A Universal Carrier Test for the Long Tail of Mendelian Disease

Balaji S. Srinivasan[1,2,3,†], Jason Flannick[1], A. Scott Patterson[1], Christopher C. Chang[1,4], Tuan Pham[1], Sharon Young[1], Amit Kaushal[5,6,7], James Lee[8], Pasquale Patrizio[9], Eric A. Evans[1,†,*]

1 Counsyl, Redwood Shores, California
2 Department of Computer Science, Stanford University, Stanford, California
3 Department of Statistics, Stanford University, Stanford, California
4 Department of Mathematics, UCSD, La Jolla, California
5 Stanford Medical School, Stanford University, Stanford, California
6 Stanford Genome Technology Center, Stanford University, Stanford, California
7 Program in Biomedical Informatics, Stanford University, Stanford, California
8 The Personal Genome Project, Harvard University, Cambridge, Massachussetts
9 Yale University Fertility Center, Yale University, New Haven, Connecticut
† These authors contributed equally to this work.
*** Address correspondence to: eaevans@counsyl.com**

Abstract

Background

Mendelian disorders are individually rare but collectively common, forming a "long tail" of genetic disease. More than 20 million people worldwide suffer from a disease in this long tail before the age of 25, with minorities and developing countries at highest risk and with the number of carriers far in excess of this figure. Importantly, the Jewish community's campaign for universal Tay-Sachs screening shows that these incurable diseases can nevertheless be prevented if carrier status is known before conception. A single highly-accurate assay for the long tail of Mendelian disease would allow us to scale this successful campaign up to the general population, thereby improving millions of lives, greatly benefiting minority health, and saving billions of dollars.

Methods and Findings

We have addressed the need for such an assay by designing the Universal Genetic Test (UNIT), a non-invasive, saliva-based carrier test for more than 100 Mendelian diseases across all major population groups. We exhaustively validated the test with a median of 147 positive and 525 negative samples per variant. By combining probes for risk alleles with family history information, we show that we can achieve extremely high levels of accuracy (median 95% CI [0.99988, 0.999999]), precision (median 95% CI [0.99993, 0.99999]), sensitivity (median 95% CI [0.99988, 0.999999]), and specificity (median 95% CI [0.99643, 1]) at the level of individual mutations. In particular, through a combination of replicated probes and confirmatory testing, we are able to reliably detect rare alleles at $q \approx 1/1000$ with positive predictive values above 0.995. To put this in context, this performance for a multiplex assay compares favorably with FDA-approved single-gene carrier tests.

Conclusions

The UNIT represents a dramatic reduction in the cost and complexity of large scale population screening. With a single inexpensive assay for a substantial fraction of the global Mendelian disease burden, an end to many preventable genetic diseases is now in sight. Moreover, given that the assay requires only a saliva sample, it is for the first time feasible to contemplate an "at-home carrier test" as a successor to the at-home pregnancy test.

Author Summary

Single-gene or "Mendelian" disorders affect more than one in 280 births, causing 10% of pediatric admissions, costing billions of dollars per year, and claiming a disproportionate number of minority lives. However, the Jewish community's successful campaign against Tay-Sachs disease demonstrates that these incurable diseases are nevertheless preventable if at-risk couples can be identified before pregnancy.

The problem is that traditional genetic tests are not scalable, accurate, or affordable enough to test the entire population. This is because single-gene disorders are individually rare but collectively frequent, forming a "long tail" of genetic disease.

To solve this problem, we developed and validated a clinical diagnostic-quality genetic test for more than 100 Mendelian conditions. This assay can be manufactured on a large scale, is applicable to every population group, and requires only a saliva sample.

Importantly, the use of saliva rather than blood makes it feasible for the first time to contemplate an "at-home carrier test" as a successor to the at-home pregnancy test, to allow for universal and rapid deployment. Widespread use of this assay by all adults before conception is thus conceivable. Such use represents a scaling up of the successful campaign against Tay-Sachs, and would promise an end to many preventable genetic diseases.

Introduction

Mendelian disease imposes a significant public health burden [1, 2] on society, with single-gene disorders accounting for at least 10% of pediatric admissions [3,4] and 20% of infant mortality [5]. Over 6,000 genetic disorders, each of which affect less than 200,000 Americans, combine to afflict 25-30 million people [6]. Because of this heterogeneity, diagnosis and treatment is difficult for the majority of individuals with a genetic disease [7, 8].

The scale of the issue can be appreciated by multiplying the North American Mendelian disease incidence [9] of 1 in 280 births by the consequent medical expenditure [10–16] of $100,000 to several million per child. The result is an average cost to the US healthcare system of at least $360 per birth, a sum which is particularly staggering in light of the relatively large [17, 18] body of knowledge about Mendelian disease.

Even this figure tends to understate the impact of Mendelian disease, as minority groups and inhabitants of developing countries have greater risks [19–23] that are not well described by average costs in the general population. For example, African Americans are far more likely to develop sickle cell anemia [24,25], Asian Americans account for the majority of thalassemia cases in America [26–28], and more than one in four members of the Jewish community possess a recessive mutation for a known Mendelian disease [29–32]. Developing countries with high rates of consanguinity or endogamous marriage traditions [23] are likewise disproportionately affected. Despite these statistics, genetic test development for minority diseases has lagged compared to that for Caucasians, in part because minorities are underrepresented in both genetic and clinical research [8, 33, 34].

This continuing impact of Mendelian disease is troubling because the conditions are preventable (Figure 1) given preconception carrier testing. Couples who test positive as carriers have several options to conceive a child without a lethal disease [35], such as preimplantation genetic diagnosis (PGD) [36–40] or donor gametes [41, 42]. With forewarning for a positive test result, couples might choose to adopt, to conceive naturally and engage in watchful waiting, or to decide not to conceive. Finally, those carrier couples who choose to conceive without any intervention at all will at a minimum benefit from knowing the diagnosis of an affected child; for some diseases ameliorative options are available [43], involving special drugs or rigorous diets from birth [44, 45] (Figure 1).

While these choices are doubtless difficult, they are generally far preferable to the decisions that must be made after a positive result during the current practice of prenatal testing. For lethal Mendelian

conditions in particular this presents a pregnant mother with a terrible choice between terminating a wanted pregnancy or losing an infant in early childhood. Empowering women and couples with access to reproductive information before pregnancy allows them to decide whether and how to prevent this tragic scenario.

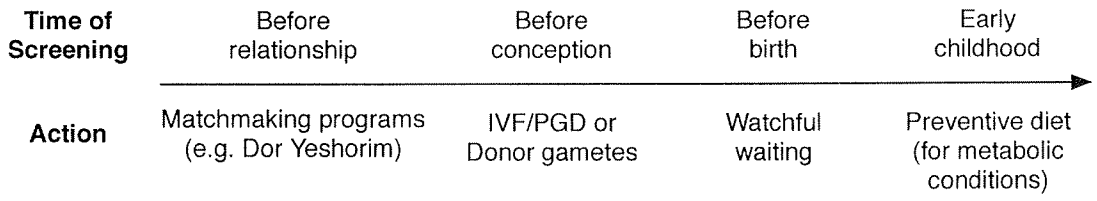


Figure 1. Pre-pregnancy Carrier Testing Allows Prevention of Mendelian Disease. The earlier an individual knows their carrier status, the more options are available for conceiving a healthy child. Before a relationship is begun, matchmaking organizations like Dor Yeshorim [46] can pair up carriers with non-carriers. If carrier status is known before conception, a couple can choose to undergo IVF/PGD [36,37] to select an embryo without the Mendelian condition, or opt to use donor sperm or eggs [35,41]. Couples who find out their carrier status during pregnancy can use amniocentesis to determine if their fetus carries a fatal genetic disease, and may opt to terminate the pregnancy if the test is positive. Finally, early diagnosis of certain heritable metabolic disorders [43,44] like PKU can alert parents to the need for preventive diets.

It is important to note that this concept of prevention is by no means theoretical, as successful campaigns have already been mounted against Tay-Sachs in the Jewish community [47,48] and beta-thalassemia in people of Mediterranean origin [48]. Ameliorative efforts such as the national newborn PKU screening campaign have also made their mark, as diagnosis has allowed many affected children to lead relatively normal lives by adhering to a highly restrictive diet; in this case what is prevented is even greater suffering because of non- or mis-diagnosed genetic syndromes.

Because of the possibilities for preventive care, many organizations have recommended that couples be offered genetic testing for specific diseases before pregnancy. For example, the ACMG recommends offering tests for cystic fibrosis [49–53] and spinal muscular atrophy [54,55] to all adults of reproductive age, with further testing indicated as a function of family history and ethnic background. Moreover, for the most common genetic diseases the public health burden is substantial enough that population screening is supported as a highly cost effective measure [13,56–58], even in developing countries [59].

Further extension of population screening is limited not so much by lack of knowledge of causal mutations [17,18] but by cost effectiveness: a disease mutation must be frequent, severe, and inexpensively assayed to be incorporated into a screening campaign. While frequency and severity are determined by the underlying biology and hence relatively fixed, recent advances in genomics have greatly reduced the cost per base and opened up new possibilities for population screening.

For this reason there have now been several calls for a significant expansion of population screening to a much wider range of Mendelian diseases [60–64]. However, building a "Universal Carrier Test" of this kind is technically challenging. First, it must have high accuracy across all assayed disease mutations [65], many of which are rare [66–70] and difficult to validate [71–75]. Next, the test should be

inexpensive [64, 76–79] enough for the cost of running the screen to be less than the financial burden of disease prevention. Finally, the test protocol should be non-invasive [80] and highly scalable [81] to avoid limits to universal uptake.

Here we describe the Universal Genetic Test (UNIT), an assay that overcomes all of these hurdles. The UNIT tests for 458 causal genetic variants for 105 Mendelian diseases with a sensitivity of >0.99988 (median 95% CI [0.99988, 0.999999]), a specificity of >0.99643 (median 95% CI [0.99643, 1]), and a positive predictive value in excess of 0.995. The test is non-invasive, requires only a saliva sample and was designed from the outset to be suitable for population screening of individuals from all ethnic groups, as a truly universal carrier test.

One of the primary purposes of the Universal Genetic Test is preventive care: by combining test results with demographic and family history, we may identify couples at risk for conceiving a baby affected with one of the 105 assayed diseases, enabling them to take preventive measures like PGD.

In this manuscript, we begin by detailing the statistical and economic constraints an assay must meet to enable a Universal Genetic Test. We then describe a design strategy that takes these constraints into account, by incorporating multiple redundant probes for every variant and using all available prior information to improve genotype calling. Next, we discuss the results of an exhaustive validation procedure, demonstrating that the assay's positive predictive value is high enough to enable population screening for the long tail of Mendelian disease. Finally, we conclude by presenting data from the use of the screen in a clinical setting at more than 100 medical centers around the country, including a number of leading fertility clinics. This data provides empirical evidence for a "long tail" [7, 82] of genetic disease, in which individuals are shown to be unlikely to carry any given mutation but surprisingly likely to carry at least one Mendelian disease allele (Figure 2).

Results

The Long Tail of Genetic Disease

We began by assembling data from many sources to demonstrate that the distribution of genetic diseases has a "long tail" – a large number of diseases, each individually rare, that collectively are surprisingly common. Figure 2A plots our estimates of the world-wide carrier frequencies of 164 debilitating diseases, in which disease prevalences and carrier frequencies for a variety of populations were curated from the literature and public databases.

Note from the figure that the 1.7% worldwide carrier frequency of a more common disease like cystic fibrosis (CF) is considerably smaller than the sum of the carrier frequencies of the less common diseases in the plot. A more sophisticated calculation that takes into account the possibility that an individual carries multiple mutations does not change this effect, and we confirm it empirically in Figure 6. The ineluctable conclusion, then, is that screening for the most common genetic diseases alone will fail to discover most of the carriers in the general population.

While seemingly surprising, this result in different form has long been known in population genetics. For example, estimates of genetic load via excess deaths from consanguineous marriages consistently produce an estimated number of recessive lethals per person of 4-5 [83]. And medical geneticist and NIH director Francis Collins [84] has noted that "Most single-gene conditions are uncommon... However, the total effect of monogenic conditions is substantial, from both the individual patient's and public health perspective".

Assaying many of these monogenic conditions simultaneously is made more challenging because the exact nature of the long tail varies by ethnic group. Figure 2B shows the carrier frequency distribution for three different populations. Although the distributions are qualitatively similar, the positions of different diseases vary. The consequence is that a universal carrier screen must assay a large number of different mutations, many of which are scarce outside of a particular subpopulation.

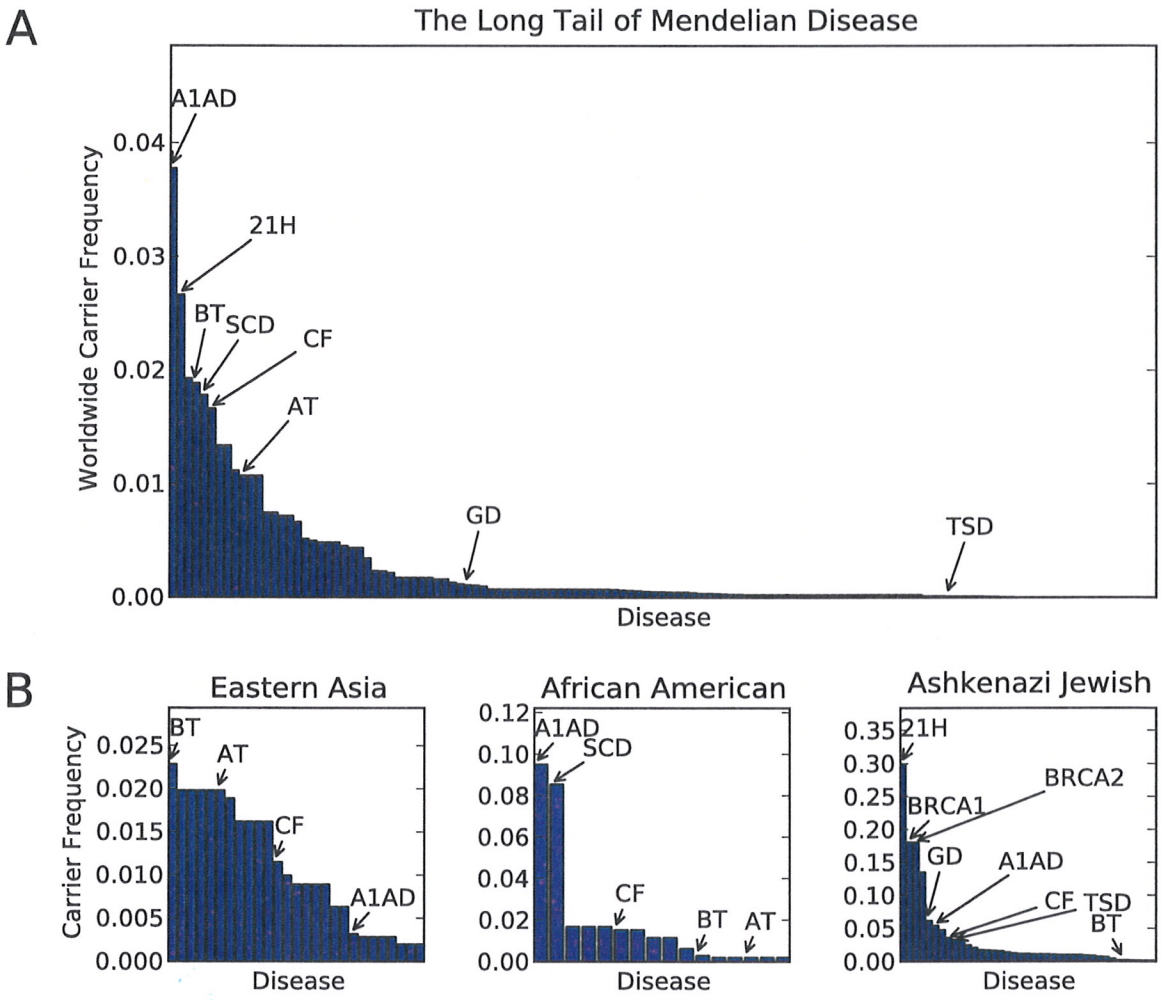


Figure 2. The Long Tail of Genetic Disease. (A) While genetic diseases are individually rare, they are collectively quite frequent. Shown are estimates of world-wide carrier frequencies for 164 debilitating diseases from the literature. Diseases are ranked by frequency on the x-axis, with their frequencies displayed on the y-axis. (B) The distributions for different ethnic groups have different rank orders of diseases. Abbreviations: 21H=21-hydroxylase deficiency nonclassic, A1AD=alpha-1 antitrypsin deficiency, AT=alpha thalassemia, BT=beta thalassemia, CF=cystic fibrosis, GD=Gaucher's disease, SCD=sickle cell disease, TSD=Tay-Sachs disease.

Statistical and Economic Requirements for a Universal Carrier Screen

A universal carrier screen for the long tail of genetic disease must have both a low false negative rate (FNR) to reliably identify carriers and a low false positive rate (FPR) to reduce the rate of unnecessary preventive measures. Moreover, for each disease the screen should have a high mutation detection frequency (MDF), corresponding to the fraction of causal mutations for the disease detectable by the assay.

All of these values should be achieved by the most cost-effective test possible, as high cost has been a major impediment to screening uptake [76–78]. In particular, for the test to be covered by third-party payers, the savings that result from early identification of a disease must be larger than the cost of the screen. These cost savings are a function of the frequency of each disease, its cost of treatment and prevention, and the accuracy and completeness of the screen. To quantify these savings, consider Figure 3, which illustrates the medical outcomes for the simplest possible case of a perfect screen for an autosomal recessive single-gene disease with complete penetrance.

