

Functional Characterisation of Alpha-Galactosidase A Mutations as a Basis for a New Classification System in Fabry Disease

Jan Lukas¹, Anne-Katrin Giese¹, Arseni Markoff², Ulrike Grittner³, Ed Kolodny⁴, Hermann Mascher⁵, Karl J. Lackner⁶, Wolfgang Meyer⁷, Phillip Wree¹, Viatcheslav Saviouk⁸, Arndt Rolfs^{1*}

1 Albrecht-Kossel-Institute for Neuroregeneration, Centre for Mental Health, University of Rostock, Rostock, Germany, **2** Institute of Medical Biochemistry and IZKF, University of Muenster, Muenster, Germany, **3** Department for Biostatistics and Clinical Epidemiology, Charité-University Medicine, Berlin, Germany, **4** Department of Neurology, New York University School of Medicine, New York, New York, United States of America, **5** pharm-analyt, Labor GmbH, Baden, Austria, **6** Institute for Clinical Chemistry and Laboratory Medicine, University of Mainz, Mainz, Germany, **7** Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom, **8** Institute for Molecular Diagnostics, Centogene GmbH, Rostock, Germany

Abstract

Fabry disease (FD) is an X-linked hereditary defect of glycosphingolipid storage caused by mutations in the gene encoding the lysosomal hydrolase α -galactosidase A (GLA, α -gal A). To date, over 400 mutations causing amino acid substitutions have been described. Most of these mutations are related to the classical Fabry phenotype. Generally in lysosomal storage disorders a reliable genotype/phenotype correlation is difficult to achieve, especially in FD with its X-linked mode of inheritance. In order to predict the metabolic consequence of a given mutation, we combined in vitro enzyme activity with in vivo biomarker data. Furthermore, we used the pharmacological chaperone (PC) 1-deoxygalactonojirimycin (DGJ) as a tool to analyse the influence of individual mutations on subcellular organelle-trafficking and stability. We analysed a significant number of mutations and correlated the obtained properties to the clinical manifestation related to the mutation in order to improve our knowledge of the identity of functional relevant amino acids. Additionally, we illustrate the consequences of different mutations on plasma lyso-globotriaosylsphingosine (lyso-Gb3) accumulation in the patients' plasma, a biomarker proven to reflect the impaired substrate clearance caused by specific mutations. The established system enables us to provide information for the clinical relevance of PC therapy for a given mutant. Finally, in order to generate reliable predictions of mutant GLA defects we compared the different data sets to reveal the most coherent system to reflect the clinical situation.

Citation: Lukas J, Giese A-K, Markoff A, Grittner U, Kolodny E, et al. (2013) Functional Characterisation of Alpha-Galactosidase A Mutations as a Basis for a New Classification System in Fabry Disease. *PLoS Genet* 9(8): e1003632. doi:10.1371/journal.pgen.1003632

Editor: Atul Mehta, University College London, United Kingdom

Received: January 12, 2013; **Accepted:** May 14, 2013; **Published:** August 1, 2013

Copyright: © 2013 Lukas et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The study has been supported partially by an unrestricted scientific grant from Shire Human Genetic Therapies (Germany). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: arndt.rolfs@med.uni-rostock.de

Introduction

With a suspected prevalence of 1:3,100 to 13,341 [1],[2], Fabry disease (FD, OMIM #301500) is the second most frequent lysosomal storage disorder. FD causes the accumulation of intracellular/lysosomal, plasma and urinary globotriaosylceramide (Gb3) which, due to mutations within the α -galactosidase A (*GLA*, Xq22) gene, cannot be cleared. Defects of the *GLA* gene product, caused mostly by single amino acid substitutions, lead to its early degradation within the endoplasmic reticulum [3] and prohibit intracellular trafficking of the enzyme to the destination organelle, the lysosome. In affected patients, typical FD presents as a multisystemic disorder and in classic cases stroke, acroparaesthesia, hypohidrosis, angiokeratoma, cornea verticillata, cardiac and kidney disease [4],[5] develop. However, milder mono- or oligosymptomatic cases have been reported [6]–[9]. Oligosymptomatic cases further impede the diagnosis, e.g., large numbers of abnormal variant forms of *GLA* found in newborn-screenings are either linked to a monosymptomatic and late onset of the disease

[1],[2] or are coincidental findings that can be considered as variants without significant metabolic consequences. Some other mutations seem to be related to a certain phenotype with a predominantly single organ involvement, hence, both a cardiac variant [10]–[12] and a cerebrovascular variant [13] have been described. Typically these mutations are associated with a late onset phenotype. There appears to be at least one mutation (*p.D313Y*), most likely a polymorphism [14],[15] but which can also be found in stroke of unexplained aetiology [13],[16].

Due to the X-linked mode of inheritance, genetic sequencing in females is the only valid tool to diagnose FD, since enzyme activity in patient leucocytes can be in the normal range in a high number of female heterozygotes [17], even in those severely affected. Because of the X-linked inheritance, two thirds of all Fabry patients can be expected to be females.

Globotriaosylsphingosine (lyso-Gb3), a deacylated metabolite of Gb3 has been described as a useful biomarker to quantify the burden of FD. Lyso-Gb3 was shown to be the storage material in many cells accumulating to high levels in vasoendothelial cells of

Author Summary

Fabry disease is caused by a single gene deficiency. It is the second most common lysosomal storage disorder and the result is a build-up of glycosphingolipids in different areas of the body (kidneys, intestine, etc). It is an important consideration for clinicians in the diagnosing of stroke, kidney and cardiovascular diseases. Many symptoms of Fabry are seen in other diseases as well (both inherited and non-inherited), which makes diagnosis difficult. We observed numerous novel mutations in patients that displayed a monosymptomatic, however life-threatening course of Fabry disease. This prompted us to study and characterise those mutations with regard to their biochemical and clinical consequences. Overall, 171 Fabry mutations were considered in an overexpression system for the prediction of the clinical course of Fabry disease. Furthermore, we highlight the usefulness of the *in vitro* system that we developed which will help with therapeutic decisions, by testing the responsiveness of mutant enzymes to the pharmacological chaperone DGJ. This work aims to draw the attention of clinicians and researchers to milder forms of Fabry disease which might at first appear unrelated to this clinically heterogeneous disease.

blood vessels. Lyso-Gb3 was proven to be more specifically increased in FD patients [18] than Gb3 and demonstrated to be reduced in mice under enzyme replacement therapy (ERT). Moreover, lyso-Gb3 has been proposed to take part in the development of FD nephropathy [19] and general inflammatory processes. Even though lyso-Gb3 can be used as a reliable biomarker for FD, it is only reliable for monitoring as no absolute correlation has been described between the nature of the mutation and lyso-Gb3 levels. It was concluded that long periods of exposure to this agent lead to the development of symptoms [20].

Several recent studies tried to generate more impartial data about the cellular and structural consequence of a given *GLA* mutation. They were typically based on overexpression systems and analysed either the residual *GLA* enzyme activity [21] or focussed on enzyme trafficking and stability aspects [3]. Other studies tried to establish genotype/phenotype correlations in clinical cases of Fabry patients [22]. The elucidation of the crystal structure of the enzyme [23] was the basis for the establishment of structure/function correlation models [24]–[26].

In a recent study, newly described mutations were analysed in an overexpression system. Enzyme activity was measured and responsiveness to 1-deoxygalactonojirimycin (DGJ) was tested in Western blot assay and in patient's T-lymphocytes [27]. This approach was used to determine the likelihood of a newly identified *GLA* variant leading to FD. Wider studies with T-lymphocytes have previously been undertaken [28],[29]. A systematic approach to examine a broad range of mutations has been developed in order to facilitate diagnosis and therapeutic decisions [30].

In the present study 171 *GLA* mutations were characterised biochemically for residual enzyme activity, degradation status and behaviour towards pharmacological chaperone (PC) DGJ. Moreover, the obtained data sets were compared with patient plasma lyso-Gb3 levels, the computational prediction algorithm PolyPhen2 (Polymorphism Phenotype v2) [31] and the clinical phenotype. We paid particular attention to mutations believed to be on the borderline for causing the disease.

Results

Determination of single amino acid substitution mediated α -galactosidase A damage

We tested 171 mutations with regard to enzyme activity and degradation status to evaluate which mutation leads to diminished enzyme activity. Among the mutations tested, 50 had not been described previously. In sum, 158 missense, 6 nonsense mutations causing immediate polypeptide abort, 4 small deletions (1–2 nucleotides) and 3 small insertions (1–4 nucleotides) were examined for potential damage (Table S1). Mutant enzymes were studied with respect to their effect on protein degradation and pharmacological chaperone responsiveness (Figure S1).

Enzyme activity and intracellular enzyme level were correlated for mutations natively displaying activities above 6% (Spearman correlation coefficient $r_s = 0.866$, $p < 0.001$) (Figure 1B). Enzyme levels were not associated with residual activity since highly defective enzymes seemed to lack the capacity to process the substrate (Spearman correlation coefficient $r_s = 0.128$, $p = 0.272$) (Figure 1A). Figure S1 shows enzyme content for some strongly affected enzymes (e.g. *p.D231N*, *p.L275F* and *p.L415P*) with no measurable activity. Overexpressed mutant α -Gal A was abundantly present in the cells, notably above the endogenous level at the expected size of 46 kDa. Thus, the protein was evidently processed [32], but it is not catalytically active. A closer look at position *p.Arg118* (buried residue) revealed that the actual described mutation for this moiety (*p.R118C*) displayed the lowest detected enzyme activity at about 20% of the wild type enzyme (Figure 2). All other mutations at this position (e.g. R118G, R118S, see Table S1) were less severely affected (range 20–70.9%). Vice versa, at position S126 (stabilising loop), the described mutation S126G displayed the mildest effect on activity loss. The strong decrease of activity for S126C is explained by the chemical nature of the amino acid exchange from serine to cysteine, predicted to disrupt protein conformation (due to auxiliary disulphide bond introduction). As expected when changing the important aspartate residue at position 264 (close active site) [33], the enzyme lost its activity. The only exception was *p.D264N* which maintained about 37.9% of its activity. We arranged 4 classes of mutations based upon the enzyme activity data: class I: 0%, class II: >0%–20%, class III: $\geq 20\%$ –60%, class IV: $\geq 60\%$ for further analyses.

Responsiveness to pharmacological chaperone DGJ

In order to examine the importance of enzyme activity for clinical and therapeutic decisions, we analysed enzyme activity changes after addition of the pharmacological chaperone DGJ. In order to simplify the outcome measures, a responder was defined as a mutation whose activity was increased 1.5-fold or >5% compared to the untreated value. Under this directive, 42.8% of the missense mutants were responsive to DGJ. Mutations in sites *p.Arg118*, *p.Ser126* and *p.Asp264* gave rise to enzymes that displayed an overall minor (if any) increase of activity through DGJ treatment indicating that the moieties are generally situated in sites where treatment with the PC does not result in significant elevation of α -Gal A activity and that the type of amino acid substitution is biochemically less relevant (exception: *p.D264N*; see also Table S1). Furthermore, DGJ analysis showed that active-site-associated amino acid substitutions (*p.D93E/Y*, *p.D170N*, *p.R227Q* and *p.D231N*) could not retain lost activity under treatment.

DGJ responsiveness was highly associated with residual activity of the mutant enzyme *in vitro* ($p < 0.001$ in linear trend test, Table S2). Class I mutants were less likely to respond to PC treatment (14.5%) than class II mutants (82.8%) with a less severe impact on

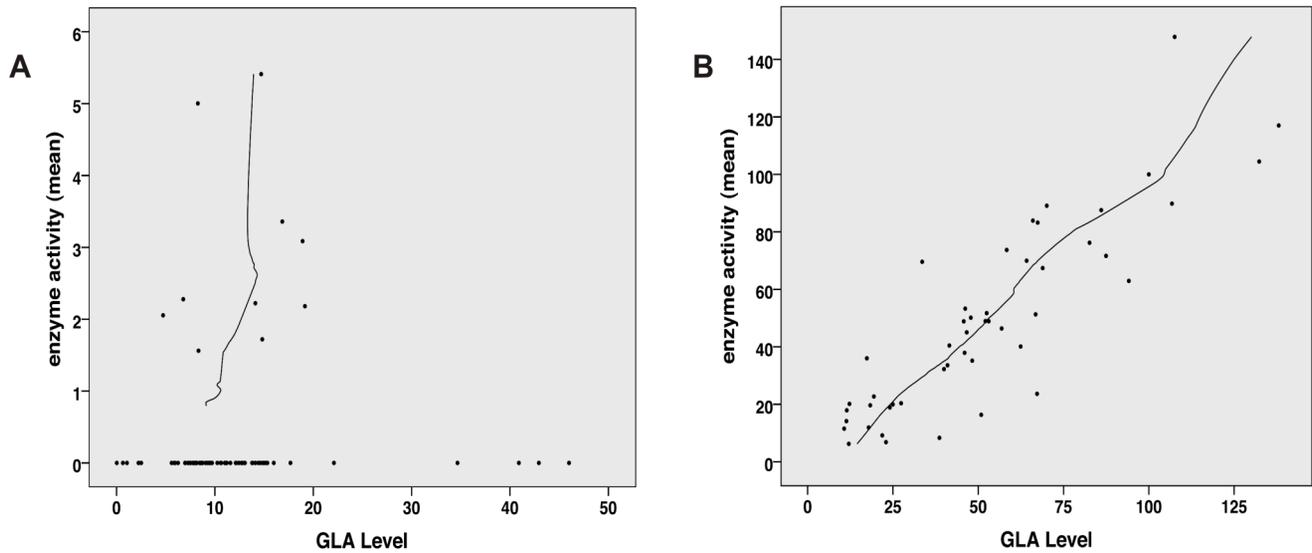


Figure 1. Correlation analysis of α -Gal A level (semi-quantitative Western Blot) and activity. **A:** No correlation between GLA level and residual activity for mutations possessing less than 6% residual activity ($n = 76$, Spearman correlation coefficient $r_s = 0.128$, $p = 0.272$). This implies that the catalytic unit is affected by the mutation and thus high amount of enzyme cannot compensate for the loss of activity. **B:** For mutation possessing more than 6% residual GLA activity ($n = 48$), the *in vitro* enzyme activity and GLA levels correlate with each other, indicating that the catalytic core is still intact and mutation most likely affect protein trafficking (Spearman correlation coefficient $r_s = 0.866$, $p < 0.001$).
doi:10.1371/journal.pgen.1003632.g001

