

# Clinical genomics can facilitate countrywide estimation of autosomal recessive disease burden

Mohamed Abouelhoda<sup>1,2</sup>, Turki Sobahy<sup>1,2</sup>, Mohamed El-Kalioby<sup>1,2</sup>, Nisha Patel<sup>1</sup>, Hanan Shamseldin<sup>1</sup>, Dorota Monies<sup>1,2</sup>, Nada Al-Tassan<sup>1,2</sup>, Khushnooda Ramzan<sup>1</sup>, Faiqa Imtiaz<sup>1</sup>, Ranad Shaheen<sup>1</sup> and Fowzan S. Alkuraya<sup>1-3</sup>

**Background:** Most autosomal recessive diseases are rare, but they collectively account for a substantial proportion of disease burden, especially in consanguineous populations. Estimation of this disease burden, however, is hampered by many factors, including lack of countrywide registries. Establishing carrier frequency can be a practical surrogate to estimate disease burden, although the requirement of a large representative cohort may be challenging.

**Purpose:** We propose that the application of clinical genomics in the diagnostic setting offers a unique opportunity to estimate carrier frequency in the population as a secondary benefit.

**Methods:** We used a data set of ~7,100 patients who underwent genomic testing for various Mendelian disorders to estimate the carrier frequency.

**Results:** We were able to calculate the frequency of 259 confirmed founder recessive mutations. We found the corresponding disease burden to be, at minimum, ~7 per 1,000 children born to first-cousin parents, with disorders related to intellectual disability and vision impairment being the most common.

**Conclusion:** Our approach can be utilized to inform the design of new policies for the prevention of genetic disorders and highlights an important secondary benefit of clinical genomics.

*Genet Med* advance online publication 28 April 2016

**Key Words:** carrier frequency; consanguinity; disease burden; genomics; Mendeliome

## INTRODUCTION

Diseases that can be attributed to mutations in single genes are often referred to as Mendelian, and their phenotypic spectrum essentially spans all structural and functional aspects of the human body.<sup>1,2</sup> Despite the potential of these diseases to reveal the medical relevance of individual human genes, their rarity makes them less familiar to clinical practice and more difficult to study epidemiologically compared to their common multifactorial counterparts.

The rarity of Mendelian diseases is the consequence of a strong negative selection that operates on causative alleles because many result in phenotypes that are detrimental to the reproductive success of the affected individuals.<sup>3</sup> In the case of dominant disorders, the negative selection operates conspicuously; i.e., the defective allele is eliminated by the failure of the patient to reproduce. Dominant lethal alleles are thereby replenished through their introduction anew in the genetic pool, i.e., de novo. Recessive alleles, however, tend to persist in the genetic pool even when they result in lethal phenotypes because carriers typically have normal reproductive health.<sup>4</sup>

Although individual Mendelian diseases tend to be rare, some can achieve a relatively high frequency.<sup>5</sup> This may be due to an unusual “heterozygotes advantage” that endows health benefits to carriers, leading to reproductive success, as in the classic case of carriers of *HBB* mutations and resistance of malaria.<sup>6</sup> Founder effect is another major factor that can inflate the frequency of recessive diseases because it increases the proportion of carriers; classic examples include Finnish and Ashkenazi Jewish diseases.<sup>7,8</sup> Consanguinity is a special case of founder effect whereby the carrier frequency is not increased per se but the very limited genetic pool within any given family increases the probability of a mate to be a carrier of the same recessive allele.<sup>9</sup> It should be noted that, although individually rare, Mendelian diseases collectively account for significant morbidity and mortality, especially in the pediatric population.<sup>10</sup>

Countrywide estimates of burden of diseases—an essential metric that informs public health policies—are routinely obtained for a wide variety of disease conditions.<sup>11</sup> Although Mendelian diseases are not rare in aggregate, estimating their disease burden at a national level is cumbersome and faces

<sup>1</sup>Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia; <sup>2</sup>Saudi Human Genome Project, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia; <sup>3</sup>Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh, Saudi Arabia. Correspondence: Fowzan S. Alkuraya (falkuraya@kfshrc.edu.sa)