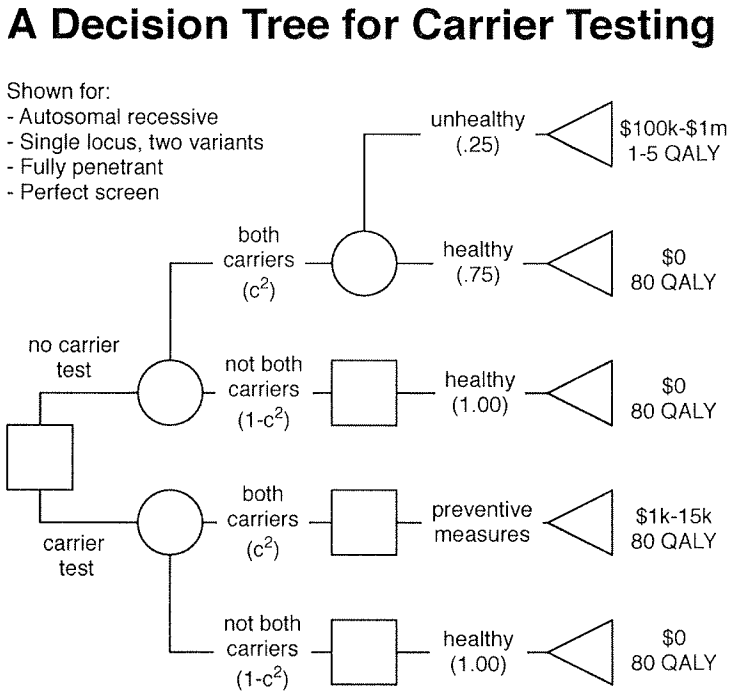


Figure 3. A Decision Tree for Carrier Testing. Without screening, a child has a non-negligible chance of suffering from a preventable genetic disease (top branch). With screening, this chance may be substantially reduced (bottom branch). For simplicity, this decision tree depicts a perfect screen for a single-gene fully penetrant autosomal recessive disease. It is a jumping off point to introduce non-idealities (as in Figure 4).

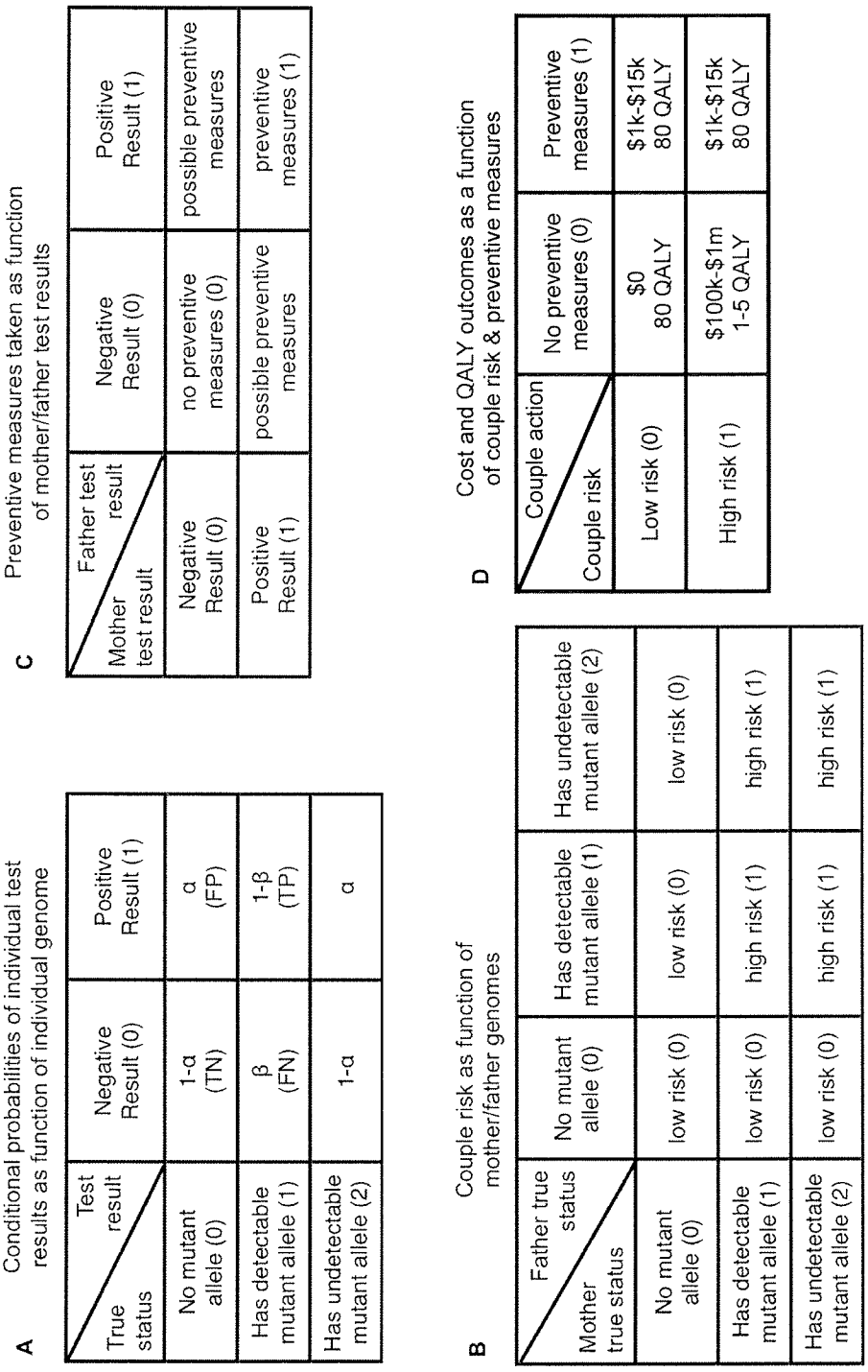
Figure 3 is a starting point to calculate the statistical and economic requirements for a universal carrier screen. To proceed we need to introduce two non-idealities: the possibility of false positive/negative results and the fact that some causal mutations may be absent from the screen (i.e., the MDF may not be 1.0).

First suppose the carrier frequency of a disease within a population is q, the cost of treatment is C_t, the cost of prevention after a positive test is C_p, the cost of the screen is C_s, and the screen's MDF, FPR, and FNR are given by m, α, and β respectively.

Next, the cases enumerated in the decision tree of Figure 3 must be augmented to accomodate the possibility of false positives, false negatives, and untested mutations; these new cases combinatorially expand the decision space and are shown in Figure 4.

In Figure 4a, there are three possibilities for the mother of a given child: she does not carry any mutant

alleles, she carries a mutant allele which is present in the assay, or she carries a mutant allele which is not detected by the assay. For each of these possibilities, there are two outcomes: the assay produces a negative result (allele is not detected) or a positive result (the allele is detected). Doing a complete enumeration over both mother and father pairs, we have $3 \times 2 \times 3 \times 2 = 36$ cases to consider (Figure 4b-c). For each of these cases, we can score the couple as "at risk" or not (corresponding to whether the mother and father both carry a mutant allele) and as "using prevention" or not (indicating whether the mother and father take preventive measures like IVF/PGD. There are $2 \times 2 = 4$ such outcomes, each with a different resulting expected cost of treatment (Figure 4d).



A Conditional probabilities of individual test results as function of individual genome

True status \ Test result	Negative Result (0)	Positive Result (1)
No mutant allele (0)	1-α (TN)	α (FP)
Has detectable mutant allele (1)	β (FN)	1-β (TP)
Has undetectable mutant allele (2)	1-α	α

B Couple risk as function of mother/father genomes

Mother true status \ Father true status	No mutant allele (0)	Has detectable mutant allele (1)	Has undetectable mutant allele (2)
No mutant allele (0)	low risk (0)	low risk (0)	low risk (0)
Has detectable mutant allele (1)	low risk (0)	high risk (1)	high risk (1)
Has undetectable mutant allele (2)	low risk (0)	high risk (1)	high risk (1)

C Preventive measures taken as function of mother/father test results

Mother test result \ Father test result	Negative Result (0)	Positive Result (1)
Negative Result (0)	no preventive measures (0)	possible preventive measures
Positive Result (1)	possible preventive measures	preventive measures (1)

D Cost and QALY outcomes as a function of couple risk & preventive measures

Couple risk \ Couple action	No preventive measures (0)	Preventive measures (1)
Low risk (0)	$0 80 QALY	$1k-$15k 80 QALY
High risk (1)	$100k-$1m 1-5 QALY	$1k-$15k 80 QALY

Figure 4. Carrier Screening Outcome Tables. The numerical codes for outcomes are used in Table 1. (A) The screening outcome for an individual at a given locus depends on both genetics and the sensitivity and specificity properties of the test. (B) A couple's true risk is a function of their respective genetics. (C) Whether a couple takes preventive measures is a function of their test results. (D) Cost and quality adjusted life years (QALY) outcomes for the child of a couple are functions of both their true risk and whether or not they take preventive measures. Note that here we consider an autosomal recessive locus; in general the probability of having at least one child suffer from a genetic disease varies as a function of the genetic architecture of the trait (e.g., penetrance, dominance) and the number of children born by the carrier couple.

Mother	Father	Mother result	Father result	At risk?	Prevent?	Outcome probability	Cost	QALY
0	0	0	0	0	0	(1-q)(1-q)(1-α)(1-α)	0	Lp
0	0	0	1	0	0	(1-q)(1-q)(1-α)(α)	0	Lp
0	0	1	0	0	0	(1-q)(1-q)(α)(1-α)	0	Lp
0	0	1	1	0	1	(1-q)(1-q)(α)(α)	Cp	Lp
0	1	0	0	0	0	(1-q)(qm)(1-α)(β)	0	Lp
0	1	0	1	0	0	(1-q)(qm)(1-α)(1-β)	0	Lp
0	1	1	0	0	0	(1-q)(qm)(α)(β)	0	Lp
0	1	1	1	0	1	(1-q)(qm)(α)(1-β)	Cp	Lp
0	2	0	0	0	0	(1-q)(q(1-m))(1-α)(1-α)	0	Lp
0	2	0	1	0	0	(1-q)(q(1-m))(1-α)(α)	0	Lp
0	2	1	0	0	0	(1-q)(q(1-m))(α)(1-α)	0	Lp
0	2	1	1	0	1	(1-q)(q(1-m))(α)(α)	Cp	Lp
1	0	0	0	0	0	(qm)(1-q)(β)(1-α)	0	Lp
1	0	0	1	0	0	(qm)(1-q)(β)(α)	0	Lp
1	0	1	0	0	0	(qm)(1-q)(1-β)(1-α)	0	Lp
1	0	1	1	0	1	(qm)(1-q)(1-β)(α)	Cp	Lp
1	1	0	0	1	0	(qm)(qm)(β)(β)	Cd	Ld
1	1	0	1	1	0	(qm)(qm)(β)(1-β)	Cd	Ld
1	1	1	0	1	0	(qm)(qm)(1-β)(β)	Cd	Ld
1	1	1	1	1	1	(qm)(qm)(1-β)(1-β)	Cp	Lp
1	2	0	0	1	0	(qm)(q(1-m))(β)(1-α)	Cd	Ld
1	2	0	1	1	0	(qm)(q(1-m))(β)(α)	Cd	Ld
1	2	1	0	1	0	(qm)(q(1-m))(1-β)(1-α)	Cd	Ld
1	2	1	1	1	1	(qm)(q(1-m))(1-β)(α)	Cp	Lp
2	0	0	0	0	0	(q(1-m))(1-q)(1-α)(1-α)	0	Lp
2	0	0	1	0	0	(q(1-m))(1-q)(1-α)(α)	0	Lp
2	0	1	0	0	0	(q(1-m))(1-q)(α)(1-α)	0	Lp
2	0	1	1	0	1	(q(1-m))(1-q)(α)(α)	Cp	Lp
2	1	0	0	1	0	(q(1-m))(qm)(1-α)(β)	Cd	Ld
2	1	0	1	1	0	(q(1-m))(qm)(1-α)(1-β)	Cd	Ld
2	1	1	0	1	0	(q(1-m))(qm)(α)(β)	Cd	Ld
2	1	1	1	1	1	(q(1-m))(qm)(α)(1-β)	Cp	Lp
2	2	0	0	1	0	(q(1-m))(q(1-m))(1-α)(1-α)	Cd	Ld
2	2	0	1	1	0	(q(1-m))(q(1-m))(1-α)(α)	Cd	Ld
2	2	1	0	1	0	(q(1-m))(q(1-m))(α)(1-α)	Cd	Ld
2	2	1	1	1	1	(q(1-m))(q(1-m))(α)(α)	Cp	Lp

Table 1. Enumerating the possible outcomes of tandem carrier testing for a couple. Tabulating all possible outcome cases for pairs of individuals using the tables and numerical codes of Figure 4 allows us to put a probability distribution over couple screening outcomes. Note that we record "possible preventive measures" as "no preventive measures" to be conservative. Note also that many of these outcomes have vanishingly small probabilities but are included for completeness. Using this exhaustive outcome enumeration, we can obtain Equation 1 and evaluate it for diseases with different carrier frequencies to determine the necessary properties of a cost-saving universal carrier screen.

With this figure as guidance, we can derive an equation to estimate the cost savings of a carrier

test. We stress that this calculation considers only the economics of a screen and does not include the psychological and human cost of disease, which is almost incalculable. For simplicity, we consider the case of fully penetrant autosomal recessive diseases; other inheritance patterns have a similar cost analysis. We assume that a disease d has carrier frequency $q_d = 1 - p_d$, that the treatment cost for an individual affected with disease d is C_t^d, and that the prevention cost faced by a couple with a positive test result is C_p^d. Furthermore, we assume the screen has a mutation detection frequency for d of m_d and that the overall FPR and false negative rate of the screen are α and $1 - \beta$.

The expected economic cost faced by a couple who does not take a carrier screen is $C_t^d q_d^2$. They incur a cost only if they conceive an affected child, which can occur if they are both carriers. We address the possibility that a carrier couple does not conceive an affected child by conservatively setting C_t^d to no more than half the true expected treatment cost. For simplicity we assume random mating with respect to the disease.

By multiplying the probability by the expected cost and summing over each row, we can obtain an expression with 36 terms which expresses the expected cost savings of the screen. Using this expression, we can posit q, C_t, C_p as fixed parameters of the disease to derive bounds on the screen parameters C_s, m, α, β (Figure 1). A screen which satisfies these stringent bounds will be cost effective at scale.

The economic cost faced by a couple who take the screen varies based on their carrier status. If neither are carriers, which occurs with probability q_d^2, they face cost C_p^d with probability α^2; the screen must produce two false positives. If exactly one is a carrier, which occurs with probability $2p_d q_d$, they face cost C_p^d with probability $m_d \beta \alpha$; the screen must produce one false positive and one true positive.

If both couples are carriers, the cost depends on whether they carry the same disease mutation. We denote the probability that both carry the same mutation as s_d; this probability depends on the allelic spectrum of the disease [85]. If both carry the same mutation, which occurs with probability s_d, they face cost C_t^d with probability $(1-m_d)+m_d(1-\beta^2)$ (if the screen does not both assay the mutation and produce two true positives), and they face cost C_p^d with probability $m_d\beta^2$ (if the screen assays the mutation and produces two true positives). If they carry different mutations, which occurs with probability $1 - s_d$, they face cost C_t^d with probability $\left((1-m_d)^2 + 2m_d(1-m_d)(1-\beta) + m_d^2(1-\beta^2)\right)$ (if the screen does not assay either mutation, if it assays exactly one mutation but produces a false negative, or if it assays both mutations and produces two false positives), and they face cost C_p^d with probability $m_d^2\beta^2$ (if the screen assays both mutations and produces two true positives).

Combining these equations (and dropping d subscripts and superscripts for clarity), the cost faced by a couple that takes the screen is

$$p^2\alpha^2 C_p + 2pqm\beta\alpha C_p + q^2\Big(s\left(\left(1-m\beta^2\right)C_t + m\beta^2 C_p\right) + (1-s)\left(\left((1-m) + 2(1-m)m(1-\beta) + m^2(1-\beta^2)\right)C_t + m^2\beta^2 C_p\right)\Big) \quad (1)$$

A screen will be cost effective if this cost, summed over all diseases on the screen and added to the cost of the screen C_s, is less than $\sum_d C_t^d q_d^2$.

We compared the cost of a universal screen for the diseases in Figure 2 to (1) the cost faced without a screen and (2) the combined cost of separate screens for each disease. We used values of $C_t^d = \$750,000$, $C_s = \$700$ (cost of the UNIT for a couple), and $C_p^d = \$10,000$ (assuming a significant number at-risk couples choose PGD). We used our curated world-wide carrier frequencies for q_d; we assumed (extremely conservatively but for the sake of comparison) that disease-specific tests had perfect MDF values ($m_d = 1$). This assumption is quite conservative as it is currently not possible to achieve 100% mutation detection with any clinical genotyping assay.

We continued by setting the postulated universal screen's MDF values to those of the UNIT. For both the universal screen and the disease specific screens we used values of $\beta = 0.001$ and $\alpha = 0.004$, highly conservative estimates of the UNIT's performance. We estimated s_d as $\sum_i f(v_d^i)^2$, where $f(v_d^i)$ is

the frequency of the i^{th} causal variant for d in carriers of d. We used curated values from the literature for $f(v_d^i)$; this results in the implicit assumption that if an individual has a mutation not known in the literature there is a negligible probability that the other individual has the same mutation. This assumption is quite reasonable for most familial mutations, whose danger is likely to be caused by compound heterozygosity when encountered in non-consanguineous contexts.

Figure 5 plots these costs and suggests several points about the cost-effectiveness of a universal screen. First, a combination of disease-specific tests is far too expensive to screen all of the diseases in the long tail; for rare diseases the cost of the screen far outweighs the cost of treatment. Second, a universal screen becomes increasingly cost-effective as it includes more diseases; any reduction in treatment cost is beneficial because the incremental screening cost for each disease is very low. Third, the overall conclusion is that a sufficiently accurate UNIT will result in health care savings when applied at the population level.

