



# Frequency of genomic secondary findings among 21,915 eMERGE network participants

A full list of authors and their affiliations appears at the end of the paper.

**Purpose:** Discovering an incidental finding (IF) or secondary finding (SF) is a potential result of genomic testing, but few data exist describing types and frequencies of SFs likely to appear in broader clinical populations.

**Methods:** The Electronic Medical Records and Genomics Network Phase III (eMERGE III) developed a CLIA-compliant sequencing panel of 109 genes and 1551 variants of clinical relevance or research interest and deployed this panel at ten clinical sites. We evaluated medically actionable SFs across 67 genes and 14 single-nucleotide variants (SNVs) in a diverse cohort of 21,915 participants drawn from a variety of settings (e.g., primary care, biobanks, specialty clinics).

**Results:** Correcting for testing indication, we found a 3.02% overall frequency of SFs; 2.54% from 59 genes the American College of Medical Genetics and Genomics recommends for SF return, and

0.48% in other genes, primarily *HFE* and *PALB2*. SFs associated with cancer susceptibility were most frequent (1.38%), followed by cardiovascular diseases (0.87%), and lipid disorders (0.50%). After removing *HFE*, the frequency of SFs and proportion of pathogenic versus likely pathogenic SFs did not differ in those self-identifying as White versus others.

**Conclusion:** Here we present frequencies and types of medically actionable secondary findings to support informed decision making by patients, participants, and practitioners engaged in genomic medicine.

*Genetics in Medicine* (2020) 22:1470–1477; <https://doi.org/10.1038/s41436-020-0810-9>

**Keywords:** personal genomics; clinical sequencing; incidental findings secondary findings; eMERGE

## INTRODUCTION

Increasing numbers of patients are undergoing panel, exome, or genome sequencing testing as a part of routine clinical care; consequently, genetic variation that is medically actionable, but not related (secondary) to the indication for testing will become increasingly common. Though there are recommendations regarding the return of these findings,<sup>1–3</sup> challenges remain regarding the identification, interpretation, and return of these results to patients. To begin to address these challenges, the American College of Genetics and Genomics (ACMG) published guidelines listing 56 genes<sup>1</sup> from which secondary findings (SFs, sometimes called incidental or additional findings) should be reported when genomic sequencing is used for clinical purposes, and the types of variants, per interpretation guidelines, that would qualify as true SFs. In 2017, the ACMG working group published an updated version of these guidelines, adding 4 genes and removing 1 gene to make up the current list of 59 genes.<sup>2</sup>

Despite continued development and uptake of this list by the community, the most recent update to the ACMG guideline for secondary findings return still refers to pathogenicity attributes from the 2008 ACMG variant interpretation guidelines.<sup>4</sup> However, many labs return SFs classified as likely pathogenic (LP) or pathogenic (P) based on the 2015 ACMG/Association for Molecular Pathology

(ACMG/AMP) interpretation framework,<sup>5</sup> which is not a direct match to the known pathogenic (KP) and expected pathogenic (EP) categories recommended by current ACMG secondary findings guidelines. One major challenge to the application of the ACMG secondary findings guidelines is that, to date, much of the expert opinion regarding actionability of these gene–disease pairs does not derive from the experience of testing the broader patient populations that these guidelines are primarily intended to support. Specifically, prior estimates of SF frequency have been limited in the age, ancestry, and phenotypic diversity of the sample, restricted to a subset of all known actionable genes (e.g., hereditary cancer only), and often constrained to a single health-care entity or testing lab.<sup>6–9</sup>

Providing data to guide the process and inform the rationale for or against SF return was a major goal of the third phase of the National Human Genome Research Institute's (NHGRI) Electronic Medical Records and Genomics Network (eMERGE III). eMERGE III aims to study and improve standards and methods for delivery of clinical and research data across a multisite cohort, while providing actionable genetic results derived from a next-generation sequencing platform to eMERGE III research participants. This current phase builds on the network's decade of experience in genomics and return of results in the context

Correspondence: Adam S. Gordon ([adam.gordon@northwestern.edu](mailto:adam.gordon@northwestern.edu))

Submitted 30 November 2019; revised 24 February 2020; accepted: 6 April 2020

Published online: 17 June 2020

of health systems, clinical data abstraction from the electronic health record (EHR), linking health-care processes that inform care with those that catalyze research, and studying multiple models of participant consent.

To this end, we developed eMERGEseq, a sequencing panel focused on 109 genes of network interest, and then selected a subset deemed actionable by network consensus including all 56 genes on the initial ACMG list alongside an additional 11 genes and 14 individual variants in 11 additional genes. The panel was deployed by two CLIA sequencing labs on samples from 25,015 participants drawn from ten different clinical sites across the United States,<sup>10–12</sup> the 21,915 participants in the SF cohort described here are participants eligible for results from the entire consensus return list who were not ascertained based on prior sequencing results. Here we describe the frequency of actionable SFs within this cohort to better inform researchers, patients, and providers seeking an understanding of the spectrum of results that can be expected from personal genomic testing.

