polyposis	<i>MUTYH</i>	c.452A>G	p.Tyr151Cys	P/LP	P and LP (2 var)	missense	0	2	
		<i>MUTYH</i>	c.1103G>A	p.Gly368Asp	P/LP	P and LP (2 var)	missense	0	2	
		<i>MUTYH</i>	c.1130C>T	p.Pro377Leu	pathogenic	P and LP (2 var)	missense	0	2	
metabolic	Wilson disease	<i>ATP7B</i>	c.122A>G	p.Asn41Ser	P/LP	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.1745_1746del	p.Ile582fs	P/LP	P and LP (2 var)	frameshift	1	2	
		<i>ATP7B</i>	c.1772G>A	p.Gly591Asp	P/LP	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.2605G>A	p.Gly869Arg	P/LP	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.2662A>C	p.Thr888Pro	P/LP	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.2930C>T	p.Thr977Met	pathogenic	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.3007G>A	p.Ala1003Thr	P/LP	P and LP (2 var)	missense	0	2	
		<i>ATP7B</i>	c.3083_3085delinsG	p.Lys1028fs	pathogenic	P and LP (2 var)	frameshift	1	2	
	biotinidase deficiency	<i>BTD</i>	c.1308A>C	p.Gln436His	P/LP	P and LP (2 var)	missense	0	2	
		<i>BTD</i>	c.1552C>T	p.Arg518Cys	pathogenic	P and LP (2 var)	missense	0	2	
	hereditary haemochromatosis	<i>HFE</i>	c.845G>A	p.Cys282Tyr	pathogenic	C282Y homs only	missense	0	2	
	Pompe disease	<i>GAA</i>	c.841C>T	p.Arg281Trp		LP	P and LP (2 var)	missense	0	3
		<i>GAA</i>	c.1194+3G>C		.	LP	P and LP (2 var)	splice region	0	3
		<i>GAA</i>	c.1210G>A	p.Asp404Asn		pathogenic	P and LP (2 var)	missense	0	3
<i>GAA</i>		c.2238G>C	p.Trp746Cys		pathogenic	P and LP (2 var)	missense	0	3	

**Table 2 Summary of non-singleton autosomal recessive actionable variants in the cohort of 4,198 participants.** The diseases are grouped into the categories of cancer and metabolic. c., coding sequence change; p., protein sequence change; clin sig, clinical significance according to ClinVar; LP, likely pathogenic; P/LP, pathogenic/likely

pathogenic; actionable, variant categories considered to be actionable; LOF, loss of function; consequence, the molecular consequence of the variant; star, ClinVar star status; 2 var, 2 variants must be present, so either homozygous or compound heterozygous for the P/LP variant; homs only indicates that only homozygotes for the named variant should be reported.



Gene 1	variant ID	inheritance	genotype	Gene 2	variant ID	inheritance	genotype
<i>BRCA1</i>	c.5207T>C	AD	het	<i>PMS2</i>	c.137G>T	AD	het
<i>BRCA2</i>	c.6275_6276del	AD	het	<i>MSH6</i>	c.742C>T	AD	het
<i>BRCA2</i>	c.517-2A>G	AD	het	<i>GAA</i>	c.841C>T	AR	hom
<i>BRCA2</i>	c.5073dupA	AD	het	<i>HFE</i>	c.845G>A	AR	hom

**Table 3. Actionable genotype co-occurrences.** AD, autosomal dominant; het, heterozygous; hom, homozygous.

gene	c.	p.	disease	inher	maf.ork	ork/ukb	maf.shet	shet/ukb	group	origin
<i>BRCA1</i>	c.5207T>C	p.Val1736Ala	Hereditary breast ovarian cancer	AD	0.00478	470	0	0	Orkney	Westray
<i>BRCA2</i>	c.517-2A>G	.		AD	0	0	0.00237	155	Shetland	Whalsay
<i>ATP7B</i>	c.3083_3085delinsG	p.Lys1028fs	Wilson disease	AR	0	0	0.00475	180	Shetland	Burra
<i>ATP7B</i>	c.3007G>A	p.Ala1003Thr		AR	0	0	0.00356	350	Shetland	Sandsting/Lunnasting/Delting
<i>ATP7B</i>	c.1772G>A	p.Gly591Asp		AR	0.00215	66	0	0	Orkney	Westray
<i>TTN</i>	c.93396_93400del	p.Ala31133_Trp31134insTer	Dilated cardiomyopathy	AD	0	0	0.00380	3700	Shetland	Yell
<i>TTN</i>	c.67519C>T	p.Gln22507Ter		AD	0.00048	700	0.00071	470	Orkney	West Mainland
<i>KCNH2</i>	c.1750G>A	p.Gly584Ser	Long QT syndrome	AD	0	0	0.00119	90	Shetland	Aithsting, Unst
<i>MUTYH</i>	c.1205C>T	p.Pro377Leu	<i>MYH</i> -associated polyposis	AR	0.00431	70	0.00047	8	Orkney	Westray
<i>GAA</i>	c.1210G>A	p.Asp404Asn	Pompe disease	AR	0.00287	74	0	0	Orkney	St. Andrews

**Table 4 Founder effects on ten actionable variants in Orkney and Shetland.** Drifted variants were defined as actionable variants with an increase in frequency over 50-fold higher than UKB, a pedigree origin before c.1800 CE and a minor allele count of  $\geq 5$ . For UKB, we used the allele frequency browser (based on ~490k WGS). c., complementary DNA position and nucleotide change; p., protein position and amino-acid change; inher, inheritance; maf.ork, minor allele frequency in Orkney; ork/ukb, fold uplift in frequency in Orkney over UKB; maf.shet, minor allele frequency in Shetland; shet/ukb, fold uplift in frequency in Shetland over UKB; group, island group the variant has drifted upwards in; origin, subpopulation (parish, isle or region) where the variant arose (by tracing back pedigrees) or has become particularly common.

Actionable genotype carriers	Category
64	Results letters sent, including one to Next of Kin
33	Not consented - did not reply to Return of Results consent invitation
8	Deceased and no request from Next of Kin
2	Answered no to Return of Results consent invitation
10	Not deemed returnable by NHS
117	TOTAL ACTIONABLE GENOTYPE CARRIERS, including <i>BRCA1</i> males and accounting for co-occurrences

**Table 5 Summary of carriers of actionable genotypes and results returned.** Consent was only sought 8-18 years after recruitment, which is probably why 28% of actionable genotype carriers did not reply to the invitation. A further 11% were deceased, while 2.7% replied choosing not to have return of results. NHS, National Health Service.