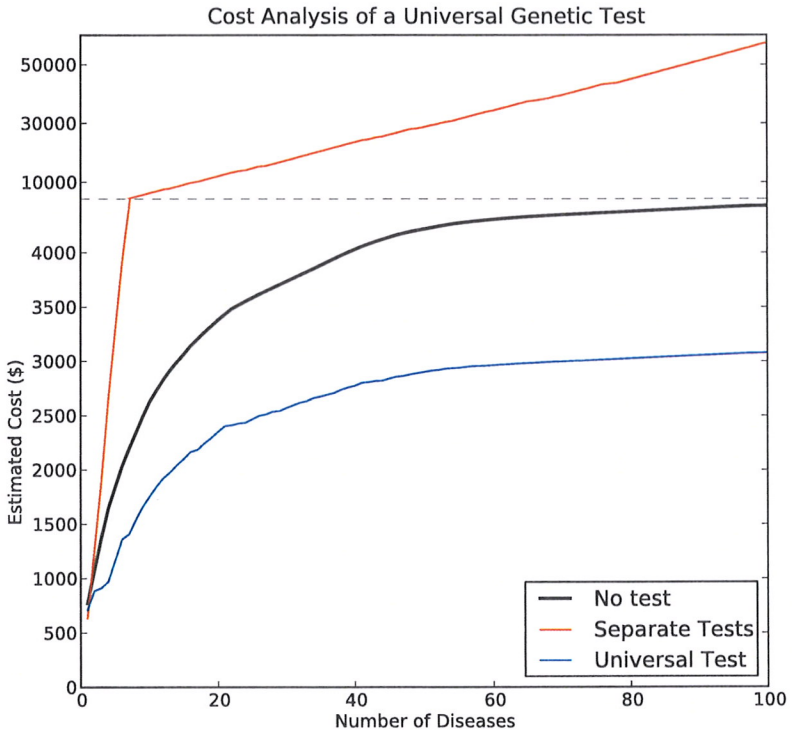


Figure 5. Cost effectiveness of a universal carrier test. This curve shows the cumulative expected cost of three different treatment paradigms: no test, a separate test for each disease, and a universal carrier test. As we proceed down the list of diseases with the highest mutation detection frequencies on the UNIT, the universal screen becomes increasingly cost-effective compared to the absence of screening. In contrast, separate tests for each disease are clearly not cost-effective.

Universal Genetic Test Performance Statistics

We exhaustively validated the UNIT on a combination of reference gDNA samples [86], synthetic DNA samples [87–89], and clinical DNA samples. In brief, we achieved extremely high sensitivity, specificity, and positive predictive values by combining multiple redundant probes, a triage strategy (Table 4), and two-stage followup testing for positive carrier couples. Probes in the assay which passed triage had

essentially digital accuracy, with either complete success or else no-call on hundreds of control samples per variant. The overall performance results are summarized in Table 2.

We constructed a large gold standard reference database by combining samples from public sources and sequence verified samples, providing large numbers of labeled positive and negative samples for each variant. When used in conjunction with domain knowledge that gave us *a priori* information on the number of clusters for each variant, we could establish very robust call boundaries with strong separation between genotypes.

For example, consider the representative plot in Figure 7. Domain knowledge for this variant (CFTR deltaF508) tells us that there are only two clusters expected: the heterozygote (carrier) for the deleterious recessive deletion and the homozygote (wildtype). The plot also allows us to intuitively understand the idea of a false negative (a labeled red positive that lands in the blue wildtype homozygote cluster) and a false positive (a labeled blue negative which lands in the red carrier heterozygote cluster). For this and many other variants, no false negatives or positives were observed in our entire sample dataset. Full details on individual variants are in Table 6.

Accuracy and precision are the most informative estimates of the total aggregate error rate of the assay. Both statistics indicate an overall average error rate of approximately 1 in 50,000. To put this performance in context, this means the Universal Genetic Test is a highly multiplex assay which nevertheless compares very favorably to the reported accuracy and precision of FDA-approved single gene assays for cystic fibrosis (Table 3). It is also within the range of the top single gene DNA based tests for Tay-Sachs mutations [90], which had 4 errors (false positives + false negatives) per 100,000 couples.

Metric	Value	95% CI
Precision	0.99997	[0.99993, 0.99999]
Accuracy	0.99998	[0.99988, 0.999999]
Sensitivity	0.99998	[0.99988, 0.999999]
Specificity	>0.996	[0.99643, 1]
False positive rate	<0.004	[0, 0.00357]
False negative rate	0.00002	[0.000001, 0.00012]
Positive predictive value	>0.995	[0.99992, 1]
Negative predictive value	0.99907	[0.99474, 0.99995]

Table 2. Performance summary of the Universal Genetic Test. Quantitative definitions of each parameter are given in Table 5 and Materials and Methods.

Test	Accuracy	95% CI	Precision	95% CI	Reference
eSensor® Cystic Fibrosis Carrier Detection System	99.97%	[0.99924, 0.99991]	99.9%	n/a	510(k): k060543
Tag-It™ Cystic Fibrosis Kit (TM Bioscience Corporation)	100%	[0.99869, 1]	>99.99%	[0.99980, 0.99998]	510(k): k043011
Cystic Fibrosis Genotyping Assay (Celera)	>99.99%	[0.99977, 1]	100%	[0.99990, 1]	510(k): k062028
InPlex CF Molecular Test (Third Wave Technology)	99.96%	[0.99782, 0.99998]	99.987%	[0.99984, 0.99990]	510(k): k063787
Universal Genetic Test (Counsyl)	99.998%	[0.99993, 0.99999]	99.997%	[0.99993, 0.99999]	Present study

[n/a] eSensor reported a contradictory call rate of 0.008%.

Table 3. Performance comparison to four IVDs used in cystic fibrosis carrier screening. The Universal Genetic Test has accuracy and precision levels comparable to FDA-approved in vitro diagnostic (IVD) devices for cystic fibrosis testing. In other words, this multiplex saliva-based diagnostic has performance comparable to that of traditional single-gene tests. Here, accuracy and precision point estimates are taken directly as reported from the regulatory filings. 95% confidence intervals were re-calculated based on reported counts to serve as a consistent basis for comparison.

Two-Stage Testing

To further ensure the highest possible accuracy, the Universal Genetic Test process includes verification of carrier couple results with two-stage testing. For an assay with a sufficiently low FPR, this is cost-effective and can boost accuracy substantially. For example, the FPR of the UNIT is less than 0.004. For rare alleles with frequencies around 0.001, the positive predictive value (PPV) of the first stage of the test is better than $0.001/(0.001 + 0.004) = 0.20$. This represents at least a 200-fold enrichment over background frequency, which is exactly the purpose of a screening assay. Using a biochemically independent followup test with a FPR of <0.001 increases the overall PPV beyond $0.20/(0.20 + 0.001)$, which means it is >0.995 (for the empirical point estimate, see Table 2). As individual loci can now be inexpensively assayed by a variety of methods, even a high overall carrier rate does not significantly increase cost as there are only a few loci (usually just one) on which to perform confirmatory followup.

In summary, it is clear from Table 2 that the test satisfies the rigorous statistical and economic criteria discussed earlier and is hence suitable for a cost-effective population-wide screen.

The Empirical Clinical Long Tail of Genetic Disease

The empirical distribution of carrier frequencies was calculated from clinical samples tested in our laboratory. As shown in Figure 6, the theoretical predictions of a long-tail of genetic disease were validated by this empirical data. In aggregate, approximately 35% of samples are found to be carriers of at least one disease.

In addition to individual carrier frequencies, our clinical samples allow us to calculate the rate of carrier couples. We find this rate to be approximately 0.6-0.8%. Importantly, our clinical test results are highly enriched with patient samples originating from fertility clinics, which include both patients seeking fertility treatments and gamete donors. It may be the case that carriers for some diseases are at an increased risk of fertility problems, similar to the relationship of CFTR mutations to congenital absence of the vas deferens [91]. Also, couples previously identified as carriers or with known family history of disease might be retested while seeking fertility treatments. Thus, we cannot rule out the possibility that this enrichment has resulted in an increased frequency of carriers and carrier couples.

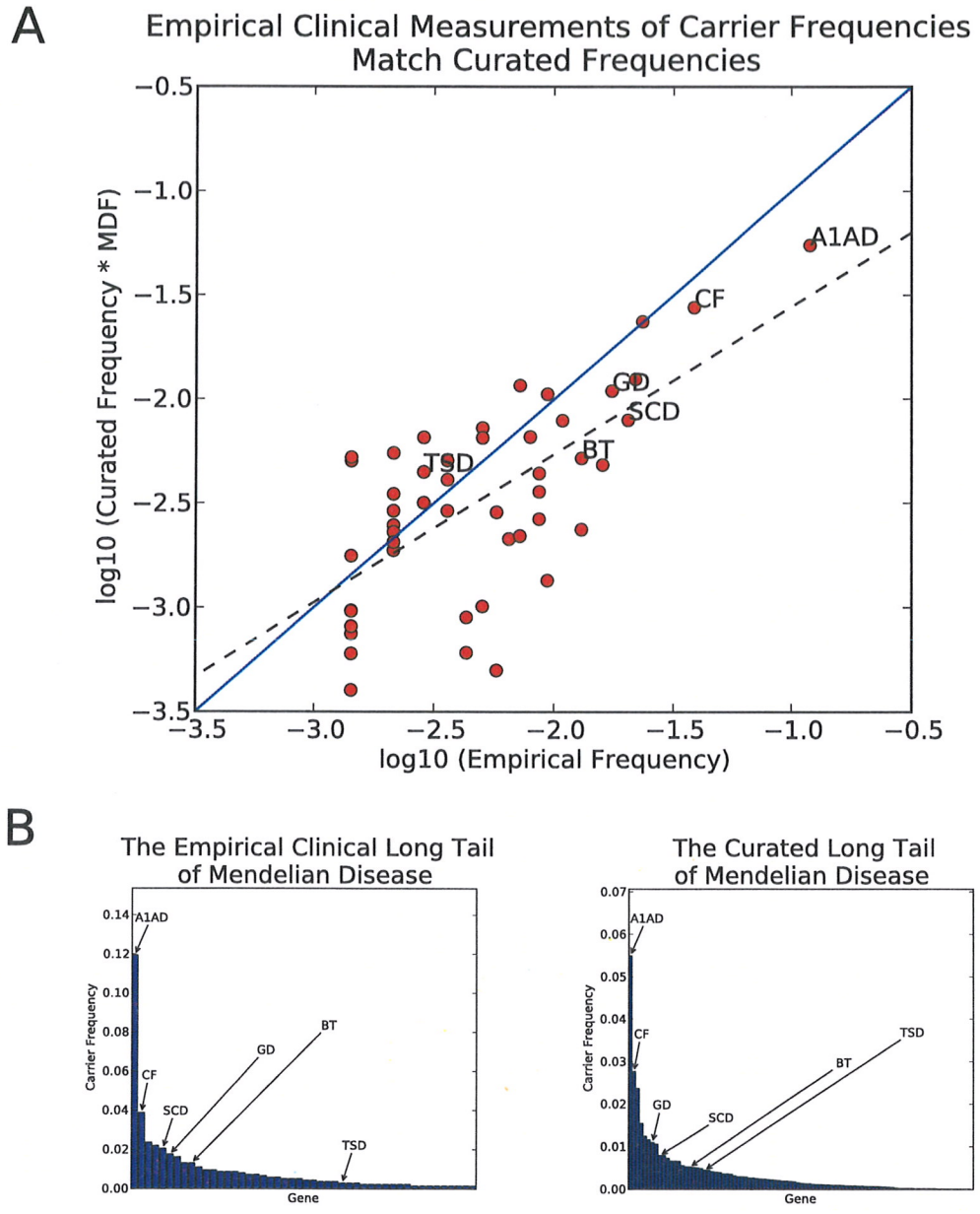


Figure 6. The Empirical Long Tail of Genetic Disease. (A) Log-log plot of observed vs. curated carrier frequency for major diseases. The solid blue line is the $y = x$ line, representing a perfect match between experimental and curated frequency data. The dashed line is a robust regression line for curated vs. experimental data. Not only is the overall match quite good, the heteroskedascity is in the direction we expect: diseases of lower frequency show more scatter, as the values stated by experimentalists from the literature are more uncertain (e.g. "the prevalence is approximately 1 in 20000"). (B) Empirical and curated long tail plots associated with (A). Abbreviations: A1AD=alpha-1 antitrypsin deficiency, AT=alpha thalassemia, BT=beta thalassemia, CF=cystic fibrosis, GD=Gaucher's disease, SCD=sickle cell disease, TSD=Tay-Sachs disease.

Discussion

Inspired by the success of the universal Tay-Sachs screening program in the Jewish community, we have developed a single, non-invasive, inexpensive, highly-accurate assay for the long tail of Mendelian disease. The assay has been thoroughly validated on reference genomic DNA from biobanks as well as synthetic heterozygotes, and has been successfully used for patient testing in our clinical reference laboratory. We have further shown that the assay is extremely sensitive and specific, and that the empirical carrier frequencies detected by the assay correspond well to those predicted by theory. The assay is currently in use at more than one hundred clinics [92] across the United States and has already identified several confirmed carrier couples.

Importantly, the assay makes the prospect of a universal at-home carrier testing feasible for the first time and has several interesting implications. We discuss these in turn below.

Feasibility of an At-Home Carrier Test

Similarities to the At-Home Pregnancy Test

The introduction of the first at-home pregnancy test was marked by controversy [93]. Critics noted that the test was performed at home, without the supervision of a trained medical professional. Most patients who took it were otherwise healthy individuals who tested negative, and those who tested positive experienced a life changing event requiring significant medical followup.

The fact that the Universal Genetic Test requires only a saliva sample makes it possible to contemplate an at-home carrier test which is structurally similar to the at-home pregnancy test. Just like the at-home pregnancy test, most patients who take such a test would be otherwise healthy individuals who test negative, with couples who test positive experiencing a life changing event requiring significant medical followup.

Urine and saliva samples collected by laymen may be analyzed in a lab

The only logistical difference between the at-home carrier test and the at-home pregnancy test is the sample collection procedure. Importantly, the actual analysis for an at-home carrier test would not be performed on the premises but by a laboratory scientist in a clincial reference laboratory. Just like a urine test, the laboratory scientist does not need to be present for the process of urine or saliva sample collection. With urine testing, the patient collects their urine sample and leaves it in a tray for later analysis by a licensed laboratory scientist, which may occur up to 24-48 hours later. Similarly, with an at-home carrier test a patient collects their saliva sample and puts it in a remailing envelope for later analysis by a licensed laboratory scientist, which may occur up to 24-48 hours later.

Indeed, the parallel is particularly strong for two reasons. First, there are already at-home urine tests [94]. Second, several studies of the reliability of at-home pregnancy test kits [95, 96] found that not only were at-home tests highly sensitive and specific in the hands of trained operators, but that the difficulties a few laypersons encountered were primarily related to educational level rather than test characteristics. These issues were largely solved with the introduction of a "thin blue line" binary assay readout [93]; the promulgation of reproductive education in schools likely assisted as well.

The proposed at-home carrier test does not have this difficulty, as the 'at-home' part relates solely to sample collection and the readout is already quite binary (e.g. carrier couple or not) rather than continuous (e.g. diabetes risk factor).

Logistics of test provisioning are separate from molecular etiology

Though convincing arguments have been voiced against premature testing for complex diseases [97], it is of crucial importance to separate the logistics of test provisioning (at-home vs. office visit) from

the molecular etiology of the disease (Mendelian vs. complex). An attempt to predict complex disease susceptibility from genes will fail regardless of whether a test is obtained over the internet or in a clinical setting because the requisite signal is simply not present with today's technology [98].

By contrast, the signal for predicting Mendelian disorders does not change as a function of whether a sample is collected at home or in the clinic. The only question is the narrow technical issue of whether mailing a saliva sample significantly degrades signal vs. sample collection in a clinical setting; both internal data and external references indicate this is not the case, and that saliva samples are as reliable as blood [99–102] for the purposes of genotyping.

Benefits of an at-home carrier test for privacy, logistics, equitable access, and reproductive freedom

Many of the same arguments that ultimately proved convincing during the introduction of the at-home pregnancy test apply to the at-home carrier test.

First, we must consider privacy. Just as many women may not want to disclose their pregnancy to outside parties (even including physicians), many individuals and couples desire the minimum outside knowledge of genetic test information.

Second, if current guidelines were implemented and every adult in the United States was actually offered a carrier test before pregnancy (as recommended by the ACMG and ACOG), we would face the burden of millions of otherwise healthy people seeking physican office visits to receive a test for which they already effectively have a prescription. By contrast, the at-home carrier test is a simple preventive measure that has the potential to significantly reduce the rate of high-risk pregnancies, as well as being the first large scale implementation of genomics in preventive medicine. Given the acknowledged bipartisan consensus on the need to control health care costs, the logistical argument for such a test is quite compelling.

Third, an inflexible requirement that the test only be offered at the physican's office will necessarily raise costs, reduce accessibility, and increase health care disparities. Underserved groups in rural communities are often located geographically far from the kinds of major medical centers that are generally the first to adopt new technology. However, even rural areas now have reasonable access to broadband technology, and the Obama administration has made the expansion of this access a legislative priority [103]. By making the at-home carrier test available via the internet, we can provide the test universally without implicit discrimination against members of rural communities.

Fourth and finally, an at-home carrier test would be a major victory for reproductive freedom. The at-home pregnancy tests played a major role in giving women greater control over their reproductive lives, allowing them to avoid unwanted pregnancies while avoiding stigma [93]. Similarly, the at-home carrier test would allow a woman and her partner to confidentially decide which reproductive options to take in the event of a positive result, allowing them to prevent their children from suffering from genetic disease while avoiding stigma.

The apparent conclusion, then, is that the President of the ACMG was prescient in his recent comments [104]:

> Korf said that DTC testing could be considered a "disruptive technology" that arguably has its share of faults now, but "the danger is that by turning our backs, little by little, as the quality improves, it could become a very powerful approach" and the clinical genetics community will have missed its chance to play a role.

That is, the prospect of a diagnostic grade at-home carrier test for Mendelian diseases will likely reshape the debate [105, 106] over so-called "DTC" genetic testing, which has to this point conflated test provisioning (clinical vs. mail-in sample collection) with disease etiology (complex vs. Mendelian). By focusing on a medical diagnostic for well-understood Mendelian disease that has as its precedents

the at-home pregnancy test and the successful campaign for Tay-Sachs screening, the emphasis turns to technical issues (sample collection, diagnostic performance) rather than basic scientific questions about complex disease.