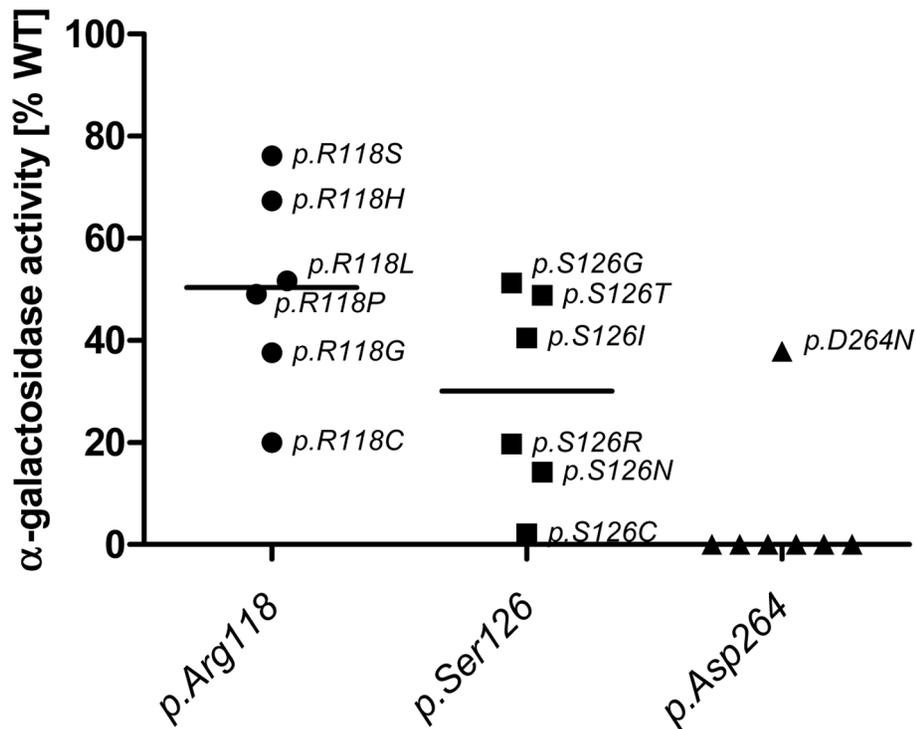


Figure 2. *In vitro* activity of specific GLA site mutations. Note that mutations at position *p.Asp264* almost always lead to a loss of GLA activity, while the same does not hold true for *p.Arg118* and *p.Ser126*. Interestingly mutations in *p.R118* do not lead to a loss of activity below 20% of WT and range from 20% to 80% while *p.Ser126* can lose all activity with certain mutations and retains no more than 60% activity. This highlights the differential effects of the mutational site and amino acid change on α -Gal A activity. Given is the median activity of all mutations in each position (horizontal mark).
doi:10.1371/journal.pgen.1003632.g002

enzyme function. Most of the best responders belong to this category (*p.A156V*, *p.I253S*, *p.R301Q*, etc.). In classes III and IV, a high percentage of responders was present (71.4 and 63.2% respectively), though with a high number of only weakly stimulated enzymes.

Association of different parameters

In a first approach, we compared residual enzyme activity to biomarker levels obtained from patients with the according mutations to investigate whether *in vitro* data reflect lyso-Gb3 as a marker for clinical severity.

To make sure that lyso-Gb3 was an adequate FD measure for testing of *in vitro* enzyme activity accuracy, we ascertained that lyso-Gb3 in male and female Fabry patients (Figure 3) were considerably higher than lyso-Gb3 in healthy controls [34]. The pathological cut-off for lyso-Gb3 measurements was set to 0.9 ng/ml (95th percentile of healthy individuals).

The gender-dependence of lyso-Gb3 values in hemizygotic males and heterozygotic females carrying the same mutation (compare Table S1) reflects the shortcoming of patient-derived data. Still, the median values of the patient/proband cohorts showed values above the pathological cut-off in both males and females (5.5 ng/ml for female individuals, 28.6 ng/ml for males). In figure 3 coloured dots represent patients with the mutations *p.S126G*, *p.A143T* and *p.D313Y*, including most of the values not detected as pathological. Strikingly, the overall inconspicuousness of mutations such as *p.A143T* or *p.D313Y* (see mean values in

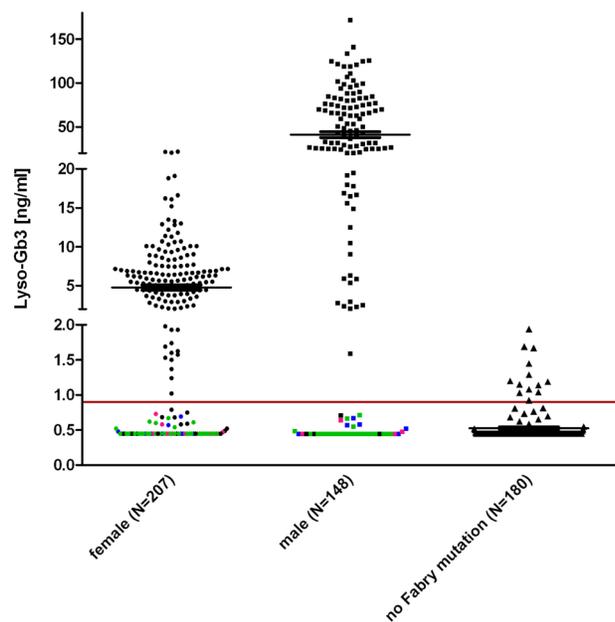


Figure 3. Lyso-Gb3 values for female and male Fabry patients compared to control. The horizontal mark indicated the median. It is noteworthy that lyso-Gb3 levels in males are ~10 times higher than in females. Each data point represents one patient. Indicated in pink are patients with the mutation *p.S126G* (8f/4m), in blue *p.A143T* (10f/8m) and in green *p.D313Y* (33f/24m) to illustrate that most found non-pathogenic mutations belong to either one or the other patient cohort. Other exceptions are: *p.R118C*, *p.V316I*, *p.E418G* (one male patient each) and *p.A20P*, *p.D83N*, *p.I91T*, *p.S102L*, *p.R112C*, *p.R118C*, *p.D175E*, *p.G325S*, *p.A368T*, *p.T385A*, *p.W399**, *c.1208delT*, *p.L415F*, (one female patient each) and *p.R252T* (4×), *p.N215S* (3×). About 180 healthy probands were tested with no Fabry gene variation and had values of 0.9 ng/ml (95th percentile calculation). doi:10.1371/journal.pgen.1003632.g003

Table S1) was not gender-dependent, meaning that those patients would remain undetected in males and females alike. However, Table S3 shows that lyso-Gb3 displays sensitivity for classic mutations. Depicted is another set of mutations found in patients where we measured plasma lyso-Gb3. In this table, most of the mutations led to a complete loss of the enzyme (truncations, splice mutations, etc). The positive predictive value for lyso-Gb3 (for all 355 FD patients and 180 control individuals, Figure 3) was 95%. The negative predictive value was 60.9%.

Female and male biomarker data were analysed separately to evaluate the association between enzyme activity and levels of the biomarker lyso-Gb3. For the categorisation of lyso-gb3-values we used the following cut-points: class I: lyso-gb3 $\geq 80/10$ ng/ml for males/females, class II: lyso-gb3 20–80/5–10 ng/ml for males/females, class III: lyso-gb3 0.9–20/0.9–5 ng/ml for males/females, class IV: lyso-gb3 <0.9 ng/ml (for males/females). The enzyme classes were strongly associated with the biomarker levels of the patients ($p < 0.001$ for males and for females, linear trend test, Table 1). Mutations with an activity lower than 20% showed gender-independent elevated lyso-Gb3 values.

Trend test analysis revealed statistically significant associations with enzyme activity ($p < 0.001$, linear trend test, Table 1). The *in silico* prediction tool PolyPhen2 revealed strong association with the enzyme activity classes as well ($p < 0.001$)

In addition, we also considered the correlation of *in vitro* enzyme activity with accessible surface area of the mutated amino acid residues as a possible determinant for the biological consequence of a given mutation as proposed previously [33]. However, we found only a weak linear trend (for details see Table S2).

Association of enzyme activity reduction to clinical phenotype

Table 2 shows the proportion of mutations correctly classified into the clinical phenotype groups by using lyso-Gb3 values of females, males, *in vitro* enzyme activity and PolyPhen2 scores. Enzyme activity showed the highest rates of correct classification both for 72 mutations where data were available and for 21 mutations for which we had values in all four measures (PolyPhen2 scores, lyso-Gb3 for females, lyso-Gb 3 for males and enzyme activity). The classification rates were 83% for 72 mutations and 86% for the subset of 21 mutations. Lyso-Gb3 for males and PolyPhen2 scores showed similar prediction quality: PolyPhen2 scores led to a slightly higher classification rate of 76% in the subset compared to 71% for lyso-Gb3 values for males. With 67% in the subset of mutations lyso-Gb3 values for females showed the lowest rate of correctly classified mutations.

Discussion

The validity of potential biomarker lyso-Gb3

Until now, biomarker data for novel mutations have been regarded as the gold standard for diagnosis (5) immediately following genetic diagnosis. This may not be accurate for milder cases of the disease, for example when the patient is tested early in the development of the disease (even though an age dependent increase is still speculative) or the mutation leads to a minor catalytic defect, since we failed to detect our *p.S126G*, *p.A143T* and *p.D313Y* patients (mean values are non-pathological, see Table S1) as well as three (female) *p.N215S* cases. This is in accordance with recent findings [35]. We reported a family with only female mutation carriers [36]. In the case of an unknown mutation this is a difficult situation, because lyso-Gb3 analysis is less strongly associated with disease phenotype in females (Table 2). For some newly described mutations, we lacked a detailed clinical description, however the

Table 1. Association of *in vitro* enzyme activity and clinical and computational parameters.

	enzyme activity				p (for linear trend test)
	0%	>0%–20%	≥20%–60%	≥60%	
disease phenotype					
classic	42	8	2	0	<0.001
classic/variant	3	1	3	0	
variant	1	2	6	4	
lyso-gb3 male					
I (≥80 ng/ml)	9	0	0	1	<0.001
II (20–80 ng/ml)	16	3	0	0	
III (0.9–20 ng/ml)	4	2	3	1	
IV (<0.9 ng/ml)	0	0	3	3	
lyso-gb 3 female					
I (≥10 ng/ml)	5	0	0	0	<0.001
II (5–10 ng/ml)	13	1	0	0	
III (0.9–5 ng/ml)	9	4	4	1	
IV (<0.9 ng/ml)	0	0	4	7	
PolyPhen2					
benign	5	2	12	13	<0.001
possibly damaging	8	7	4	1	
probably damaging	68	14	11	2	

Statistical association of enzyme activity and other parameters indicative for Fabry disease. The matrix confirms the high degree of translatability of *in vitro* data to the individual patients' biomarker phenotype. Lyso-Gb3-based classes of the mutants are fitting the enzyme activity classes.

doi:10.1371/journal.pgen.1003632.t001

phenotype and hence rationale for FD testing was either stroke (e.g. *p.D83N*, *p.S102L*, *p.N139S*, *p.R252T*, *p.V316I*, *p.L415F*, *p.E418G*), kidney disease (e.g. *p.A37T*, *p.H225D*, *p.E398A*) or in one case cardiological symptoms (*p.R220Q*) of unexplained etiology. A surprisingly high number of these missense mutations was found in oligosymptomatic patients and in the event of low lyso-Gb3 and high residual activity values. About 7.1% of males and 16.7% of females lyso-Gb3 (mean) of all mutations examined (Table S1) were within the non-pathogenic range. However, if we subtract all mutations related to mono- or oligosymptomatic FD, lyso-Gb3 is pathogenic for 100% of males and 96.9% of females. The remaining 3.1% of females harbour the following mutations: *p.A20P*, *p.I91T*, *p.W262** and *p.W399**. The truncating mutation *p.W262** has not

been described before but is expected to cause classic FD. The biomarker lyso-Gb3 demonstrated a weaker association to clinical phenotype than *in vitro* enzyme activity for each mutation studied. Table S3 shows additional 52 mutations for which biomarker data was collected. All classical mutations showed elevated values above the normal range in males.

