Brief Report

In Silico Analysis of Missense Mutations as a First Step in Functional Studies: Examples from Two Sphingolipidoses

Ana Joana Duarte [1,2,3], **Diogo Ribeiro** [1,3], **Luciana Moreira** [1,3] and **Olga Amaral** [1,3,*]

1. Instituto Nacional de Saúde Dr Ricardo Jorge (INSA, IP), Departamento de Genética Humana, Unidade de Investigação e Desenvolvimento, Rua Alexandre Herculano 321, 4000-055 Porto, Portugal; ana.duarte@insa.min-saude.pt (A.J.D.); diogo.ribeiro@insa.min-saude.pt (D.R.); luciana.moreira@insa.min-saude.pt (L.M.)
2. ICBAS, University of Porto, 4099-002 Porto, Portugal
3. CECA, ICETA, University of Porto, 4099-002 Porto, Portugal
* Correspondence: olga.amaral@insa.min-saude.pt; Tel.: +351-223401100

check for
updates

Abstract: In order to delineate a better approach to functional studies, we have selected 23 missense mutations distributed in different domains of two lysosomal enzymes, to be studied by in silico analysis. In silico analysis of mutations relies on computational modeling to predict their effects. Various computational platforms are currently available to check the probable causality of mutations encountered in patients at the protein and at the RNA levels. In this work we used four different platforms freely available online (Protein Variation Effect Analyzer- PROVEAN, PolyPhen-2, Swiss-model Expert Protein Analysis System—ExPASy, and SNAP2) to check amino acid substitutions and their effect at the protein level. The existence of functional studies, regarding the amino acid substitutions, led to the selection of the distinct protein mutants. Functional data were used to compare the results obtained with different bioinformatics tools. With the advent of next-generation sequencing, it is not feasible to carry out functional tests in all the variants detected. In silico analysis seems to be useful for the delineation of which mutants are worth studying through functional studies. Therefore, prediction of the mutation impact at the protein level, applying computational analysis, confers the means to rapidly provide a prognosis value to genotyping results, making it potentially valuable for patient care as well as research purposes. The present work points to the need to carry out functional studies in mutations that might look neutral. Moreover, it should be noted that single nucleotide polymorphisms (SNPs), occurring in coding and non-coding regions, may lead to RNA alterations and should be systematically verified. Functional studies can gain from a preliminary multi-step approach, such as the one proposed here.

Keywords: lysosomal glucocerebrosidase; lysosomal alpha-galactosidase; sphingolipidoses; functional studies; in silico analysis; *GBA1*; *GLA*; *CSTB*; *ARSA*; *GALC*

1. Introduction

Lysosomal storage diseases (LSDs) are a large group of inherited disorders leading to various clinical symptoms, caused by defects in lysosomal enzymes, transporter proteins, activator proteins, or other proteins involved in lysosomal function or biogenesis. Such defects lead to total or partial loss of enzyme activity and consequent accumulation of substrate, which results in impaired organelle function, leading to subsequent multi-organ dysfunction. The enzymes involved in two of the less rare LSDs are lysosomal glucocerebrosidase (GlcCerase, glucosylceramidase or acid-β-glucosidase, EC 3.2.1.45), and lysosomal acid-α-galactosidase (α-GAL or α-Gal A, EC 3.2.1.22). Most commonly,

mutations in the genes coding for the aforementioned enzymes, *GBA1* (ID: 2629) and *GLA* (ID: 2717) lead to the development of Gaucher [1] and Fabry diseases, respectively [2]. Gaucher disease (GD, MIM #230800, 230900, 231000) is the most common lysosomal storage disorder [1], and Fabry disease (FD, MIM #301500) is an X-linked disorder that has a large range of phenotypes [2]. In the case of both shingolipidoses, therapeutic approaches based on enzyme replacement or small-molecule compounds have been successfully developed. Interestingly, both hydrolytic reactions require activation by saposins; whereas globotriaosylceramide binds saposin B (SapB) prior to presentation to α-GAL, saposin C (SapC) enhances GlcCerase activity [3]. The enzyme GlcCerase is a peripheral membrane protein that catalyzes the hydrolysis of glucosylceramide (GlcCer) to ceramide and glucose. The three-dimensional (3D) structure of human GlcCerase was discovered by Dvir et al. in 2003; it comprises three non-contiguous domains, with the catalytic site located on the $(\alpha/\beta)_8$ TIM barrel in domain III [4]. The 3D structure of human α-GAL was found by Garman and Garboczi in 2004; it is composed of two domains and shares with GlcCerase a $(\alpha/\beta)_8$ TIM barrel motif in the catalytic domain and an immunoglobulin fold in the same relative position [5].

To date, hundreds of mutations in *GBA1* and *GLA* genes have been identified (http://www.hgmd.cf.ac.uk), and new sequencing techniques continuously identify variants in large numbers (http://gnomad.broadinstitute.org/). Large sets of information were also recently assembled through the 1000 genomes project (http://www.internationalgenome.org/1000-genomes-browsers/).

