IDENTIFICATION OF MODIFIER GENES/NETWORKS OF LYSOSOMAL BIOLOGY

Presented by: ANYELO ALBERTO DURAN MOJICA

A thesis submitted to the School of Medicine, Clínica Alemana-Universidad del Desarrollo to attain the degree of Doctor (PhD) in Sciences and Innovation in Medicine

TUTOR: PhD. ANDRES DAVID KLEIN POSTERNACK

CO-TUTOR: PhD. JUAN FRANCISCO CALDERON GIADROSIC

May, 2023

Santiago, Chile

© Excerpts from this work may be reproduced for academic or research purposes, provided that the bibliographic reference is included.

Acknowledgments

First, I would like to thank my thesis supervisors, Profs. Andrés Klein, and Juan Francisco Calderón, for the opportunity to work in their laboratories and for the timely corrections, supervision, and guidance throughout my doctorate studies. Secondly, I would like to thank Prof. Frances Platt at the University of Oxford, an inspiring scientist with extraordinary professional and personal skills. I thank all the members of the Platt Lab, especially (Dave Smith, Allie Colaco, Dawn Shepherd, Kerri-Lee Wallom, María Fernández-Suárez, Claire Fletcher, Maysa Falah, Hsintsung Chen, Yuzhe Weng, Kellie Harkin, Lara Scofano, Thamali Ayagama, Krushni Lalji, Silvia Sultana, Carla Da Silva, Nick Platt, and David Priestman). To the James and Allie family, Kerri Family and David family, and the support staff of the Pharmacology Department - Oxford (Carolyn Thackrah, Jimmy Martin, James Weller, and Graeme Laidler) who in one way or another were always there to support me in any way I needed. I would also like to thank Prof. Silvana Zanlungo (UC), Dra. María José Yáñez (USS), Prof. Jaime Gutiérrez (USS), and Prof. Andrea Leiva (USS), as well as the members of their laboratories, for all the scientific and technical support provided to achieve this thesis. I want to thank Prof. Boris Rebolledo, Dra. Grazielle Esteves Ribeiro, Dr. Rodrigo Martínez, Mr. Marcelo Rojas, and Mr. Wagner Giovanni Guzmán for their continuous support. I want to thank the institutional support of the Faculty of Medicine, Clínica Alemana Universidad del Desarrollo, for their financial support. I want to give special thanks to the Institute of Science and Innovation in Medicine (ICIM) and all the staff associated with each of its Centres and units for allowing me to develop, grow as a person and broaden my horizons. A special thank you to MgSc. María Elena Pérez, for all the logistical support in the laboratory for

using equipment and quoting of reagents. To all the members of the Klein Lab (Juan Carlos Rubilar, Macarena Las Heras, Valeria Olguin, Pinkus Benjamín Szenfeld, Pablo Carrillo, Wagner Giovanni Guzmán, and Ana González), for all the shared moments. To my classmates (Eduardo Durán, Macarena Las Heras, Rodrigo Martínez, and Ainoa Fernández), with whom we embarked on this journey five years ago. To the funding institutions (ANID-CHILE, LysoMod) for supporting our work. To my professors in Venezuela, Dra. Nereida Valero and Dr. Jesús Mosquera, thank you for all the support, guidance, and learning throughout my life; wherever I go, I always have them present. To my family, who have always supported me from afar, they are my source of pride and joy, to my four mothers and my father, who, from the highest, you are a worthy being and an example of brotherhood and friendship. And finally, to Lineth my dearest love, thank you for being a pillar of support and my pilot-co-pilot on the journey.

Index

Chapter 1: Introduction ………..8-10

Chapter 2: Related articles, and derivatives of this doctoral thesis…………….11-12

Article 1: Genetic Background Matters: Population-Based Studies in Model Organisms for Translational Research..13-26

Article 2: Identification of genetic modifiers of murine hepatic β-glucocerebrosidase activity...27-32

Article 3: A Mouse Systems Genetics Approach Reveals Common and Uncommon Genetic Modifiers of Hepatic Lysosomal Enzyme Activities and Glycosphingolipids……………………………………………………………….33-55