For each mutation studied mean lyso-Gb3 values were always higher in male than in female patients (higher sensitivity in males). However the values obtained in females also accurately reflect disease pathology (similar specificity). It has to be emphasised that every mutation causing a pathological lyso-Gb3 mean value in males also demonstrates with a pathologically elevated mean value in females. In the same vein, the mutation N215S that has a higher

Table 2. *In vitro* enzyme activity reflects clinical phenotype and is a predictor for FD.

	Lyso-Gb3 for females	Lyso-Gb3 for males	<i>In vitro</i> enzyme activity	PolyPhen2 classes
	N = 33	N = 32	N = 72	N = 72
Nagelkerkes R ²	0.44	0.41	0.53	0.12
–2 Log Likelihood (df)	12.44 (3)	14.43 (3)	15.41 (3)	16.13 (2)
Mutations correct classified	69.7%	78.1%	83.3%	73.6%
	N = 21	N = 21	N = 21	N = 21
Nagelkerkes R ²	0.45	0.40	0.82	0.53
–2 Log Likelihood (df)	12.91 (3)	13.44 (3)	4.18 (3)	7.80 (2)
Mutations correct classified	66.7%	71.4%	85.7%	76.2%

Comparison of prediction quality for different parameters with regard to clinical phenotype (classic, variant, classic/variant, see Table S1); variance estimates from ordinal regressions for the outcome 'clinical phenotype' with different covariates, 1. for all data available in the particular dimension, 2. for a subset of 21 mutations that have values in all dimensions.

doi:10.1371/journal.pgen.1003632.t002

frequency in the population manifests with a mean value of 4.2 ng/ml (male) obtained from 7 different patients displaying increased lyso-Gb3 and 1.1 ng/ml (female) obtained from 6 different patients. However, 3 of the females showed normal values of 0.79 ng/ml, 0.68 ng/ml, one individual had a value below the limit of quantitation (<LLOQ), respectively. Essentially, lyso-Gb3 is a marker of classic rather than uncertain FD, (i.e. monosymptomatic or mild cases) reflecting disease severity after all. However, to overcome the limitation of individual patient constitution, we propose a new classification system based on *in vitro* enzyme activity.

In vitro enzyme activity as a tool for diagnostic management

Previous studies have focused on mutant enzyme activity, but this has never been proposed for use as a diagnostic tool. In the present study we utilised a cell culture model to generate enzymatic data for a large subset of *GLA* mutations. As stated before, Fabry disease-causing mutations are not limited to active site residues [37], but rather distributed across the entire protein. We therefore chose mutations in domains across the entire protein (see also criteria in the Materials and Methods section). Herein, we report and characterise 171 mutations on the basis of *in vitro* overexpression and enzyme activity. Among the tested mutations, 50 (29.2%) had not been described previously (note: the *p.Arg118*, *p.Ser126* and *p.Asp264* substitutions had not been found in patients yet, see Table S1). A strong association to biomarker plasma lyso-Gb3, the PolyPhen2 *in silico* prediction and the DGJ responsiveness of the mutation identifies those parameters as potential predictors of the clinical phenotype of the mutations. However, *in vitro* enzyme activity measurement was the most accurate predictor of disease phenotype assigning 85.7% of the mutants to the correct disease group whereas biomarker lyso-Gb3 for females and males and PolyPhen2 analysis only predicted 66.7, 71.4 and 76.2%, respectively, (as indicated by the highest explained variance in regression (R^2), lowest -2Log-Likelihood , and highest rate of correct classified mutations, Table 2). The ability of the mutations to respond to PC treatment classified 76.2% of the mutation correctly (data not shown).

The predictive quality of PolyPhen2 is hampered especially when it comes to mutations that are in unknown molecular interaction sites or mutations in domains with unclear structural significance, such as the N-terminal signal peptide region (i.e. *p.A15E* is incorrectly predicted to display a benign impact on enzyme function). However, the precision could assumingly be enhanced using a specifically constructed algorithm as has been conducted for the prediction of DGJ responding mutants in FD [26]. The pharmacological chaperone DGJ can rescue unstable, degradation-prone mutants, therefore responsiveness to DGJ could be correlated with the level of damage of the mutation as proposed previously [26]. Since only 14.5% of class I mutants could be rescued with the pharmacological chaperone DGJ and the proportion of responding mutations is much higher within the other classes, residual activity is associated with whether a mutation can be rescued or not ($p < 0.001$, Table 1). As mentioned before, the drawback for the lyso-Gb3 is reasoned by clinical variability of the individual patients, especially in females.

Unknown and mild GLA mutations - the challenge to establish pathogenicity

Due to the fact that genotype/phenotype correlations only exist for severe *GLA* mutations and are mostly linked to the classic phenotype, clinical decision-making in diagnosis and therapy in

Fabry disease is challenging. Further, the genotype/phenotype correlation depends on the patients' gender. However, there is no doubt that clinicians would apply an ERT to a female patient with a classic FD mutation as well as to males in pre-clinical or milder forms of the disease, since therapeutic success is closely connected to the correct treatment schedule and a timely onset of the treatment to improve the probability of a complete symptom reversal. Before the year 2000, *p.D313T* had been described as the only enzyme variant with high residual activity and was therefore classified as a SNP. In contrast *p.A143T* which also has significant residual activity [28],[38], was classified as a disease-causing mutation [39] and also appears more frequent than usual FD mutations. The number of mutations resulting in a milder disease phenotype increased in the course of systematic screening programs identifying mutations in patient cohorts with symptoms common for other diseases. This highlights the need to re-consider the approach to Fabry diagnosis.

Mutations with residual activity undergo early degradation

Due to the fact that FD mutations can affect both the stability and catalytic function of the enzyme we correlated the level and activity of *GLA* enzyme (Figure S1, Figure 1). The mutations *p.D93E* and *p.D231N* had *GLA* levels of 8.0% and 46.0% of wild type respectively, but no residual activity. Both mutants responded to the stabilising effect of DGJ, increasing *GLA* levels to 24.7% and 66.7% respectively, but no increase in activity was detected (Figure S1). In contrast to this finding class III and IV mutations as well as milder affecting class II mutations (activity >6%) displayed a strong correlation between *GLA* level and activity (Spearman correlation coefficient $r_s = 0.866$, $p < 0.001$). This emphasises the assumption that class III–IV (and the majority of class II) mutations mainly demonstrate an early degradation defect that retains enzyme activity. To test if mutations can affect stability alone, we performed a kinetic assay. Many of the *GLA* variants (*p.A143T*, *p.A156V*, *p.R301Q*, *p.L310F*) had unchanged kinetic properties (Table S4) indicating that the main reason for the molecular defect is early degradation. At least 3-fold higher activity above endogenous *GLA* was necessary to calculate kinetic parameters. Some of the mutations failed this criterion (*p.R49G*, *p.S65I*, *p.D231N* and *p.L415P* had activities around 200 nmol 4-MU/mg protein/hr which corresponds to endogenous *GLA*) and were therefore excluded from the analysis. However, the finding that even large amounts of enzyme could not increase activity above the endogenous level of the HEK293H cells indicates a severe impairment of the kinetic properties. Furthermore, this analysis revealed one mutation provoking only a mild decrease in enzyme activity, strong DGJ responsiveness and a benign *in silico* prediction that might nevertheless significantly change the kinetic properties of the enzyme (*p.H46P*). This could explain the disease pathology in patients with this mutation, since an increased K_M value could lead to diminished intra-lysosomal substrate turnover. Evidence rises that classification of mutants as either active site mutations or enzyme stability-abolishing mutations may not explain all aspects of the underlying defect. This was further explored by analysing a possible impact on the splicing behaviour resulting from the respective mutation. Supplementary Tables S5 A and B report on a few variants where the effect on splicing at exon-intron junctions as well as on cryptic splicing sites may contribute to their pathogenicity (for method see Text S1 in the supplementary material). However, the majority of missense mutations exhibit no noticeable effect on splicing.

Altogether, we propose to conduct all available testing (i.e. lyso-Gb3, *in silico* and *in vitro* overexpression analysis) to determine the

characteristics of novel *GLA* mutations even though in severe cases it might be sufficient to test just one of the discussed parameters, but detailed testing might reveal unforeseen damage to the enzyme and its function. We point out the usefulness of a readily available *in vitro* assay in deciding whether an ERT or alternative treatment strategies (DGJ is currently being tested in clinical trials) should be started and whether a patient should be treated after pre-symptomatic genetic diagnosis at all, showing increased FD risk (harbouring a mutation with residual activity lower 20%) or if a frequent follow-up on the patient is initially more reasonable. However, only epidemiologic studies could determine to which extent mutations with a substantially lowered, yet high, enzyme activity (e.g. *p.D83N*, *p.S126G*, *p.A143T*, *p.D3131Y*) contribute to the symptoms of Fabry disease and what role the degree of variation plays among individuals.

Materials and Methods

Patients and blood samples

Blood samples were obtained from patients undergoing biochemical analysis or genetic testing for verification of Fabry disease by the Albrecht-Kossel-Institute for Neuroregeneration (AKos). All patients agreed for testing of their blood samples. The project was in concordance with the regulations of the local Ethical Committee of the University Rostock.

Cell culture

HEK293H cells were maintained in DMEM (Dulbecco's Modified Eagle Medium, Invitrogen, Karlsruhe, Germany) supplemented with 10% FBS (fetal bovine serum; PAA Laboratories, Pasching, Austria) and 1% penicillin/streptomycin (Invitrogen, Karlsruhe, Germany). All cells were incubated in a water-jacket incubator (Binder, Tuttlingen, Germany) at 37°C under 5% CO₂ atmosphere. DGJ (Sigma Aldrich, Munich, Germany) was added to the culture medium from an aqueous stock solution (10 mM).

Cloning of α -galactosidase A in an overexpression vector

A plasmid containing the full length cDNA of α -galactosidase A (IRAU_{p969H0320D}, aligned to accession no. NM_000169.2) was obtained from ImaGenes GmbH, Berlin, Germany. Amplification for subcloning was performed using cloned *Pfu* DNA polymerase (Stratagene, La Jolla CA, USA), with the primers 5'-AGGTCCG-GATCCG ACAATGCAGCTGAGGAACC-3' (forward) and 5'-GGTGTTCGAATTAAAGTAAGT CTTTAAATGACATCT-GCA-3' (reverse) introducing unique restriction sites for *Bam*HI and *Bst*BI. The amplicon was inserted into mammalian expression vector pcDNA3.1/V5-His₆ (Invitrogen, Karlsruhe, Germany).

Site-directed mutagenesis of α -galactosidase A

Expression vectors harbouring α -Gal A mutations were generated by site-directed PCR mutagenesis (3,28) using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). Nucleotide exchanges, deletions or insertions were individually introduced by PCR amplification with *Pfu*Ultra DNA polymerase, the pcDNA3.1/*GLA* plasmid vector containing the wild type sequence was used as template and a 27–37-mer primer set, with sense and antisense primers carrying one of the respective sequence modifications central to their length. Each mutant plasmid was sequenced on a 3130 xl Genetic Analyzer (Applied Biosystems, Darmstadt, Germany).

Transient expression of mutant enzymes in HEK293H cells

1.5 × 10³ cells were seeded 24 hours before transfection in each well of a 24-well culture plate using 500 μ l DMEM medium

(Invitrogen, Karlsruhe, Germany) supplemented with 10% Fetal bovine serum (PAA Laboratories, Pasching, Austria). Transient expression of mutant enzymes in HEK293H cells was carried out using Lipofectamine 2000 transfection reagent (Invitrogen, Karlsruhe, Germany), according to the manufacturer's protocol. Typically, prior to transfection, a mixture of plasmid DNA (0.8 μ g) and Lipofectamine 2000 transfection reagent (2 μ l) in 100 μ l of serum-free DMEM or Opti-MEM medium (Invitrogen, Karlsruhe, Germany) was incubated at room temperature for 20 min and applied to the cells thereafter. The cell layer was subsequently incubated for 6 hours at 37°C, the medium containing the transfection reagent was removed and 500 μ l fresh DMEM was added. In this step, the DGJ was added where intended. The cells were incubated for another 60 hours and harvested.

Enzymatic measurement of α -galactosidase A

Cell pellets obtained from confluent 24-well cell culture plates were homogenised in 200 μ l water and subjected to 5 freeze-thaw cycles using liquid nitrogen. The supernatant collected after centrifugation of the homogenate at 10000 × g for 5 min was used in enzyme assays. Protein concentration was measured with the BCA protein assay kit (Thermo Scientific, Braunschweig, Germany) according to the manufacturer's manual. 10 μ l of the cell lysates at a concentration of 50 μ g/ml were assayed with 20 μ l of 4-MU- α -D-galactopyranoside (2 mM, Sigma Aldrich, Munich, Germany) in 0.06 M phosphate citrate buffer (pH 4.7) with some adaptations from the original method described by Desnick et al. [40]. Enzyme reactions were terminated by the addition of 0.2 ml of 1.0 M glycine buffer (pH 10.5), prepared by adjusting the pH using 1.0 M NaOH. The released 4-MU was determined by fluorescence measurement at 360 and 465 nm as the excitation and emission wavelengths respectively, using a microplate fluorescence reader (Tecan, Männedorf, Switzerland). The measured enzyme activity was calculated as nmol 4-MU/mg protein and normalised to one hundred percent wild-type activity.