The Limitations of Targeted Mutation Analysis

It is well known that no medical test is 100% accurate. In this connection is important to recall that the Universal Genetic Test is risk-reducing rather than risk-eliminating, and that a particularly important source of false negatives are genetic mutations that absent from the panel (either because they are as yet unknown, recently discovered, or resistant to genotyping based analysis).

That said, in every case, the mutations assayed by the Universal Genetic Test are currently used as part of a targeted mutation panel offered by at least one clinic [18, 107, 108], as shown in Table 7. In other words, any arguments about mutation coverage apply equally to current clinical practice, as no one contends that perfect panels for all tested diseases are currently available.

The ultimate solution to the issue of mutations which are absent from a mutation panel is likely to be diagnostic-quality resequencing. However, that approach presents its own set of concerns. Diagnostic resequencing causes us to immediately move from a situation in which the problem is that some mutations are absent to a situation in which thousands of alterations are present. Given that everyone's genome sequence is effectively unique, and given the inherently noisy nature of next-generation sequencing technology, it is still highly technologically nontrivial to reliably identify heretofore unobserved deleterious variants.

The conclusion, then, is that while the prospect of an idealized multiplexed sequencing assay is attractive, the technology simply does not yet exist to make a diagnostic-quality sequencing assay. Criticisms of a multiplex test which center on mutation detection frequency must take this issue into account; while no test is perfect, the cost-benefit analysis of Figure 5 clearly shows that highly accurate detection of the most common mutations at low cost is far superior to no screening at all.

Implications of a Universal Carrier Screen

There are several novel aspects of the Universal Genetic Test. First and foremost, the public health consequences are significant: this assay enables scaling up the success of Tay-Sachs screening to screen the general population for a wide variety of preventable genetic diseases.

Second, by shifting as much testing as possible to the pre-pregnancy rather than prenatal stage, more preventive options become available. This is interesting in that it simultaneously reduces the number of terminations while expanding choice, and has the potential to be a significant milestone for reproductive health and women's rights.

Third, it is a concrete initial step towards routine use of a genome sequence in medicine. ACMG currently recommends that *all* adults of reproductive age be offered carrier testing for cystic fibrosis [49] and spinal muscular atrophy [54], and both NHGRI [109] and ACMG [61] have anticipated a scaled up carrier screen similar in many ways to our assay. It is thus not unreasonable to postulate that an offer of universal carrier screening will become a routine part of medical care.

Fourth, the assay is an important tool for closing health disparities. By making the test available via the internet, we can make sure that rural communities have access to the latest technology at the same time as wealthy areas with expensive medical centers. And by manufacturing a single inclusive assay for all populations at scale, by working with insurers to cover it as preventive care, and by providing financial aid for the needy [110], we can strive to ensure that minority groups benefit equally from the promise of universal carrier screening.

Fifth, this assay will likely increase the demand for genetic information as people seek to learn more about their test results. This is not necessarily a negative eventuality. Just as the computer revolution increased the demand for computer scientists and promoted computer literacy, so too will increased use

of the fruits of the Human Genome Project increase the demand for medical geneticists and promote genetic literacy.

Sixth and finally, clinical data from the assay provides the first genome-wide, multivariate dataset on carrier frequencies. The evidence is consistent with the theoretical prediction that each person carries 4-5 recessive lethals on average [83].

This last point deserves some elaboration. At first it seems quite surprising to find that 35% of people are carriers for at least one disease in our panel. However, this equates to an expectation of only .35 recessive lethal alleles per person, only a fraction of the 4-5 predicted recessive lethals per person. In other words, more than 90% of the Mendelian disease burden remains to be accounted for. Testing for this enormous remainder is clearly a desirable direction for future work, and will necessarily involve both a transition from targeted genotyping to diagnostic resequencing and an effort to systematically map the hundreds of unmapped Mendelian loci in OMIM [17].

Materials and Methods

Literature curation

We used a systematic curation of the medical genetics literature and databases [18, 107, 108] to identify clinically significant variants associated with single-gene disorders. We selected variants that are (1) currently tested by at least one other clinical laboratory using different genotyping technologies and (2) suitable for population screening. We further focused on diseases and variants where mutation detection was amenable to highly multiplex genotyping methods. For each variant we recorded the associated disease, as well as the sequence of the disease-causing and wildtype alleles. The frequencies of the disease-causing and wildtype alleles in all populations for which data was available were also recorded.

Disease severity was categorized as either mild or severe. For each disease, a genotype-to-phenotype map was constructed. These maps capture both the general case of autosomal recessive inheritance and exceptions, such as the importance of the cis/trans relationships of the R117H and IVS8-5T variants of CFTR. The determination of carrier and affected status was made by reference to these maps.

Genotyping technology

The UNIT uses a customized multiple Molecular Inversion Probe (MIP) assay [111–113] to convert the information content of a genetic variant into fluorescently-labeled tag sequences. The system was modified to accommodate a number of variants beyond biallelic SNPs, including insertions, deletions, triallelic SNPs, copy number variants, and nearby polymorphisms [114, 115].

Probes that test each curated disease-causing variant were included. These probes were designed to detect both the mutant and wildtype alleles of each variant. Thus, heterozygous genotypes are determined by the positive detection of both a wildtype and mutant allele. The median number of probes tested per disease-causing variant is 3 for insertions/deletions and 2 for SNPs. In turn, each probe is measured 3 times. Overall, 105 genetic diseases are represented in the UNIT panel. Additional probes were included as quality control indicators, including markers for sex determination.

To maximize accuracy of the assay, we used a multi-stage approach to design the probes. We began with a large set of potential probes for a comprehensive set of variants and pursued a triage strategy, removing (in order) probes that did not format properly, probes that did not convert, and probes that were callable but had low accuracies. Table 4 summarizes the results of each stage of validation. The result of this approach was a set of 925 probes for 458 causal disease variants.

Stage	Probes Filtered	Probes Remaining
Initial	—	2096
Format Correctly	664	1432
Convert Correctly	292	1140
High Accuracy/Precision	215	925
Final	—	925

Table 4. A Multi-Stage Approach to Assay Design. We used a multi-stage approach to design the UNIT. We began with 2096 probes to assay a large number of disease-causing variants. We then applied a series of increasingly stringent quality control (QC) metrics to obtain the final set of 925 probes for the assay.

Validation study design

To assess the performance of UNIT, a panel of clinically-characterized reference samples was tested. Because patient samples were not available for all targeted mutations, synthetic patient samples were created which contain each mutant sequence of interest. [87,88] Each reference sample was tested at least 3 times over the course of the validation study.

Reference samples representing 131 variants in the UNIT were obtained from the Coriell cell repository, a biobank of genomic DNA reference materials. DNA samples representing all mutant alleles, including rarer alelles for which genomic DNA reference materials were unavailable were generated by DNA synthesis. Double-stranded DNA was synthesized and cloned into standard high-copy plasmid vectors. The integrity of each insert was confirmed by bi-directional sequencing. Plasmids were combined with human reference gDNA NA10838 in an approximately equal-molar ratio. A dilution series of plasmid in gDNA was used to set the final concentration of plasmid in gDNA that achieves a synthetic heterozygote genotype in the combined samples.

Performance metrics

Formulas for the performance metrics are provided in Table 5.

Metric	Description
TM	# replicate calls that equal the modal genotype
FM	# replicate calls that do not equal the modal genotype
TP	# true positive calls
TN	# true negative calls
FP	# false positive calls
FN	# false negative calls
TC	TP + TN
FC	FP + FN
Precision	$\frac{TM}{TM+FM}$
Accuracy	$\frac{TC}{TC+FC}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
False positve rate	$\frac{FP}{FP+TN}$
False negative rate	$\frac{FN}{FN+TP}$
Positive predictive value	$\frac{TP}{TP+FP}$
Negative predictive value	$\frac{TN}{TN+FN}$

Table 5. Performance Metrics for Validating the Universal Genetic Test. Assay level performance metrics are in Table 2. Variant level performance metrics are in Table 6.

Empirical carrier frequency determination

Empirical carrier frequency estimates were exported from the Counsyl clinical testing result database.

Acknowledgments

Critically reviewed the manuscript: Rishi Kacker, Ramji Srinivasan, and Balaji K. Srinivasan.

References

1. Grosse SD, Wordsworth S, Payne K (2008) Economic Methods for Valuing the Outcomes of Genetic Testing: Beyond Cost-effectiveness Analysis. Genet Med 10: 648–654.

2. Costa T, Scriver CR, Childs B (1985) The Effect of Mendelian Disease on Human Health: A Measurement. Am J Med Genet 21: 231–242.

3. Scriver CR, Neal JL, Saginur R, Clow A (1973) The Frequency of Genetic Disease and Congenital Malformation Among Patients in a Pediatric Hospital. Can Med Assoc J 108: 1111–1115.

4. Kumar P, Radhakrishnan J, Chowdhary MA, Giampietro PF (2001) Prevalence and Patterns of Presentation of Genetic Disorders in a Pediatric Emergency Department. Mayo Clin Proc 76: 777–783.

5. Berry RJ, Buehler JW, Strauss LT, Hogue CJ, Smith JC (1987) Birth Weight-specific Infant Mortality Due to Congenital Anomalies, 1960 and 1980. Public Health Rep 102: 171–181.

6. National Organization for Rare Diseases. URL `http://www.rarediseases.org/`.

7. Walker J. Healthcare and the Long Tail. URL `http://thehealthcareblog.com/the_health_care_blog/2006/07/techqualityphys.html`.

8. Risch N, Burchard E, Ziv E, Tang H (2002) Categorization of Humans in Biomedical Research: Genes, Race and Disease. Genome Biol 3: 2007.

9. Emery A, Rimoin D (2007) Principles and Practice of Medical Genetics. New York: Churchill Livingstone, fifth edition.

10. Ghotbi N, Tsukatani T (2005) Evaluation of the National Health Policy of Thalassaemia Screening in the Islamic Republic of Iran. East Mediterr Health J 11: 308–318.

11. Karnon J, Zeuner D, Brown J, Ades AE, Wonke B, et al. (1999) Lifetime Treatment Costs of Beta-thalassaemia Major. Clin Lab Haematol 21: 377–385.

12. Ho WL, Lin KH, Wang JD, Hwang JS, Chung CW, et al. (2006) Financial Burden of National Health Insurance for Treating Patients With Transfusion-dependent Thalassemia in Taiwan. Bone Marrow Transplant 37: 569–574.

13. Warren E, Anderson R, Proos AL, Burnett LB, Barlow-Stewart K, et al. (2005) Cost-effectiveness of a School-based Tay-Sachs and Cystic Fibrosis Genetic Carrier Screening Program. Genet Med 7: 484–494.

14. Kauf TL, Coates TD, Huazhi L, Mody-Patel N, Hartzema AG (2009) The Cost of Health Care for Children and Adults With Sickle Cell Disease. Am J Hematol 84: 323–327.

15. Baumann U, Stocklossa C, Greiner W, von der Schulenburg JM, von der Hardt H (2003) Cost of Care and Clinical Condition in Paediatric Cystic Fibrosis Patients. J Cyst Fibros 2: 84–90.

16. Lieu TA, Ray GT, Farmer G, Shay GF (1999) The Cost of Medical Care for Patients With Cystic Fibrosis in a Health Maintenance Organization. Pediatrics 103: e72.

17. Wilkie AO, Amberger JS, McKusick VA (1994) A Gene Map of Congenital Malformations. J Med Genet 31: 507–517.

18. Pagon RA (2006) GeneTests: An Online Genetic Information Resource for Health Care Providers. J Med Libr Assoc 94: 343–348.

19. Need AC, Goldstein DB (2009) Next Generation Disparities in Human Genomics: Concerns and Remedies. Trends Genet 25: 489–494.

20. Simeonsson RJ (1991) Early Prevention of Childhood Disability in Developing Countries. Int J Rehabil Res 14: 1–12.

21. El-Hazmi MA (1997) Early Recognition and Intervention for Prevention of Disability and Its Complications. Eastern Mediterranean Health Journal 3: 154–161.

22. Burchard EG, Ziv E, Coyle N, Gomez SL, Tang H, et al. (2003) The Importance of Race and Ethnic Background in Biomedical Research and Clinical Practice. N Engl J Med 348: 1170–1175.

23. Zlotogora J (1994) High Frequencies of Human Genetic Diseases: Founder Effect With Genetic Drift or Selection? Am J Med Genet 49: 10–13.

24. National Heart, Lung, and Blood Institute (1996) Sickle Cell Anemia. Technical Report 96-4057, National Institutes of Health.

25. Lenker DP, Mills EJ (2005) Professional Guide to Diseases. Philadelphia: Lippincott Williams & Wilkins, eighth edition.

26. Hofgrtner WT, West Keefe SF, Tait JF (1997) Frequency of Deletional Alpha-thalassemia Genotypes in a Predominantly Asian-American Population. Am J Clin Pathol 107: 576–581.

27. Lau YL, Chan LC, Chan YY, Ha SY, Yeung CY, et al. (1997) Prevalence and Genotypes of Alpha- and Beta-thalassemia Carriers in Hong Kong – Implications for Population Screening. N Engl J Med 336: 1298–1301.

28. Vichinsky EP, MacKlin EA, Waye JS, Lorey F, Olivieri NF (2005) Changes in the Epidemiology of Thalassemia in North America: A New Minority Disease. Pediatrics 116: e818–825.

29. Gross SJ, Pletcher BA, Monaghan KG (2008) Carrier Screening in Individuals of Ashkenazi Jewish Descent. Genet Med 10: 54–56.

30. Goodman R, Motulsky AG (1979) Genetic Diseases Among Ashkenazi Jews. Philadelphia: Lippincott Williams & Wilkins.

31. Motulsky AG (1995) Jewish Diseases and Origins. Nat Genet 9: 99–101.

32. Lehavi O, Aizenstein O, Bercovich D, Pavzner D, Shomrat R, et al. (2003) Screening for Familial Dysautonomia in Israel: Evidence for Higher Carrier Rate Among Polish Ashkenazi Jews. Genet Test 7: 139–142.

33. Gifford AL, Cunningham WE, Heslin KC, Andersen RM, Nakazono T, et al. (2002) Participation in Research and Access to Experimental Treatments by HIV-infected Patients. N Engl J Med 346: 1373–1382.

34. Sugarman EA, Rohlfs EM, Silverman LM, Allitto BA (2004) CFTR Mutation Distribution Among U.S. Hispanic and African American Individuals: Evaluation in Cystic Fibrosis Patient and Carrier Screening Populations. Genet Med 6: 392–399.

35. Gardner DK, Weissman A, Howles CM, Shoham Z (2004) Textbook of Assisted Reproductive Techniques: Laboratory and Clinical Perspectives. Informa HealthCare, second edition.

36. Rechitsky S, Verlinsky O, Kuliev A, Ozen RS, Masciangelo C, et al. (2003) Preimplantation Genetic Diagnosis for Familial Dysautonomia. Reprod Biomed Online 6: 488–493.

37. Sermon K, Van Steirteghem A, Liebaers I (2004) Preimplantation Genetic Diagnosis. Lancet 363: 1633–1641.

38. Handyside AH, Lesko JG, Tarn JJ, Winston RM, Hughes MR (1992) Birth of a Normal Girl After in Vitro Fertilization and Preimplantation Diagnostic Testing for Cystic Fibrosis. N Engl J Med 327: 905–909.

39. Unsal E, Akta Y, Uner O, BaltacI A, Ozcan S, et al. (2008) Successful Application of Preimplantation Genetic Diagnosis for Leigh Syndrome. Fertil Steril 90: 11–13.

40. Munn S, Magli C, Cohen J, Morton P, Sadowy S, et al. (1999) Positive Outcome After Preimplantation Diagnosis of Aneuploidy in Human Embryos. Hum Reprod 14: 2191–2199.

41. Jones SL, Fallon LA (2002) Reproductive Options for Individuals at Risk for Transmission of a Genetic Disorder. J Obstet Gynecol Neonatal Nurs 31: 193–199.

42. Blakeslee S (1984) Infertile Woman Has Baby Through Embryo Transfer. NY Times .

43. Gilbert-Barness E, Barness LA (2000) Metabolic Diseases: Foundations of Clinical Management, Genetics and Pathology. Westborough, MA: Eaton Publishing Company.

44. Cabalska MB, Nowaczewska I, Sendecka E, Zorska K (1996) Longitudinal Study on Early Diagnosis and Treatment of Phenylketonuria in Poland. Eur J Pediatr 155 Suppl 1: S53–55.

45. Axelrod FB, Goldberg JD, Ye XY, Maayan C (2002) Survival in Familial Dysautonomia: Impact of Early Intervention. J Pediatr 141: 518–523.

46. Ekstein J, Katzenstein H (2001) The Dor Yeshorim Story: Community-based Carrier Screening for Tay-Sachs Disease. Adv Genet 44: 297–310.

47. Kaback M, Lim-Steele J, Dabholkar D, Brown D, Levy N, et al. (1993) Tay-Sachs Disease–carrier Screening, Prenatal Diagnosis, and the Molecular Era. An International Perspective, 1970 to 1993. The International TSD Data Collection Network. JAMA 270: 2307–2315.

48. Mitchell JJ, Capua A, Clow C, Scriver CR (1996) Twenty-year Outcome Analysis of Genetic Screening Programs for Tay-Sachs and Beta-thalassemia Disease Carriers in High Schools. Am J Hum Genet 59: 793–798.

49. Watson MS, Cutting GR, Desnick RJ, Driscoll DA, Klinger K, et al. (2004) Cystic Fibrosis Population Carrier Screening: 2004 Revision of American College of Medical Genetics Mutation Panel. Genet Med 6: 387–391.

50. Merck Carrier Screening Guidelines. URL `www.merck.com/mmhe/sec22/ch256/ch256b.html`.

51. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, et al. (2001) Laboratory Standards and Guidelines for Population-based Cystic Fibrosis Carrier Screening. Genetics in Medicine 3: 149–154.

52. Grody WW, Desnick RJ (2001) Cystic Fibrosis Population Carrier Screening: Here at Last–Are We Ready? Genetics in Medicine 3: 87–90.