Submitted 11 December 2015; accepted 10 February 2016; advance online publication 28 April 2016. doi:10.1038/gim.2016.37

many challenges. This is particularly concerning when one considers that most Mendelian diseases are incurable and many rank highly in clinical severity, thus implying prevention as the best health-care policy to combat them.<sup>12</sup> In the absence of reasonable estimates of the overall burden that these disorders pose to a country's state of health, well-informed prevention strategies will be difficult to design. A major obstacle in estimating disease burden for Mendelian diseases is diagnosis, not only because they are unfamiliar to many physicians but also because their presentation can vary greatly such that they may not be diagnosed clinically, even by trained medical geneticists.<sup>13</sup> Although this can be mitigated by adopting genomic molecular approaches that are agnostic to clinical judgment,<sup>14</sup> the logistic challenge of a national registry that logs all molecularly diagnosed cases can be formidable.

In this study, we show that estimating the disease burden for recessive Mendelian diseases can be informed by clinical genomics, a growing trend in medical genetics whereby an individual patient is tested for all or many genes at once.<sup>15</sup> In a large data set of exomes and gene panels performed on 7,101 patients from all regions of Saudi Arabia, we were able to calculate the carrier frequency for confirmed pathogenic pediatric-onset recessive conditions. The overall disease burden inferred from this set of data reveals surprising patterns that we hope will inform policymakers locally and encourage others to explore this promising application of clinical genomics.

## MATERIALS AND METHODS

As part of our ongoing effort to characterize autosomal recessive diseases in Saudi Arabia, we have genomically analyzed 7,101 patients: 1,549 by whole-exome sequencing and the rest by targeted gene panel sequencing. The assignment of a panel to a patient was based on the phenotype and diagnosis, as explained before in our report of the "Mendeliome" study.<sup>15</sup> It is important to mention that in this study we used more samples than in the Mendeliome study.

The sequencing work was carried out using the setup of the Saudi Genome Project based on Ion Torrent sequencing technology. The content and design of the multigene panels were described previously,<sup>15</sup> as was the commercial Inherited Disease Panel of LifeTech, together comprising 3,085 unique known disease genes across all panels. Whole-exome sequencing was performed using primer-based sequencing, targeting all the genes in the genome (approximately 25,000 genes and open-reading frames). The sequencing and data analysis procedures are the same as described previously.<sup>15</sup>

Our analysis of this large number of patients with suspected Mendelian diseases revealed a large number of likely causal variants, but we chose a subset of 618 autosomal recessive variants from 357 genes that can be assigned pathogenic status with the highest confidence for subsequent analysis (**Supplementary Tables S1 and S2** online). **Supplementary Table S3** online includes gene-per-gene sequencing quality information in terms of depth and coverage. The overall average coverage of all genes is 98%, and the overall average depth is 701×

Obviously, the homozygous or compound heterozygous state of these alleles is highly biased in our cohort and cannot be used to directly estimate disease frequency. However, we reasoned that their presence in the heterozygous state is unrelated to the disease state of the tested individual. For example, the probability of being a carrier for a hemoglobinopathy mutation is not influenced by being homozygous for a Bardet-Biedl disease mutation. Therefore, we screened our cohort for these 618 variants but counted their occurrence only in the heterozygous state (excluding compound heterozygosity). The minimal number of samples screened for carrier status per variant is 1,549 (1,549 exomes + 0 multigene panels), the maximum is 4,975 (1,549 exomes + 3,426 multigene panels), and the average was 2,322 samples tested per variant. The number of samples screened for each variant is shown in **Supplementary Tables S1 and S2** online.