Western blot analysis

Western blot analysis for the detection of α -Gal A protein was performed using a custom made rabbit anti- α -Gal A polyclonal antibody from Eurogentec, Cologne, Germany (animals were immunised against peptides QRDSEGR_{LQ}ADPQRFP (corresponding to amino acids 99–114) and KQGYQLRQGDN-FEVWE (corresponding to amino acids 326–341). Furthermore, a mouse GAPDH monoclonal antibody 6C5 (Abcam, Cambridge, UK) was used to visualise GAPDH as an internal loading control.

HEK293H cell lysates were generated by aspirating the media from the 24-well culture plates, washing the cells once with 1 × PBS (Biochrom AG, Berlin, Germany) and directly applying 200 μ l ice cold RIPA buffer supplemented with protease inhibitor cocktail tablets (Roche Applied Science, Penzberg, Germany) prior to a 20 minute incubation on ice. The cells were then rinsed from the wells, transferred to microcentrifuge tubes and spun at 14000 × g for 10 minutes at 4°C to pellet debris. The supernatant was used for the analysis. 50 μ g protein were mixed with a suitable volume of 5 × Laemmli loading buffer, boiled for 5 minutes on a thermo shaker, centrifuged at 14000 × g for 10 minutes at 4°C and loaded on a Criterion precast 4–15% Tris-HCl gel (Bio-Rad, Munich, Germany). Proteins were transferred electrophoretically to a nitrocellulose (Amersham Hybond ECL) membrane (GE Healthcare, Munich, Germany). The membrane was blocked with 5% (w/v) non-fat dried skimmed milk in TBS-Tween 20 [10 mM Tris/HCl (pH 7.5) with 150 mM NaCl and 0.1% Tween 20] at room temperature for 1 hour, and then treated with a primary antibody against GAPDH diluted 1:10,000 in a milk/blot solution

[3% (w/v) non-fat dried skimmed milk in TBS-Tween 20] at 4°C overnight. The blot was then washed three times with excess TBS-Tween 20 and treated with a primary antibody against α -Gal A diluted 1:500 in the 3% milk/blot solution for 1 hour at room temperature. After another wash procedure, a secondary antibody mix of an Alexa Fluor labelled 680 goat anti-rabbit IgG antibody (Molecular Probes, Karlsruhe, Germany) and an IRDye800 conjugated goat anti-mouse IgG antibody (Rockland-Biomol, Hamburg, Germany) both diluted 1:10,000 in the 3% milk/blot solution was applied to the membrane. Following extensive washing with TBS-Tween 20, protein bands were visualised by an Odyssey Infrared Imager (Li-Cor Biosciences, Lincoln, NE, USA). Quantification and protein size determination were performed using the Odyssey software.

Determination of kinetic properties of α -galactosidase A mutants

α -Gal A mutants were overexpressed in HEK293H cells. 60 hours after transfection in 24-well plates, the cells were harvested in 100 μ l H₂O and the contents of two wells were pooled. Cells were lysed by 5 freeze-thaw cycles and subsequent centrifugation to obtain a cell-debris free lysate. Quantitative Western blot analysis was carried out to reveal the ratio of the mutant enzyme : wild type. The amount of lysate required to obtain equal amounts of enzyme was calculated. Respective quantities of wild type and mutant enzyme were subjected to the kinetic evaluation described earlier [41]. Substrate concentrations of 0.5, 1, 2, 4, 8 and 28 mM were assayed with constant amount of enzyme in a Lineweaver Burk plot in order to obtain kinetic parameters. To prove that equal amounts of input enzyme (mutants and wild type) were employed for the assay, α -Gal A content of the lysates was controlled with another Western blot before applying the substrate. (Figure S2A). For standard curve, 5, 10, 20, 50 and 100 ng of agalsidase alpha (Replagal, Shire Human Genetic Therapeutics, Dublin, Ireland) were subjected to Western blot analysis and fluorescence values were plotted (see Figure S2B). In the individual experiments between 2 and 6 ng of enzyme were employed.

Lyso-Gb3 determination

As reference standards (Matreya LLC, USA; purity >98%) we used lyso-ceramide trihexoside (= lyso-globotriaosylsphingosine = lyso-Gb3). Internal standards were D5-fluticasone propionate (EJY Tech., Inc. Rockville, MD, USA).

We used a gradient HPLC method on a reversed phase column (ACE 3 C8, 50 \times 2.1 mm). The two HPLC pumps and the column oven PE Series 200 were provided by Perkin Elmer, USA. The mass spectrometer used was an API 4000 Q-Trap supplied by Applied Biosystems, USA. The following experimental conditions were used: column temperature 60°C, flow at 0.9 mL/min, injection volume 10 μ L, mobile phase with 50 mM formic acid in water (A) and 50 mM formic acid in acetonitrile/acetone (1/1 = v/v; B), gradient at 5% B from 0 to 0.3 minutes, followed by a linear gradient up to 73% B (0.3 to 2.6 minutes) and further on to 100% B (2.6 to 5.7 minutes). From 5.7 to 6.7 minutes 100% B was used. Re-equilibration was done from 6.7 to 7.5 minutes at 5% B. ESI in positive mode was used for peak detection.

The detection mode was MRM, the vaporizer temperature was set at 500°C, ionisation voltage was 5.5 kV, curtain gas pressure was 40 psi. Lyso-Gb3 quantifier was 786.6 to 282.2 m/z and 506.3 to 313.0 m/z for the internal standard D5-fluticasone propionate.

For the sample analysis, 50 μ l aliquots were used. 100 μ l of internal standard working solution (in ethanol) were added.

Samples were mixed for about 30 seconds and centrifuged at 4,000 rpm for 2 minutes. The clear supernatant was transferred into appropriate auto sampler vials which were closed thereafter with crimp caps. Normal and pathological values were determined in 145 healthy controls and 275 Fabry patients with genetically confirmed disease.

Selection of mutations

The following criteria influenced the decisions whether a given mutant was to be included in the study:

1. The mutation was novel or has not been well described (no biochemical activity data were available; N = 138)
2. The mutation lead to a single amino acid substitution (N = 147)
3. Patient-derived data (phenotype severity, biomarker lyso-Gb3) are available (N = 100)
4. Three sites (R118, S126, D264) were selected for “permuting” the native amino acid into every possible residue resulting from single nucleotide exchanges. R118C and S126G are involved in variant forms of the FD phenotype and found to be more prevalent (9, own data). D264V/Y are involved in classical FD (N = 15)

Statistical analysis

We used Spearman's rho (r_s) rank correlation coefficient to test associations between GLA level and enzyme activity (Figure 1). Linear trend tests were used to test associations between two ordinal variables, for example enzyme activity class and DGJ responsiveness (Table 1). To compare the predictive values of female lyso-Gb3 values, male lyso-Gb3 values, *in vitro* enzyme activity and PolyPhen2-scores with regard to clinical phenotype on the mutation level we used ordinal regression with only one of the four different prediction measures as the independent variable and with ‘clinical phenotype’ as the outcome. This regression analysis was performed for all mutations where values on specific prediction measure and the outcome were available. For comparison we repeated the regression analysis for the 21 mutations of which values for all four measures were available (Table 2). We used Nagelkerke's R-Square as a measure of explained variance, the -2 Log-Likelihood as a measure of Goodness of Fit of the model and the proportion of correct classified mutations. All analyses were done using IBM SPSS Statistics, Release 20.0.0 (SPSS, Inc., 2011, Chicago, IL, www.spss.com).

Supporting Information

Figure S1 Detailed view of the employed analysis. Applied to analysis were residual enzyme activity with (black bars) or without (grey bars) the addition of 20 μ M DGJ to the medium and the respective change in cellular protein level semi-quantitatively calculated on a fluorescence reader. The Western Blot method used here was not able to detect endogenously expressed α -Gal A enzyme. Values are mean \pm SEM. * p <0.05, ** p <0.01, *** p <0.005. (TIF)

Figure S2 **A.** Control Western Blot. Semiquantitative analysis of α -Gal A level determined by Western Blot were used to examine kinetic properties of the mutant enzymes compared to the wild type. **B.** Recombinant enzyme agalsidase alpha used for Enzyme Replacement Therapy was kindly provided by Shire Human Genetics Therapeutics to estimate total amount subjected to the kinetic assay [see Material and Methods section] (TIF)

Table S1 Overview of mutations tested *in vitro* for enzyme activity. Number of experiments (enzyme activity) and patient numbers (lyso-Gb3) are indicated in brackets. The α -Gal A activity limit of quantification (LOQ) in HEK293H cells was defined as 235.3 nmol 4-MU/mg protein for untreated mutations and 292.5 nmol 4-MU/mg protein for 20 μ M DGJ treated mutations, respectively, which accounts for 95% of the values obtained from empty vector only transfections. Number of experiments (enzyme activity) and patient numbers (lyso-Gb3) are indicated in brackets. Note that even though females have much lower values of lyso-Gb3 all 6 mutations that caused no elevated lyso-Gb3 in males likewise caused no elevation in females (where applicable), indicating that these mutations may not lead to an accumulation. [†]Disease phenotype is conventionally divided into classic, variant and classic/variant. The latter is used for mutations where variant and classic types of FD are reported or a mild classical phenotype is observed. [‡]Mutations have not been described in patients yet. [§]*F396Y was terminated from HGMD. Not a genomic mutation is responsible for the finding. As underlying mechanism RNA editing was proposed. (DOC)

Table S2 Association of *in vitro* enzyme activity with biochemical and crystallographic data. *In vitro* enzyme activity is associated with the responsiveness to pharmacological chaperone DGJ. As another biochemical parameter DGJ responsiveness is demonstrated to be associated to residual enzyme activity. Enzyme activity shows only a weak linear trend with the parameter “accessible surface area” obtained from crystallographic studies. Accessible surface area is defined as the “average accessibility of each atom in the residue” [33]. However, this model does not take active site residues with a usually high surface accessibility that display low residual activity into consideration. Cut points for accessible surface area were extracted from Garman (2007) [33]. (DOC)

Table S3 Lyso-Gb3 values of classic or presumed classic mutations. Lyso-Gb3 was measured in male and female Fabry patients. The mean is displayed in the table. Generally males have

much higher lyso-Gb3-levels than females. With three exceptions (data obtained from one female patient harbouring the mutation *p.A20P*, *p.W262**, *p.W399**, respectively) all of the mutations shown here caused elevated lyso-Gb3 values above the pathological cut-off of 0.9 ng/ml. * All females are heterozygotes. (DOC)

Table S4 Kinetic properties of α -Gal A mutants. Asterisks indicate a significant change towards the wild type enzyme ($p < 0.05$). Agalsidase alfa has been tested to validate comparability of the assay. (DOC)

Table S5 Putative impact on splicing. **A:** Effect of missense mutations on natural acceptor and donor splicing sites. Analysis of the splice-sites revealed 3 possible splice-site abolishments. **B:** Effect of missense mutations on cryptic and novel acceptor and donor sites. The analysis suggested possible changes in 3 sites involving the activation of cryptic and novel donor/acceptor sites. Note: the range of values is given in parenthesis for each splicing algorithm. (DOC)

Text S1 Additional method. Description of the method applied to reveal alterations of mRNA splicing. (DOC)

Acknowledgments

The authors would like to thank Mandy Löbert, Tina Czajka and Michael Klingenhäger for their important technical support. Daniel Mascher was supportive in the quantification of lyso-Gb3 in plasma. Replagal used in this study was a kind gift of Dr. Werner Foeller (Shire Human Genetics Therapies).

Author Contributions

Conceived and designed the experiments: JL AKG AM UG EK VS AR. Performed the experiments: JL AKG UG HM KJL PW VS. Analyzed the data: JL AM UG VS WM. Contributed reagents/materials/analysis tools: AM UG HM KJL VS AR. Wrote the paper: JL AKG AM UG WM AR.