The term in silico goes back to the 1970s and is related to the computer component silicon. In silico methods are based on computational approaches for the prediction of effects prior to development of laboratory methods [6] In vitro functional studies involve laboratory assays for testing various types of functions and usually rely on cell based overexpression systems and cell based "test-tube" assays. The in vitro studies require many consumables, specific laboratory apparatus and time-consuming optimizations and laboratory procedures. Computational modeling is an important tool to deal with the rapid increase in bioinformatics information. Presently, the most rapid and inexpensive way to predict whether a Single-Nucleotide Polymorphism (SNP) will potentially cause disease is by performing a computational analysis. Different parameters, such as clinical, populational, structural, and bioinformatics, need to be considered when analyzing results. The increasing number of computational tools available, and the rapidly growing number of available crystal structures, has turned in silico modeling into an accurate complementary, and often crucial, prediction methodology. Different computational platforms take into account, to various degrees, factors such as the general rules of protein chemistry, 3D structure, and homologies in amino acid sequences among various species or related proteins [7]. Most of the *GBA1* and *GLA* mutations are point mutations leading to missense substitutions, and functional studies exist in several of those mutations. The existence of in vitro data adds relevant information for the computational study of the molecular basis of the protein impairment.

In this report 23 single nucleotide alterations in the *GBA1* and *GLA* genes, leading to amino acid substitutions with protein functional studies, were re-examined through in silico analysis and their effects were predicted and compared using different platforms available online. The analyzed amino acid substitutions were individually mapped into the available 3D enzyme model, using the PyMOL tool [8]. Comparison of the computational data obtained with previous in vitro expression, or functional, data was carried out to evaluate the prediction accuracy regarding the establishment of genotype/phenotype correlation. In silico prediction of the amino acid substitution impact at the protein level may, sometimes, be considered as an alternative to in vitro expression or as a pre-study indicator of the need for research at the functional level. Attentive in silico analysis is a potentially valuable option for immediate guidance regarding patient care and counseling, being able to confer an indication of prognosis value to the genotyping results.

Functional effects of mutations were predicted with different tools in an attempt to distinguish between variant amino acid substitutions, considering evolutionary information, structural features, and other relevant information. Tools that combine different types of existing information are more

complete and were used in the present study [9–11]. Missense mutations were selected in different domains of the GlcCerase and of the α-GAL proteins with different types of functional evidence of causality. In order to broaden the scope of the study, mutations in three other genes related to neurodegenerative diseases were also added to the present study.

2. Results

The aim of this work was to investigate the prediction value of different bioinformatics tools, applying them to single amino acid substitutions in the *GBA1* and *GLA* genes.

GlcCerase and α-GAL structures were obtained from the Protein Data Bank (PDB). *GBA1* mutations (p.F109V, p.P182L, p.D140H, p.K157Q, p.W184R, p.N188S, p.E326K, p.R359Q, p.G377S, p.R395P, p.N396T, p.P415R, and p.L444P) and *GLA* mutations (p.D33G, p.M42V, p.R112C, p.F113L, p.R118C, p.C142W, p.D231G, p.D266N, p.S297F and p.D313Y) were mapped into 3D GlcCerase and α-GAL structures: the first X-ray human GlcCerase to be solved (PDB code 1OGS) [4] and into the first 3D α-GAL structure (PDB code 1R46) [5] (Figures 1 and 2). Three-dimensional structures were designed using PyMOL (http://www.pymol.org) in order to visualize how these alterations could affect enzyme structure.
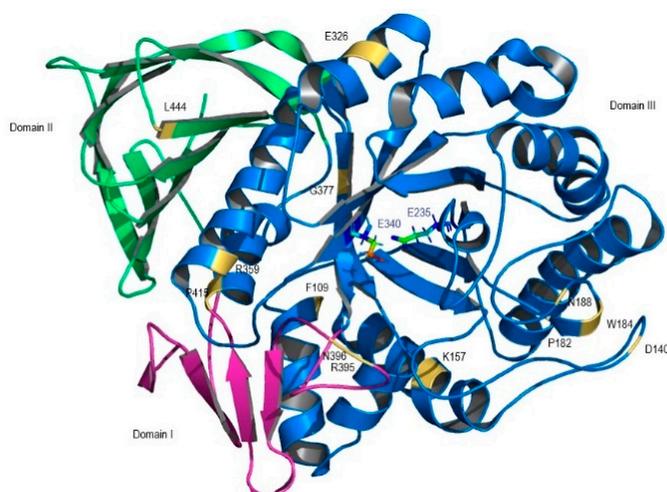