53. NIH (1999) Genetic Testing for Cystic Fibrosis. April 14–16, 1997. Arch Intern Med 159: 1529–1539.

54. College News (2008) Genetics in Medicine 10: 917–918.

55. Prior TW (2008) Carrier Screening for Spinal Muscular Atrophy. Genet Med 10: 840–842.

56. Wei S, Quigg MH, Monaghan KG (2007) Is Cystic Fibrosis Carrier Screening Cost Effective? Community Genet 10: 103–109.

57. Wildhagen MF, van Os TA, Polder JJ, ten Kate LP, Habbema JD (1998) Explorative Study of Costs, Effects and Savings of Screening for Female Fragile X Premutation and Full Mutation Carriers in the General Population. Community Genet 1: 36–47.

58. Tur-Kaspa I, Rechitsky S, Aljadeff G, Grotjan E, Verlinsky Y (2006) O-139: Preimplantation genetic diagnosis (PGD) for all cystic fibrosis (CF) carrier couples: Strategy and cost analysis. Fertility and Sterility 86: S59 - S59.

59. Samavat A, Modell B (2004) Iranian National Thalassaemia Screening Programme. BMJ 329: 1134–1137.

60. Calonge N (2008) The Genomic Era: A Public Health Perspective. NHGRI.

61. Pletcher BA, Gross SJ, Monaghan KG, Driscoll DA, Watson MS (2008) The Future Is Now: Carrier Screening for All Populations. Genet Med 10: 33–36.

62. MacNeil JS (2005) Offer Early Genetic Screening, Prompt Disclosure. OB/GYN News .

63. Khoury MJ, McCabe LL, McCabe ER (2003) Population Screening in the Age of Genomic Medicine. N Engl J Med 348: 50–58.

64. Jones S (2000) Genetics in Medicine: Real Promises, Unreal Expectations. One Scientist's Advice to Policymakers in the United Kingdom and the United States. Technical report, Milbank Memorial Fund. URL `http://www.milbank.org/000712genetics.html`.

65. Cunningham G (2002) The Science and Politics of Screening Newborns. N Engl J Med 346: 1084–1085.

66. Bekker H, Denniss G, Modell M, Bobrow M, Marteau T (1994) The Impact of Population Based Screening for Carriers of Cystic Fibrosis. J Med Genet 31: 364–368.

67. Jeffreys AJ, May CA (2003) DNA Enrichment by Allele-specific Hybridization (DEASH): A Novel Method for Haplotyping and for Detecting Low-frequency Base Substitutional Variants and Recombinant DNA Molecules. Genome Res 13: 2316–2324.

68. Druley TE, Vallania FL, Wegner DJ, Varley KE, Knowles OL, et al. (2009) Quantification of Rare Allelic Variants From Pooled Genomic DNA. Nat Methods 6: 263–265.

69. Applied Biosystems (2007). Targeted Resequencing and Rare Allele Detection.

70. Owen AB (2007) Infinitely Imbalanced Logistic Regression. Journal of Machine Learning Research 8: 761–773.

71. Lin J, Kaltoft MS, Brandao AP, Echaniz-Aviles G, Brandileone MC, et al. (2006) Validation of a Multiplex Pneumococcal Serotyping Assay With Clinical Samples. J Clin Microbiol 44: 383–388.

72. Strom CM, Janeszco R, Quan F, Wang SB, Buller A, et al. (2006) Technical Validation of a TM Biosciences Luminex-based Multiplex Assay for Detecting the American College of Medical Genetics Recommended Cystic Fibrosis Mutation Panel. J Mol Diagn 8: 371–375.

73. Weidmann M, Armbruster K, Hufert FT (2008) Challenges in Designing a Taqman-based Multiplex Assay for the Simultaneous Detection of Herpes Simplex Virus Types 1 and 2 and Varicella-zoster Virus. J Clin Virol 42: 326–334.

74. Daly TM, Dumaual CM, Miao X, Farmen MW, Njau RK, et al. (2007) Multiplex Assay for Comprehensive Genotyping of Genes Involved in Drug Metabolism, Excretion, and Transport. Clin Chem 53: 1222–1230.

75. Nakai K, Habano W, Fujita T, Nakai K, Schnackenberg J, et al. (2002) Highly Multiplexed Genotyping of Coronary Artery Disease-associated SNPs Using MALDI-TOF Mass Spectrometry. Hum Mutat 20: 133–138.

76. McGhee SA, Stiehm ER, McCabe ER (2005) Potential Costs and Benefits of Newborn Screening for Severe Combined Immunodeficiency. J Pediatr 147: 603–608.

77. Elsas L (2008) An Overview of Issues Facing Carrier Screening in Large Populations. NHGRI.

78. Hoffman E (2008) An Update on Technologies Relevant to Carrier Screening. NHGRI.

79. Rogowski WH, Grosse SD, Khoury MJ (2009) Challenges of Translating Genetic Tests Into Clinical and Public Health Practice. Nat Rev Genet 10: 489–495.

80. Gason AA, Sheffield E, Bankier A, Aitken MA, Metcalfe S, et al. (2003) Evaluation of a Tay-Sachs Disease Screening Program. Clin Genet 63: 386–392.

81. Mefford HC, Cooper GM, Zerr T, Smith JD, Baker C, et al. (2009) A Method for Rapid, Targeted CNV Genotyping Identifies Rare Variants Associated With Neurocognitive Disease. Genome Res 19: 1579–1585.

82. Anderson C (2006) The Long Tail: Why the Future of Business Is Selling Less of More. 77 West 66th Street, New York, New York 10023: Hyperion.

83. McConkey E (1993) Human Genetics: The Molecular Revolution. Sudbury, MA: Jones and Bartlett, First edition.

84. Guttmacher AE, Collins FS (2002) Genomic Medicine: A Primer. The New England Journal of Medicine 347: 1512–1520.

85. Reich DE, Lander ES (2001) On the Allelic Spectrum of Human Disease. Trends Genet 17: 502–510.

86. Coriell Institute for Medical Research. URL `http://www.coriell.org/`.

87. Christensen TM, Jama M, Ponek V, Lyon E, Wilson JA, et al. (2007) Design, Development, Validation, and Use of Synthetic Nucleic Acid Controls for Diagnostic Purposes and Application to Cystic Fibrosis Testing. The Journal of Molecular Diagnostics 9: 315–9.

88. Berwouts S, Gordon JT, Rundell CA, Barton DE, Dequeker E (2008) Evaluation and Use of a Synthetic Quality Control Material, Included in the European External Quality Assessment Scheme for Cystic Fibrosis. Hum Mutat 29: 1063–70.

89. Blue Heron Biotechnology. URL `http://blueheronbio.com/`.

90. Bach G, Tomczak J, Risch N, Ekstein J (2001) Tay-Sachs Screening in the Jewish Ashkenazi Population: DNA Testing is the Preferred Procedure. American Journal of Medical Genetics 99: 70–75.

91. Zielenski J, Patrizio P, Corey M, Handelin B, Markiewicz D, et al. (1995) CFTR Gene Variant for Patients with Congenital Absence of Vas Deferens. American Journal of Human Genetics 57: 958–960.

92. Map of Counsyl Clinics. URL `https://counsyl.com/map`.

93. A Thin Blue Line: The History of the Pregnancy Test Kit. URL `http://history.nih.gov/exhibits/thinblueline/`.

94. Litholink Mail-in Urine Sample Collection Process. URL `http://www.litholink.com/downloads/Stone_FAQAboutCollection.pdf`.

95. Hicks JM, Iosefsohn M (1989) Reliability of Home Pregnancy-test kits in the Hands of Laypersons. The New England journal of medicine 320: 320–321.

96. Hicks JM Home Testing: to Do or Not to Do?, url = http://view.ncbi.nlm.nih.gov/pubmed/8419061, volume = 39, year = 1993. Clinical chemistry : 7–8.

97. McGuire AL, Burke W (2008) An Unwelcome Side Effect of Direct-to-Consumer Personal Genome Testing: Raiding the Medical Commons. JAMA 300: 2669–2671.

98. Goldstein DB (2009) Common Genetic Variation and Human Traits. N Engl J Med 360: 1696–1698.

99. Iwasiow R, Desbois A, Birnboim H (2004) Long-term Stability of DNA From Saliva Samples Stored in OrageneDNA. Technical report.

100. Rylander-Rudqvist T, Hakansson N, Tybring G, Wolk A (2006) Quality and Quantity of Saliva DNA Obtained From the Self-administrated Oragene Method–A Pilot Study on the Cohort of Swedish Men. Cancer Epidemiol Biomarkers Prev 15: 1742-1745.

101. Genotek D (2004) Long-term Storage of OrageneDNA Samples. Technical report.

102. Cormier TA, Tran K, Forman JE, Jackson A, Lem P, et al. SNP Genotyping of Saliva DNA Using Affymetrix GeneChip Targeted Genotyping System .

103. Commerce Departments' NTIA And USDA'S RUS Announce Availability of $4.8 Billion In Recovery Act Funding To Bring Broadband To More Americans. URL `http://www.ntia.doc.gov/press/2010/BTOP_BIP_NOFAII_100115.html`.

104. Though Wary of DTC Testing, Some Clinical Geneticists See Opportunity for Education. URL `genomeweb.com/dxpgx/though-wary-dtc-testing-some-clinical-geneticists-see-opportunity-education`.

105. Magnus D, Cho MK, Cook-Deegan R (2009) Direct-to-Consumer Genetic Tests: Beyond Medical Regulation? Genome Medicine 1.

106. Evans JP, Green RC (2009) Direct to Consumer Genetic Testing: Avoiding a Culture War. Genetics in medicine 11: 568–569.

107. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, et al. (2003) Human Gene Mutation Database (HGMD): 2003 Update. Hum Mutat 21: 577–581.

108. Cystic Fibrosis Mutation Database. URL `http://www.genet.sickkids.on.ca/cftr/app`.

109. Population-based Carrier Screening for Single Gene Disorders (Feb. 2008). URL `www.genome.gov/27026048`.

110. Counsyl Offers Financial Aid. URL `https://www.counsyl.com/order/aid`.

111. Hardenbol P, Baner J, Jain M, Nilsson M, Namsaraev EA, et al. (2003) Multiplexed Genotyping With Sequence-tagged Molecular Inversion Probes. Nat Biotech 21: 673–678.

112. Hardenbol P, Yu F, Belmont J, Mackenzie J, Bruckner C, et al. (2005) Highly Multiplexed Molecular Inversion Probe Genotyping: Over 10,000 Targeted SNPs Genotyped in a Single Tube Assay. Genome Research 15: 269–75.

113. Moorhead M, Hardenbol P, Siddiqui F, Falkowski M, Bruckner C, et al. (2006) Optimal Genotype Determination in Highly Multiplexed SNP Data. Eur J Hum Genet 14: 207–15.

114. Daly TM, Dumaual CM, Miao X, Farmen MW, Njau RK, et al. (2007) Multiplex Assay for Comprehensive Genotyping of Genes Involved in Drug Metabolism, Excretion, and Transport. Clin Chem 53: 1222–30.

115. Wang Y, Moorhead M, Karlin-Neumann G, Wang NJ, Ireland J, et al. (2007) Analysis of Molecular Inversion Probe Performance for Allele Copy Number Determination. Genome Biology 8: R246.

Supplementary Tables and Figures

Table 6. Variant Statistics. Performance statistics for the disease causing variants of the UNIT. Column headers are NP (number of samples that tested positive), NN (number of samples that tested negative), ACC (accuracy), PRC (precision). Note that the vast majority of variants have digital accuracy, corresponding to highly separated clusters of the form seen in Figure 7. The variants that had one error are FANCC 322delG, HBB CAP+1 A>C, HBB Glu6fs, and MEFV V726A; these errors appear to be caused by unusually low titration ratios during sample preparation of synthetic heterozygotes (see Materials and Methods) rather than by intrinsic assay properties.