References

- Spada M, Pagliardini S, Yasuda M, Tükel T, Thiagarajan G, et al. (2006) High incidence of Later-Onset Fabry Disease Revealed by Newborn Screening. *Am J Hum Genet* 79:31–40.
- Wittmann J, Karg E, Turi S, Legnini E, Wittmann G, et al. (2012) Newborn Screening for Lysosomal Storage Disorders in Hungary. *JIMD Rep* 6: 117–125.
- Ishii S, Chang HH, Kawasaki K, Yasuda K, Wu HL, et al. (2007) Mutant α -galactosidase A enzymes identified in Fabry disease patients with residual enzyme activity: biochemical characterization and restoration of normal intracellular processing by 1-deoxygalactonojirimycin. *Biochem J* 406 (Pt 2):285–295.
- Kotanko P, Kramar R, Devrnja D, Paschke E, Voigtländer T, et al. (2004) Results of a Nationwide Screening for Anderson-Fabry Disease among Dialysis Patients. *J Am Soc Nephrol* 15:1323–1329.
- Mehta A, Beck M, Eyskens F, Feliciani C, Kantola I, et al. (2010) Fabry disease: a review of current management strategies. *QJM* 103:641–59.
- Rolfs A, Boettcher T, Zschiesche M, Morris P, Winchester B, et al. (2005) Prevalence of Fabry disease in patients with cryptogenic stroke: a prospective study. *Lancet* 366:1794–6.
- Nakao S, Takenaka T, Maeda M, Kodama C, Tanaka A, et al. (1995) An atypical variant of Fabry's disease in men with left ventricular hypertrophy. *N Engl J Med* 333:288–93.
- Kase R, Bierfreund U, Klein A, Kolter T, Utsumi K, et al. (2000) Characterization of two alpha-galactosidase mutants (Q279E and R301Q) found in an atypical variant of Fabry disease. *Biochim Biophys Acta* 1501:227–35.
- Sims K, Politei J, Banikazemi M, Lee P (2009) Stroke in Fabry disease frequently occurs before diagnosis and in the absence of other clinical events: natural history data from the Fabry Registry. *Stroke* 40:788–94.
- Sachdev B, Takenaka T, Teraguchi H, Tei C, Lee P, et al. (2002) Prevalence of Anderson-Fabry disease in male patients with late onset hypertrophic cardiomyopathy. *Circulation* 105:1407–11.
- Eng CM, Resnick-Silverman LA, Nichaus DJ, Astrin KH, Desnick RJ (1993) Nature and frequency of mutations in the alpha-galactosidase A gene that cause Fabry disease. *Am J Hum Genet* 53:1186–97.
- Topaloglu AK, Ashley GA, Tong B, Shabbeer J, Astrin KH, et al. (1999) Twenty Novel Mutations in the α -Galactosidase A Gene Causing Fabry Disease. *Mol Med* 5:806–811.
- Brouns R, Thijs V, Eyskens F, Van den Broeck M, Belachew S, et al. (2010) Belgian Fabry Study: Prevalence of Fabry Disease in a Cohort of 1000 Young Patients with Cerebrovascular Disease. *Stroke* 41:863–8.
- Froissart R, Guffon N, Vanier MT, Desnick RJ, Maire I (2003) Fabry disease: D313Y is an alpha-galactosidase A sequence variant that causes pseudodeficient activity in plasma. *Mol Genet Metab* 80:307–14.
- Yasuda M, Shabbeer J, Benson SD, Maire I, Burnett RM, et al. (2003) Fabry Disease: Characterization of α -Galactosidase A Double Mutations and the D313Y Plasma Enzyme Pseudodeficiency Allele. *Hum Mutat* 22:486–492.
- Baptista MV, Ferreira S, Pinho-E-Melo T, Carvalho M, Cruz VT, et al. (2010) Mutations of the GLA gene in young patients with stroke: the PORTY-STROKE study—screening genetic conditions in Portuguese young stroke patients. *Stroke* 41:431–6.
- Linthorst GE, Poorthuis BJ, Hollak CE (2008) Enzyme Activity for Determination of Presence of Fabry Disease in Women Results in 40% False-Negative Results. *J Am Coll Cardiol* 51:2082; author reply 2082–3.
- Aerts JM, Groener JE, Kuiper S, Donker-Koopman WE, Strijland A, et al. (2008) Elevated globotriaosylsphingosine is a hallmark of Fabry disease. *Proc Natl Acad Sci U S A* 105:2812–7.
- Sanchez-Niño MD, Sanz AB, Carrasco S, Saleem MA, Mathieson PW, et al. (2011) Globotriaosylsphingosine actions on human glomerular podocytes: implications for Fabry nephropathy. *Nephrol Dial Transplant* 26:1797–802.
- Rombach SM, Dekker N, Bouwman MG, Linthorst GE, Zwinderman AH, et al. (2010) Plasma globotriaosylsphingosine: Diagnostic value and relation to clinical manifestations of Fabry disease. *Biochim Biophys Acta* 1802:741–8.

21. Shimotori M, Maruyama H, Nakamura G, Suyama T, Sakamoto F, et al. (2008) Novel mutations of the GLA gene in Japanese patients with Fabry disease and their functional characterization by active site specific chaperone. *Hum Mutat* 29:331.
22. Altarescu GM, Goldfarb LG, Park KY, Kaneski C, Jeffries N, et al. (2001) Identification of fifteen novel mutations and genotype-phenotype relationship in Fabry disease. *Clin Genet* 60:46–51.
23. Garman SC, Garbozi DN (2004) The Molecular Defect Leading to Fabry Disease: Structure of Human α -Galactosidase. *J Mol Biol* 337:319–335.
24. Matsuzawa F, Aikawa S, Doi H, Okumiya T, Sakuraba H (2005) Fabry disease: correlation between structural changes in α -galactosidase, and clinical and biochemical phenotypes. *Hum Genet* 117:317–328.
25. Saito S, Ohno K, Sese J, Sugawara K, Sakuraba H (2010) Prediction of the clinical phenotype of Fabry disease based on protein sequential and structural information. *J Hum Genet* 55:175–8.
26. Andreotti G, Guarradino MR, Cammisa M, Correria A, Cubellis MV (2010) Prediction of the responsiveness to pharmacological chaperones: lysosomal human alpha-galactosidase, a case of study. *Orphanet J Rare Dis* 5:36.
27. Filoni C, Caciotti A, Carraresi L, Cavicchi C, Parini R, et al. (2010) Functional studies of new *GLA* gene mutations leading to conformational Fabry disease. *Biochim Biophys Acta* 1802:247–52.
28. Shin SH, Kluepfel-Stahl S, Cooney AM, Kaneski CR, Quirk JM, et al. (2008) Prediction of response of mutated alpha-galactosidase A to a pharmacological chaperone. *Pharmacogenet Genomics* 18:773–80.
29. Benjamin ER, Flanagan JJ, Schilling A, Chang HH, Agarwal L, et al. (2009) The pharmacological chaperone 1-deoxygalactonojirimycin increases alpha-galactosidase A levels in Fabry patient cell lines. *J Inher Metab Dis* 32:424–40.
30. Wu X, Katz E, Della Valle MC, Mascioli K, Flanagan JJ, et al. (2011) A Pharmacogenetic Approach to Identify Mutant Forms of α -Galactosidase A that Respond to a Pharmacological Chaperon for Fabry Disease. *Hum Mutat* 32:965–77.
31. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–9.
32. Desnick RJ, Ioannou YA, Eng CM (2001) The metabolic and molecular basis of inherited disease, ed 8, (New York, USA, McGraw-Hill), pp. 3733–3774.
33. Garman SC (2007) Structure–function relationships in α -galactosidase A. *Acta Paediatrica Suppl* 96:6–16.
34. van Breemen MJ, Rombach SM, Dekker N, Poorthuis BJ, Linthorst GE, et al. (2011) Reduction of elevated plasma globotriaosylsphingosine in patients with classic Fabry disease following enzyme replacement therapy. *Biochim Biophys Acta* 1812:70–6.
35. Niemann M, Rolfs A, Giese A, Mascher H, Breunig F, Ertl G, Wanner C, Weidemann F (2013) Lyso-Gb3 Indicates that the Alpha-Galactosidase A Mutation D313Y is not Clinically Relevant for Fabry Disease. *JIMD Rep* 7:99–102.
36. Lukas J, Torras J, Navarro I, Giese A-K, Böttcher T, et al. (2012) Broad spectrum of Fabry disease manifestation in an extended Spanish family with a new deletion in the *GLA* gene. *Clin Kidney J* 5: 395–400.
37. Sakuraba H, Oshima A, Fukuhara Y, Shimmoto M, Nagao Y, et al. (1990) Identification of point mutations in the α -Galactosidase A Gene in Classical and Atypical hemizygotes with Fabry Disease. *Am J Hum Genet* 47:784–789.
38. Fan JQ, Ishii S (2003). Cell-Based Screening of Active-Site Specific Chaperone for the Treatment of Fabry Disease. *Methods Enzymol* 363:412–420.
39. Eng CM, Ashley GA, Burgert TS, Enriquez AL, D'Souza M, et al. (1997) Fabry disease: thirty-five mutations in the alpha-galactosidase A gene in patients with classic and variant phenotypes. *Mol Med* 3:174.
40. Desnick RJ, Allen KY, Desnick SJ, Raman MK, Bernlohr RW, et al. (1973) Fabry's disease: enzymatic diagnosis of hemizygotes and heterozygotes. Alpha-galactosidase activities in plasma, serum, urine, and leukocytes. *J Lab Clin Med* 81:157–71.
41. Beutler E and Kuhl W (1972) Purification and Properties of Human α -Galactosidases. *J Biol Chem* 217:7195–7200.

cDNA change	AA change	MaxEnt	NNSPLICE	HSF	Comment
c194G>T	<i>p.S65I</i>	-100%	-99%	-14.10%	Predicted change at donor site 1 bps downstream: -71.0%
c548G>T	<i>p.G183V</i>	-18.80%	-6.50%	-4.60%	Predicted change at acceptor site 1 bps upstream: -9.9%
c.638A>T	<i>p.K213M</i>	16.70%	-4.90%	-5.00%	Predicted change at donor site 2 bps downstream: -8.9%

Supplementary Table S5A

cDNA change	AA change	MaxEnt (0-12)	NNSPLICE (0-1)	HSF (0-100)	Comment
c358C>G	<i>p.L120V</i>	0 ⇒ 6.52	0 ⇒ 0.45	0 ⇒ 79.65	New donor site may be activated
c1025G>T	<i>p.R342L</i>	0 ⇒ 1.93	0 ⇒ 0.43	87.37 ⇒ 90.53 (+3.6%)	Cryptic acceptor site may be activated
c1115T>C	<i>p.L372P</i>	4.20 ⇒ 6.23 (+48.5%)	0 ⇒ 0.71	79.06 ⇒ 81.04 (+2.5%)	Cryptic donor site may be activated

Supplementary Table S5B

Supplementary Table S4:

Mutation	K_M (mM)	V_{max} (mmol MU/mg protein/h)
Agalsidase alfa	3.7	1.8
Wild-type	2.9	2.0
H46P	8.6*	2.8
R49G	n.d.	n.d.
E59K	5.0*	1.6
S65I	n.d.	n.d.
A143T	3.6	3.7
A156V	1.8	1.1
D231N	n.d.	n.d.
R301P	2.4	0.3*
R301Q	3.5	1.7
L310F	4.6	1.0
L415P	n.d.	n.d.

Supplementary Table S3:

amino acid	cDNA	Lyso Gb3 (mean) Norm < 0.9 ng/ml		mutation class	clinical phenotype	Reference
		male	female*			
<i>p.Q2*</i>	c4C>T	13.5 (1)		nonsense		own data
	c35_58del	26.9 (1)	5.3 (3)	small deletion		own data
<i>p.A20P</i>	c58G>C	8.5 (1)	0.7 (1)		classic	Nakao (1995) N Engl J Med 333:288
<i>p.W24*</i>	c71G>A		7.4 (2)	nonsense		own data
<i>p.L45P</i>	c134T>C	42.7 (2)	9.8 (1)			own data
	c162delT	16.9 (1)		small deletion		Schafer (2005) Hum Mutat 25:412
<i>p.C56Y</i>	c167G>A	33.1 (1)	3.4 (1)			Davies (1996) Eur J Hum Genet 4:219
<i>p.I91T</i>	c272T>C		0.7 (1)			
Exon 2 del	g.100545168_100545720del		8.0 (6)	gross deletion		Lukas (2012) Clin Kidney J 5: 395–400
	c354_368del		10.1 (1)	small deletion		own data
	c369+1G>A	65.1 (1)	5.7 (1)	splice		Ashton-Prolla (2000) J Investig Med 48:227
<i>p.Y151*</i>	c453C>A	126.0 (1)		nonsense		Shabbeer (2005) Hum Mutat 25:299
<i>p.W157*</i>	c469C>T	16.5 (1)		nonsense	classic	Eng (1994) Hum Mol Genet 3:1795
<i>p.W162C</i>	c486G>T	28.5 (1)	5.6 (1)			Germain (1996) Hum Genet 98:719
<i>p.V164G</i>	c491T>G	19.5 (1)	2.1 (1)			own data
<i>p.L167Q</i>	c500T>A		10.7 (1)			own data
	c547+1G>A	82.7 (3)	7.3 (3)	splice		Ashton-Prolla (2000) J Investig Med 48:227
	c568delG	42.9 (2)	12.9 (1)	small deletion	classic	Shabbeer (2002) Mol Genet Metab 76:23
	c571delC	83.2 (1)	10.6 (1)	small deletion		own data
<i>p.C202Y</i>	c605G>A	31.8 (1)	5.5 (1)			Eng (1997) Mol Med 3:174
<i>p.W204*</i>	c611G>A	102.0 (1)	6.4 (1)	nonsense		Ashton-Prolla (2000) J Investig Med 48:227

	c639+1G>A		5.2 (1)	splice	classic	Rodriguez-Mari (2003) Hum Mutat 22:258
<i>p.R220X</i>	c658C>T		3.7 (2)	nonsense	classic	Meaney (1994) Hum Mol Genet 3:1019
<i>p.R227X</i>	c679C>T	50.3 (3)	6.0 (2)	nonsense	classic	Davies (1993) Hum Mol Genet 2:1051
	c718_719delAA	62.6 (1)	5.4 (1)	small deletion	classic	Davies (1994) Hum Mol Genet 3:667
	c.719delA	172.0 (1)		small deletion		own data
	c742_743delTT	84.9 (1)	4.9 (2)	small deletion		own data
	c776delC	30.7 (1)		small deletion		own data
	c777delA	45.4 (1)	13.0 (1)	small deletion		Eng (1997) Mol Med 3:174
<i>p.W262*</i>	c785G>A		0.9 (1)	nonsense		Shabbeer (2006) Hum Genomics 2:297
	c801+2T>C		16.6 (1)	splice	.	Blaydon (2001) Hum Mutat 18:459
	c802delT	80.5 (1)	13.3 (3)	small deletion		own data
<i>p.Q280*</i>	c838C>T		6.5 (2)	nonsense		own data
<i>p.R301*</i>	c901C>T	16.4 (2)	6.3 (3)	nonsense	classic	Eng (1994) Hum Mol Genet 3:1795
	c.913delC		5.4 (1)	small deletion		own data
<i>p.Q312*</i>	c934C>T	71.7 (2)	9.3 (2)	nonsense		own data
<i>p.Q321*</i>	c961C>T		6.4 (1)	nonsense		Schafer (2005) Hum Mutat 25:412
	c993_994insA	68.8 (1)	5.8 (3)	small insertion		own data
<i>p.Q333*</i>	c997C>T	25.7 (1)		nonsense		Lorenz (2003) Wien Klin Wochenschr 115:235
<i>p.W340*</i>	c1020G>A		5.5 (1)	nonsense	classic	Eng (1993) Am J Hum Genet 53:1186
<i>p.R342*</i>	c1025G>A	82.3 (2)	4.8 (2)	nonsense	classic	Ploos van Amstel (1994) Hum Mol Genet 3:503
<i>p.R342P</i>	c1025G>C		3.3 (1)			own data
	c1029-1030delTC	94.2 (1)		small deletion		own data
<i>p.S345*</i>	c1034C>A	119.0 (1)		nonsense		own data
<i>p.W349*</i>	c1046G>A	66.2 (1)		nonsense		Ashley (2001) J Hum Genet 46:192
<i>p.Q357*</i>	c1069C>T	122.0 (1)		nonsense		Cooper (2000) Hum Genet 107:535
<i>p.E358*</i>	c1072G>T		7.6 (1)	nonsense		own data