Figure 1. Location of GlcCerase missense mutations considered in this study presented in a solid ribbon model. Domain I is shown in pink, domain II is in green, and domain III is shown is blue. Mutations are identified in yellow. Active site residues (E235 and E340, in blue letters) are shown as ball-and-stick model, with different colors.
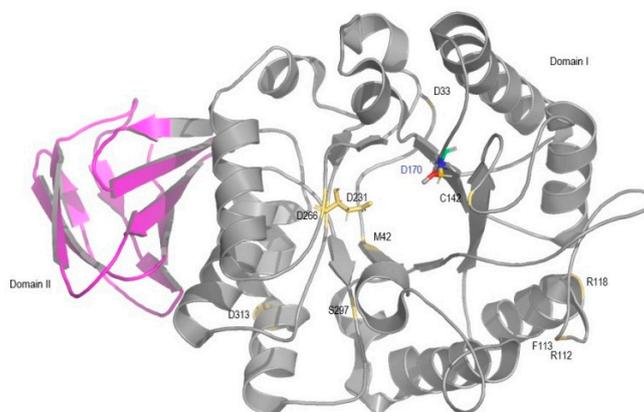

Figure 2. Location of α-GAL missense mutations considered in this study presented in a solid ribbon model. Domain I is shown in grey and domain II is shown in pink. Mutations are identified in yellow, and active site residues (D170 and D231, the non-mutated active site residue in blue letters) are shown as ball-and-stick model, with different colors.

The missense mutations, depicted in Figures 1 and 2, were computationally analyzed and the retrieved results were compared with the in vitro results and functional data (Table 1), in order to ascertain the validity of the platforms used and evaluate the prediction accuracy regarding the establishment of genotype/phenotype correlation. Differences in the results obtained reflect the different types of algorithms used in the computational platforms.

Table 1. In silico analysis comparison of 23 missense mutations in the GBA1 and GLA genes.

Gene Mutants	PROVEAN	PolyPhen-2	SNAP2	ExPASy	Protein Function and Structure
			Prediction Expected Accuracy		
GBA1 Mutants					
F109V CM005404	Deleterious	Probably damaging	Effect 75%	NA	15% of wt activity; weakly conserved; domain III; stable protein [12]
P182L rs80205046	Deleterious	Probably damaging	Effect 85%	NA	Near null activity; buried site; domain III; unstable protein [13]
D140H rs147138516	Neutral	Benign	Neutral 57%	Disease	73% of wt activity; domain III [14]
K157Q rs121908297	Deleterious	Probably damaging	Neutral 57%	Disease	9.7% of wt activity; domain III; conserved region; unstable protein [14]
W184R rs61748906	Deleterious	Probably Damaging	Effect 66%	Disease	Inactive enzyme; domain III periphery; alteration of enzyme geometry; unstable protein [12]
N188S rs364897	Deleterious	Benign	Effect 66%	Disease	66.6% of wt activity; domain III periphery; stable protein [13]
E326K rs2230288	Neutral	Benign	Neutral 56%	Disease	42.7–25% of wt activity; domain III; stable protein [15]
R359Q rs74979486	Deleterious	Probably Damaging	Effect 75%	Disease	4.5% of wt activity; domain III stable protein; highly conserved region [12]
G377S rs121908311	Deleterious	Probably Damaging	Effect 91%	Disease	17% of wt activity; domain III; stable protein [12]
R395P	Deleterious	Benign	Effect 75%	NA	4.5% of wt activity; domain I, loop 2; stable protein [12]
N396T rs75385858	Deleterious	Probably Damaging	Effect 85%	NA	14% of wt activity; domain I; stable protein [12]
P415R rs121908295	Deleterious	Probably damaging	Effect 59%	Disease	Near null activity; conserved region; unstable protein [16]
L444P rs421016	Deleterious	Possibly damaging	Effect 91%	NA	5.7–9% of wt activity; unstable protein [12,17].
GLA Mutants					
D33G rs869312136	Deleterious	Possibly damaging	Effect 75%	Unclassified	37% of wt activity; periphery of domain I [18]
M42V	Deleterious	Probably damaging	Effect 85%	Disease	7% of wt activity; domain I; unstable protein [19]
R112C rs104894834	Deleterious	Probably damaging	Effect 91%	Disease	5% of wt activity; periphery of domain I; unstable protein [19]
F113L rs869312142	Deleterious	Probably damaging	Effect 91%	Disease	20% of wt activity; periphery of domain I; altered alpha-GAL surface; unstable protein [19]
R118C rs148158093	Deleterious	Probably Damaging	Effect 53%	NA	29–32% of wt activity; periphery of domain I; unstable protein [20,21].
C142W	Deleterious	Probably damaging	Effect 95%	NA	5% of wt activity; domain I; near active site pocket; unstable protein [19].
D231G	Deleterious	Probably damaging	Effect 95%	NA	4% of wt activity; domain I, active site pocket; stable protein [19]
D266N rs869312407	Deleterious	Probably damaging	Effect 95%	Disease	5% of wt activity; domain I, near the active site pocket; buried; unstable protein [19]
S297F rs28935489	Deleterious	Probably damaging	Effect 95%	Disease	5% of wt activity; unstable protein [19]
D313Y rs28935490	Deleterious	Probably damaging	Effect 95%	Disease	76% of wt activity; in domain I periphery; stable protein [22,23]

Legend: wt—wild-type; NA—Results not available with that computational tool.