The carrier frequency for variant  $X$   $CF(X)$  is computed according to the following equation:

$$CF(X) = \frac{\text{No. occurrences of } X \text{ in heterozygous state}}{(\text{No. screened samples for this variant} - \text{No. occurrences of } X \text{ in homozygous state})}$$

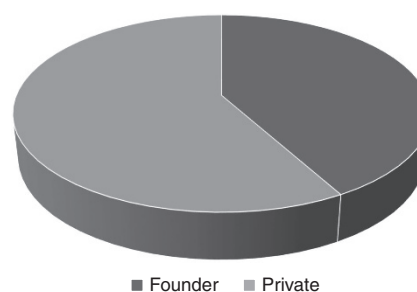
## RESULTS

### Founder mutations account for a minority of recessive mutations in Saudi Arabia

We have previously suggested that the remarkable genetic and allelic heterogeneity we observe in Saudi Arabia is due, at least in part, to the effect of consanguinity, which theoretically can render de novo recessive mutations homozygous in a span of three generations only. In other words, consanguinity can inflate the contribution of young mutations to the mutational spectrum of autosomal recessive diseases.<sup>16</sup> Consistent with this, we noticed that 58% (359/618) of the disease-causing mutations encountered did not appear to have a founder effect beyond the family in which they were identified, as inferred by their absence in the heterozygous state in any of more than 7,000 samples tested; i.e., they were "private" mutations (**Figure 1** and **Supplementary Table S2** online).

### Wide range of carrier frequencies of founder mutations

We identified 259 autosomal recessive mutations for which at least one carrier was encountered in our cohort of more



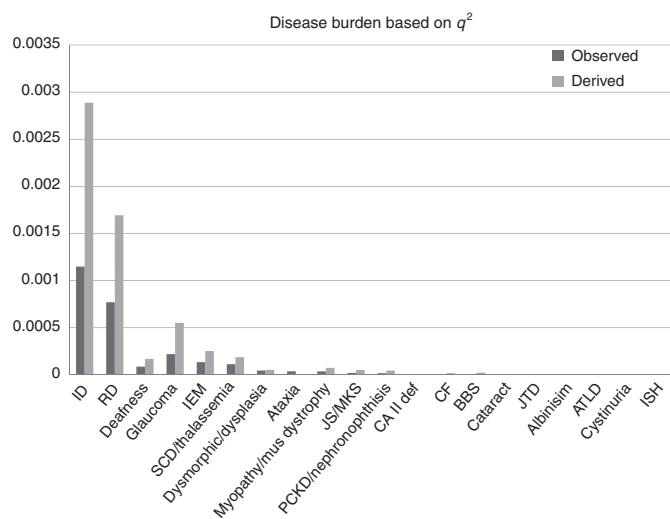
**Figure 1** Relative distribution of founder and private variants.