<i>p.Y397*</i>	c1191T>G		9.4 (1)	nonsense		own data
<i>p.W399*</i>	c1196G>A	88.5 (3)	0.5 (1)	nonsense	classic	Eng (1997) Mol Med 3:174
	c1208delT	66.2 (1)	6.1 (2)	small deletion	classic	Germain (2002) Mol Med 8:306
	c1235-1236delCT	21.3 (4)	5.2 (4)	small deletion		Blaydon (2001) Hum Mutat 18:459
	c1277-1278delAA		6.5 (1)	small deletion		Yasuda (2003) Hum Mutat 22:486

Supplementary Table S2:

	enzyme activity				p (for linear trend test)
	0%	>0% - 20%	≥20%-60%	≥60%	
DGJ responsiveness					
not responsive (n=90)	69	5	8	8	<0.001
responsive (n=69)	14	24	20	11	
accessible surface area [A²]					
0-1 (buried, 56)	38	11	6	1	0.121
1-10 (partially buried, 35)	20	7	5	3	
10-40 (exposed)	4	0	1	1	

Supplementary Table S1:

amino acid	cDNA	<i>in vitro</i> enzyme activity [% WT] in mean ± SEM (N)		Ratio DGJ/unreated	Lyso-Gb3 [mean (N)]		clinical phenotype ^ϕ	PolyPhen2	Reference
		-DGJ	+DGJ		male	female			
<i>p.E7*</i>	c19G>T	0(3)	0(3)	n.d.					Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.A15E</i>	c44C>A	0(3)	0(3)	n.d.	40.7 (2)			benign	own data
<i>p.W24C</i>	c72G>T	45.6±11.2(3)	89.8±21.0(3)	2.0				benign	own data
<i>p.G35R</i>	c103G>C	17.9±2.1(3)	72.3±4.7(3)	4.0	68.9 (3)	6.8 (4)	variant	probably damaging	Davies (1994) Hum Mol Genet 3:667
<i>p.L36S</i>	c107T>C	0(3)	0(3)	n.d.				probably damaging	Erdos (2008) Mol Genet Metab 95:224
<i>p.A37T</i>	c109G>A	69.6±13.6(3)	132.9±27.8(3)	1.9	1.6 (1)			probably damaging	own data
<i>p.M42V</i>	c124A>G	0(3)	11.9±2.0(3)	n.d.	111.0 (1)		classic	probably damaging	Davies (1996) Eur J Hum Genet 4:219
<i>p.G43S</i>	c127G>A	0(3)	0(4)	n.d.				probably damaging	own data
<i>p.H46P</i>	c137A>C	40.1±1.2(3)	98.8±2.1(3)	2.5				benign	Hwu (2009) Hum Mutat 30:1397
<i>p.R49G</i>	c145C>G	0(3)	0(3)	n.d.	75.8 (1)	7.3 (3)	classic	probably damaging	Germain (2002) Mol Med 8:306
<i>p.R49C</i>	c145C>T	0(3)	5.1±0.5(3)	n.d.	89.9 (1)	21.5 (1)		probably damaging	own data
<i>p.M51K</i>	c152T>A	0(4)	8.7±4.4(4)	n.d.			classic	benign	Ashley (2001) J Hum Genet 46:192
<i>p.M51I</i>	c153G>A	37.4±5.9(7)	62.0±16.2(3)	1.7			variant	benign	Spada (2006) Am J Hum Genet 79:31
<i>p.C52W</i>	c156C>G	0(3)	0(3)	n.d.				probably damaging	own data
<i>p.E59K</i>	c175G>A	2.2±0.5(8)	18.5±2.9(4)	14.2			classic	benign	Eng (1994) Hum Mol Genet 3:1795
<i>p.C63Y</i>	c188G>A	0(3)	0(3)	n.d.	77.3 (1)	5.4 (2)		probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.S65I</i>	c194G>T	0(4)	11.2±1.2(4)	n.d.				probably damaging	own data
<i>p.E66K</i>	c196G>A	6.8±0.3(3)	18.3±2.3(3)	2.7	6.1 (2)		classic	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.L68F</i>	c202C>T	0(3)	4.5±1.8(3)	n.d.	68.4 (1)		classic	probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.A73V</i>	c218C>T	44.0±8.0(8)	64.7±8.3(3)	1.5			variant	probably damaging	Spada (2006) Am J Hum Genet 79:31
<i>p.K82*</i>	c244A>T	0(3)	0(4)	n.d.					Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.D83N</i>	247G>A	62.9±12.8(4)	71.6±9.7(4)	1.1		0.5 (1)	variant	probably damaging	own data
<i>p.Y86D</i>	c256T>G	0(3)	0(4)	n.d.				probably damaging	Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.L89H</i>	c266T>A	0(3)	0(3)	n.d.				possibly damaging	own data

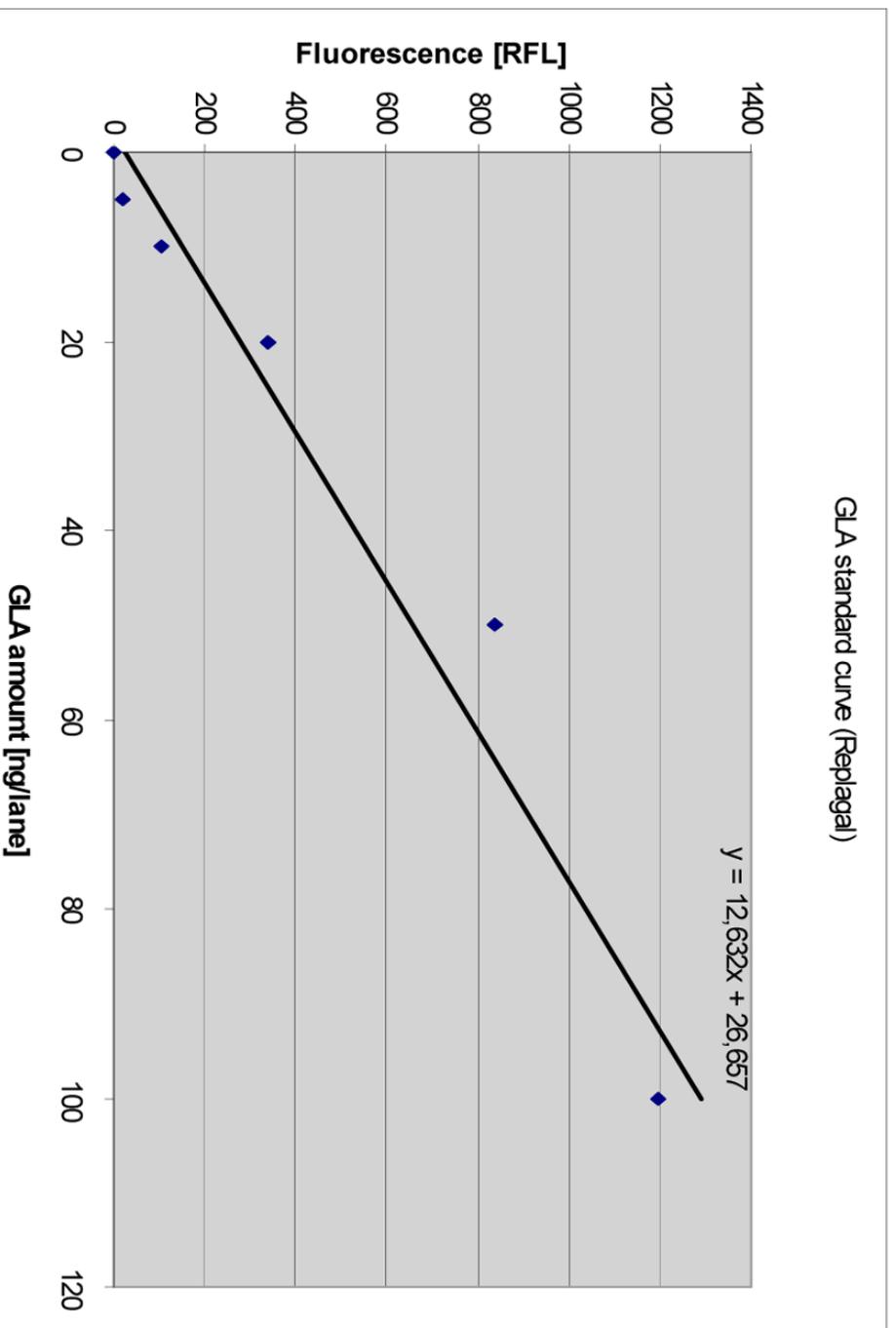
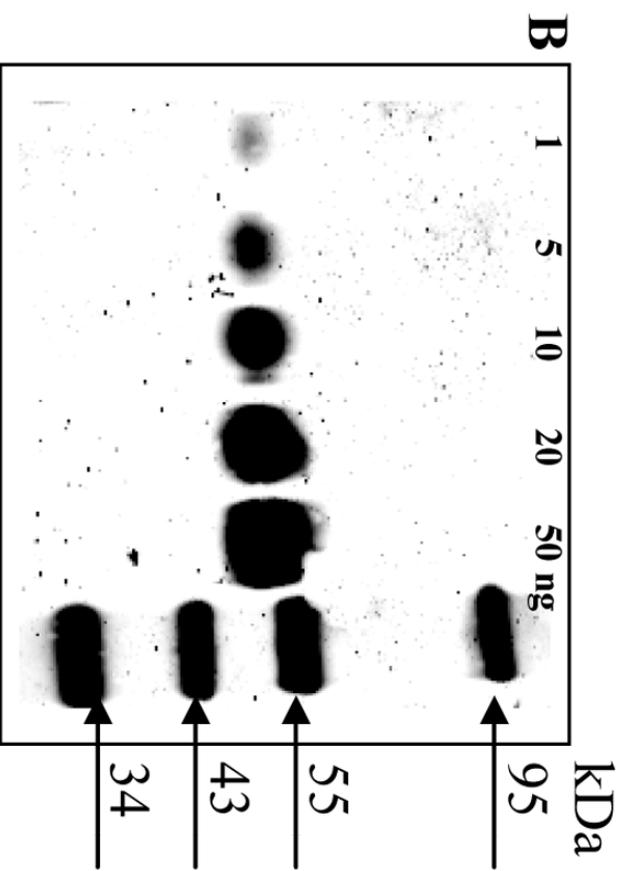
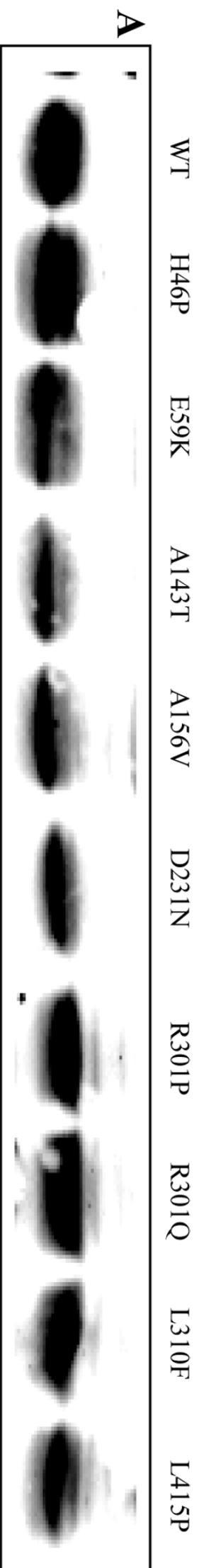
	c276_277in sTGAT	0(3)	0(4)	n.d.						own data
<i>p.D93Y</i>	c277G>T	0(3)	0(3)	n.d.					probably damaging	own data
<i>p.D93E</i>	c279C>G	0(3)	0(3)	n.d.					possibly damaging	own data
<i>p.W95L</i>	c284G>T	0(3)	0(3)	n.d.					probably damaging	Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.W95C</i>	c285G>T	0(3)	0(3)	n.d.					probably damaging	own data
<i>p.R100T</i>	c299G>C	0(4)	0(4)	n.d.					probably damaging	Eng (1997) Mol Med 3:174
<i>p.S102L</i>	c305C>T	71.6±4.4(4)	78.9±8.3(4)	1.1		0,5 (1)			benign	own data
<i>p.E103*</i>	c307G>T	0(3)	0(4)	n.d.						Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.R112C</i>	c334C>T	0(5)	0(3)	n.d.	20.2 (1)	3,6 (5)	classic		probably damaging	Ishii (1992) Hum Genet 89:29
<i>p.R112H</i>	c335G>A	1.6±0.6(3)	19.4±2.3(3)	n.d.	2.2 (2)	1.0 (1)	variant		probably damaging	Eng (1994) Hum Mol Genet 3:1795
<i>p.R118S</i>	c352C>A	76.2±9.5(3)	70.9±2.6(3)	n.d.					benign	Ψ
<i>p.R118G</i>	c352C>G	37.6±4.1(3)	44.0±5.0(3)	1.2					benign	Ψ
<i>p.R118C</i>	352C>T	20.0±1.3(3)	23.7±2.9(4)	1.2	0.6 (1)	0.5 (1)	variant		possibly damaging	Spada (2006) Am J Hum Genet 79:31
<i>p.R118H</i>	c353G>A	67.4±3.4(3)	66.9±3.1(3)	n.d.					benign	Ψ
<i>p.R118P</i>	c353G>C	49.0±4.9(3)	47.7±2.0(3)	n.d.					benign	Ψ
<i>p.R118L</i>	c353G>T	51.7±3.8(3)	51.4±4.0(3)	n.d.					benign	Ψ
<i>p.L120V</i>	c358C>G	50.1±5.0(3)	62.0±2.6(3)	1.2					probably damaging	Chien (2009) Hum Genet 125:336 [2]
<i>p.H125P</i>	c374A>C	0(3)	0(3)	n.d.					probably damaging	Auray-Blais (2008) Mol Genet Metab 93:331
<i>p.S126R</i>	c376A>C	19.7±1.2(6)	23.5±2.3(4)	1.2					possibly damaging	Ψ
<i>p.S126G</i>	c376A>G	51.3±9.6(4)	67.4±11.9(4)	1.3	0.5 (4)	0.5 (8)	variant		benign	Branton (2002) Medicine (Baltimore) 81:122
<i>p.S126C</i>	c376A>T	2.1± 0.4(3)	2.7±0.2(3)	n.d.					possibly damaging	Ψ
<i>p.S126N</i>	c377G>A	14.2±2.4(3)	17.0±2.3(3)	1.2					benign	Ψ
<i>p.S126T</i>	c377G>C	48.9±11.4(3)	42.8±7.9(4)	n.d.					benign	Ψ
<i>p.S126I</i>	c377G>T	40.5±8.1(4)	43.2±7.7(4)	1.1					possibly damaging	Ψ
<i>p.L129P</i>	c386T>C	0(3)	0(3)	n.d.	72.1 (2)	7.9 (2)	classic/variant		probably damaging	Whybra (2001) J Inherit Metab Dis 24:715
<i>p.L131P</i>	c392T>C	0(3)	0(3)	n.d.	68.6 (4)	7.0 (2)	classic		probably damaging	Eng (1994) Hum Mol Genet 3:1795
<i>p.G132R</i>	c394G>A	0(3)	0(3)	n.d.			classic		probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.G132E</i>	c395G>A	0(3)	0(3)	n.d.	119.0 (1)	12.3 (1)			probably damaging	Kwan (2007) Am Soc Hum Genet Meet Abs .:1537
<i>p.A135V</i>	c404C>T	0(3)	6.9±2.4(3)	n.d.	104.2 (2)		classic		benign	Shabbeer (2005) Hum Mutat 25:299
<i>p.D136E</i>	c408T>A	0(3)	31.3±3.2(3)	n.d.	42.3 (1)	6.6 (2)			probably damaging	own data
<i>p.G138R</i>	c412G>A	0(3)	0(3)	n.d.		3.6 (1)	classic		probably damaging	Eng (1997) Mol Med 3:174