In order to broaden the scope of the present study, we analyzed 14 additional mutations in other genes involved in neurodegenerative lysosomal-related disorders (Table 2). In all cases, functional studies were available. These previous mutations provided an ampler comparison between in vitro and in silico results.

Table 2. In silico analysis of 14 other single nucleotide mutations in the genes *ARSA* (MIM ID 607574) and *GALC* (MIM ID 606890), *CSTB* (MIM ID 601145).

Gene Mutants	PROVEAN	PolyPhen-2	SNAP2	ExPASy	Protein Function and Structure
			Prediction Expected Accuracy		
ARSA mutants					
G86D rs74315460	Deleterious	Probably damaging	Effect 95%	Disease	Null activity; unstable protein [24]
C156R rs199476348	Deleterious	Probably Damaging	Effect 59%	Disease	50% of wt activity [25]
T274M rs74315472	Deleterious	Probably Damaging	Effect 95%	Disease	35% of wt activity [26]
C300F rs74315484	Deleterious	Probably Damaging	Effect 95%	Disease	Null activity; disruption of disulfide bond linking major and minor β-sheets [27,28]
T409I rs74315481	Neutral	Possibly damaging	Effect 75%	Disease	60% of wt activity [29]
GALC Mutants					
I82M without reference SNP (rs)	Deleterious	Probably Damaging	Neutral 57%	Disease	Normal activity [30]
G286D rs199847983	Deleterious	Probably Damaging	Effect 71%	Disease	17.5% of wt activity [31]
Y335C rs757407613	Deleterious	Probably Damaging	Effect 75%	Disease	10% of wt activity [32]
G553R rs748573754	Deleterious	Probably Damaging	Effect 91%	Disease	1.8% of wt activity [31]
L634S rs138577661	Deleterious	Probably Damaging	Effect 95%	Disease	12% of wt activity [30]
CSTB mutants					
Q22Q rs386833443	Neutral	NA	Neutral 82%	NA	Expected abnormal peptide with premature truncation [33]
G4R rs74315443	Deleterious	Probably Damaging	Effect 85%	Disease	Binding pocket modification; interaction properties compromised [34]
G50E rs312262708	Deleterious	Possibly Damaging	Effect 95%	NA	Altered stability and interaction with target proteins [35,36]
Q71P rs796052392	Deleterious	Possibly Damaging	Effect 75%	NA	Changes in second binding loop; altered binding affinities [37]

Legend: wt—wild-type; NA—Results not available with that computational tool.

3. Discussion

In vitro mutagenesis and subsequent expression of mutant proteins, or functional studies and characterization, is a cumbersome task in terms of time, workload, and cost. For these reasons, in silico analysis is a desirable, fast, inexpensive, and reliable way to boost our understanding of how an amino acid substitution could affect the protein structure and function. Availability of 3D protein structures enables the mapping of amino acid substitutions and, therefore, helps complement the information acquired from different computational platforms. These aspects facilitate preliminary research in the biomedical field. As observed with the tools used here, the incorporation of more data increases the accuracy of the results, and thus makes predictions more reliable.

When a novel missense mutation is detected in a disease context, and its polymorphic nature has been excluded by population studies, it is possible to predict its outcome through in silico analyses, by first performing computational SNP evaluation followed by modeling the amino acid substitution into the 3D protein structure. In silico analysis is necessary to predict the impact of novel mutations in diseases such as the lysosomal disorders analyzed here. However, general limitations exist—for instance, the structural-based prediction tools may be unable to accurately predict mutation effects due to a lack of homologous structures in the databases. In such cases, functional analysis studies should be performed to elucidate how the missense mutation affects the protein function and contributes to the patient phenotype.

Overall, the retrieved results from the different computational platforms were rather similar, although they use different data sources and algorithms. The biggest difference observed seemed to be between PROVEAN and the other platforms, since it takes into account fewer variables. On the other hand, SNAP2 relies on protein and DNA data, as well as evolutionary and conservation information, and therefore is able to check more aspects regarding the impact of amino acid substitutions. In the case of α-GAL and GlcCerase mutations p.F113L and p.W184R, the location on the periphery of the proteins could suggest that they did not have a significant effect on enzyme activity and stability. However, the computational studies indicate them as damaging missense mutations and in vitro studies confirm that the respective proteins are unstable with reduced activity (p.F113L) or even inactive (p.W184R) [12,19]. These types of mutations usually occur in specific protein binding sites. These specific amino acids can be located on sites that are vital for the dimerization in α-GAL or tetramer formation in GlcCerase [38], or be located in sites where the activator proteins (Saposin B in α-GAL and Saposin C in GlcCerase) binds. Binding disruption will lead to partial or total loss of protein function.