## MATERIALS AND METHODS

### Panel development and actionability assessment

The process to determine the eMERGE III sequencing platform content and consensus actionability list, described elsewhere,<sup>10</sup> was led by the eMERGE III Sequencing Centers and Clinical Annotation Working Group (ClinAnn WG), with input from the Return of Results Working Group and all eMERGE network members. At the time the eMERGEseq platform was designed, the ACMG recommended return of pathogenic variants in 56 gene–disease pairs<sup>1</sup> as secondary findings when a genomic sequencing test was ordered for a clinical indication. These genes were all included on the platform, in addition to 53 site-nominated additional genes and single-nucleotide variants (SNVs) for inclusion on the panel, to support both discovery-focused and clinical return of result aims.

Pathogenic variants in all 56 ACMG gene–disease pairs were determined to be returnable, as well as variants in these genes pathogenic for other conditions (e.g., *WT1* NM\_024426.6:c.1447+5G>A, pathogenic for Fraser syndrome in addition to Wilms tumor, its ACMG pair). Additional gene actionability recommendations from the literature were also considered for the remaining 53 genes; these were then individually reviewed as previously described,<sup>10</sup> considering both evidence for the gene–disease association and nature of the recommended clinical action or intervention. Additional consensus genes included *BMPRIA*, *PALB2*, *POLD1*, *POLE*, *SMAD4*, *COL5A1*, *KCNE1*, *KCNJ2*, *HNF1A*, *HNF1B*, *CACNA1A*, and *OTC*; genes with individual actionable variants included *ACADM* (NM\_000016.5:c.985A>G [p.Lys329Glu]), *ALDOB* (NM\_000035.4:c.356\_359CAA[1] [p.Asn120fs]), *BCKDHB* (NM\_183050.4:c.832G>A [p.Gly278Ser] and NM\_183050.4:c.548G>C [p.Arg183Pro]), *FAH* (NM\_000137.3:c.782C>T [p.Pro261-Leu]), *G6PC* (NM\_000151.4:c.247C>T [p.Arg83Cys]), *CPT2* (NM\_000098.3:c.1239\_1240del [p.Lys414fs]), *BLM* (NM\_000057.4:c.2207\_2212delinsTAGATTC [p.Tyr736fs]),

*CYP21A2* (NM\_000500.9:c.293-13C>G), *F5* (NM\_000130.4:c.1601G>A [p.Arg534Gln]), *HFE* (NM\_000410.3:c.845G>A [p.Cys282Tyr], homozygotes only), and *MEFV* (NM\_000243.2:c.2177T>C [p.Val726Ala] and NM\_000243.2:c.2080A>G [p.Met694Val]).<sup>10</sup> Though all sites participated in this review, latitude was given to individual sites regarding returning these results to account for differences in study population (e.g., pediatrics sites), study aims, consent, and institutional review.

The ACMG updated its recommendation to 59 genes in 2017, dropping 1 (*MYLK*) and adding 4 (*ATP7B*, *BMPRIA*, *SMAD4*, and *OTC*).<sup>2</sup> Three of these new genes were included on the eMERGEseq panel and had already been deemed actionable by the eMERGE III network, and the fourth was represented by a single SNV (*ATP7B*; NM\_000053.4:c.3207C>A [p.His1069Gln]), which was returned only by two clinical sites. *MYLK* was specifically discussed and, given its calculated score per the ClinGen gene–disease validity framework<sup>13,14</sup> did not differ from other ACMG-included genes associated with thoracic aortic aneurysms and aortic dissections (TAAD), it was retained as actionable by the eMERGE network. Given the research nature of the project, the consent process, and the participants' accessibility to medical care, after some discussion it was decided by network consensus to return both pathogenic (P) and likely pathogenic (LP) variants in all 68 genes.

### Cohort description

The eMERGE III network consists of ten clinical study sites, two sequencing centers (SCs), and a coordinating center. A primary goal of eMERGE III is to harmonize and implement an entire, unified pipeline simulating real-life clinical genomic testing within a broad spectrum of participants and clinical settings, with each clinical site's cohort of ~2500 participants being, by design, markedly different in ascertainment strategy and demographics. The study adhered to the principles set out in the Declaration of Helsinki, with informed consent from all participants, as approved by individual sites' institutional review boards (IRBs); a specific discussion of the ethical considerations across the eMERGE III study is described elsewhere.<sup>15</sup> One clinical site that did not return the consensus list to participants could not be included in this analysis. Another clinical site could not be included as they ascertained participants based on previous non-CLIA sequencing results enriching for putative pathogenic variants, which would bias the analysis; results from this site have been previously described elsewhere.<sup>9</sup> The resulting SF cohort ( $N = 21,915$ ) includes a diverse range of participants drawn from across the lifespan and across the health system, from primary care to specialty clinics; Table 1 details the demographic makeup of both the overall and individual site cohorts.