**Table 1** Estimated burden of disease categories based on derived and observed combined carrier frequency

Disease	Combined carrier frequency		Disease burden based on $q^2$		Disease burden based on $qf$	
	Observed	Derived	Observed	Derived	Observed	Derived
Albinism	0.00258	0.00224	$1.66708 \times 10^6$	$1.25842 \times 10^6$	$3.11168 \times 10^5$	$2.70353 \times 10^5$
Alport	0.00065	0.00156	$1.04193 \times 10^7$	$6.07731 \times 10^7$	$7.77921 \times 10^6$	$1.87877 \times 10^5$
Arthropathy	0.00129	0.00065	$4.16771 \times 10^7$	$1.04462 \times 10^7$	$1.55584 \times 10^5$	$7.78927 \times 10^6$
Ataxia	0.01227	0.00139	$3.76136 \times 10^5$	$4.85376 \times 10^7$	0.000147805	$1.67902 \times 10^5$
ATLD	0.00258	0.00204	$1.66708 \times 10^6$	$1.0455 \times 10^6$	$3.11168 \times 10^5$	$2.46421 \times 10^5$
BBS	0.00387	0.00958	$3.75094 \times 10^6$	$2.29267 \times 10^5$	$4.66753 \times 10^5$	0.000115395
CA II deficiency	0.00516	0.00303	$6.66833 \times 10^6$	$2.28937 \times 10^6$	$6.22337 \times 10^5$	$3.64649 \times 10^5$
Cataract	0.00387	0.00215	$3.75094 \times 10^6$	$1.15642 \times 10^6$	$4.66753 \times 10^5$	$2.59164 \times 10^5$
CF	0.00452	0.00858	$5.10544 \times 10^6$	$1.83866 \times 10^5$	$5.44545 \times 10^5$	0.00010334
Corneal dystrophy	0.00065	0.00210	$1.04193 \times 10^7$	$1.1051 \times 10^6$	$7.77921 \times 10^6$	$2.53348 \times 10^5$
Cystinuria	0.00258	0.00377	$1.66708 \times 10^6$	$3.55615 \times 10^6$	$3.11168 \times 10^5$	$4.54472 \times 10^5$
DBSA	0.00065	0.00085	$1.04193 \times 10^7$	$1.80463 \times 10^7$	$7.77921 \times 10^6$	$1.02379 \times 10^5$
Deafness	0.0187	0.026	0.0000874225	0.000169	0.000225335	0.0003133
Dysmorphic/dysplasia	0.01356	0.01454	$4.5949 \times 10^5$	$5.28274 \times 10^5$	0.000163363	0.000175165
FRAM	0.00194	0.00365	$9.37734 \times 10^7$	$3.33239 \times 10^6$	$2.33376 \times 10^5$	$4.39942 \times 10^5$
Genodermatosis	0.00194	0.00357	$9.37734 \times 10^7$	$3.18817 \times 10^6$	$2.33376 \times 10^5$	$4.30316 \times 10^5$
Glaucoma	0.02970	0.04697	0.000220472	0.000551545	0.000357844	0.000565989
GrisCELLI syndrome	NA	0.00231	NA	$1.3285 \times 10^6$	NA	$2.77778 \times 10^5$
HSAN2A	0.00194	0.00229	$9.37734 \times 10^7$	$1.30712 \times 10^6$	$2.33376 \times 10^5$	$2.75534 \times 10^5$
ID	0.06779	0.10746	0.001148724	0.002886843	0.000816817	0.001294877
IEM	0.02324	0.03182	0.000135034	0.0002531	0.000280052	0.00038341
ISH	0.00258	0.00397	$1.66708 \times 10^6$	$3.93726 \times 10^6$	$3.11168 \times 10^5$	$4.78205 \times 10^5$
Jaundice	0.00258	0.00354	$1.66708 \times 10^6$	$3.12823 \times 10^6$	$3.11168 \times 10^5$	$4.26252 \times 10^5$
JTD	0.00323	0.00238	$2.60482 \times 10^6$	$1.41051 \times 10^6$	$3.88961 \times 10^5$	$2.86223 \times 10^5$
Mitochondrial disorder	0.00194	0.00378	$9.37734 \times 10^7$	$3.57127 \times 10^6$	$2.33376 \times 10^5$	$4.55437 \times 10^5$
JS/MKS	0.00904	0.01435	$2.04218 \times 10^5$	$5.1452 \times 10^5$	0.000108909	0.000172869
Myopathy/muscular dystrophy	0.01227	0.01722	$3.76136 \times 10^5$	$7.41658 \times 10^5$	0.000147805	0.000207548
Myopia	0.00065	0.00065	$1.04193 \times 10^7$	$1.04462 \times 10^7$	$7.77921 \times 10^6$	$7.78927 \times 10^6$
PCKD/nephronophthisis	0.00839	0.01344	$1.76086 \times 10^5$	$4.51385 \times 10^5$	0.00010113	0.000161916
Nephrotic syndrome	0.00065	0.00154	$1.04193 \times 10^7$	$5.9661 \times 10^7$	$7.77921 \times 10^6$	$1.8615 \times 10^5$
PALS	0.00258	0.00303	$1.66708 \times 10^6$	$2.29013 \times 10^6$	$3.11168 \times 10^5$	$3.64709 \times 10^5$
PCD	0.00129	0.00065	$4.16771 \times 10^7$	$1.04462 \times 10^7$	$1.55584 \times 10^5$	$7.78927 \times 10^6$
Postmicrophthalmos	0.00129	0.00171	$4.16771 \times 10^7$	$7.34274 \times 10^7$	$1.55584 \times 10^5$	$2.06512 \times 10^5$
RD	0.05552	0.08231	0.000770609	0.001693693	0.000669012	0.000991824
Recurrent fetal loss	0.00129	0.00129	$4.16771 \times 10^7$	$4.18932 \times 10^7$	$1.55584 \times 10^5$	$1.55987 \times 10^5$
SSS	NA	0.00427	NA	$4.55862 \times 10^6$	NA	$5.14557 \times 10^5$
Temtamy syndrome	0.00129	0.00217	$4.16771 \times 10^7$	$1.17708 \times 10^6$	$1.55584 \times 10^5$	$2.61469 \times 10^5$
SCD/thalassemia	0.02130	0.02740	0.000113466	0.000187724	0.000256714	0.0003302
Tufting enteropathy	NA	0.00239	NA	$1.43083 \times 10^6$	NA	$2.88278 \times 10^5$
Wilson disease	NA	0.00157	NA	$6.18701 \times 10^7$	NA	$1.89565 \times 10^5$
PID	0.00129	0.00996	$4.16771 \times 10^7$	$2.4806 \times 10^5$	$1.55584 \times 10^5$	0.000120031
Renal agenesis/hypoplasia/dysplasia	0.00174	0.00174	$7.56144 \times 10^7$	$7.56144 \times 10^7$	$2.09565 \times 10^5$	$2.09565 \times 10^5$