A major limitation of this study is that there are few neutral, or low-score, variants to be analyzed. This problem arises because studies are not exhaustive enough and mutations that may look neutral are often not sufficiently investigated. Particular attention should be given to mutations in the "milder" spectrum. In addition to amino acid substitutions, SNPs may alter RNA processing by interfering with consensus sequences. A silent mutation or a "neutral" amino acid substitution may alter consensus sequences involved in splicing and lead to abnormal transcripts. Such mutations risk being overlooked and labeled as non-causal. An example to take into account is that of an apparently neutral/silent mutation on the *CSTB* gene (p.Q22Q in Table 2), which affected RNA processing and was proven to be causal only by functional studies [39,40].

Limitations of in silico analysis also arise since mutations (in the patient) may have additive or compensatory effects and the tools used only predict single protein changes. Besides the wide range of mutant variants and clinical phenotypes, in some cases, mutations in the same gene may be associated with more than one disease. Certain *GBA1* mutations are known to be associated with Gaucher Disease (GD) and with Parkinson's disease (PD) [41,42]. A good example of this association is mutation p.E326K, which has been repeatedly investigated [14,38,43]. The association of a single protein with different diseases is an additional limitation for in vivo and in vitro assays. Recently, a complex integration of in silico computational analysis has been used for the understanding of the association of *GBA1* mutations in GD and PD [44]. This latter approach, integrating multiple parameters, namely molecular dynamics, seems to pave the way for the development of more dependable in silico computational modeling approaches.

In general, it is possible to conclude that in silico methods remain an accurate way to make a rapid analysis regarding the expected effect of mutations. Nonetheless, the more factors that are taken into account, the more accurate the prediction will be. In order to take the best advantage of in silico analysis, different computational platforms should be used, trying to cover the major factors influencing protein structure and function. RNA processing alterations should also be routinely investigated by in silico analysis. The SNP impact at the RNA level can be investigated by using some of the various RNA assessment tools, such as Human Splicing Finder [45], GeneSplicer [46], NetGene2 [47], or Berkeley Drosophila Genome Project (BDGP) Splice Prediction by Neural Network [48].

4. Materials and Methods

4.1. Data Identifiers

The UniProt database was used to retrieve information about the protein sequence of *GBA* (UniProtKB P04062), *GLA* (UniProtKB P06280), *CSTB* (UniProtKB P04080), *ARSA* (UniProtKB P15289) and *GALC* (UniProtKB P54803). The three-dimensional structure of these proteins was obtained from the Protein Data Bank (PDB, http://www.rcsb.org/pdb/home/home.do) with the following 3D reference IDs: GlcCerase (1OGS), α-GAL (1R46), Cystatin-B (2OCT) for CSTB gene, Arylsulfatase A (1AUK) for ARSA gene, and Galactocerebrosidase (3ZR5) for GALC gene.

4.2. In Silico Methods

Twenty-three missense mutations were analyzed using four different computational platforms freely available online (Table 1). PROVEAN (Protein Variation Effect Analyzer) (http://provean.jcvi.org/) is a computational tool that predicts whether an amino acid substitution or indel will have an impact on the biological function of a protein. PROVEAN is useful for filtering sequence variants to identify nonsynonymous or indel variants that are predicted to be functionally important. Results are given as "deleterious" or "neutral", according to scores [49,50]. The PolyPhen-2 (Polymorphism Phenotyping v2) program (http://genetics.bwh.harvard.edu/pph2/) uses the sequence homology and knowledge of 3D structures; it predicts the possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. The results are classified as "benign", "possibly damaging", "probably damaging", or "unknown" [7,51]. The ExPASy Swiss-model [52] is a fully automated protein structure modeling server, accessible via the ExPASy web page (https://swissmodel.expasy.org/), and was also used in this study [53]. The SNAP2 (screening for non-acceptable polymorphisms) program (www.rostlab.org/services/SNAP/) incorporates evolutionary information, predicted aspects of protein structure, and other relevant information in order to make predictions regarding the functionality of mutated proteins [11]. The results are retrieved as "having an effect" or "being neutral", and a score, correlated with the severity of the change, is given for each substitution along with the percentage of expected accuracy [10].

5. Conclusions

In the present work, we show that a comparison of the results between various platforms is crucial and, in the case of the most deleterious mutants, the results are generally clear. In the case of the more neutral mutations, functional studies and more refined in silico approaches are fundamental for the understanding of the mutation's impact on the RNA processing, protein function, and pathophysiology of the disease.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Beutler, E.; Grabowski, G.A. Gaucher disease. In *The Metabolic and Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Eds.; McGraw-Hill: New York, NY, USA, 2001; pp. 3635–3668.

2. Desnick, R.J.; Ioannou, Y.A.; Eng, C.M. α-Galactosidase A deficiency: Fabry disease. In *The Metabolic and Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Eds.; McGraw-Hill: New York, NY, USA, 2001; pp. 3733–3774.