### Clinical sequencing and interpretation pipeline

Details of the eMERGEseq clinical sequencing pipeline are described in detail elsewhere.<sup>10</sup> Briefly, participants were

# ARTICLE

**Table 1** Demographic characteristics of the eMERGE III secondary findings cohort and nine contributing clinical sites.

eMERGE site	N	Self-reported race/ethnicity						% Female	Median age (years)	Age range (years)
		Hispanic or Latinx	Black or African American	Asian	American Indian, Alaska Native, or Pacific Islander	White	Unknown/not reported			
CCHMC	2956	24	1041	16	2	1697	176	34	20	4–85
Harvard	2492	146	145	72	0	2038	91	57	61	22–100
KPW/UW	2495	43	51	996	51	1297	57	61	65	32–104
<i>SC:LMM</i>	<i>7943</i>									
CHOP	2972	126	1171	42	3	1524	106	21	18	7–32
Columbia	2566	636	302	198	9	1072	349	55	55	21–97
Mayo	3018	500	11	18	2	2390	97	48	65	29–74
Northwestern	2984	167	405	132	7	2257	16	62	65	21–98
Vanderbilt	2432	24	153	23	3	2205	24	47	69	27–97
<i>SC:BCM-HGSC</i>	<i>13972</i>									
<b>Total</b>	<b>21915</b>	<b>1666</b>	<b>3279</b>	<b>1497</b>	<b>77</b>	<b>14480</b>	<b>916</b>	<b>53.0</b>	<b>62</b>	<b>4–104</b>

Self-reported race/ethnicity derives from multiple-choice selection of US Census–standardized terms and does not fully capture the complexity of participant racial identity. Sequencing centers are shown in italics.

*BCM-HGSC* Baylor College of Medicine Human Genome Sequencing Center, *CCHMC* Cincinnati Children’s Hospital Medical Center, *CHOP* Children’s Hospital of Philadelphia, *KPW/UW* Kaiser Permanente Washington/University of Washington, *LMM* Broad Institute and Partners Laboratory for Molecular Medicine, *SC* sequencing center.

enrolled at each site, blood collected, DNA extracted locally, then sent to one of two assigned SCs for targeted sequencing in a College of American Pathologists (CAP)/CLIA setting: the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC), Houston Texas; or the Broad Institute and Partners Laboratory for Molecular Medicine (LMM), Cambridge, Massachusetts. Variants identified through this sequencing test were interpreted by the assigned SC in the context of the site-provided disease status and test indication (if present). Variant classifications from both laboratories were based primarily on ACMG/AMP criteria,<sup>5</sup> supplemented by recommendations from ClinGen’s expert panels and Sequence Variant Interpretation Working Group. Final interpretation decisions incorporated these recommendations along with evidence from the literature and, when present, additional SC internal data accrued from previous testing. Evidence summaries from the laboratories supporting these interpretations were included on the clinical reports. Variant classifications along with evidence and interpretive summaries for all reported variants were submitted directly to ClinVar by both BCM-HGSC and LMM (submitter IDs 500199 and 21766, respectively). Though there are subtle differences between the sequencing and variant analysis approaches at the two SCs, an extensive preliminary harmonization effort between the two SCs at all critical components of the sequencing and interpretation workflow ensured that test quality and accuracy were highly concordant.<sup>10</sup>

## Working definition and identification of secondary findings

Although there is some debate as to the specific terminology of each case, the rate of clinically actionable findings unrelated to any participant ascertainment indication can be estimated. In this study, for the purpose of clarity, a secondary finding

(SF) is defined as a pathogenic (P) or likely pathogenic (LP) variant in a gene on the consensus list with an associated disease phenotype unrelated to any participant indication that was provided to the SCs at sample intake. For individuals without an indication for testing, any positive (P or LP) finding was considered secondary. Thus, SFs in this cohort derive from individuals both with and without a reported indication for testing. Supplementary Table 1 lists these indications and their frequency in this cohort.

Given the mix of ascertainment protocols, the distribution of those eligible to receive primary findings (i.e., those with an indication for testing) varies considerably across site and disease phenotype; this could artificially skew overall estimates of SF frequency for phenotypes that were also indications for testing. To correct for this, we adjusted the total number of participants contributing to each frequency estimate based on indication for testing. For example, in calculating the SF rate for *LDLR*, we excluded the 1626 participants indicated to have a lipid disorder, as these participants by definition could not have received an SF in *LDLR*. Although a limited set of pharmacogenetic markers were also assayed and reported by the SCs, these results, known to be relatively common,<sup>16</sup> are not included in this analysis. Likewise, carrier status for conditions with autosomal recessive inheritance is not included in this analysis.

## RESULTS

### Overall SF rates

Among the 1166 positive reports across the entire eMERGE III-SF cohort ( $N = 21,915$ ), we identified 661 actionable findings unrelated to participant test indication, resulting in an overall SF rate of 3.02% (Fig. 1). Of these, 556 findings

(2.54%) are variants in genes recommended by ACMG to be reported as SFs, while the remaining 105 findings (0.48%) are in other consensus actionable loci. Ten individuals were found to have two independent SFs, each unrelated to participant test indication. Ten variants were each reported in  $\geq 5$  participants (Table 2). Three of these ten variants, in *MYBPC3*, *KCNQ1*, and *RYR1* respectively, were reported as LP. A full list of variants determined to be SFs and their rates within this cohort can be found in Supplementary Table 2. Additionally, two participants were found to have the unanticipated SF of acquired mosaic trisomy 12, a marker in chronic lymphocytic leukemia,<sup>17</sup> after consultation between the relevant SC and clinical site, these results were included on the participant's clinical report.