Gene	Variant	NP	NN	ACC	PRC
ABCC8	3992-9G>A	147	525	1.000	1.000
ABCC8	F1388del	147	525	1.000	1.000
ABCC8	V187D	147	525	1.000	1.000
ACADM	G170R	147	525	1.000	1.000
ACADM	G242R	147	525	1.000	1.000
ACADM	K304E	154	518	1.000	1.000
ACADM	L59F	73	525	1.000	1.000
ACADM	R181C	74	525	1.000	1.000
ACADM	R181H	73	525	1.000	1.000
ACADM	Y42H	147	525	1.000	1.000
ACADS	G185S	238	287	1.000	1.000
ACADS	R107C	150	522	1.000	1.000
AGA	199_200delGA	147	525	1.000	1.000
AGA	C163S	147	525	1.000	1.000
AGL	1484delT	147	525	1.000	1.000
AGL	17delAG	73	525	1.000	1.000
AGL	Q6X	74	525	1.000	1.000
AGXT	F152I	147	525	1.000	1.000
AGXT	G170R	147	525	1.000	1.000
AGXT	I244T	147	525	1.000	1.000
AIRE	R257X	147	525	1.000	1.000
AIRE	Y85C	147	525	1.000	1.000
ALDH3A2	P315S	147	525	1.000	1.000
ALDOB	A149P	153	519	1.000	1.000
ALDOB	Delta4E4	147	525	1.000	1.000
ALDOB	N334K	147	525	1.000	1.000
ALDOB	Y204X	147	525	1.000	1.000
ALPL	1559delT	147	525	1.000	1.000
ALPL	D361V	147	525	1.000	1.000
ALPL	E174K	147	525	1.000	1.000
ALPL	F310L	147	525	1.000	1.000
ALPL	G317D	147	525	1.000	1.000
ARSA	IVS2+1G>A	147	525	1.000	1.000
ARSA	P377L	147	525	1.000	1.000
ARSA	P426L	147	525	1.000	1.000
ARSA	T274M	147	525	1.000	1.000
ASPA	A305E	150	522	1.000	1.000
ASPA	E285A	155	517	1.000	1.000
ASPA	IVS2-2A>G	147	525	1.000	1.000
ASPA	Y231X	151	521	1.000	1.000
ATM	R35X	150	522	1.000	1.000
ATP7B	1340del4	147	525	1.000	1.000
ATP7B	2337delC	147	525	1.000	1.000
ATP7B	H1069Q	147	525	1.000	1.000
ATP7B	R778G	48	525	1.000	1.000
ATP7B	W779X	47	525	1.000	1.000
BBS1	M390R	151	521	1.000	1.000
BBS10	C91fs	149	523	1.000	1.000
BCKDHB	E322X	147	525	1.000	1.000
BCKDHB	G278S	147	525	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
BCKDHB	R183P	149	523	1.000	1.000
BCS1L	S78G	150	522	1.000	1.000
BLM	2281del6ins7	150	522	1.000	1.000
BLM	2407insT	147	525	1.000	1.000
BTD	A171T	147	525	1.000	1.000
BTD	D252G	147	525	1.000	1.000
BTD	D444H	97	502	1.000	1.000
BTD	F403V	147	525	1.000	1.000
BTD	G98:d7i3	147	525	1.000	1.000
BTD	Q456H	74	524	1.000	1.000
BTD	R538C	147	525	1.000	1.000
CBS	G307S	147	525	1.000	1.000
CBS	I278T	147	525	1.000	1.000
CFTR	1078delT	77	522	1.000	1.000
CFTR	1161delC	18	525	1.000	1.000
CFTR	1288insTA	147	525	1.000	1.000
CFTR	1609delCA	13	525	1.000	1.000
CFTR	1677delTA	9	525	1.000	1.000
CFTR	1717-1G>A	28	522	1.000	1.000
CFTR	1811+1.6kbA>G	147	525	1.000	1.000
CFTR	1812-1G>A	150	522	1.000	1.000
CFTR	1898+1G>A	55	522	1.000	1.000
CFTR	1898+1G>T	47	525	1.000	1.000
CFTR	1898+5G>T	48	525	1.000	1.000
CFTR	1949del84	147	525	1.000	1.000
CFTR	2043delG	74	525	1.000	1.000
CFTR	2055del9>A	73	525	1.000	1.000
CFTR	2105-2117del13insAGAAA	22	525	1.000	1.000
CFTR	2183AA>G	33	518	1.000	1.000
CFTR	2184delA	29	521	1.000	1.000
CFTR	2184insA	22	525	1.000	1.000
CFTR	2307insA	147	525	1.000	1.000
CFTR	2789+5G>A	150	522	1.000	1.000
CFTR	2869insG	147	525	1.000	1.000
CFTR	296+12T>C	147	525	1.000	1.000
CFTR	3120+1G>A	79	520	1.000	1.000
CFTR	3120G>A	73	525	1.000	1.000
CFTR	3171delC	74	525	1.000	1.000
CFTR	3199del6	73	525	1.000	1.000
CFTR	3272-26A>G	152	520	1.000	1.000
CFTR	3659delC	77	522	1.000	1.000
CFTR	3667del4	73	525	1.000	1.000
CFTR	3821delT	35	525	1.000	1.000
CFTR	3849+10kbC>T	152	520	1.000	1.000
CFTR	3876delA	41	523	1.000	1.000
CFTR	3905insT	41	522	1.000	1.000
CFTR	394delTT	24	522	1.000	1.000
CFTR	405+1G>A	22	525	1.000	1.000
CFTR	405+3A>C	18	525	1.000	1.000
CFTR	406-1G>A	21	525	1.000	1.000
CFTR	457TAT>G	22	525	1.000	1.000
CFTR	574delA	73	525	1.000	1.000
CFTR	621+1G>T	162	510	1.000	1.000
CFTR	663delT	73	525	1.000	1.000
CFTR	711+1G>T	77	522	1.000	1.000
CFTR	711+5G>A	73	525	1.000	1.000
CFTR	712-1G>T	74	525	1.000	1.000
CFTR	935delA	74	525	1.000	1.000
CFTR	936delTA	73	525	1.000	1.000
CFTR	A455E	150	522	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
CFTR	A559T	9	525	1.000	1.000
CFTR	C524X	9	525	1.000	1.000
CFTR	D1152H	151	521	1.000	1.000
CFTR	E60X	151	521	1.000	1.000
CFTR	E92X	22	525	1.000	1.000
CFTR	F311del	73	525	1.000	1.000
CFTR	F508C	18	520	1.000	1.000
CFTR	F508del	77	461	1.000	1.000
CFTR	G178R	78	521	1.000	1.000
CFTR	G330X	17	525	1.000	1.000
CFTR	G480C	13	525	1.000	1.000
CFTR	G542X	16	522	1.000	1.000
CFTR	G551D	18	520	1.000	1.000
CFTR	G622D	147	525	1.000	1.000
CFTR	G85E	29	518	1.000	1.000
CFTR	G91R	21	525	1.000	1.000
CFTR	I148T	77	522	1.000	1.000
CFTR	I506V	39	525	1.000	1.000
CFTR	I507del	15	522	1.000	1.000
CFTR	IVS8-5T	37	488	1.000	1.000
CFTR	K710X	73	525	1.000	1.000
CFTR	L206W	73	525	1.000	1.000
CFTR	M1101K	41	522	1.000	1.000
CFTR	N1303K	77	522	1.000	1.000
CFTR	P574H	150	522	1.000	1.000
CFTR	Q1238X	35	525	1.000	1.000
CFTR	Q493X	17	521	1.000	1.000
CFTR	Q552X	9	525	1.000	1.000
CFTR	Q890X	147	525	1.000	1.000
CFTR	R1066C	74	525	1.000	1.000
CFTR	R1070Q	73	525	1.000	1.000
CFTR	R1158X	76	522	1.000	1.000
CFTR	R1162X	79	520	1.000	1.000
CFTR	R117C	18	525	1.000	1.000
CFTR	R117H	25	522	1.000	1.000
CFTR	R1283M	73	525	1.000	1.000
CFTR	R334W	24	522	1.000	1.000
CFTR	R347H	23	522	1.000	1.000
CFTR	R347P	20	522	1.000	1.000
CFTR	R352Q	18	525	1.000	1.000
CFTR	R553X	16	522	1.000	1.000
CFTR	R560T	17	521	1.000	1.000
CFTR	R709X	74	525	1.000	1.000
CFTR	R75X	46	522	1.000	1.000
CFTR	S1196X	52	525	1.000	1.000
CFTR	S1235R	55	509	1.000	1.000
CFTR	S1251N	35	525	1.000	1.000
CFTR	S1255X	35	525	1.000	1.000
CFTR	S364P	18	525	1.000	1.000
CFTR	S549I	13	525	1.000	1.000
CFTR	S549N	16	522	1.000	1.000
CFTR	S549R(A>C)	13	525	1.000	1.000
CFTR	S549R(T>G)	16	522	1.000	1.000
CFTR	T338I	18	525	1.000	1.000
CFTR	V520F	16	522	1.000	1.000
CFTR	W1089X	35	525	1.000	1.000
CFTR	W1204X(c.3611G>A)	47	525	1.000	1.000
CFTR	W1204X(c.3612G>A)	48	525	1.000	1.000
CFTR	W1282X	79	520	1.000	1.000
CFTR	Y1092X	79	520	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
CFTR	Y122X	3	522	1.000	1.000
CFTR	dele2-3 21kb	147	525	1.000	1.000
CHM	IVS13+2dupT	147	525	1.000	1.000
CLN5	2467AT	147	525	1.000	1.000
CLN8	R24G	147	525	1.000	1.000
CLRN1	N48K	150	522	1.000	1.000
CNGB3	819_826del8	147	525	1.000	1.000
CNGB3	886-896del11insT	148	524	1.000	1.000
CNGB3	E336X	74	525	1.000	1.000
CNGB3	IVS8-3T>G	73	525	1.000	1.000
CNGB3	R403Q	147	525	1.000	1.000
CNGB3	T383fs	154	518	1.000	1.000
CPT1A	G710E	147	525	1.000	1.000
CPT1A	P479L	147	525	1.000	1.000
CPT2	G549D	74	525	1.000	1.000
CPT2	Leu178_Ile186delinsPhe	147	525	1.000	1.000
CPT2	P227L	147	525	1.000	1.000
CPT2	P50H	73	525	1.000	1.000
CPT2	P604S	147	525	1.000	1.000
CPT2	Q413fs	147	525	1.000	1.000
CPT2	Q550R	73	525	1.000	1.000
CPT2	R124X	147	525	1.000	1.000
CPT2	R503C	147	525	1.000	1.000
CPT2	R631C	73	525	1.000	1.000
CPT2	S113L	152	520	1.000	1.000
CPT2	S38fs	74	525	1.000	1.000
CPT2	Y628S	74	525	1.000	1.000
CTNS	537del21	147	525	1.000	1.000
CTNS	D205N	147	525	1.000	1.000
CTNS	L158P	147	525	1.000	1.000
CTNS	W138X	147	525	1.000	1.000
CTSK	X330W	147	525	1.000	1.000
DHCR7	C380Y	147	525	1.000	1.000
DHCR7	IVS8-1G>C	98	500	1.000	1.000
DHCR7	L109P	147	525	1.000	1.000
DHCR7	L157P	22	525	1.000	1.000
DHCR7	R352Q	74	525	1.000	1.000
DHCR7	R352W	73	525	1.000	1.000
DHCR7	R404C	147	525	1.000	1.000
DHCR7	T93M	147	525	1.000	1.000
DHCR7	V326L	74	525	1.000	1.000
DHCR7	W151X(c.452G>A)	65	521	1.000	1.000
DHCR7	W151X(c.453G>A)	64	525	1.000	1.000
DLD	105insA	147	525	1.000	1.000
DLD	G229C	147	525	1.000	1.000
DPYD	IVS14+1G>A	150	522	1.000	1.000
F11	E117X	147	525	1.000	1.000
F11	F283L	149	523	1.000	1.000
F11	IVS14+1G>A	70	525	1.000	1.000
F11	IVS14del14	77	525	1.000	1.000
F5	D2222G	210	462	1.000	1.000
F5	H1299R	215	457	1.000	1.000
F5	R506Q	33	492	1.000	1.000
FAH	E357X	147	525	1.000	1.000
FAH	IVS12+5G>A	151	521	1.000	1.000
FAH	IVS8-1G>C	147	525	1.000	1.000
FAH	P261L	79	520	1.000	1.000
FAH	W262X	73	525	1.000	1.000
FANCC	322delG	76	523	0.987	1.000
FANCC	IVS4+4A>T	147	525	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
FANCC	Q13X	73	525	1.000	1.000
FANCC	R548X	147	525	1.000	1.000
FH	1431_1433dupAAA	151	521	1.000	1.000
G6PC	459insTA	147	525	1.000	1.000
G6PC	727G>T	147	525	1.000	1.000
G6PC	F327del	147	525	1.000	1.000
G6PC	G188R	147	525	1.000	1.000
G6PC	G270V	147	525	1.000	1.000
G6PC	Q242X	147	525	1.000	1.000
G6PC	Q27fsdelC	147	525	1.000	1.000
G6PC	Q347X	3	522	1.000	1.000
G6PC	R83C	78	521	1.000	1.000
G6PC	R83H	73	525	1.000	1.000
G6PD	N126D	147	525	1.000	1.000
G6PD	R459L	73	525	1.000	1.000
G6PD	R459P	74	525	1.000	1.000
G6PD	S188F	152	520	1.000	1.000
G6PD	V68M	147	525	1.000	1.000
G6PT1	1211delCT	147	525	1.000	1.000
G6PT1	A367T	147	525	1.000	1.000
G6PT1	G339C	73	525	1.000	1.000
G6PT1	G339D	74	525	1.000	1.000
GAA	D645E	147	525	1.000	1.000
GALC	Ex11-17del	149	523	1.000	1.000
GALC	G270D	147	525	1.000	1.000
GALC	R168C	208	464	1.000	1.000
GALT	F171S	150	522	1.000	1.000
GALT	IVS2-2A>G	147	525	1.000	1.000
GALT	K285N	150	522	1.000	1.000
GALT	L195P	150	522	1.000	1.000
GALT	Q169K	147	525	1.000	1.000
GALT	Q188R	156	516	1.000	1.000
GALT	S135L	77	522	1.000	1.000
GALT	T138M	73	525	1.000	1.000
GALT	X380R	147	525	1.000	1.000
GALT	Y209C	74	525	1.000	1.000
GBA	1035insG	76	522	1.000	1.000
GBA	D409V	74	525	1.000	1.000
GBA	IVS2+1G>A	59	667	1.000	1.000
GBA	L444P	71	633	1.000	1.000
GBA	N370S	166	506	1.000	1.000
GBA	R463C	73	525	1.000	1.000
GBA	R463H	74	525	1.000	1.000
GBA	R496H	147	525	1.000	1.000
GBA	V394L	150	522	1.000	1.000
GCDH	A421V	147	525	1.000	1.000
GCDH	R402W	147	525	1.000	1.000
GJB2	167delT	150	522	1.000	1.000
GJB2	313del14	150	522	1.000	1.000
GJB2	35delG	39	512	1.000	1.000
GJB2	E120del	73	525	1.000	1.000
GJB2	M34T	44	507	1.000	1.000
GJB2	Q124X	74	525	1.000	1.000
GJB2	R184P	147	525	1.000	1.000
GJB2	V37I	28	522	1.000	1.000
GJB2	W24X	70	525	1.000	1.000
GJB2	W77R	48	525	1.000	1.000
GJB2	W77X	47	525	1.000	1.000
GNE	M712T	147	525	1.000	1.000
GRHPR	103delG	147	525	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
HADHA	E474Q	150	522	1.000	1.000
HADHA	Q342X	147	525	1.000	1.000
HBB	-28A>G	52	525	1.000	1.000
HBB	-29A>G	47	525	1.000	1.000
HBB	-30T>A	48	525	1.000	1.000
HBB	-87C>G	77	522	1.000	1.000
HBB	-88C>T	73	525	1.000	1.000
HBB	619 bp deletion	147	525	1.000	1.000
HBB	CAP+1 A>C	148	524	1.000	0.997
HBB	Glu6fs	17	520	1.000	0.992
HBB	Gly16fs	4	525	1.000	1.000
HBB	Gly24 T>A	18	525	1.000	1.000
HBB	Hb C	27	522	1.000	1.000
HBB	Hb D-Punjab	74	525	1.000	1.000
HBB	Hb E	26	525	1.000	1.000
HBB	Hb O-Arab	73	525	1.000	1.000
HBB	Hb S	14	522	1.000	1.000
HBB	IVS-I-1(G>A)	29	522	1.000	1.000
HBB	IVS-I-1(G>T)	26	525	1.000	1.000
HBB	IVS-I-110	153	519	1.000	1.000
HBB	IVS-I-5	26	525	1.000	1.000
HBB	IVS-I-6	31	519	1.000	1.000
HBB	IVS-II-654	152	520	1.000	1.000
HBB	IVS-II-705	73	525	1.000	1.000
HBB	IVS-II-745	77	522	1.000	1.000
HBB	IVS-II-844	48	525	1.000	1.000
HBB	IVS-II-849(A>C)	52	525	1.000	1.000
HBB	IVS-II-849(A>G)	47	525	1.000	1.000
HBB	IVS-II-850	76	522	1.000	1.000
HBB	K17X	4	525	1.000	1.000
HBB	Lys8fs	10	525	1.000	1.000
HBB	Phe41fs	73	525	1.000	1.000
HBB	Phe71fs	152	520	1.000	1.000
HBB	Poly A: AATAAA->AATAAG	74	525	1.000	1.000
HBB	Poly A: AATAAA->AATGAA	73	525	1.000	1.000
HBB	Pro5fs	24	525	1.000	1.000
HBB	Q39X	80	519	1.000	1.000
HBB	Ser9fs	25	525	1.000	1.000
HBB	W15X	5	525	1.000	1.000
HEXA	1278insTATC	150	522	1.000	1.000
HEXA	G269S	84	515	1.000	1.000
HEXA	IVS12+1G>C	157	515	1.000	1.000
HEXA	IVS7+1G>A	73	525	1.000	1.000
HEXA	IVS9+1G>A	154	518	1.000	1.000
HEXA	R178C	73	525	1.000	1.000
HEXA	R178H	77	522	1.000	1.000
HEXA	R247W	78	521	1.000	1.000
HFE	C282Y	92	485	1.000	1.000
HFE	E168Q	38	525	1.000	1.000
HFE	E168X	35	525	1.000	1.000
HFE	H63D	127	398	1.000	1.000
HFE	H63H	36	525	1.000	1.000
HFE	Q127H	147	525	1.000	1.000
HFE	Q283P	95	525	1.000	1.000
HFE	S65C	45	517	1.000	1.000
HFE	V53M	17	525	1.000	1.000
HFE	V59M	36	525	1.000	1.000
HFE	W169X	35	525	1.000	1.000
HGD	G161R	152	520	1.000	1.000
HGD	G270R	147	525	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
HGD	IVS1-1G>A	147	525	1.000	1.000
HGD	IVS5+1G>A	147	525	1.000	1.000
HGD	M368V	147	525	1.000	1.000
HGD	P230S	147	525	1.000	1.000
HGD	S47L	73	525	1.000	1.000
IDUA	A327P	147	525	1.000	1.000
IDUA	W402X	147	525	1.000	1.000
IKBKAP	IVS20+6T>C	151	521	1.000	1.000
IKBKAP	P914L	147	525	1.000	1.000
IKBKAP	R696P	147	525	1.000	1.000
IVD	A311V	147	525	1.000	1.000
LAMA3	R650X	147	525	1.000	1.000
LAMB3	3024delT	147	525	1.000	1.000
LAMB3	Q243X	147	525	1.000	1.000
LAMB3	R144X	147	525	1.000	1.000
LAMB3	R42X	147	525	1.000	1.000
LAMB3	R635X	147	525	1.000	1.000
LAMC2	R95X	147	525	1.000	1.000
LRPPRC	A354V	147	525	1.000	1.000
MCOLN1	511_6944del	55	669	1.000	1.000
MCOLN1	IVS3-2A>G	150	522	1.000	1.000
MEFV	A744S	149	523	1.000	1.000
MEFV	F479L	147	525	1.000	1.000
MEFV	I692del	26	525	1.000	1.000
MEFV	K695R	43	508	1.000	1.000
MEFV	M680I	44	525	1.000	1.000
MEFV	M694I	25	525	1.000	1.000
MEFV	M694V	26	525	1.000	1.000
MEFV	P369S	153	519	1.000	1.000
MEFV	R408Q	152	520	1.000	1.000
MEFV	R653H	147	525	1.000	1.000
MEFV	R761H	147	525	1.000	1.000
MEFV	T267I	147	525	1.000	1.000
MEFV	V726A	148	524	1.000	0.997
MPI	R295H	147	525	1.000	1.000
MUTYH	Y165C	151	521	1.000	1.000
NBN	657del5	152	520	1.000	1.000
NPC1	I1061T	154	518	1.000	1.000
NPHS1	121_122del	147	525	1.000	1.000
NPHS1	R1109X	147	525	1.000	1.000
PAH	G272X	39	525	1.000	1.000
PAH	I65T	154	518	1.000	1.000
PAH	IVS-10int-546	147	525	1.000	1.000
PAH	IVS12+1G>A	149	523	1.000	1.000
PAH	L48S	147	525	1.000	1.000
PAH	R158Q	147	525	1.000	1.000
PAH	R252W	70	525	1.000	1.000
PAH	R261Q	38	525	1.000	1.000
PAH	R408Q	47	525	1.000	1.000
PAH	R408W	49	524	1.000	1.000
PAH	Y414C	57	520	1.000	1.000
PCDH15	R245X	147	525	1.000	1.000
PEX1	G843D	147	525	1.000	1.000
PEX7	G217R	74	525	1.000	1.000
PEX7	L292X	147	525	1.000	1.000
PKHD1	9689delA	147	525	1.000	1.000
PKHD1	Leu1965fs	147	525	1.000	1.000
PKHD1	R496X	147	525	1.000	1.000
PKHD1	T36M	147	525	1.000	1.000
PKHD1	V3471G	147	525	1.000	1.000

Continued on next page

Table 6 – continued from previous page

Gene	Variant	NP	NN	ACC	PRC
PMM2	F119L	147	525	1.000	1.000
PMM2	R141H	150	522	1.000	1.000
POMGNT1	IVS17+1G>A	147	525	1.000	1.000
PPT1	L10X	147	525	1.000	1.000
PPT1	R122W	147	525	1.000	1.000
PPT1	R151X	147	525	1.000	1.000
PPT1	T75P	147	525	1.000	1.000
PYGM	G204S	147	525	1.000	1.000
PYGM	K542T	74	525	1.000	1.000
PYGM	K542X	73	525	1.000	1.000
PYGM	R49X	154	518	1.000	1.000
RMRP	262G>T	147	525	1.000	1.000
RMRP	g.70A>G	147	525	1.000	1.000
RS1	E72K	73	525	1.000	1.000
RS1	G109R	147	525	1.000	1.000
RS1	G74V	74	525	1.000	1.000
SACS	5254C>T	147	525	1.000	1.000
SACS	6594delT	147	525	1.000	1.000
SERPINA1	S allele	34	491	1.000	1.000
SERPINA1	Z allele	155	517	1.000	1.000
SGCB	S114F	147	525	1.000	1.000
SLC12A6	R675X	147	525	1.000	1.000
SLC12A6	Thr813fsX813	152	520	1.000	1.000
SLC17A5	Leu336fsX13	147	525	1.000	1.000
SLC17A5	R39C	147	525	1.000	1.000
SLC25A15	F188del	147	525	1.000	1.000
SLC26A2	C653S	147	525	1.000	1.000
SLC26A2	IVS1+2T>C	147	525	1.000	1.000
SLC26A2	R178X	147	525	1.000	1.000
SLC26A2	R279W	148	524	1.000	1.000
SLC26A2	V340del	147	525	1.000	1.000
SLC26A4	E384G	147	525	1.000	1.000
SLC26A4	L236P	148	524	1.000	1.000
SLC26A4	T416P	147	525	1.000	1.000
SMN1	Exon 7 deletion	64	826	1.000	1.000
SMPD1	L302P	150	522	1.000	1.000
SMPD1	R496L	152	520	1.000	1.000
SMPD1	fsP330	50	640	1.000	1.000
TH	R233H	147	525	1.000	1.000
TPP1	G284V	147	525	1.000	1.000
TPP1	IVS5-1G>A	75	523	1.000	1.000
TPP1	IVS5-1G>C	74	525	1.000	1.000
TPP1	R208X	147	525	1.000	1.000
TTPA	744delA	147	525	1.000	1.000

Table 7. Genotype/phenotype association references for each variant in UNIT.