Derived combined carrier frequency refers to the simple addition of the carrier frequency of each mutation for a given disease category. Because this method may inflate the combined carrier frequency because an individual can be a carrier for more than one mutation, we directly queried whole exomes to calculate the "observed" combined carrier frequency.

ATLD, ataxia telangiectasia-like disorder; BBS, Bardet-Biedl syndrome; CA II, carbonic anhydrase; CF, cystic fibrosis; DBSA, Diamond-Bodian-Shwachman syndrome; FRAM, familial retina artery macroaneurysm; HSAN2A, hereditary sensory and autonomic neuropathy type IIA; ID, intellectual disability; IEM, inborn errors of metabolism (organic acidemias, amino acidopathies, carbohydrate metabolic disorders, lysosomal storage disorders, and peroxisomal diseases); ISH, infantile systemic hyalinosis; JS/MKS, Joubert syndrome/Meckel Gruber syndrome; JTD, Jeune asphyxiating thoracic dystrophy; NA, not applicable; PALS, Papillon-Lefevre syndrome; PCD, primary ciliary dyskinesia; PCKD, polycystic kidney disease; PID, primary immunodeficiency; RD, retinal dystrophy; SCD, sickle cell disease; SSS, Sanjad-Sakati syndrome.



**Figure 2** Top 10 diseases with respect to their burden based on  $q^2$ . For abbreviation key, see the [Table 1](#) footnote

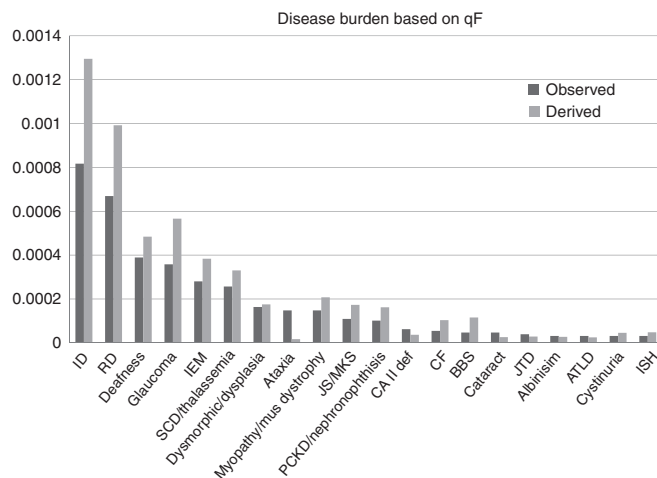
than 7,000 unrelated patients (excluding parents and other relatives), which enabled us to calculate the corresponding carrier frequency ([Supplementary Table S1](#) online). [Supplementary Table S3](#) online shows the combined carrier frequency for each screened gene. The highest carrier frequency for a single mutation was 0.0218 for the known founder mutation *CYP1B1*:NM\_000104.3:c.1103G>A:p.Arg368His, which causes congenital glaucoma, a highly endemic disease in Saudi Arabia, albeit with reduced penetrance.<sup>17</sup> The combined carrier frequency for the three founder mutations in *CYP1B1* was 0.047. The lowest carrier frequency was 0.0002 for ADA:NM\_000022.2:c.385G>A:p.Val129Met, which causes severe combined immunodeficiency due to lack of adenosine deaminase. However, the combined carrier frequency for all immunodeficiency founder mutations was high, approaching 0.01.