### SF rates by gene and disease domain

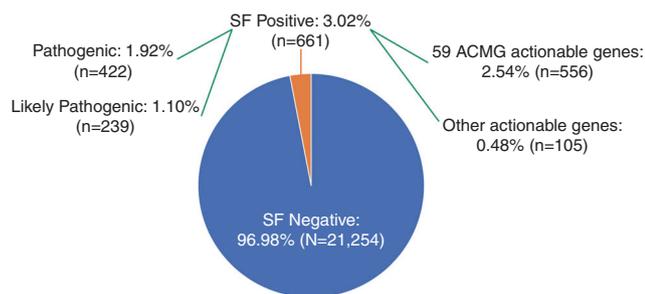
SF counts by disease domain are illustrated in Fig. 2a. Most SFs identified in this study were in cancer-associated genes, with cardiovascular disease-associated genes being the next most common group; these two domains represent the majority of genes on the SF list. For cancer-associated genes, most SFs were P (1.12% P vs. 0.26% LP), while for the cardiac-associated genes most were LP (0.36% P vs. 0.51% LP). Figure 2b illustrates the SFs reported by gene. SFs were most commonly reported in *BRCA2*, *LDLR*, *HFE*, *MYBPC3*, and

*BRCA1*, in that order. Of note, for *HFE*, only homozygosity for p.Cys282Tyr (NM\_001300749.2:c.845G>A) was returned.

### SF rates across self-reported ancestry groups

Table 3 summarizes the rate of SFs by self-reported ancestry. Self-reported Caucasian/White participants had the highest rate, 3.10% (95% confidence interval [CI] 2.84–3.40,  $N = 14,480$ ); followed by Asian, 2.74% (2.02–3.69,  $N = 1497$ ); Black/African American, 2.29% (1.83–2.86,  $N = 3279$ ); Hispanic/Latinx, 1.98% (1.41–2.77,  $N = 1666$ ); and American Indian, Alaska Native, or Native Pacific Islander, 1.30% (0.23–6.97,  $N = 77$ ). In participants where self-reported race was known, there was a significant excess of SFs in those who self-reported as White versus those who did not when *HFE* p.Cys282Tyr (NM\_001300749.2:c.845G>A) results were included (chi-square  $p = 0.003$ , 1 df). This allele is known to be common only in European ancestry individuals<sup>18</sup> and indeed, 65/68 participants in our cohort with this finding identified as White; of the remaining 3, one identified as Black/African American and the final 2 were unknown. After excluding this allele, frequency of SFs did not differ significantly between those self-reporting their race as White (2.68%) versus all other groups combined (2.36%; chi-square  $p = 0.20$ , 1 df). The fraction SFs classified as LPs also did not differ between those self-reporting as White (38.1%) and other groups (41.0%; chi-square,  $p = 0.50$ ).

We did not have access to data on whether participants self-reported Ashkenazi Jewish (AJ) ancestry. These individuals, representing roughly 1.73% of the US population, are known to have a 1/40 rate of harboring a pathogenic *BRCA1/2* founder allele.<sup>19</sup> Based on these rates, we would expect to find ~10 individuals with one of these variants across our cohort of 21,915; we found 26 participants with one of these variants (NM\_000059.3[*BRCA2*]:c.5946del [p.Ser1982fs], NM\_007294.3[*BRCA1*]:c.66\_67AG[1] [p.Glu23fs]), suggesting enrichment of individuals of AJ ancestry in our cohort. This enrichment was not driven by one clinical site alone. However, excluding participants with these variants from analyses does not change our finding that the most frequent disease category of s was