Gene	Variant	References
ABCC8	3992-9G>A	PMID: 7716548
ABCC8	F1388del	PMID: 8923011
ABCC8	V187D	PMID: 10334322
ACADM	G170R	PMID: 7929823
ACADM	G242R	PMID: 1684086
ACADM	K304E	PMID: 2393404
ACADM	L59F	PMID: 16291504

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
ACADM	R181C	PMID: 15832312
ACADM	R181H	PMID: 16291504
ACADM	Y42H	PMID: 11409868
ACADS	G185S	PMID: 8725270
ACADS	R107C	PMID: 1692038
AGA	199_200delGA	PMID: 7627186
AGA	C163S	PMID: 1904874
AGL	1484delT	PMID: 9412782
AGL	17delAG	PMID: 8755644
AGL	Q6X	PMID: 8755644
AGXT	F152I	PMID: 8101040
AGXT	G170R	PMID: 1703535
AGXT	I244T	PMID: 9192270
AIRE	R257X	PMID: 9398840
AIRE	Y85C	PMID: 10677297
ALDH3A2	P315S	PMID: 9204959
ALDOB	A149P	PMID: 3383242
ALDOB	Delta4E4	PMID: 2339710
ALDOB	N334K	PMID: 2336380
ALDOB	Y204X	PMID: 8438046 PMID: 15880727
ALPL	1559delT	PMID: 7833929
ALPL	D361V	PMID: 1409720
ALPL	E174K	PMID: 1409720
ALPL	F310L	PMID: 8954059
ALPL	G317D	PMID: 8406453
ARSA	IVS2+1G>A	PMID: 1670590
ARSA	P377L	PMID: 7866401
ARSA	P426L	PMID: 7866401
ARSA	T274M	PMID: 8104633
ASPA	A305E	PMID: 8023850
ASPA	E285A	PMID: 8252036
ASPA	IVS2-2A>G	PMID: 8023850
ASPA	Y231X	PMID: 8023850
ATM	R35X	PMID: 8968760
ATP7B	1340del4	PMID: 9311736
ATP7B	2337delC	PMID: 8298641
ATP7B	H1069Q	PMID: 8298641
ATP7B	R778G	PMID: 8533760
ATP7B	W779X	PMID: 8938442
BBS1	M390R	PMID: 12118255
BBS10	C91fs	PMID: 16582908
BCKDHB	E322X	PMID: 11509994
BCKDHB	G278S	PMID: 11509994
BCKDHB	R183P	PMID: 11509994
BCS1L	S78G	PMID: 12215968
BLM	2281del6ins7	PMID: 7585968
BLM	2407insT	PMID: 17407155
BTD	A171T	PMID: 10206677
BTD	D252G	PMID: 10400129
BTD	D444H	PMID: 10206677
BTD	F403V	PMID: 10400129
BTD	G98: d7i3	PMID: 7550325
BTD	Q456H	PMID: 9232193
BTD	R538C	PMID: 9099842
CBS	G307S	PMID: 7506602
CBS	I278T	PMID: 1301198

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
CFTR	1078delT	PMID: 1379211 genet.sickkids.on.ca
CFTR	1161delC	PMID: 9482579 genet.sickkids.on.ca
CFTR	1288insTA	PMID: 15365999 genet.sickkids.on.ca
CFTR	1609delCA	PMID: 1284477 genet.sickkids.on.ca
CFTR	1677delTA	PMID: 1710601 genet.sickkids.on.ca
CFTR	1717-1G>A	PMID: 2236053 genet.sickkids.on.ca
CFTR	1811+1.6kbA>G	PMID: 7534040 genet.sickkids.on.ca
CFTR	1812-1G>A	PMID: 7517264 genet.sickkids.on.ca
CFTR	1898+1G>A	PMID: 1284540 genet.sickkids.on.ca
CFTR	1898+1G>T	PMID: 7537147 genet.sickkids.on.ca
CFTR	1898+5G>T	PMID: 7543385 genet.sickkids.on.ca
CFTR	1949del84	PMID: 1373934 genet.sickkids.on.ca
CFTR	2043delG	PMID: 1379210 genet.sickkids.on.ca
CFTR	2055del9>A	PMID: 9298826 genet.sickkids.on.ca
CFTR	2105-2117del13insAGAAA	PMID: 11668613 genet.sickkids.on.ca
CFTR	2183AA>G	PMID: 7513889 genet.sickkids.on.ca
CFTR	2184delA	PMID: 7525963 genet.sickkids.on.ca
CFTR	2184insA	PMID: 7525450 genet.sickkids.on.ca
CFTR	2307insA	PMID: 7686423 genet.sickkids.on.ca
CFTR	2789+5G>A	PMID: 15698946 genet.sickkids.on.ca
CFTR	2869insG	PMID: 1373935 genet.sickkids.on.ca
CFTR	296+12T>C	PMID: 9482579 genet.sickkids.on.ca
CFTR	3120+1G>A	PMID: 9150159 genet.sickkids.on.ca
CFTR	3120G>A	genet.sickkids.on.ca PMID: 16436646
CFTR	3171delC	PMID: 10794365 genet.sickkids.on.ca
CFTR	3199del6	PMID: 15371908 genet.sickkids.on.ca
CFTR	3272-26A>G	PMID: 1379210 genet.sickkids.on.ca
CFTR	3659delC	PMID: 2236053 genet.sickkids.on.ca
CFTR	3667del4	PMID: 7517264 genet.sickkids.on.ca

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
CFTR	3821delT	PMID: 1710600 genet.sickkids.on.ca
CFTR	3849+10kbC>T	PMID: 7521937 genet.sickkids.on.ca
CFTR	3876delA	PMID: 10777364 genet.sickkids.on.ca
CFTR	3905insT	PMID: 7525450 genet.sickkids.on.ca
CFTR	394delTT	PMID: 7691344 genet.sickkids.on.ca
CFTR	405+1G>A	PMID: 7506605 genet.sickkids.on.ca
CFTR	405+3A>C	PMID: 9150159 genet.sickkids.on.ca
CFTR	406-1G>A	PMID: 10798368 genet.sickkids.on.ca
CFTR	457TAT>G	PMID: 7691352 genet.sickkids.on.ca
CFTR	574delA	PMID: 1379210 genet.sickkids.on.ca
CFTR	621+1G>T	PMID: 1710599 genet.sickkids.on.ca
CFTR	663delT	PMID: 10993719 genet.sickkids.on.ca
CFTR	711+1G>T	PMID: 1710599 genet.sickkids.on.ca
CFTR	711+5G>A	PMID: 7526928 genet.sickkids.on.ca
CFTR	712-1G>T	PMID: 9439669 genet.sickkids.on.ca
CFTR	935delA	PMID: 10798368 genet.sickkids.on.ca
CFTR	936delTA	PMID: 8064813 genet.sickkids.on.ca
CFTR	A455E	PMID: 2236053 genet.sickkids.on.ca
CFTR	A559T	PMID: 1695717 genet.sickkids.on.ca
CFTR	C524X	PMID: 1284466 genet.sickkids.on.ca
CFTR	D1152H	PMID: 7739684 genet.sickkids.on.ca
CFTR	E60X	PMID: 1284534 genet.sickkids.on.ca
CFTR	E92X	PMID: 7512993 genet.sickkids.on.ca
CFTR	F311del	PMID: 7509232 genet.sickkids.on.ca
CFTR	F508C	PMID: 1977306 genet.sickkids.on.ca
CFTR	F508del	PMID: 2475911 genet.sickkids.on.ca
CFTR	G178R	PMID: 1710599 genet.sickkids.on.ca
CFTR	G330X	PMID: 9150159 genet.sickkids.on.ca
CFTR	G480C	PMID: 1284534 genet.sickkids.on.ca

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
CFTR	G542X	PMID: 2236053 genet.sickkids.on.ca
CFTR	G551D	PMID: 2236053 genet.sickkids.on.ca
CFTR	G622D	PMID: 9736778 genet.sickkids.on.ca
CFTR	G85E	PMID: 1710599 genet.sickkids.on.ca
CFTR	G91R	PMID: 7682984 genet.sickkids.on.ca
CFTR	I148T	PMID: 1284534 genet.sickkids.on.ca
CFTR	I506V	PMID: 1977306 genet.sickkids.on.ca
CFTR	I507del	PMID: 2236053 genet.sickkids.on.ca
CFTR	IVS8-5T	PMID: 12843327 PMID: 7739684 genet.sickkids.on.ca
CFTR	K710X	PMID: 1379210 genet.sickkids.on.ca
CFTR	L206W	PMID: 7691344 genet.sickkids.on.ca
CFTR	M1101K	PMID: 7680525 genet.sickkids.on.ca
CFTR	N1303K	PMID: 1998343 genet.sickkids.on.ca
CFTR	P574H	PMID: 2236053 genet.sickkids.on.ca
CFTR	Q1238X	PMID: 7683952 genet.sickkids.on.ca
CFTR	Q493X	PMID: 2236053 genet.sickkids.on.ca
CFTR	Q552X	PMID: 1709778 genet.sickkids.on.ca
CFTR	Q890X	PMID: 1284534 genet.sickkids.on.ca
CFTR	R1066C	PMID: 1379210 genet.sickkids.on.ca
CFTR	R1070Q	PMID: 7683628 genet.sickkids.on.ca
CFTR	R1158X	PMID: 1371265 genet.sickkids.on.ca
CFTR	R1162X	PMID: 2045102 genet.sickkids.on.ca
CFTR	R117C	PMID: 7525450 genet.sickkids.on.ca
CFTR	R117H	PMID: 2344617 genet.sickkids.on.ca
CFTR	R1283M	PMID: 1284468 genet.sickkids.on.ca
CFTR	R334W	PMID: 2045102 genet.sickkids.on.ca
CFTR	R347H	PMID: 1284538 genet.sickkids.on.ca
CFTR	R347P	PMID: 2344617 genet.sickkids.on.ca

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
CFTR	R352Q	PMID: 1284538 genet.sickkids.on.ca
CFTR	R553X	PMID: 1695717 genet.sickkids.on.ca
CFTR	R560T	PMID: 2236053 genet.sickkids.on.ca
CFTR	R709X	PMID: 7535742 genet.sickkids.on.ca
CFTR	R75X	PMID: 7525450 genet.sickkids.on.ca
CFTR	S1196X	PMID: 7681034 genet.sickkids.on.ca
CFTR	S1235R	PMID: 7508414 genet.sickkids.on.ca
CFTR	S1251N	PMID: 1284535 genet.sickkids.on.ca
CFTR	S1255X	PMID: 1284534 genet.sickkids.on.ca
CFTR	S364P	PMID: 9150159 genet.sickkids.on.ca
CFTR	S549I	PMID: 2236053 genet.sickkids.on.ca
CFTR	S549N	PMID: 1695717 genet.sickkids.on.ca
CFTR	S549R(A>C)	PMID: 1903761
CFTR	S549R(T>G)	PMID: 2236053 genet.sickkids.on.ca
CFTR	T338I	PMID: 7505693 genet.sickkids.on.ca
CFTR	V520F	PMID: 1284466 genet.sickkids.on.ca
CFTR	W1089X	PMID: 1284534 genet.sickkids.on.ca
CFTR	W1204X(c.3611G>A)	PMID: 1284534 genet.sickkids.on.ca PMID: 7522211
CFTR	W1204X(c.3612G>A)	PMID: 7522211
CFTR	W1282X	PMID: 2236053 genet.sickkids.on.ca
CFTR	Y1092X	PMID: 1284534 PMID: 16049310 genet.sickkids.on.ca
CFTR	Y122X	PMID: 1284471 genet.sickkids.on.ca
CFTR	dele2-3 21kb	PMID: 10798353 genet.sickkids.on.ca
CHM	IVS13+2dupT	PMID: 1302003
CLN5	2467AT	PMID: 9662406
CLN8	R24G	PMID: 10508524
CLRN1	N48K	PMID: 12080385
CNGB3	819_826del8	PMID: 10888875 PMID: 15657609
CNGB3	886-896del11insT	PMID: 15657609
CNGB3	E336X	PMID: 10958649
CNGB3	IVS8-3T>G	PMID: 15459792 PMID: 15657609
CNGB3	R403Q	PMID: 15161866

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
CNGB3	T383fs	PMID: 10888875 PMID: 15657609
CPT1A	G710E	PMID: 10607472
CPT1A	P479L	PMID: 11441142
CPT2	G549D	PMID: 10090476
CPT2	Leu178_Ile186delinsPhe	PMID: 9758712
CPT2	P227L	PMID: 10090476
CPT2	P50H	PMID: 7711730
CPT2	P604S	PMID: 9758712
CPT2	Q413fs	PMID: 10090476
CPT2	Q550R	PMID: 9758712
CPT2	R124X	PMID: 9562964
CPT2	R503C	PMID: 10090476
CPT2	R631C	PMID: 1528846
CPT2	S113L	PMID: 8358442
CPT2	S38fs	PMID: 10862092
CPT2	Y628S	PMID: 8651281
CTNS	537del21	PMID: 9792862
CTNS	D205N	PMID: 9792862
CTNS	L158P	PMID: 10482956
CTNS	W138X	PMID: 9537412
CTSK	X330W	PMID: 8703060
DHCR7	C380Y	PMID: 10677299
DHCR7	IVS8-1G>C	PMID: 9653161
DHCR7	L109P	PMID: 10677299
DHCR7	L157P	PMID: 9653161
DHCR7	R352Q	PMID: 10677299
DHCR7	R352W	PMID: 9653161
DHCR7	R404C	PMID: 9653161
DHCR7	T93M	PMID: 9653161
DHCR7	V326L	PMID: 9653161
DHCR7	W151X(c.452G>A)	PMID: 9653161
DHCR7	W151X(c.453G>A)	PMID: 10677299
DLD	105insA	PMID: 8968745
DLD	G229C	PMID: 9934985
DPYD	IVS14+1G>A	PMID: 8892022
F11	E117X	PMID: 2813350
F11	F283L	PMID: 2813350
F11	IVS14+1G>A	PMID: 2813350
F11	IVS14del14	PMID: 8807341
F5	H1299R	PMID: 9375735
F5	R506Q	PMID: 8164741
FAH	E357X	PMID: 8318997
FAH	IVS12+5G>A	PMID: 8318997
FAH	IVS8-1G>C	PMID: 9633815
FAH	P261L	PMID: 9633815
FAH	W262X	PMID: 8162054
FANCC	322delG	PMID: 8128956
FANCC	IVS4+4A>T	PMID: 8348157
FANCC	Q13X	PMID: 8128956
FANCC	R548X	PMID: 8103176 PMID: 8882868
FH	1431_1433dupAAA	PMID: 9635293
G6PC	459insTA	PMID: 8211187
G6PC	727G>T	PMID: 7668282
G6PC	F327del	PMID: 7814621
G6PC	G188R	PMID: 8733042

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
G6PC	G270V	PMID: 7573034
G6PC	Q242X	PMID: 7573034
G6PC	Q27fsdelC	PMID: 7573034
G6PC	Q347X	PMID: 8182131
G6PC	R83C	PMID: 8211187
G6PC	R83H	PMID: 7655466
G6PD	N126D	PMID: 3393536
G6PD	R459L	PMID: 2263506
G6PD	R459P	PMID: 8447319
G6PD	S188F	PMID: 3393536
G6PD	V68M	PMID: 3393536
G6PT1	1211delCT	PMID: 9781688 PMID: 9758626
G6PT1	A367T	PMID: 10518030 PMID: 15906092
G6PT1	G339C	PMID: 9428641 PMID: 9758626
G6PT1	G339D	PMID: 11949931
GAA	D645E	PMID: 8094613
GALC	Ex11-17del	PMID: 7581365
GALC	G270D	PMID: 9272171
GALC	R168C	PMID: 7581365
GALT	F171S	PMID: 1610789
GALT	IVS2-2A>G	PMID: 11754113
GALT	K285N	PMID: 1427861
GALT	L195P	PMID: 1373122
GALT	Q169K	PMID: 10649501
GALT	Q188R	PMID: 1897530
GALT	S135L	PMID: 7887417
GALT	T138M	PMID: 7887416
GALT	X380R	PMID: 10408771
GALT	Y209C	PMID: 10408771
GBA	1035insG	PMID: 1961718
GBA	D409V	PMID: 8118460
GBA	IVS2+1G>A	PMID: 1589760
GBA	L444P	PMID: 2880291
GBA	N370S	PMID: 3353383
GBA	R463C	PMID: 1972019
GBA	R463H	PMID: 17427031
GBA	R496H	PMID: 8432537
GBA	V394L	PMID: 2508065
GCDH	A421V	PMID: 8900227
GCDH	R402W	PMID: 8900227
GJB2	167delT	PMID: 9285800
GJB2	313del14	PMID: 9529365
GJB2	35delG	PMID: 9285800
GJB2	E120del	PMID: 10544226
GJB2	M34T	PMID: 9139825 PMID: 17041943
GJB2	Q124X	PMID: 9600457
GJB2	R184P	PMID: 10544226
GJB2	V37I	PMID: 10633133
GJB2	W24X	PMID: 9139825
GJB2	W77R	PMID: 9328482
GJB2	W77X	PMID: 12792423
GNE	M712T	PMID: 11528398
GRHPR	103delG	PMID: 10484776