3. Lieberman, R.L.; D'Aquino, J.A.; Ringe, D.; Petsko, G.A. Effects of pH and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability. *Biochemistry* **2009**, *48*, 4816–4827. [CrossRef] [PubMed]

4. Dvir, H.; Harel, M.; McCarthy, A.A.; Toker, L.; Silman, I.; Futerman, A.H.; Sussman, J.L. X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease. *EMBO Rep.* **2003**, *4*, 704–709. [CrossRef] [PubMed]

5. Garman, S.C.; Garboczi, D.N. The molecular defect leading to Fabry disease: Structure of human alpha-galactosidase. *J. Mol. Biol.* **2004**, *337*, 319–335. [CrossRef] [PubMed]

6. Amberg, A. In Silico Methods. In *Drug Discovery and Evaluation: Safety and Pharmacokinetic Assays*; Vogel, H.G., Maas, J., Hock, F.J., Mayer, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2013. [CrossRef]

7. Tchernitchko, D.; Goossens, M.; Wajcman, H. In silico prediction of the deleterious effect of a mutation: Proceed with caution in clinical genetics. *Clin. Chem.* **2004**, *50*, 1974–1978. [CrossRef] [PubMed]

8. Rigsby, R.E.; Parker, A.B. Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem. Mol. Biol. Educ.* **2016**, *44*, 433–437. [CrossRef] [PubMed]

9. Hecht, M.; Bromberg, Y.; Rost, B. News from the protein mutability landscape. *J. Mol. Biol.* **2013**, *425*, 3937–3948. [CrossRef] [PubMed]

10. Hecht, M.; Bromberg, Y.; Rost, B. Better prediction of functional effects for sequence variants. *BMC Genom.* **2015**, *16* (Suppl. 8), S1. [CrossRef] [PubMed]

11. Bromberg, Y.; Rost, B. SNAP: Predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res.* **2007**, *35*, 3823–3835. [CrossRef] [PubMed]

12. Amaral, O.; Marcao, A.; Sa Miranda, M.; Desnick, R.J.; Grace, M.E. Gaucher disease: Expression and characterization of mild and severe acid beta-glucosidase mutations in Portuguese type 1 patients. *Eur. J. Hum. Genet.* **2000**, *8*, 95–102. [CrossRef] [PubMed]

13. Montfort, M.; Chabas, A.; Vilageliu, L.; Grinberg, D. Functional analysis of 13 GBA mutant alleles identified in Gaucher disease patients: Pathogenic changes and "modifier" polymorphisms. *Hum. Mutat.* **2004**, *23*, 567–575. [CrossRef] [PubMed]

14. Ron, I.; Dagan, A.; Gatt, S.; Pasmanik-Chor, M.; Horowitz, M. Use of fluorescent substrates for characterization of Gaucher disease mutations. *Blood Cells Mol. Dis.* **2005**, *35*, 57–65. [CrossRef] [PubMed]

15. Chabas, A.; Gort, L.; Diaz-Font, A.; Montfort, M.; Santamaria, R.; Cidras, M.; Grinberg, D.; Vilageliu, L. Perinatal lethal phenotype with generalized ichthyosis in a type 2 Gaucher disease patient with the [L444P;E326K]/P182L genotype: Effect of the E326K change in neonatal and classic forms of the disease. *Blood Cells Mol. Dis.* **2005**, *35*, 253–258. [CrossRef] [PubMed]

16. Grace, M.E.; Berg, A.; He, G.S.; Goldberg, L.; Horowitz, M.; Grabowski, G.A. Gaucher disease: Heterologous expression of two alleles associated with neuronopathic phenotypes. *Am. J. Hum. Genet.* **1991**, *49*, 646–655. [PubMed]

17. Liou, B.; Kazimierczuk, A.; Zhang, M.; Scott, C.R.; Hegde, R.S.; Grabowski, G.A. Analyses of variant acid beta-glucosidases: Effects of Gaucher disease mutations. *J. Biol. Chem.* **2006**, *281*, 4242–4253. [CrossRef] [PubMed]

18. Lukas, J.; Giese, A.K.; Markoff, A.; Grittner, U.; Kolodny, E.; Mascher, H.; Lackner, K.J.; Meyer, W.; Wree, P.; Saviouk, V.; et al. Functional characterisation of alpha-galactosidase a mutations as a basis for a new classification system in fabry disease. *PLoS Genet.* **2013**, *9*, e1003632. [CrossRef] [PubMed]

19. Park, J.Y.; Kim, G.H.; Kim, S.S.; Ko, J.M.; Lee, J.J.; Yoo, H.W. Effects of a chemical chaperone on genetic mutations in alpha-galactosidase A in Korean patients with Fabry disease. *Exp. Mol. Med.* **2009**, *41*, 1–7. [CrossRef] [PubMed]

20. Spada, M.; Pagliardini, S.; Yasuda, M.; Tukel, T.; Thiagarajan, G.; Sakuraba, H.; Ponzone, A.; Desnick, R.J. High incidence of later-onset fabry disease revealed by newborn screening. *Am. J. Hum. Genet.* **2006**, *79*, 31–40. [CrossRef] [PubMed]