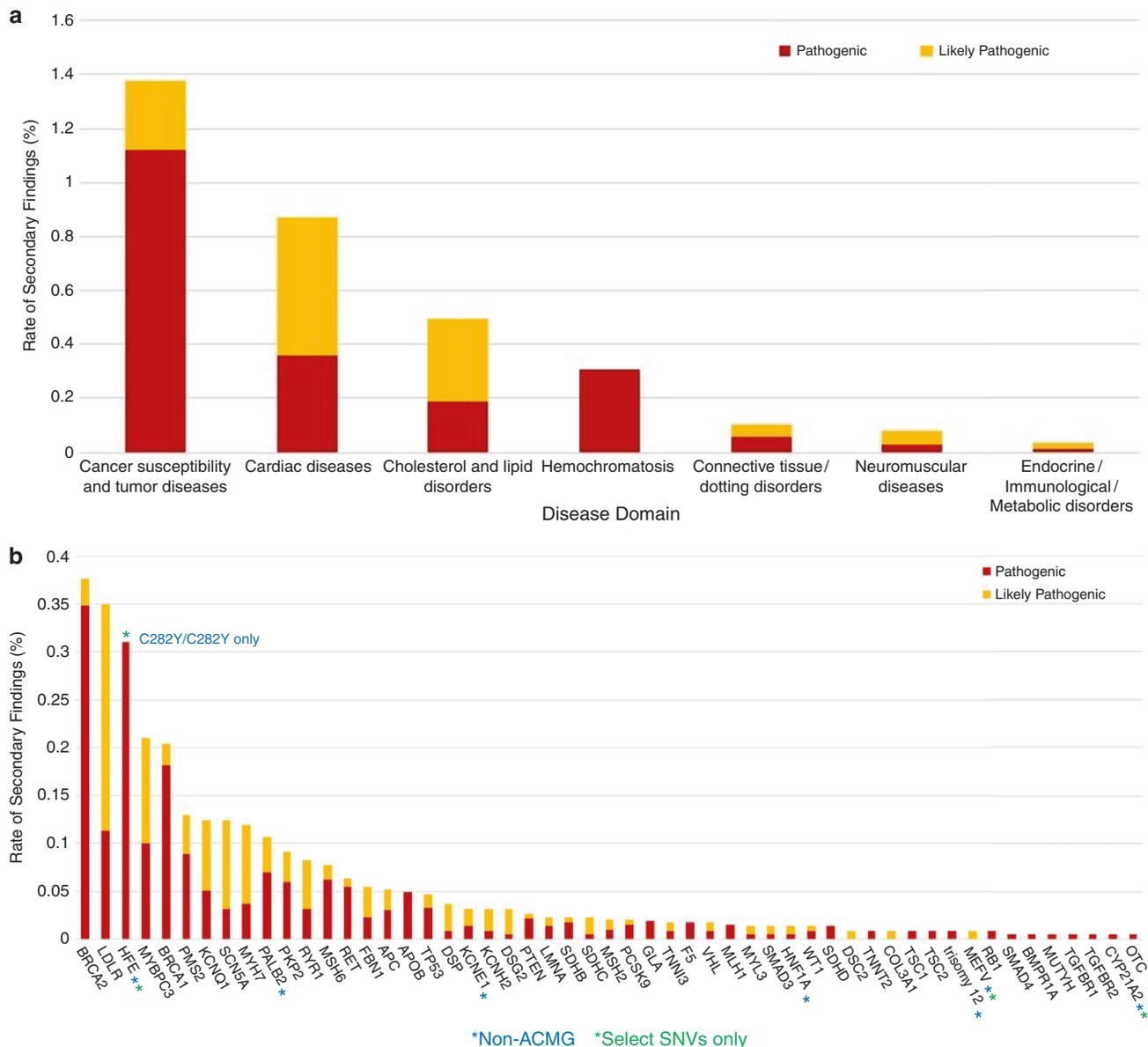
**Fig. 1 Overall frequency of secondary findings (SFs) across the eMERGE III-SF cohort.** Ten individuals were found to have two different SFs. ACMG American College of Medical Genetics and Genomics.

**Table 2 Actionable secondary findings variants identified secondarily in  $\geq 5$  participants.**

Gene	DNA change	Protein consequence	Interpretation	Participant count	rsID	ClinVar variant ID
<i>BRCA2</i>	c.5946del	p.Ser1982Argfs*22	P	17	rs80359550	9325
<i>MYBPC3</i>	c.3628–41_3628–17del	Splice	LP	13	rs36212066	177677
<i>APOB</i>	c.10580G>A	p.Arg3527Gln	P	12	rs5742904	17890
<i>BRCA1</i>	c.68_69del	p.Glu23Valfs*17	P	9	rs80357914	17662
<i>RET</i>	c.2410G>A	p.Val804Met	P	7	rs79658334	37102
<i>KCNQ1</i>	c.1085A>G	p.Lys362Arg	LP	6	rs12720458	52953
<i>MYBPC3</i>	c.1504C>T	p.Arg502Trp	P	6	rs375882485	42540
<i>PMS2</i>	c.137G>T	p.Ser46Ile	P	5	rs121434629	9245
<i>RYR1</i>	c.1840C>T	p.Arg614Cys	LP/P <sup>a</sup>	5	rs118192172	12964
<i>PKP2</i>	c.2146–1G>C	Splice	P	5	rs193922674	6756

LP likely pathogenic, P pathogenic.

<sup>a</sup>Sequencing centers differ in their interpretation of this variant.



**Fig. 2 Secondary finding (SF) rates.** SF rates grouped by (a) associated disease domain and (b) gene. Interpretations as reported by the sequencing centers. Rates are calculated using the proportion of participants who do not have a related indication as enumerated in Supplementary Table 1. Hemochromatosis findings are *HFE* p.C282Y homozygotes only (NM\_001300749.2:c.845G>A). ACMG American College of Medical Genetics and Genomics, SNV single-nucleotide variant.

genes associated with cancer susceptibility (1.38%; 1.24% without AJ founder alleles), though previous studies found the highest rate to be cardiac disease-associated genes<sup>8,20</sup> (0.87% in our cohort).