Chapter 3: Conclusions and future projections...56-58

Additional references...59-65

Abstract

Lysosomal storage diseases (LSDs) are a heterogeneous group of ~70 rare inherited metabolic diseases caused by loss-of-function variants in genes encoding for lysosomal enzymes, their activators, or transport proteins. Clinical symptoms manifest during early childhood or adolescence, causing varying degrees of disability and short life expectancies. At a cellular level, LSD cells show a progressive accumulation of undegraded substrates. In a subset of LSDs, called sphingolipidosis, the primary buildup material corresponds to lipids. These diseases can affect several organs, including the liver, brain, heart, peripheral nervous system, haematologic, skeletal, gastrointestinal system, lung, muscle, and others. On the other hand, defects in processing sphingolipids have also been observed in patients with common diseases such as neurodegenerative disorders and cancer. To date, the phenotypic variability observed in monogenic conditions is thought to be influenced by genomic loci variation other than in the primary disease locus. These genes are called modifiers. We harnessed the natural genetic variation between different strains of healthy mice to identify modifier genes/networks of lysosomal biology. We used a systems genetics approach. We measured the hepatic activity of 12 lysosomal enzymes and several of their natural substrates in livers derived from a panel of inbred mouse strains, followed by genetic regulators mapping by genome-wide association studies (GWAS), transcriptome associations, Bayesian integration, and pathway enrichment analysis. The GWA study identified 137 non-redundant genes associated with changes in lysosomal enzyme activities and 1744 modifiers for GSLs. Among these genes, 30 are shared between the enzymes and lipid

groups. They are clustered into three pathways and associated with other non-LSD diseases. Surprisingly, they are under the regulation of ten common transcription factors, suggesting a common regulation of sphingolipid metabolism. In summary, we have identified novel regulators of lysosomal enzymes and GSL levels that may serve as therapeutic targets for LSD and implicated GSL metabolism in other diseases.

Keywords: systems genetics, inbred strains, modifier genes, lysosomal enzyme activity, glycosphingolipids, disorders with lysosomal dysfunction.

Chapter 1: Introduction

The lysosome is a specialized organelle with a highly dynamic and regulated structure that plays critical roles in cellular catabolism [1,2]. Lysosomes were initially thought of as a sac of enzymes for degrading macromolecules uptake through the endocytic pathway [1,3]. Recently, other functions have been uncovered, such as metabolic regulators [4,5], cholesterol transport, ion homeostasis, immune response, bone mineralization, plasma membrane repair, and transcriptional regulation [6–11]. Given their multifunctional character, lysosomes are considered central organelles for maintaining and regulating cellular homeostasis [1,2,12].

Defects in genes encoding for lysosomal proteins can cause ~70 rare monogenic congenital conditions known as lysosomal storage diseases (LSDs) [13–16]. Sphingolipidoses correspond to a subset of the LSD group, characterized by the intralysosomal build-up of partially digested sphingolipids [17,18], a fascinating group of lipids that combines ceramide with one or more sugar residues [19]. Sphingolipids are involved in various biological processes, including cell migration, apoptosis, senescence, growth regulation, and inflammatory responses [20]. Thus, intracellular accumulation of sphingolipids leads to cellular dysfunction, tissue damage, and organ failure [21,22].

Variants in genes associated with lysosomal storage disorders (LSDs) have been identified in several prevalent conditions such as Parkinson's disease (PD) and Alzheimer's disease (AD) [23–26], cancer [27], non-alcoholic fatty liver disease (NAFLD) [28], cardiovascular disease [29,30], and others [31,32]. This observation reinforces the notion that lysosomal dysfunction is crucial in the pathogenesis of these diseases.

Currently, there is no known cure for LSDs, and therapeutic options are limited [16]. Some available approaches include enzyme replacement therapy (ERT) [33,34], substrate reduction therapy (SRT) [35,36], cell therapies [37,38], and gene therapy [39,40], among others [41–44]. However, these treatments are often expensive, invasive, and not universally effective [45–49]. Therefore, there is a great need to develop new therapies for better management and ultimately cure these disorders [50–53].

The use of lysosomal modifying genes to restore the activity of the deficient enzyme or reduce the stored lipids is an underexplored therapeutic strategy [54,55]. A modifier gene refers to a locus in the DNA sequence that alters the typical phenotypes associated with target genes and affects organelle function [56]. The main challenge lies in discovering these modifier genes.