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
HADHA	E474Q	PMID: 7811722
HADHA	Q342X	PMID: 7846063
HBB	-28A>G	PMID: 6308558 globin.bx.psu.edu
HBB	-29A>G	PMID: 6583702 globin.bx.psu.edu
HBB	-30T>A	PMID: 3382401 globin.bx.psu.edu
HBB	-87C>G	PMID: 6280057 globin.bx.psu.edu
HBB	-88C>T	PMID: 6086605 globin.bx.psu.edu
HBB	619 bp deletion	PMID: 287080 globin.bx.psu.edu
HBB	CAP+1 A>C	globin.bx.psu.edu PMID: 3683554
HBB	Glu6fs	PMID: 6310991 globin.bx.psu.edu
HBB	Gly16fs	PMID: 2064964 globin.bx.psu.edu
HBB	Gly24 T>A	globin.bx.psu.edu
HBB	Hb C	PMID: 8294201 globin.bx.psu.edu
HBB	Hb D-Punjab	PMID: 2307460 globin.bx.psu.edu
HBB	Hb E	PMID: 7177196 globin.bx.psu.edu
HBB	Hb O-Arab	PMID: 7908281 globin.bx.psu.edu
HBB	Hb S	PMID: 3267215 globin.bx.psu.edu
HBB	IVS-I-1(G>A)	PMID: 1634236 globin.bx.psu.edu
HBB	IVS-I-1(G>T)	PMID: 6714226 globin.bx.psu.edu
HBB	IVS-I-110	PMID: 6264477 globin.bx.psu.edu
HBB	IVS-I-5	PMID: 6188062 globin.bx.psu.edu
HBB	IVS-I-6	PMID: 6280057 globin.bx.psu.edu
HBB	IVS-II-654	PMID: 6585831 globin.bx.psu.edu
HBB	IVS-II-705	PMID: 6298782 globin.bx.psu.edu
HBB	IVS-II-745	PMID: 7177196 globin.bx.psu.edu
HBB	IVS-II-844	PMID: 2001456 globin.bx.psu.edu
HBB	IVS-II-849(A>C)	PMID: 2424301 globin.bx.psu.edu
HBB	IVS-II-849(A>G)	PMID: 6583702 globin.bx.psu.edu
HBB	IVS-II-850	PMID: 7558878 globin.bx.psu.edu
HBB	K17X	PMID: 88735 globin.bx.psu.edu

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
HBB	Lys8fs	PMID: 3828533 globin.bx.psu.edu
HBB	Phe41fs	PMID: 6826539 globin.bx.psu.edu
HBB	Phe71fs	PMID: 6585831 globin.bx.psu.edu
HBB	Poly A: AATAAA->AATAAG	PMID: 3048433 globin.bx.psu.edu
HBB	Poly A: AATAAA->AATGAA	PMID: 2375910 globin.bx.psu.edu
HBB	Pro5fs	PMID: 2606727 globin.bx.psu.edu
HBB	Q39X	PMID: 6457059 globin.bx.psu.edu
HBB	Ser9fs	PMID: 6714226 globin.bx.psu.edu
HBB	W15X	PMID: 6714226 globin.bx.psu.edu
HEXA	1278insTATC	PMID: 2848800
HEXA	G269S	PMID: 2522679
HEXA	IVS12+1G>C	PMID: 2837213
HEXA	IVS7+1G>A	PMID: 2220821
HEXA	IVS9+1G>A	PMID: 1837283
HEXA	R178C	PMID: 2137287
HEXA	R178H	PMID: 2961848
HEXA	R247W	PMID: 1384323
HFE	C282Y	PMID: 8696333
HFE	E168Q	PMID: 10953950
HFE	E168X	PMID: 10930379
HFE	H63D	PMID: 8696333 PMID: 18566337
HFE	H63H	PMID: 9490291 PMID: 15863206
HFE	Q127H	PMID: 10401000
HFE	Q283P	PMID: 12737937
HFE	S65C	PMID: 10194428
HFE	V53M	PMID: 10401000
HFE	V59M	PMID: 10401000
HFE	W169X	PMID: 10930379
HGD	G161R	PMID: 9154114 PMID: 10482952
HGD	G270R	PMID: 10482952
HGD	IVS1-1G>A	PMID: 10205262
HGD	IVS5+1G>A	PMID: 10482952
HGD	M368V	PMID: 9529363
HGD	P230S	PMID: 8782815
HGD	S47L	PMID: 10970188
IDUA	A327P	PMID: 7550242
IDUA	W402X	PMID: 1301196
IKBKAP	IVS20+6T>C	PMID: 11179008
IKBKAP	P914L	PMID: 12687659
IKBKAP	R696P	PMID: 11179008
IVD	A311V	PMID: 9665741
LAMA3	R650X	PMID: 8530087
LAMB3	3024delT	PMID: 11023379
LAMB3	Q243X	PMID: 8824879
LAMB3	R144X	PMID: 8824879

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
LAMB3	R42X	PMID: 7706760
LAMB3	R635X	PMID: 7698759
LAMC2	R95X	PMID: 8012394
LRPPRC	A354V	PMID: 12529507
MCOLN1	511_6944del	PMID: 10973263
MCOLN1	IVS3-2A>G	PMID: 10973263
MEFV	A744S	PMID: 9668175
MEFV	F479L	PMID: 9668175
MEFV	I692del	PMID: 9668175
MEFV	K695R	PMID: 9668175
MEFV	M680I	PMID: 10090880 PMID: 9288758
MEFV	M694I	PMID: 9288094
MEFV	M694V	PMID: 9288758
MEFV	P369S	PMID: 10090880
MEFV	R408Q	PMID: 10364520
MEFV	R653H	PMID: 11470495
MEFV	R761H	PMID: 9668175
MEFV	T267I	PMID: 9668175
MEFV	V726A	PMID: 9288758
MPI	R295H	PMID: 12414827
MUTYH	Y165C	PMID: 11818965
NBN	657del5	PMID: 9590180
NPC1	I1061T	PMID: 10480349
NPHS1	121_122del	PMID: 9660941
NPHS1	R1109X	PMID: 9660941
PAH	G272X	PMID: 1975559
PAH	I65T	PMID: 1301187
PAH	IVS-10int-546	PMID: 1769645
PAH	IVS12+1G>A	PMID: 10598814 PMID: 3008810
PAH	L48S	PMID: 1679030
PAH	R158Q	PMID: 2606484
PAH	R252W	PMID: 2574153
PAH	R261Q	PMID: 2574153
PAH	R408Q	PMID: 1312992
PAH	R408W	PMID: 2884570
PAH	Y414C	PMID: 2014036
PCDH15	R245X	PMID: 12711741
PEX1	G843D	PMID: 9398847
PEX7	G217R	PMID: 9090381
PEX7	L292X	PMID: 9090381
PKHD1	9689delA	PMID: 12846734
PKHD1	Leu1965fs	PMID: 11919560
PKHD1	R496X	PMID: 12506140
PKHD1	T36M	PMID: 11919560
PKHD1	V3471G	PMID: 12506140
PMM2	F119L	PMID: 9140401
PMM2	R141H	PMID: 9140401
POMGNT1	IVS17+1G>A	PMID: 11709191
PPT1	L10X	PMID: 9425237
PPT1	R122W	PMID: 7637805
PPT1	R151X	PMID: 9425237
PPT1	T75P	PMID: 9425237
PYGM	G204S	PMID: 8316268
PYGM	K542T	PMID: 8316268
PYGM	K542X	PMID: 16786513

Continued on next page

Table 7 – continued from previous page

Gene	Variant	References
PYGM	R49X	PMID: 8316268
RMRP	262G>T	PMID: 11207361
RMRP	g.70A>G	PMID: 11207361
RS1	E72K	PMID: 9618178
RS1	G109R	PMID: 9326935
RS1	G74V	PMID: 9618178
SACS	5254C>T	PMID: 10655055
SACS	6594delT	PMID: 10655055
SERPINA1	S allele	PMID: 2567291
SERPINA1	Z allele	PMID: 6306478
SGCB	S114F	PMID: 9032047
SLC12A6	R675X	PMID: 12368912
SLC12A6	Thr813fsX813	PMID: 12368912
SLC17A5	Leu336fsX13	PMID: 10947946
SLC17A5	R39C	PMID: 10581036
SLC25A15	F188del	PMID: 10369256
SLC26A2	C653S	PMID: 11241838
SLC26A2	IVS1+2T>C	PMID: 10482955
SLC26A2	R178X	PMID: 8528239
SLC26A2	R279W	PMID: 8571951
SLC26A2	V340del	PMID: 8528239
SLC26A4	E384G	PMID: 9618167
SLC26A4	L236P	PMID: 9618166
SLC26A4	T416P	PMID: 9618166
SMN1	Exon 7 deletion	PMID: 18941424
SMPD1	L302P	PMID: 1391960
SMPD1	R496L	PMID: 2023926
SMPD1	fsP330	PMID: 8401540
TH	R233H	PMID: 9703425
TPP1	G284V	PMID: 11339651
TPP1	IVS5-1G>A	PMID: 10330339
TPP1	IVS5-1G>C	PMID: 9295267
TPP1	R208X	PMID: 9295267
TTPA	744delA	PMID: 7719340

Table 8. Diseases in the UNIT. Different diseases have different modes of inheritance: AR (autosomal recessive) or XL (X-linked recessive).

Disease	Inheritance
ABCC8-related Hyperinsulinism	AR
Achromatopsia	AR
Andermann Syndrome	AR
Alkaptonuria	AR
ARSACS	AR
Aspartylglycosaminuria	AR
Ataxia With Vitamin E Deficiency	AR
Ataxia-telangiectasia	AR
Polyglandular Autoimmune Syndrome Type 1	AR
Limb-girdle Muscular Dystrophy Type 2E	AR
Biotinidase Deficiency	AR
Bloom Syndrome	AR
Canavan Disease	AR
Carnitine Palmitoyltransferase IA Deficiency	AR

Continued on next page

Table 8 – continued from previous page

Disease	Inheritance
Cartilage-hair Hypoplasia	AR
Cystic Fibrosis	AR
Choroideremia	XL
CLN5-related Neuronal Ceroid Lipofuscinosis	AR
Northern Epilepsy	AR
Congenital Disorder of Glycosylation Type Ia	AR
Congenital Disorder of Glycosylation Type Ib	AR
Congenital Finnish Nephrosis	AR
Cystinosis	AR
Factor v Leiden Thrombophilia	AR
Factor XI Deficiency	AR
Familial Dysautonomia	AR
Familial Mediterranean Fever	AR
Fanconi Anemia Type C	AR
Salla Disease	AR
Fumarase Deficiency	AR
Gaucher Disease	AR
GJB2-related DFNB 1 Nonsyndromic Hearing Loss and Deafness	AR
Glucose-6-phosphate Dehydrogenase Deficiency	XL
Glutaric Acidemia Type 1	AR
Glycogen Storage Disease Type Ia	AR
Glycogen Storage Disease Type Ib	AR
Pompe Disease	AR
Glycogen Storage Disease Type III	AR
Glycogen Storage Disease Type V	AR
Inclusion Body Myopathy 2	AR
GRACILE Syndrome	AR
Sickle Cell Disease	AR
Hereditary Fructose Intolerance	AR
Hereditary Thymine-uraciluria	AR
HFE-associated Hereditary Hemochromatosis	AR
Hyperornithinemia-hyperammonemia-homocitrullinuria Syndrome	AR
Primary Hyperoxaluria Type 1	AR
Primary Hyperoxaluria Type 2	AR
Isovaleric Acidemia	AR
Leigh Syndrome, French-Canadian Type	AR
Maple Syrup Urine Disease Type 3	AR
Long Chain 3-hydroxyacyl-CoA Dehydrogenase Deficiency	AR
Maple Syrup Urine Disease Type 1B	AR
Medium Chain Acyl-CoA Dehydrogenase Deficiency	AR
Mucolipidosis IV	AR
Hurler Syndrome	AR
Muscle-eye-brain Disease	AR
MYH-associated Polyposis	AR
Niemann-Pick Disease Type A	AR
Niemann-Pick Disease Type C	AR
Nijmegen Breakage Syndrome	AR
Pendred Syndrome	AR
Phenylalanine Hydroxylase Deficiency	AR
Autosomal Recessive Polycystic Kidney Disease	AR
PPT1-related Neuronal Ceroid Lipofuscinosis	AR
Pycnodysostosis	AR
Rhizomelic Chondrodysplasia Punctata Type 1	AR
Short Chain Acyl-CoA Dehydrogenase Deficiency	AR
Sjogren-Larsson Syndrome	AR
Smith-Lemli-Opitz Syndrome	AR
TPP1-related Neuronal Ceroid Lipofuscinosis	AR

Continued on next page

Table 8 – continued from previous page

Disease	Inheritance
Segawa Syndrome	AR
Tyrosinemia Type I	AR
Usher Syndrome Type 3	AR
Usher Syndrome Type 1F	AR
Wilson Disease	AR
X-linked Juvenile Retinoschisis	XL
Tay-Sachs Disease	AR
Infantile Refsum Disease	AR
Galactosemia	AR
Bardet-Biedl Syndrome, BBS1-related	AR
Bardet-Biedl Syndrome, BBS10-related	AR
Herlitz Junctional Epidermolysis Bullosa, LAMA3-related	AR
Herlitz Junctional Epidermolysis Bullosa, LAMB3-related	AR
Herlitz Junctional Epidermolysis Bullosa, LAMC2-related	AR
Hypophosphatasia, Autosomal Recessive	AR
Spinal Muscular Atrophy	AR
Alpha-1-antitrypsin Deficiency, Type Z	AR
Alpha-1-antitrypsin Deficiency, Type S	AR
Krabbe Disease, Late-onset Form	AR
Beta Thalassemia Intermedia	AR
Beta Thalassemia Major	AR
Krabbe Disease, Infantile Form	AR
Carnitine Palmitoyltransferase II Deficiency, Lethal Neonatal Form	AR
Carnitine Palmitoyltransferase II Deficiency, Myopathic Form	AR
Homocystinuria, B6-responsive	AR
Homocystinuria, B6-non-responsive	AR
Metachromatic Leukodystrophy, Early-onset Form	AR
Metachromatic Leukodystrophy, Late-onset Form	AR
Achondrogenesis Type 1B	AR
Carnitine Palmitoyltransferase II Deficiency, Infantile Form	AR
Hexosaminidase a Deficiency, Adult-onset Form	AR
Hexosaminidase a Deficiency, Juvenile or Chronic Form	AR
Recessive Multiple Epiphyseal Dysplasia	AR
Diastrophic Dysplasia	AR

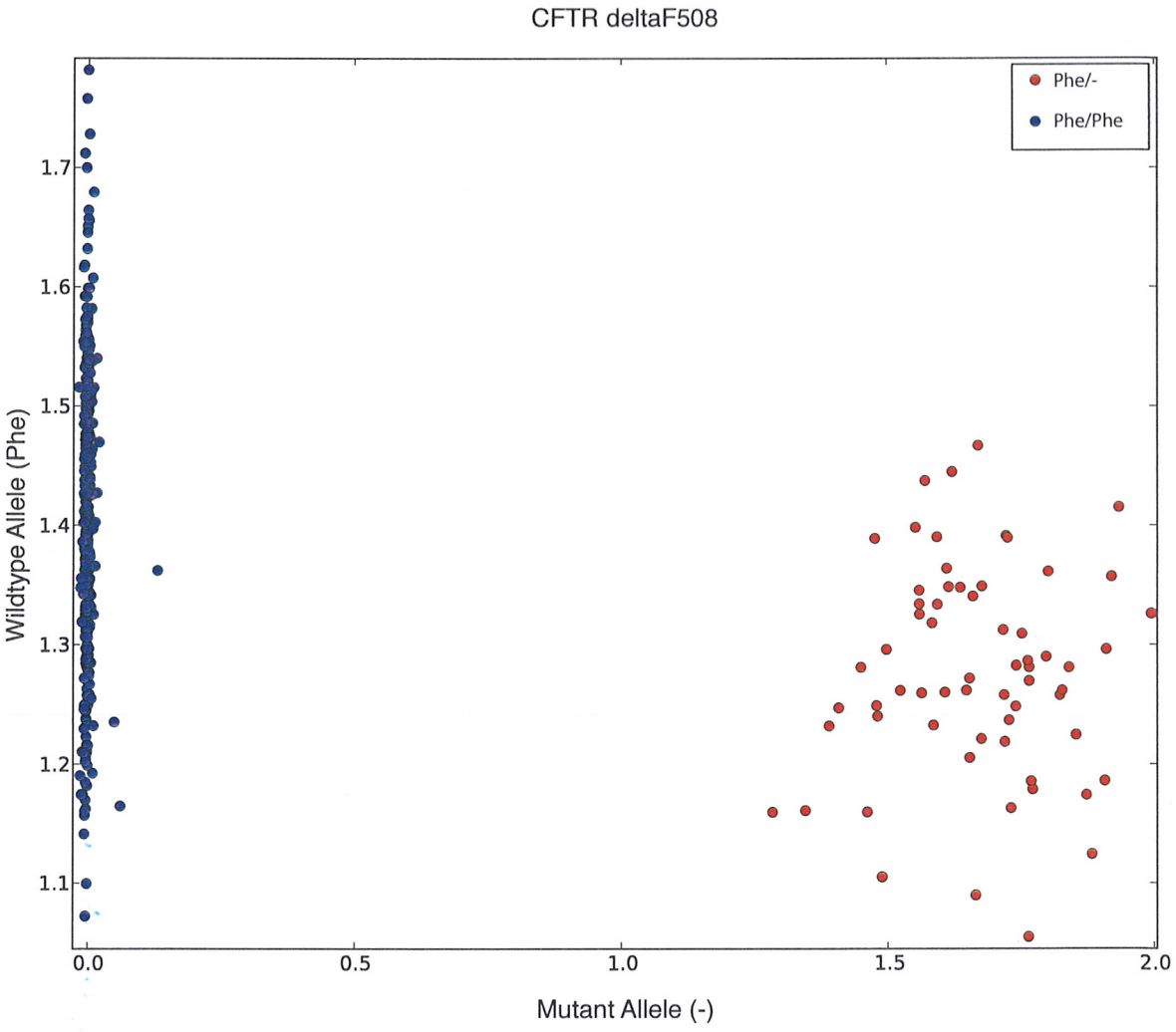


Figure 7. Sample data for the CFTR deltaF508 variant. As described in the text, red points represent heterozygotes and blue points represent homozygotes. This variant, like others, was retained after surviving the multi-stage assay design process (Table 4). Other retained variants were those with similar strong separations between clusters.