**DISCUSSION**

As patients and research participants consider the option to receive SFs as part of a genomic test, it is important that they have information on the probability and type of SFs that may be reported. Health systems must also be prepared to return these results and offer follow-up care. To our knowledge, this report is the largest study of medically actionable secondary findings to date. In our cohort, the frequency of SFs was

3.02% overall, with 2.54% in the 59 genes recommended by ACMG to be reported as SFs. This result is consistent with previous estimates of this frequency, though prior studies differ in the genes included as SFs, precede the current ACMG/AMP variant classification system, derive from a single health system, or have considerably smaller sample size.<sup>6-9,20</sup>

Our study cohort includes more individuals identifying as non-White than most studies evaluating SF rates. After removal of the *HFE* p.Cys282Tyr variant (NM\_001300749.2: c.845G>A), we did not find a higher frequency in those who self-report as White versus other race/ethnicity groups, and the proportion of SFs that were classified as LP versus P did

**Table 3** Secondary finding rates stratified by self-reported participant race/ethnicity.

Self-reported race/ethnicity	Number of participants	Number of secondary findings	Fraction of LP findings (%)	Secondary findings rate, % (95% CI)
American Indian, Alaska Native, or Pacific Islander	77	1	100	1.30 (0.23–6.97)
Hispanic or Latinx	1666	33	42	1.98 (1.41–2.77)
Black or African American	3279	75	40	2.29 (1.83–2.86)
Asian	1497	41	59	2.74 (2.02–3.69)
White	14,480	450 <sup>a</sup>	33	3.10 (2.84–3.40)
Unknown/not reported	916	55	36	6.00 (4.64–7.73)

Self-reported race/ethnicity derives from multiple-choice selection of US Census–standardized terms and does not fully capture the complexity of participant racial identity.

CI confidence interval, LP likely pathogenic.

<sup>a</sup>65 of these are *HFE* p.C282Y homozygotes (NM\_000410.3:c.845G>A).

not differ between these groups. This differs from some previous studies of SF rate, which had previously reported an excess of SFs in those of European ancestry versus African ancestry. Aside from our exclusion of *HFE*, this difference could also be potentially explained by their use of slightly different lists of genes to return, cohorts outside the realm of clinical genetic testing, and earlier versions of population databases with limited representation of non-White individuals.<sup>6,7</sup> While there may be differences in the rate of predicted deleterious variants among ancestry groups,<sup>21,22</sup> the frequency of medically actionable secondary findings in our cohort does not mirror this prediction, though our study was not powered to test differences among the individual groups comprising non-White. This limitation, and the dearth of published evidence to accurately classify variants specific to other ancestry backgrounds (e.g., cosegregation), underscores the need to amplify participant recruitment across ancestry groups historically underrepresented in genomics research to ameliorate genomic health-care disparities.<sup>23</sup>

eMERGE did consider several genes to be returnable beyond those recommended by ACMG. Of these, we returned SFs in *HFE* (associated with hemochromatosis; only in homozygotes for the higher-penetrance p.Cys282Tyr variant<sup>20</sup>) *PALB2* (breast cancer), *KCNE1* (arrhythmia), *HNFA1A* (maturity-onset diabetes of the young), and *MEFV* (familial Mediterranean fever; only in homozygotes or compound heterozygotes for NM\_000243.2:c.2080A>G [p.Met694Val] or NM\_000243.2:c.2177T>C [p.Val726Ala]). The p.Cys282Tyr variant in *HFE* (NM\_001300749.2:c.845G>A) was the third most common SF returned and contributed to the higher rate of SFs in those identifying as White. eMERGE has previously advocated for the return of this genotype as an SF,<sup>18,24</sup> reporting a penetrance rate of hemochromatosis of 24.4% in males and 14.0% in females with this genotype.<sup>18</sup>

The eMERGE III network elected to return LP SFs in this research context; 43.5% of the reported results were LP. One goal of this return was to collect further data on participants who harbor an LP variant (and potentially also their family members) to assist in any potential future reclassification. In our cohort, the proportion of variants classified as LP was

greater among cardiac genes than the proportion in cancer genes. Factors that may account for this difference are (1) lower or unknown penetrance for many cardiac genes; (2) a dramatically higher burden of missense variants among results (47% of cardiac SFs vs. 21% of cancer SFs in our study), which are generally classified with less certainty than putative truncations; and (3) the smaller evidence base available for variant interpretation in cardiac genes, which have generally been studied and reported clinically for a shorter time period and in fewer individuals than cancer genes.<sup>25</sup> While current ACMG/AMP guidelines do not explicitly recommend return of LP variants, they are often reported by clinical labs. We note, however, that even over the course of this study, some variants initially classified as LP (e.g., NM\_000256.3(*MYBPC3*):c.3628-41\_3628-17del25 and NM\_198056.2(*SCN5A*):c.3911C>T [p.Thr1304Met]) were downgraded by one of the SCs to variant of uncertain significance (VUS) or below. This highlights that the true frequency of LP will likely shift over time as new evidence emerges, new terminology for low-risk and low-penetrance loci is adopted, and the interpretation guidelines themselves continue to evolve.

Limitations of the study include smaller sample sizes for non-European ancestry groups, though this cohort is more heterogeneous than previous studies reporting SF frequencies.<sup>6–9,20</sup> Due to sample size, which limited power of interpopulation comparisons, we prespecified testing the frequency in Whites versus all other ancestries combined. As such, differences between Whites and specific non-White groups, even if small, cannot be ruled out using these data. Despite these limitations, to our knowledge, this is the largest study of SF rates in self-identified non-Whites. Though we did not have information on self-reported AJ ancestry, removing known AJ founder alleles from the analysis does not alter our findings that SFs are most likely to be in genes associated with cancer risk, though a previous study found the highest rate to be in genes associated with cardiac disease.<sup>20</sup>

An additional limitation is that we could not independently confirm patient indication for testing via EHR review. It was difficult for a minority of cases recruited in specialty clinics to

# ARTICLE

determine if a finding was truly “secondary.” In some cases, participants were assigned an indication based on the specialty clinic in which they enrolled, but it is possible that some of these participants were family members of those with the condition necessitating the visit to these clinics. Regardless, ongoing detailed EHR review for participants with positive genetics results is currently underway by the eMERGE Outcomes and Clinical Annotation workgroups to determine the spectrum of clinical features in all individuals in which a P or LP variant was identified. Finally, *ATBP7*, associated with Wilson disease, was not fully sequenced but only genotyped for a single known P variant. However, given the rarity of that disorder (estimated as 1/30,000), our estimate of the SF rate for ACMG genes is likely not affected by its omission.