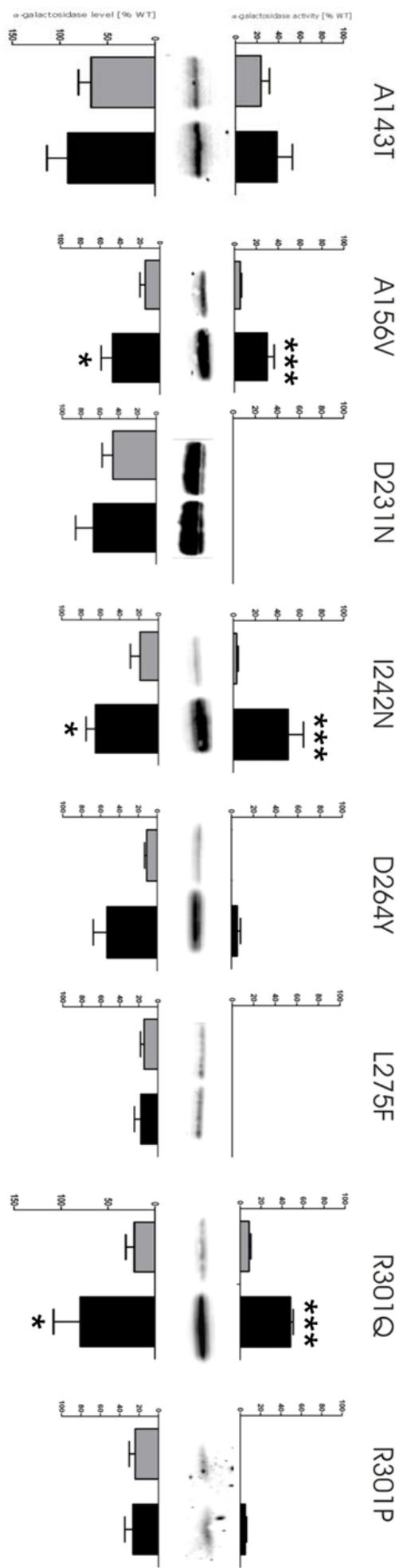
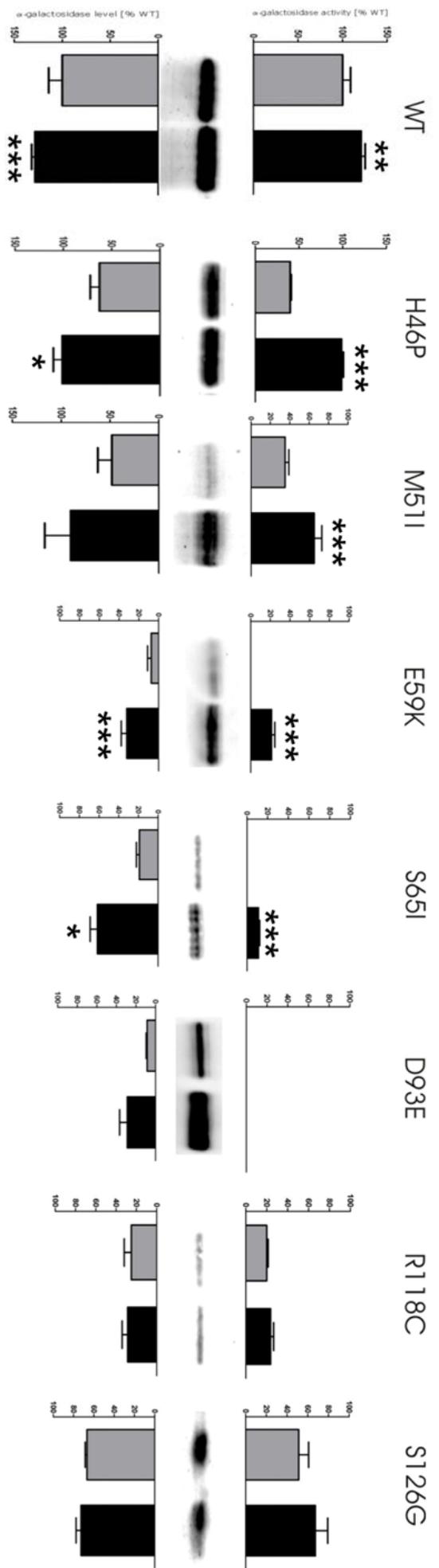
<i>p.N139S</i>	c416A>G	147.8±28.3(3)	176.4±36.4(3)	1.2	88.6 (1)	3.0 (4)	variant	benign	Havndrup (2010) Eur J Heart Fail 12:535
<i>p.T141I</i>	c422C>T	0(3)	0(3)	n.d.			classic	probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.C142R</i>	c424T>C	0(3)	0(3)	n.d.			classic	probably damaging	Topaloglu (1999) Mol Med 5:806
<i>p.A143T</i>	c427G>A	31.3±5.6(7)	49.4±4.8(6)	1.6	0.5 (8)	0.5 (12)	classic/variant	benign	Eng (1997) Mol Med 3:174
<i>p.A143P</i>	c427G>C	0(3)	0(5)	n.d.			classic	possibly damaging	Eng (1994) Hum Mol Genet 3:1795
<i>p.G147R</i>	c439G>A	0(3)	0(3)	n.d.		16.2 (1)		probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.A156D</i>	c467C>A	0(3)	0(3)	n.d.				probably damaging	own data
<i>p.A156V</i>	c467C>T	4.3±1.0(9)	16.8±2.3(5)	3.9			classic	probably damaging	Okumiya (1995) Hum Genet 95:557
<i>p.W162G</i>	c484T>G	0(3)	5.2±1.2(3)	n.d.	42.5 (3)	6.8 (3)		probably damaging	own data
<i>p.D165H</i>	c493G>C	3.4±1.1(3)	11.9±1.1(3)	3.5		3.7 (1)		probably damaging	Rakoczi (2007) Orv Hetil 148:1087
<i>p.D165Y</i>	c493G>T	0(3)	0(3)	n.d.				probably damaging	own data
<i>p.D165V</i>	c494A>T	0(4)	0(4)	n.d.	9.1 (1)		classic	probably damaging	Davies (1994) Hum Mol Genet 3:667
<i>p.D170N</i>	c508G>A	0(3)	0(3)	n.d.	72.1 (1)	4.8 (1)		probably damaging	own data
<i>p.C172G</i>	c514T>G	0(3)	4.4±2.0(3)	n.d.	59.3 (1)	21.7 (1)	classic	probably damaging	Yasuda (2003) Hum Mutat 22:486
<i>p.C172Y</i>	c515G>A	0(3)	0(3)	n.d.	80.2 (1)	5.6 (1)	classic	probably damaging	Eng (1994) Hum Mol Genet 3:1795
	c521insT	0(3)	0(3)	n.d.		5.9 (2)			own data
<i>p.D175N</i>	c523G>A	70.0±2.1(3)	84.3±8.5(3)	1.2				benign	own data
<i>p.D175E</i>	c525C>G	89.8±2.2(3)	89.0±7.3(3)	n.d.		0.5 (1)		benign	own data
	c525_526de ICA	0(3)	0(3)	n.d.					own data
<i>p.G183V</i>	c548G>T	0(3)	6.7±2.1(4)	n.d.				probably damaging	Dobrovoly (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.M187T</i>	c561G>C	0(3)	0(4)	n.d.		2.9 (1)	classic	possibly damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.L191P</i>	c572T>C	0(3)	0(3)	n.d.		4.9 (1)	classic	probably damaging	Cooper (2000) Hum Genet 107:535
<i>p.W204C</i>	c612G>C	0(3)	4.4±0.9(3)	n.d.		8.9 (1)		probably damaging	own data
<i>p.K213M</i>	c638A>T	83.4±29.6(5)	82.5±15.5(5)	n.d.	9.3 (1)			possibly damaging	own data
<i>p.N215S</i>	c644A>G	39.5±1.5(6)	63.9±3.3(6)	1.6	4.2 (7)	1.1 (6)	variant	benign	Davies (1993) Hum Mol Genet 2:1051
<i>p.I219T</i>	c656T>C	53.3±2.2(3)	85.3±10.0(3)	1.6				benign	Hwu (2009) Hum Mutat 30:1397
<i>p.R220Q</i>	c659G>A	104.5±11.3(4)	144.0±11.9(4)	1.4				benign	own data
<i>p.N224S</i>	c671A>G	31.1±2.4(3)	82.2±2.2(3)	2.6		5.3 (3)	classic	probably damaging	Ashton-Prolla (2000) J Investig Med 48:227
<i>p.H225D</i>	c673C>G	32.2±8.1(3)	60.5±19.1(3)	1.9	2.8 (1)			possibly damaging	own data
<i>p.H225R</i>	c674A>G	0(7)	3.0±0.8(4)	n.d.			classic	possibly damaging	Politei (2005) Rev Neurol 41:506
<i>p.R227Q</i>	c680G>A	0(3)	0(3)	n.d.	53.7 (5)	11.2 (3)	classic	probably damaging	Eng (1993) Am J Hum Genet 53:1186