21. Ferreira, S.; Ortiz, A.; Germain, D.P.; Viana-Baptista, M.; Caldeira-Gomes, A.; Camprecios, M.; Fenollar-Cortes, M.; Gallegos-Villalobos, A.; Garcia, D.; Garcia-Robles, J.A.; et al. The alpha-galactosidase A p.Arg118Cys variant does not cause a Fabry disease phenotype: Data from individual patients and family studies. *Mol. Genet. Metab.* **2015**, *114*, 248–258. [CrossRef] [PubMed]

22. Froissart, R.; Guffon, N.; Vanier, M.T.; Desnick, R.J.; Maire, I. Fabry disease: D313Y is an alpha-galactosidase A sequence variant that causes pseudodeficient activity in plasma. *Mol. Genet. Metab.* **2003**, *80*, 307–314. [CrossRef]

23. Yasuda, M.; Shabbeer, J.; Benson, S.D.; Maire, I.; Burnett, R.M.; Desnick, R.J. Fabry disease: Characterization of alpha-galactosidase A double mutations and the D313Y plasma enzyme pseudodeficiency allele. *Hum. Mutat.* **2003**, *22*, 486–492. [CrossRef] [PubMed]

24. Heinisch, U.; Zlotogora, J.; Kafert, S.; Gieselmann, V. Multiple mutations are responsible for the high frequency of metachromatic leukodystrophy in a small geographic area. *Am. J. Hum. Genet.* **1995**, *56*, 51–57. [PubMed]

25. Berná, L.; Gieselmann, V.; Poupetová, H.; Hrebícek, M.; Elleder, M.; Ledvinová, J. Novel mutations associated with metachromatic leukodystrophy: Phenotype and expression studies in nine Czech and Slovak patients. *Am. J. Med. Genet. A* **2004**, *129A*, 277–281. [CrossRef] [PubMed]

26. Hess, B.; Kafert, S.; Heinisch, U.; Wenger, D.A.; Zlotogora, J.; Gieselmann, V. Characterization of two arylsulfatase A missense mutations D335V and T274M causing late infantile metachromatic leukodystrophy. *Hum. Mutat.* **1996**, *7*, 311–317. [CrossRef]

27. Marcão, A.; Amaral, O.; Pinto, E.; Pinto, R.; Sá Miranda, M.C. Metachromatic leucodystrophy in Portugal-finding of four new molecular lesions: C300F, P425T, g.1190-1191insC, and g.2408delC. Mutations in brief no. 232. Online. *Hum. Mutat.* **1999**, *13*, 337–338. [CrossRef]

28. Lukatela, G.; Krauss, N.; Theis, K.; Selmer, T.; Gieselmann, V.; von Figura, K.; Saenger, W. Crystal structure of human arylsulfatase A: The aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. *Biochemistry* **1998**, *37*, 3654–3664. [CrossRef] [PubMed]

29. Hasegawa, Y.; Kawame, H.; Ida, H.; Ohashi, T.; Eto, Y. Single exon mutation in arylsulfatase A gene has two effects: Loss of enzyme activity and aberrant splicing. *Hum. Genet.* **1994**, *93*, 415–420. [CrossRef] [PubMed]

30. Furuya, H.; Kukita, Y.; Nagano, S.; Sakai, Y.; Yamashita, Y.; Fukuyama, H.; Inatomi, Y.; Saito, Y.; Koike, R.; Tsuji, S.; et al. Adult onset globoid cell leukodystrophy (Krabbe disease): Analysis of galactosylceramidase cDNA from four Japanese patients. *Hum. Genet.* **1997**, *100*, 450–456. [CrossRef] [PubMed]

31. De Gasperi, R.; Gama Sosa, M.A.; Sartorato, E.; Battistini, S.; Raghavan, S.; Kolodny, E.H. Molecular basis of late-life globoid cell leukodystrophy. *Hum. Mutat.* **1999**, *14*, 256–262. [CrossRef]

32. Fu, L.; Inui, K.; Nishigaki, T.; Tatsumi, N.; Tsukamoto, H.; Kokubu, C.; Muramatsu, T.; Okada, S. Molecular heterogeneity of Krabbe disease. *J. Inherit. Metab. Dis.* **1999**, *22*, 155–162. [CrossRef] [PubMed]

33. Matos, L.; Duarte, A.J.; Ribeiro, D.; Chaves, J.; Amaral, O.; Alves, S. Correction of a Splicing Mutation Affecting an Unverricht-Lundborg Disease Patient by Antisense Therapy. *Genes* **2018**, *9*, 455. [CrossRef] [PubMed]

34. Lalioti, M.D.; Mirotsou, M.; Buresi, C.; Peitsch, M.C.; Rossier, C.; Ouazzani, R.; Baldy-Moulinier, M.; Bottani, A.; Malafosse, A.; Antonarakis, S.E. Identification of mutations in cystatin B, the gene responsible for the Unverricht-Lundborg type of progressive myoclonus epilepsy (EPM1). *Am. J. Hum. Genet.* **1997**, *60*, 342–351. [PubMed]