To calculate the overall carrier frequency for all recessive mutations without resorting to computationally challenging combinatorial probability statistics, we queried the exomes of 1,549 unrelated patients with various Mendelian disorders and ignored the homozygous calls to avoid bias (see Materials and Methods). The probability of being a carrier for at least one of the founder mutations was approximately 0.214 (331/1549) (0.022 for at least two and 0.002 for at least three mutations).

### High burden of recessive diseases in Arabia

Although Saudi Arabia is known to be enriched for recessive diseases due to high rates of consanguinity, it has not been possible to quantify this.<sup>18</sup> Using the Hardy–Weinberg equation, we calculated the disease frequency ( $q^2$ ) based on the carrier frequency ( $2pq$ ). According to this method, we found that the most common group of diseases comprises those that can present with psychomotor delay/intellectual disability (3 per 1,000), followed by retinal dystrophies (1.7 per 1,000) and congenital glaucoma (1 per 4,500). The estimated burden of the disease categories is

## ORIGINAL RESEARCH ARTICLE



**Figure 3** Top 10 diseases with respect to their burden based on  $qF$ . For abbreviation key, see the [Table 1](#) footnote.

shown in [Table 1](#) and [Figure 2](#). However, the Hardy–Weinberg equation assumes independent mating, a condition clearly violated by the >50% rate of consanguinity in Saudi Arabia.<sup>18</sup> In the setting of consanguinity, the probability of the mutant allele  $q$  to coexist with another mutant allele is not  $qXq$  ( $q^2$ ) but  $qXF$ , where  $F$  is the inbreeding coefficient. For example, although the carrier frequency of sickle cell disease (0.008) predicts a disease frequency of 1 per 62,500, the high inbreeding coefficient in Saudi Arabia (estimated at 0.0241 (ref. 19)) puts the disease frequency at 1 per 10,300 ([Figure 3](#)). Because the latter is much closer to the estimated disease prevalence based on a previously published community-based study,<sup>20</sup> we opted to present the disease burden based on  $q^2$  and  $qF$  ([Table 1](#)). To provide a regularly updated estimate of disease burden based on newly identified mutations, a freely available database was established that can be accessed at <http://shgp.kacst.edu.sa/dbm/>.

## DISCUSSION

Genomic sequencing (whole-genome, whole-exome, and gene panels) is a new trend in diagnostics that was driven by the need to develop diagnostic tools that can overcome the limited sensitivity and specificity of clinically based diagnosis of genetic diseases.<sup>21</sup> We have previously published the results of applying this approach to more than 2,300 patients with various suspected Mendelian disorders.<sup>15</sup> Through that study and others, we were able to compile a large list of point mutations, the majority of which are recessive, consistent with the predicted effect of the high consanguinity rate in Saudi Arabia. Although these samples come from all regions in Saudi Arabia, they could not be used to estimate the disease frequency directly because of the obvious bias in sampling. However, we reasoned that each of these unrelated patients can contribute to the estimate of carrier frequency of mutations other than their causative mutations. Using this approach, we calculated the first countrywide estimate of disease burden of recessive diseases in the Middle East without resorting to independent genomic sequencing of healthy controls.

The overall disease-burden estimate we present in this work is a minimal estimate for several reasons. First, our estimate does not take into account the regional variation in carrier frequency. One good example is sickle cell disease, for which previous studies have shown a remarkably wide range of carrier frequency across Saudi provinces, from 0 in the north to 0.25 in the eastern province, where the disease is endemic.<sup>20,22</sup> Second, we relied on founder mutations to calculate carrier frequency, which account for only a minority of pathogenic mutations (most of the mutations we identified are private). Novel disease-causing mutations that have yet to be identified are obviously not included in our estimate. Finally, mutations that are not readily picked up by next-generation sequencing were not included (e.g., the common founder mutation *GORAB* NM\_152281.2:c.306dupA:p.P103Tfs\*20 (ref. 23) was consistently missed by the next-generation sequencing platform used in this study).