<i>p.D231N</i>	c691G>A	0(3)	0(3)	n.d.		5.9 (1)	classic	probably damaging	Redonnet-Vernhet (1996) J Med Genet 33:682
<i>p.I232T</i>	c695T>C	11.5±2.8(3)	61.6±7.2(5)	5.1				possibly damaging	own data
<i>p.W236C</i>	c708G>C	0(3)	0(3)	n.d.	95.8 (1)	7.8 (2)	classic	probably damaging	Davies (1996) Eur J Hum Genet 4:219
<i>p.S238N</i>	c713G>A	36.0±6.7(4)	94.3±24.5(4)	2.6			variant	probably damaging	Monserrat (2007) J Am Coll Cardiol 50:2399
<i>p.I242N</i>	c725T>A	3.1±1.4(4)	49.8±13.9(4)	16.1			classic	probably damaging	Takata (1997) Brain Dev 19:111
<i>p.R252T</i>	c755G>C	117.0±10.0(4)	134.3±22.5(4)	1.1		0.5 (4)		benign	own data
<i>p.I253S</i>	c758T>G	4.4±1.6(5)	53.4±7.1(5)	10.3				probably damaging	own data
<i>p.G261V</i>	c782G>T	0.2±1.4(5)	3.5±1.1(5)	n.d.	84.4 (2)			probably damaging	own data
<i>p.W262C</i>	c786G>C	0(3)	0(3)	n.d.	9.4 (1)	7.7 (4)	classic/variant	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.D264N</i>	c790G>A	37.9±3.7(3)	84.5±6.4(3)	2.2				probably damaging	Ψ
<i>p.D264H</i>	c790G>C	1.7±0.9(3)	10.2±2.9(3)	6.0				probably damaging	Ψ
<i>p.D264Y</i>	c790G>T	0(3)	5.4±2.9(3)	n.d.	98.9 (1)	13.7 (3)	classic	probably damaging	Shabbeer (2005) Hum Mutat 25:299
<i>p.D264A</i>	c791A>C	0(3)	0(3)	n.d.				probably damaging	Ψ
<i>p.D264G</i>	c791A>G	0(3)	0(3)	n.d.				probably damaging	Ψ
<i>p.D264V</i>	c791A>T	0(3)	0(3)	n.d.	59.3 (1)		classic	probably damaging	Eng (1993) Am J Hum Genet 53:1186
<i>p.D264E</i>	c792C>A	0(3)	0(3)	n.d.				probably damaging	Ψ
<i>p.M267T</i>	c800T>C	27.5±4.4(4)	30.5±1.8(4)	1.2	2.7 (1)			probably damaging	own data
<i>p.L268S</i>	c803T>C	0(3)	10.8±0.7(3)	n.d.	18.0 (1)	3.0 (1)	classic	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.V269M</i>	c805G>A	0(4)	17.3±1.4(4)	n.d.			classic	probably damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.V269A</i>	c806T>C	9.0±1.4(9)	45.0±4.3(4)	5.0			classic	possibly damaging	Davies (1993) Hum Mol Genet 2:1051
	c807delG	0(3)	0(3)	n.d.					own data
<i>p.I270M</i>	c810T>G	2.3±1.7(4)	33.7±16.8(4)	14.7				probably damaging	own data
<i>p.N272S</i>	c815A>G	0(3)	0(3)	n.d.			classic	probably damaging	Verovnik (2004) Eur J Hum Genet 12:678
<i>p.F273L</i>	c819T>A	0(6)	0(3)	n.d.			variant	probably damaging	Zhang. SH et al.. 2007
<i>p.L275F</i>	c823C>T	0(4)	0(3)	n.d.		3.9 (1)	classic/variant	probably damaging	Dobrovlny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.L275H</i>	c824T>A	0(5)	0(5)	n.d.				probably damaging	own data
<i>p.T282I</i>	c845C>T	5.0±0.5(3)	47.7±0.7(3)	9.5	22.6 (2)			possibly damaging	own data
<i>p.Q283P</i>	c848A>C	0(3)	0(3)	n.d.			classic	probably damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.A285D</i>	c854C>A	0(3)	0(3)	n.d.			classic	probably damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.W287*</i>	c860G>A	0(3)	0(3)	n.d.					Davies (1993) Hum Mol Genet 2:1051
<i>p.M290L</i>	c868A>T	18.5±8.2(3)	22.4±9.6(3)	1.2				probably damaging	own data

<i>p.M290I</i>	c870G>A	39.0±1.8(3)	70.5±7.2	1.8			classic	probably damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.A291T</i>	c871G>A	13.2±3.4(3)	55.7±0.7(3)	3.7				probably damaging	Zhang (2007) J Nephrol 20:716
<i>p.L294*</i>	c881T>G	0(4)	0(3)	n.d.					Blaydon (2001) Hum Mutat 18:459
<i>p.L294S</i>	c881T>C	0(3)	12.4±1.7(3)	n.d.		6.6 (2)		probably damaging	own data
	c883delT	0(3)	0(3)	n.d.					Dobrovlny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.S297C</i>	c890C>G	0(3)	3.8±0.7(3)	n.d.	121.0 (1)		classic	probably damaging	Germain (2002) Mol Med 8:306
<i>p.R301G</i>	c901C>G	19.3±4.1(7)	56.5±3.4(3)	2.9			classic	probably damaging	Lai (2001) Hum Genet 109:469
<i>p.R301Q</i>	c902G>A	8.5±1.8(9)	48.0±2.6(3)	5.6		2.2 (1)	classic/variant	probably damaging	Sakuraba (1990) Am J Hum Genet 47:784
<i>p.R301P</i>	c902G>C	0(4)	5.0±1.3(3)	n.d.		10.7 (1)	classic	probably damaging	Ashley (2001) J Hum Genet 46:192
<i>p.L310F</i>	c928C>T	0(4)	4.1±1.1(3)	41.0			classic	probably damaging	Calado (2004) Hum Genet 115:347
<i>p.L311V</i>	c931C>G	1.9±0.4(3)	40.1±4.9(3)	n.d.	53.6 (1)			probably damaging	own data
<i>p.D313Y</i>	c937G>T	83.9±21.1(4)	100.3±23.9(3)	1.2	0.5 (27)	0.5 (36)	variant	possibly damaging	Eng (1993) Am J Hum Genet 53:1186
<i>p.V316I</i>	c946G>A	65.6±22.3(3)	68.3±21.8(3)	n.d.	0.7 (1)		variant	benign	own data
<i>p.V316G</i>	c947T>G	0(3)	0(3)	n.d.	103.0 (1)	13.0 (1)		probably damaging	own data
	c948InsT	0(3)	0(3)	n.d.	75.1 (1)	5.8 (2)			own data
<i>p.I319T</i>	c956T>C	20.2±3.8(3)	58.3±12.6(3)	2.9		3.0 (2)		probably damaging	own data
<i>p.N320I</i>	c959A>T	2.0±0.5(3)	31.8±8.9(3)	17.7		2.9 (1)	classic	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.Q321H</i>	c963G>C	3.3±1.9(3)	25.3±7.1(3)	7.7				probably damaging	own data
<i>p.G325S</i>	c973G>A	25.6±9.6(3)	55.4±14.4(3)	2.2				probably damaging	own data
<i>p.Q327K</i>	c979C>A	0(4)	0(5)	n.d.			classic	probably damaging	Davies (1993) Hum Mol Genet 2:1051
<i>p.Q327E</i>	c979C>G	21.5±2.8(3)	80.9±10.0(3)	3.8		1.7 (1)	classic/variant	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.G328A</i>	c983G>C	6.2±1.0(3)	30.0±6.4(3)	4.8			classic	probably damaging	Eng (1993) Am J Hum Genet 53:1186
<i>p.E341K</i>	c1021G>A	0(3)	0(3)	n.d.	64.4 (2)		classic	probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.E341D</i>	c1023A>C	0(3)	0(3)	n.d.			classic	probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.R342Q</i>	c1025G>A	0(4)	0(3)	n.d.		2.6 (3)	classic	possibly damaging	Ploos van Amstel (1994) Hum Mol Genet 3:503
<i>p.R342L</i>	c1025G>T	0(3)	0(3)	n.d.	48.3 (2)	3.0 (2)		possibly damaging	own data
<i>p.L344P</i>	c1031T>C	0(3)	0(3)	n.d.	24.9 (1)		classic	probably damaging	Schafer (2005) Hum Mutat 25:412
<i>p.S345P</i>	c1033T>C	0(3)	13.3±3.8(3)	n.d.	16.7 (1)		classic	benign	Schafer (2005) Hum Mutat 25:412
<i>p.W349R</i>	c1045T>A	1.6±1.4(4)	8.9±3.9(4)	5.6				probably damaging	own data
<i>p.R356W</i>	c1066C>T	16.9±2.3(9)	62.7±3.9(3)	3.7		2.3 (2)	classic	possibly damaging	Bernstein (1989) J Clin Invest 83:1390

<i>p.R356Q</i>	c1067G>A	89.1±5.0(3)	99.4±4.3(3)	1.1				benign	Hwu (2009) Hum Mutat 30:1397
<i>p.G360C</i>	c1078G>T	11.9±0.5(3)	26.5±2.7(3)	2.2				probably damaging	Lin (2007) Hum Genet 122:212
<i>p.G361R</i>	c1081G>C	0(3)	3.7±1.1(3)	n.d.			classic	probably damaging	Davies (1993) Hum Mol Genet 2:1051
<i>p.R363H</i>	c1088G>A	31.9±2.9(8)	57.9±5.5(3)	1.8			classic/variant	possibly damaging	Cooper (2000) Hum Genet 107:535
<i>p.A368T</i>	c1102G>A	103.7±33.6(3)	93.3± 5.0(3)	n.d.		0.5 (3)		benign	own data
<i>p.L372P</i>	c1115T>C	0(3)	2,6±0,7(4)	n.d.				probably damaging	own data
<i>p.L372R</i>	c1115T>G	0(3)	0(3)	n.d.				probably damaging	Dobrovolny (2008) Am Soc Hum Genet Meet Abs . 150
<i>p.G373D</i>	c1118G>A	0(4)	0(4)	n.d.				probably damaging	Germain (2001) Hum Mutat 17:353
<i>p.C378R</i>	c1132T>C	0(3)	0(3)	n.d.				probably damaging	Shin (2001) ASHG Annual Meeting Abstracts. 1799
<i>p.I384N</i>	c1151T>A	0(3)	0(3)	n.d.	25.1(1)	1.5 (1)	classic	probably damaging	Shabbeer (2002) Mol Genet Metab 76:23
<i>p.T385A</i>	c1153A>G	45.0±3.7(3)	48.9±7.5(3)	1.1		0.5 (1)		benign	own data
<i>p.Q386P</i>	c1157A>C	0(3)	0(3)	n.d.			classic	possibly damaging	Shabbeer (2006) Hum Genomics 2:297
<i>p.P389L</i>	C1166C>T	0(3)	0(3)	n.d.	72.9 (1)			probably damaging	own data
<i>p.V390M</i>	c1168G>A	0(3)	0(4)	n.d.				benign	own data
<i>p.G395A</i>	c1184G>C	20.1±1.7(6)	23.1±2.0(6)	1.1				probably damaging	own data
<i>p.F396Y*</i>	c1187T>A	87.6±14.2(4)	93.8±15.7(4)	1.1				benign	Novo (1995) Nuceic Acids Research 23. 2636-40
<i>p.E398A</i>	c1193A>C	48.9±4.7(3)	62.5±7.8(3)	1.3				benign	own data
<i>p.S405R</i>	c1213A>C	91.0±10.2 (3)	92.7±11.1 (3)	n.d.		0.5 (1)			own data
	c1221delA	0(4)	0(3)	n.d.		8.2 (2)			own data
<i>p.T410I</i>	c1229C>T	2.3±0.8(4)	16.1±3.5(4)	n.d.				possibly damaging	own data
<i>p.L415F</i>	c1243C>T	83.2±5.9(3)	99.5±12.2(3)	1.2		0.5 (1)		benign	own data
<i>p.L415P</i>	c1244T>C	0(3)	0(3)	n.d.		4.7 (1)		probably damaging	Serebrinsky (2006) Hum Genet 119:361
<i>p.Q416X</i>	c1246C>T	0(3)	0(3)	n.d.	27.7 (3)	6.7 (1)			own data
<i>p.Q416P</i>	c1247A>C	0(3)	0(3)	n.d.					own data
<i>p.E418G</i>	c1253A>G	74.6±15.5(3)	89.1±4.9(3)	1.2		0.5 (1)		benign	own data





Text S1:

Additional method

Splicing analysis

The splicing prediction was performed using Alamut 2.2 with MaxEntScan [42], NNPLICE [43] and Human Splicing Finder (HSF) [44] modules. Average changes detected by all three algorithms at acceptor and donor sites exceeding 5% as compared to the wild type sequence were considered “interesting” and are listed in Supplementary Table 4A for changes affecting natural sites and in Supplementary Table 4B for changes leading to activation of cryptic and novel sites.

Additional References:

42. Yeo G, Burge CB (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol.* 11(2-3):377-94.
43. Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. *J Comput Biol.* 4(3):311-23.
44. Desmet FO, Hamroun D, Lalande M, Collod-Bérout G, Claustres M et al. (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 37(9):e67.