35. Auerswald, E.A.; Nägler, D.K.; Assfalg-Machleidt, I.; Stubbs, M.T.; Machleidt, W.; Fritz, H. Hairpin loop mutations of chicken cystatin have different effects on the inhibition of cathepsin B, cathepsin L and papain. *FEBS Lett.* **1995**, *361*, 179–184. [CrossRef]

36. Joensuu, T.; Kuronen, M.; Alakurtti, K.; Tegelberg, S.; Hakala, P.; Aalto, A.; Huopaniemi, L.; Aula, N.; Michellucci, R.; Eriksson, K.; et al. Cystatin B: Mutation detection, alternative splicing and expression in progressive myclonus epilepsy of Unverricht-Lundborg type (EPM1) patients. *Eur. J. Hum. Genet.* **2007**, *15*, 185–193. [CrossRef] [PubMed]

37. Alakurtti, K.; Weber, E.; Rinne, R.; Theil, G.; de Haan, G.J.; Lindhout, D.; Salmikangas, P.; Saukko, P.; Lahtinen, U.; Lehesjoki, A.E. Loss of lysosomal association of cystatin B proteins representing progressive myoclonus epilepsy, EPM1, mutations. *Eur. J. Hum. Genet.* **2005**, *13*, 208–215. [CrossRef] [PubMed]

38. Horowitz, M.; Pasmanik-Chor, M.; Ron, I.; Kolodny, E.H. The enigma of the E326K mutation in acid β-glucocerebrosidase. *Mol. Genet. Metab.* **2011**, *104*, 35–38. [CrossRef] [PubMed]

39. Pinto, E.; Freitas, J.; Duarte, A.J.; Ribeiro, I.; Ribeiro, D.; Lima, J.L.; Chaves, J.; Amaral, O. Unverricht-Lundborg disease: Homozygosity for a new splicing mutation in the cystatin B gene. *Epilepsy Res.* **2012**, *99*, 187–190. [CrossRef] [PubMed]

40. Duarte, A.J.; Ribeiro, D.; Chaves, J.; Amaral, O. Characterization of a rare Unverricht-Lundborg disease mutation. *Mol. Genet. Metab. Rep.* **2015**, *4*, 68–71. [CrossRef] [PubMed]

41. O'Regan, G.; de Souza, R.M.; Balestrino, R.; Schapira, A.H. Glucocerebrosidase Mutations in Parkinson Disease. *J. Parkinsons Dis.* **2017**, *7*, 411–422. [CrossRef] [PubMed]

42. Migdalska-Richards, A.; Schapira, A.H. The relationship between glucocerebrosidase mutations and Parkinson disease. *J. Neurochem.* **2016**, *139* (Suppl. 1), 77–90. [CrossRef] [PubMed]

43. Berge-Seidl, V.; Pihlstrøm, L.; Maple-Grødem, J.; Forsgren, L.; Linder, J.; Larsen, J.P.; Tysnes, O.B.; Toft, M. The GBA variant E326K is associated with Parkinson's disease and explains a genome-wide association signal. *Neurosci. Lett.* **2017**, *658*, 48–52. [CrossRef] [PubMed]

44. Thirumal Kumar, D.; Eldous, H.G.; Mahgoub, Z.A.; George Priya Doss, C.; Zayed, H. Computational modelling approaches as a potential platform to understand the molecular genetics association between Parkinson's and Gaucher diseases. *Metab. Brain Dis.* **2018**. [CrossRef] [PubMed]

45. Desmet, F.O.; Hamroun, D.; Lalande, M.; Collod-Beroud, G.; Claustres, M.; Beroud, C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **2009**, *37*, e67. [CrossRef] [PubMed]

46. Pertea, M.; Lin, X.; Salzberg, S.L. GeneSplicer: A new computational method for splice site prediction. *Nucleic Acids Res.* **2001**, *29*, 1185–1190. [CrossRef] [PubMed]

47. Hebsgaard, S.M.; Korning, P.G.; Tolstrup, N.; Engelbrecht, J.; Rouze, P.; Brunak, S. Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.* **1996**, *24*, 3439–3452. [CrossRef] [PubMed]

48. Reese, M.G.; Eeckman, F.H.; Kulp, D.; Haussler, D. Improved splice site detection in Genie. *J. Comput. Biol.* **1997**, *4*, 311–323. [CrossRef] [PubMed]

49. Choi, Y.; Sims, G.E.; Murphy, S.; Miller, J.R.; Chan, A.P. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* **2012**, *7*, e46688. [CrossRef] [PubMed]

50. Choi, Y. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. In Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, Orlando, FL, USA, 7–10 October 2012; pp. 414–417. [CrossRef]

51. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [CrossRef] [PubMed]

52. Guex, N.; Peitsch, M.C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* **2009**, *30* (Suppl. 1), S162–S173. [CrossRef] [PubMed]

53. Artimo, P.; Jonnalagedda, M.; Arnold, K.; Baratin, D.; Csardi, G.; de Castro, E.; Duvaud, S.; Flegel, V.; Fortier, A.; Gasteiger, E.; et al. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **2012**, *40*, W597–W603. [CrossRef] [PubMed]