In summary, we evaluated the rate of SFs in a diverse cohort of 21,915 eMERGE III participants. We found a 3.02% overall frequency of SFs, of which 2.54% were in the genes deemed actionable by the ACMG. The most frequent category of SFs in our cohort were in genes associated with cancer susceptibility (1.38%), followed by cardiac diseases (0.87%) and lipid disorders (0.50%). Though the frequency of SFs and proportion of P versus LP SFs were higher in those self-identifying as White than in those identifying as another group, this difference was not significant after removing *HFE* results from the analysis. These important findings serve as a resource to inform decision making in patients and research participants undergoing genomic testing, aid the ongoing development of practice standards and guidelines in genomic medicine, and drive future research efforts in variant interpretation and SF return.

## SUPPLEMENTARY INFORMATION

The online version of this article (<https://doi.org/10.1038/s41436-020-0810-9>) contains supplementary material, which is available to authorized users.

## ACKNOWLEDGEMENTS

The eMERGE Phase III Network was initiated and funded by NHGRI through the following grants: U01HG8657 (Kaiser Permanente Washington/University of Washington); U01HG8685 (Brigham and Women’s Hospital); U01HG8672 (Vanderbilt University Medical Center); U01HG8666 (Cincinnati Children’s Hospital Medical Center); U01HG6379 (Mayo Clinic); U01HG8679 (Geisinger Clinic); U01HG8680 (Columbia University Health Sciences); U01HG8684 (Children’s Hospital of Philadelphia); U01HG8673 (Northwestern University); MD007593 (Meharry Medical College); U01HG8701 (Vanderbilt University Medical Center serving as the Coordinating Center); U01HG8676 (Partners Healthcare/Broad Institute); and U01HG8664 (Baylor College of Medicine).

## DATA ACCESS

All data sets summarized here will be publicly available in the dbGaP repository under phs001616.v1.p1 and pre-dbGaP

submission access can also be requested on the eMERGE website <https://emerge-network.org>.

## DISCLOSURE

D.R.C. serves on a consulting board for UnitedHealth Group with precision medicine efforts, which is unrelated to this paper. C.M.E. is a full-time employee/faculty member of Baylor College of Medicine. Through a professional services agreement, she serves as Chief Medical Officer and Chief Quality Officer of Baylor Genetics. R.A.G. declares that Baylor College of Medicine receives payments from Baylor Genetics Laboratories, which provides services for genetic testing; Baylor College of Medicine is part owner of Codified Genomics. H.L.R. and H.Z. are employed at Massachusetts General Hospital, which receives royalties on sales of GeneInsight Software. E.V. is a cofounder of Codified Genomics. G.L.W. is a member of the External Advisory Panel for the ClinGen Clinical Genome Resource Project. The other authors declare no conflicts of interest.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med.* 2013;15:565–574.
2. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med.* 2017;19:249–255.
3. ACMG Board of Directors. The use of ACMG secondary findings recommendations for general population screening: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2019;21:1467–1468.
4. Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. *Genet Med.* 2008;10:294–300.
5. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–424.
6. Olsson E, Cottrell CE, Davidson NO, et al. Identification of medically actionable secondary findings in the 1000 Genomes. *PLoS One.* 2015;10:e0135193.
7. Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res.* 2015;25:305–315.
8. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA.* 2014;312:1870–1879.
9. Schwartz MLB, McCormick CZ, Lazerri AL, et al. A model for genome-first care: returning secondary genomic findings to participants and their healthcare providers in a large research cohort. *Am J Hum Genet.* 2018;103:328–337.
10. The eMERGE Consortium. Harmonizing clinical sequencing and interpretation for the eMERGE III network. *Am J Hum Genet.* 2019;105:588–605.
11. Dorschner MO, Amendola LM, Turner EH, et al. Actionable, pathogenic incidental findings in 1,000 participants’ exomes. *Am J Hum Genet.* 2013;93:631–640.
12. Gordon A, Rosenthal E, Carrell D, et al. Rates of actionable genetic findings in individuals with colorectal cancer or polyps ascertained from a community medical setting. *Am J Hum Genet.* 2019;105:526–533.