This limitation also applies to large deletions/duplications and deep intronic mutations. Perhaps a good demonstration of the conservative nature of our estimate can be gleaned from local experience at the national newborn screening program at KFSHRC, where 1 out of 1,043 newborns is confirmed positive for 1 of the 17 tested conditions (phenylketonuria, maple syrup urine disease, argininosuccinase deficiency, citrullinemia, HMG-CoA lyase deficiency, isovaleric acidemia, methylmalonic acidemia, propionic acidemia,  $\beta$ -ketothiolase deficiency, methylcrotonyl-CoA carboxylase deficiency, glutaric acidemia type I, medium-chain acyl-CoA dehydrogenase deficiency 13, very-long-chain acyl-CoA dehydrogenase deficiency, congenital hypothyroidism, congenital adrenal hyperplasia, galactosemia, and biotinidase deficiency) (Alfadhel M and Al-Odaib A, personal communication) compared with the 1 out of 3,951 predicted by our conservative  $q^2$  estimate, or even compared with the less conservative 1 out of 2,608  $qF$  estimate.

Even with this minimal estimate, the figure obtained for disease burden is alarmingly high. With the probability of being a carrier for any of the founder pathogenic mutations at 0.214, we can estimate that the probability of a first-cousin union conceiving a child homozygous for the corresponding mutation is ~0.7% (6.7 per 1,000), which in itself is an underestimate given the large degree of inbreeding; i.e.,  $F$  is higher than the expected 0.0625 for first-cousin unions. This raises several issues that are of public health concern. For instance, Saudi Arabia has implemented two public health initiatives for primary and secondary prevention of genetic diseases. The first is a premarital screening program that focuses on hemoglobinopathies; the other is a newborn screening program for 17 inborn errors of metabolism.<sup>24,25</sup> Although these diseases are relatively common, the disease-burden estimate we calculate shows that other disease categories are much more common, such as autosomal recessive intellectual disability. With the increasing interest in early diagnosis of pediatric-onset diseases as a justification for inclusion in screening programs even when effective therapy is lacking,<sup>26</sup> a neonatal screening program based on these established pathogenic mutations is an appealing choice. Such genotyping-based screening will not only be less expensive but also circumvent many other

controversies surrounding neonatal genomic sequencing, such as variants of unknown significance and incidental findings.<sup>27</sup> Similarly, the premarital screening program, its controversial mandatory statute notwithstanding, that is currently based on hemoglobin electrophoresis can benefit greatly from switching to a genotyping-based platform. Again, the established pathogenic nature of the mutations we report will greatly facilitate genetic counseling surrounding this test. For comparison, the probability among Ashkenazi Jews to be a carrier for one of the “Ashkenazi Jewish mutations” is 0.3, and such a high carrier frequency was the impetus for many preventive screening programs.<sup>28,29</sup>

Our work highlights an interesting benefit of performing large-scale clinical genomics in which each tested patient serves as a source of genetic variation in the general population and carrier frequency. We hope the data we present from the largest Middle Eastern country can inform the design of next-generation prevention strategies for genetic diseases locally and beyond.

## SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

## ACKNOWLEDGMENT

This study was funded through the Saudi Human Genome Program.

## DISCLOSURE

The authors declare no conflict of interest.