Genomic modifiers have been mapped using several techniques, including genome-wide association studies (GWAS) [57], linkage analysis [58], and systems biology approaches [59]. GWAS analyzes allele frequency differences among unrelated individuals with varying phenotypic traits [60]. It is advantageous over linkage studies regarding genomic mapping resolution [61,62] and, therefore, widely applied.

The mouse is a common model organism in biomedical research due to its genetic similarity with humans (sharing 90% of the genome) and similar cellular regulations and physiology [63,64]. Furthermore, tissues can be accessed, and their genomes are already genotyped [65]. Inbred mouse strains, such as the Hybrid Mouse Diversity Panel (HMDP), provide an extraordinary source of genetic diversity for discovering modifier genes involved in physiological

processes [66–68]. However, this advantage has not been fully utilized in uncovering genetic regulators of lysosomal biology.

Hypothesis

The natural genetic variation among different strains of healthy mice can be used to identify novel putative modifiers genes/networks that regulate the activity of lysosomal enzymes and their natural substrates by genome-wide association studies.

Main aim

To identify new modifier genes/networks of lysosomal phenotypes related to diseases with lysosomal dysfunction.

Specific aims

i.- To measure the activity of twelve lysosomal enzymes and determine GSL levels in livers from a panel of inbred mice.

ii.- To identify new putative modifiers genes/networks that regulate the activity of lysosomal enzymes and GSL by GWAS, pathway enrichment analysis, and integrative networks.

iii. To determine common and unique putative regulators of lysosomal enzyme activity and GSL levels.

Chapter 2: Related articles and derivatives of this doctoral thesis

Article 1. Genetic Background Matters: Population-Based Studies in Model Organisms for Translational Research.

Olguín*, Duran*, et al. Int J Mol Sci. 2022 Jul 8;23(14):7570. doi: 10.3390/ijms23147570 (* equal contributors).

Here, we present an overview of several model organisms - *Mus musculus, Drosophila melanogaste*r, and *Saccharomyces* - extensively utilized in biomedical research and evaluate their strengths and limitations. We described the origin and applications of the HMDP mouse panel used in this thesis. We highlight successful genomic studies that have identified genomic regions contributing to observed phenotypic variability. Our emphasis is on the translational relevance of these studies to precision medicine for both rare and common diseases, achieved through the extensive use of all available omics information associated with the different model organisms.

Article 2. Identification of genetic modifiers of murine hepatic β-glucocerebrosidase (GCase) activity.

Durán et al., Biochem Biophys Rep. 2021 Aug 18;28:101105. doi: 10.1016/j.bbrep.2021.101105.

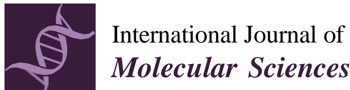
This work identified putative modifier genes of murine liver β-glucocerebrosidase activity using a systems genetics approach that allowed them to integrate intermediate molecular phenotypes. Together with the analysis with other

network-based approaches, they identified a repertoire of biological interactions based on predefined hepatic networks suggesting some functional interpretations for regulating this lysosomal enzyme.

Article 3. A Mouse Systems Genetics Approach Reveals Common and Uncommon Genetic Modifiers of Hepatic Lysosomal Enzyme Activities and Glycosphingolipids.

Durán et al., Int J Mol Sci. 2023 Mar 3;24(5):4915. doi: 10.3390/ijms24054915.

This study utilizes inbred strains from the Hybrid Mouse Diversity Panel to identify the genetic regulators influencing lysosomal enzyme activities, GSL levels, and the pathways enriched with these gene sets. Additionally, it dissects the genetic regulators that overlap between the two groups of phenotypes and identifies the transcription factors that may regulate them. The study also highlights that up-regulating a second gene can enhance the enzymatic activity associated with the primary gene, thereby improving the phenotype linked to a physiological or pathological process. These findings suggest potential new therapeutic avenues for diseases involving lysosomal dysfunction.