13. Berg JS, Foreman AK, O'daniel JM, et al. A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing. *Genet Med.* 2016;18:467–475.
14. Hunter JE, Irving SA, Biesecker LG, et al. A standardized, evidence-based protocol to assess clinical actionability of genetic disorders associated with genomic variation. *Genet Med.* 2016;18:1258–1268.
15. Fossey R, Kochan D, Winkler E, et al. Ethical considerations related to return of results from genomic medicine projects: the eMERGE Network (Phase III) experience. *J Pers Med.* 2018;8:E2.
16. Tabor HK, Auer PL, Jamal SM, et al. Pathogenic variants for Mendelian and complex traits in exomes of 6,517 European and African Americans: implications for the return of incidental results. *Am J Hum Genet.* 2014;95:183–193.
17. Autore F, Strati P, Laurenti L, Ferrajoli A. Morphological, immunophenotypic, and genetic features of chronic lymphocytic leukemia with trisomy 12: a comprehensive review. *Haematologica.* 2018;103:931–938.
18. Gallego CJ, Burt A, Sundaresan AS, et al. Penetrance of hemochromatosis in HFE genotypes resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE Network. *Am J Hum Genet.* 2015;97:512–520.
19. DellaPergola S. World Jewish population, 2012. In: Dashefsky A, Sheskin I (eds.). *American Jewish year book.* Dordrecht: Springer; 2013:279–358.
20. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med.* 2016;18:696–704.
21. Lohmueller KE, Indap AR, Schmidt S, et al. Proportionally more deleterious genetic variation in European than in African populations. *Nature.* 2008;451:994–997.
22. Fu W, O'connor TD, Jun G, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature.* 2013;493:216–220.
23. Fullerton SM, Knerr S, Burke W. Finding a place for genomics in health disparities research. *Public Health Genomics.* 2012;15:156–163.
24. Fullerton SM, Wolf WA, Brothers KB, et al. Return of individual research results from genome-wide association studies: experience of the Electronic Medical Records and Genomics (eMERGE) Network. *Genet Med.* 2012;14:424–431.
25. Harrison SM, Dolinsky JS, Knight johnson AE, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med.* 2017;19:1096–1104.

### The eMERGE Clinical Annotation Working Group

Adam S. Gordon, PhD<sup>1,2</sup>, Hana Zouk, PhD<sup>3,4</sup>, Eric Venner, PhD<sup>5</sup>, Christine M. Eng, MD<sup>5</sup>, Birgit H. Funke, PhD, FACMG<sup>3</sup>, Laura M. Amendola, MS, CGC<sup>2</sup>, David S. Carrell, PhD<sup>6</sup>, Rex L. Chisholm, PhD<sup>1</sup>, Wendy K. Chung, MD, PhD<sup>7</sup>, Joshua C. Denny, MD, MS<sup>8</sup>, Alexander Fedotov, PhD<sup>9</sup>, Hakon Hakonarson, MD<sup>10</sup>, Iftikhar J. Kullo, MD<sup>11</sup>, Eric B. Larson, MD, MPH<sup>6</sup>, Magalie S. Leduc, PhD<sup>5</sup>, Kathleen A. Leppig, MD<sup>12</sup>, Niall J. Lennon, PhD<sup>13</sup>, Jodell E. Linder, PhD<sup>8</sup>, Donna M. Muzny, MS<sup>5</sup>, Cynthia A. Prows, MSN, APRN<sup>14</sup>, Laura J. Rasmussen-Torvik, PhD<sup>15</sup>, Hila Milo Rasouly, PhD<sup>7</sup>, Dan M. Roden, MD<sup>16</sup>, Elisabeth A. Rosenthal, PhD<sup>2</sup>, Maureen E. Smith, MS<sup>1</sup>, Ian B. Stanaway, PhD<sup>17</sup>, Sara L. Van Driest, MD, PhD<sup>18</sup>, Kimberly Walker, MS<sup>5</sup>, Georgia L. Wiesner, MD<sup>19</sup>, Marc S. Williams, MD<sup>20</sup>, Leora Witkowski, PhD<sup>3</sup>, David R. Crosslin, PhD<sup>17</sup>, Richard A. Gibbs, PhD<sup>5</sup>, Heidi L. Rehm, PhD<sup>3,13</sup> and Gail P. Jarvik, MD, PhD<sup>2</sup>

<sup>1</sup>Center for Genetic Medicine, Northwestern University, Chicago, IL, USA; <sup>2</sup>Medical Genetics, University of Washington, Seattle, WA, USA; <sup>3</sup>Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, MA, USA; <sup>4</sup>Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; <sup>5</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; <sup>6</sup>Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA; <sup>7</sup>Department of Pediatrics and Medicine, Columbia University Medical Center, New York, NY, USA; <sup>8</sup>Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, TN, USA; <sup>9</sup>Irving Institute for Clinical and Translational Research, Columbia University, New York, NY, USA; <sup>10</sup>The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA; <sup>11</sup>Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA; <sup>12</sup>Genetic Services, Kaiser Permanente of Washington, Seattle, WA, USA; <sup>13</sup>The Broad Institute of MIT and Harvard, Cambridge, MA, USA; <sup>14</sup>Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; <sup>15</sup>Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; <sup>16</sup>Departments of Medicine, Pharmacology, and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA; <sup>17</sup>Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA; <sup>18</sup>Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN, USA; <sup>19</sup>Departments of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; <sup>20</sup>Genomic Medicine Institute, Geisinger, Danville, PA, USA. A full list of members and their affiliations appears in the Supplementary Information.