## REFERENCES

- Antonarakis SE, Beckmann JS. Mendelian disorders deserve more attention. *Nat Rev Genet* 2006;7:277–282.
- McKusick VA. The morbid anatomy of the human genome: the role of gene mapping in clinical medicine. *Pediatr Int* 1990;32:234–241.
- Bustamante CD, Fledel-Alon A, Williamson S, et al. Natural selection on protein-coding genes in the human genome. *Nature* 2005;437:1153–1157.
- Alkuraya FS. Human knockout research: new horizons and opportunities. *Trends Genet* 2015;31:108–115.
- Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502–510.
- Kwiatkowski DP. How malaria has affected the human genome and what human genetics can teach us about malaria. *Am J Hum Genet* 2005;77:171–192.
- Kere J. Human population genetics: lessons from Finland. *Annu Rev Genomics Hum Genet* 2001;2:103–128.
- Ostrer H. A genetic profile of contemporary Jewish populations. *Nat Rev Genet* 2001;2:891–898.
- Alkuraya FS. Autozygome decoded. *Genet Med* 2010;12:765–771.
- Newman WG, Black GC. Diagnosis of rare inherited diseases. *Orphan Drugs Rare Dis* 2014;38:35.
- Kivimäki M, Vineis P, Brunner EJ. How can we reduce the global burden of disease? *Lancet* 2015;386:2235–2237.
- Alkuraya FS. Impact of new genomic tools on the practice of clinical genetics in consanguineous populations: the Saudi experience. *Clin Genet* 2013;84:203–208.
- Alazami AM, Patel N, Shamseldin HE, et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep* 2015;10:148–161.
- Shaheen R, Patel N, Shamseldin H, et al. Accelerating matchmaking of novel dysmorphology syndromes through clinical and genomic characterization of a large cohort. *Genet Med* 2015; e-pub ahead of print 3 December 2015.
- Saudi Mendeliome Group. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol* 2015;16:134.

16. Aldahmesh MA, Abu-Safieh L, Khan AO, et al. Allelic heterogeneity in inbred populations: the Saudi experience with Alström syndrome as an illustrative example. *Am J Med Genet A* 2009;149A:662–665.
17. Bejjani BA, Stockton DW, Lewis RA, et al. Multiple CYP1B1 mutations and incomplete penetrance in an inbred population segregating primary congenital glaucoma suggest frequent de novo events and a dominant modifier locus. *Hum Mol Genet* 2000;9:367–374.
18. Alkuraya FS. Genetics and genomic medicine in Saudi Arabia. *Mol Genet Genomic Med* 2014;2:369–378.
19. el-Hazmi MA, al-Swailem AR, Warsy AS, al-Swailem AM, Sulaimani R, al-Meshari AA. Consanguinity among the Saudi Arabian population. *J Med Genet* 1995;32:623–626.
20. Al-Qurashi MM, El-Mouzan MI, Al-Herbish AS, Al-Salloum AA, Al-Omar AA. The prevalence of sickle cell disease in Saudi children and adolescents. A community-based survey. *Saudi Med J* 2008;29:1480–1483.
21. Yavarna T, Al-Dewik N, Al-Mureikhi M, et al. High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. *Hum Genet* 2015;134:967–980.
22. Lehmann H, Maranjian G, Mourant AE. Distribution of sickle-cell hemoglobin in Saudi Arabia. *Nature* 1963;198:492–493.
23. Al-Dosari M, Alkuraya FS. A novel missense mutation in SCYL1BP1 produces geroderma osteodysplastica phenotype indistinguishable from that caused by nullimorphic mutations. *Am J Med Genet A* 2009;149:2093–2098.
24. Alhamdan NA, Almazrou YY, Alswaidi FM, Choudhry AJ. Premarital screening for thalassemia and sickle cell disease in Saudi Arabia. *Genet Med* 2007;9:372–377.
25. Saadallah AA, Rashed MS. Newborn screening: experiences in the Middle East and North Africa. *J Inherit Metab Dis* 2007;30:482–489.
26. Tarini BA, Goldenberg AJ. Ethical issues with newborn screening in the genomics era. *Annu Rev Genomics Hum Genet* 2012;13:381–393.
27. ACMG Board of Directors. Points to consider in the clinical application of genomic sequencing. *Genet Med* 2012;14:759–761.
28. Levenson D. Expanded Ashkenazi Jewish prenatal testing panel well received. *Am J Med Genet A* 2010;152A:fmvii–fmviii.
29. Scott SA, Edelman L, Liu L, Luo M, Desnick RJ, Kornreich R. Experience with carrier screening and prenatal diagnosis for 16 Ashkenazi Jewish genetic diseases. *Hum Mutat* 2010;31:1240–1250.