International Journal of
Molecular Sciences

Review

Genetic Background Matters: Population-Based Studies in Model Organisms for Translational Research

Valeria Olguín [1,†], Anyelo Durán [1,†], Macarena Las Heras [1], Juan Carlos Rubilar [1], Francisco A. Cubillos [2,3], Patricio Olguín [4] and Andrés D. Klein [1,*]

1 Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago 7610658, Chile; valeria.olguin.araneda@gmail.com (V.O.); anduranm@udd.cl (A.D.); mlasherasp@udd.cl (M.L.H.); jrubilare@udd.cl (J.C.R.)
2 Departamento de Biología, Santiago, Facultad de Química y Biología, Universidad de Santiago de Chile, Santiago 9170022, Chile; francisco.cubillos.r@usach.cl
3 Millennium Institute for Integrative Biology (iBio), Santiago 7500565, Chile
4 Program in Human Genetics, Institute of Biomedical Sciences, Biomedical Neurosciences Institute, Department of Neuroscience, Facultad de Medicina, Universidad de Chile, Santiago 8380453, Chile; patricioolguin@uchile.cl
* Correspondence: andresklein@udd.cl
† These authors contributed equally to this work.

Abstract: We are all similar but a bit different. These differences are partially due to variations in our genomes and are related to the heterogeneity of symptoms and responses to treatments that patients exhibit. Most animal studies are performed in one single strain with one manipulation. However, due to the lack of variability, therapies are not always reproducible when treatments are translated to humans. Panels of already sequenced organisms are valuable tools for mimicking human phenotypic heterogeneities and gene mapping. This review summarizes the current knowledge of mouse, fly, and yeast panels with insightful applications for translational research.

Keywords: systems genetics; mouse; *Drosophila*; *Saccharomyces cerevisiae*; translational research; genetic background; precision medicine; gene mapping

Citation: Olguín, V.; Durán, A.; Las Heras, M.; Rubilar, J.C.; Cubillos, F.A.; Olguín, P.; Klein, A.D. Genetic Background Matters: Population-Based Studies in Model Organisms for Translational Research. *Int. J. Mol. Sci.* **2022**, *23*, 7570.

https://doi.org/10.3390/ijms23147570

Academic Editors: Mauricio A. Retamal, Rodrigo Del Rio and Darío Acuña-Castroviejo

Received: 23 May 2022
Accepted: 4 July 2022
Published: 8 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Precision Medicine in Humans

Precision medicine characterizes diseases at a higher resolution by genomic and other technologies, providing more accurate targeting of patient subsets with tailored therapies [1]. To make this possible, large genotyped cohorts with deep clinical annotations are required to map loci responsible for the phenotypic variability. Common approaches to gene mapping include genome-wide association studies (GWAS) and linkage analysis in families of patients with variable disease severity [1]. These studies are time-consuming and expensive due to recruiting and genotyping costs. Furthermore, it is virtually impossible with rare diseases to find large cohorts in order to assure statistical significance for the genomic mapping.

Furthermore, families presenting enough informative individuals with variable symptoms are challenging to identify [2]. Strategies using model organisms with various genetic backgrounds are valuable resources for overcoming these obstacles. In this review, we describe many panels of organisms and examples of how modeling diseases on them can accelerate the pace of discoveries toward translational research in humans.

2. Rodents as Model Organisms in Genetic Research: Advantages and Limitations

The advantages of using mouse models in biomedicine have been discussed extensively [3]. Some benefits are the following: (i) the availability of genetic tools for creating disease models by transgenic, knockout, and knock-in technologies [4–6] (https://www.

jax.org/research-and-faculty/resources/mouse-mutant-resource/available-models (accessed on 22 May 2022)); (ii) inbred mouse strains are nearly isogenic, enabling to study how the same genetic mutation modifies a phenotype of interest in different genetic backgrounds [7–11]; (iii) mouse tissues are available for omics studies which can be chal- lenging to obtain from humans [12]. Some limitations include different evolutive pressures for mice and humans; therefore, some systems, such as the immune system, do not function similarly in both species [13].

2.1. Hybrid Mouse Diversity Panel

Currently available resources in rodents to find modifiers genes by association studies can be defined in two categories: (i) reference panels, consisting of inbred strains such as the Hybrid Mouse Diversity Panel (HMDP) and the Collaborative Cross (CC); (ii) populations derived from pseudo-random breeding of inbred strains, such as the Diversity Outbred (DO) and Heterogeneous Stock (HS) (Figure 1).

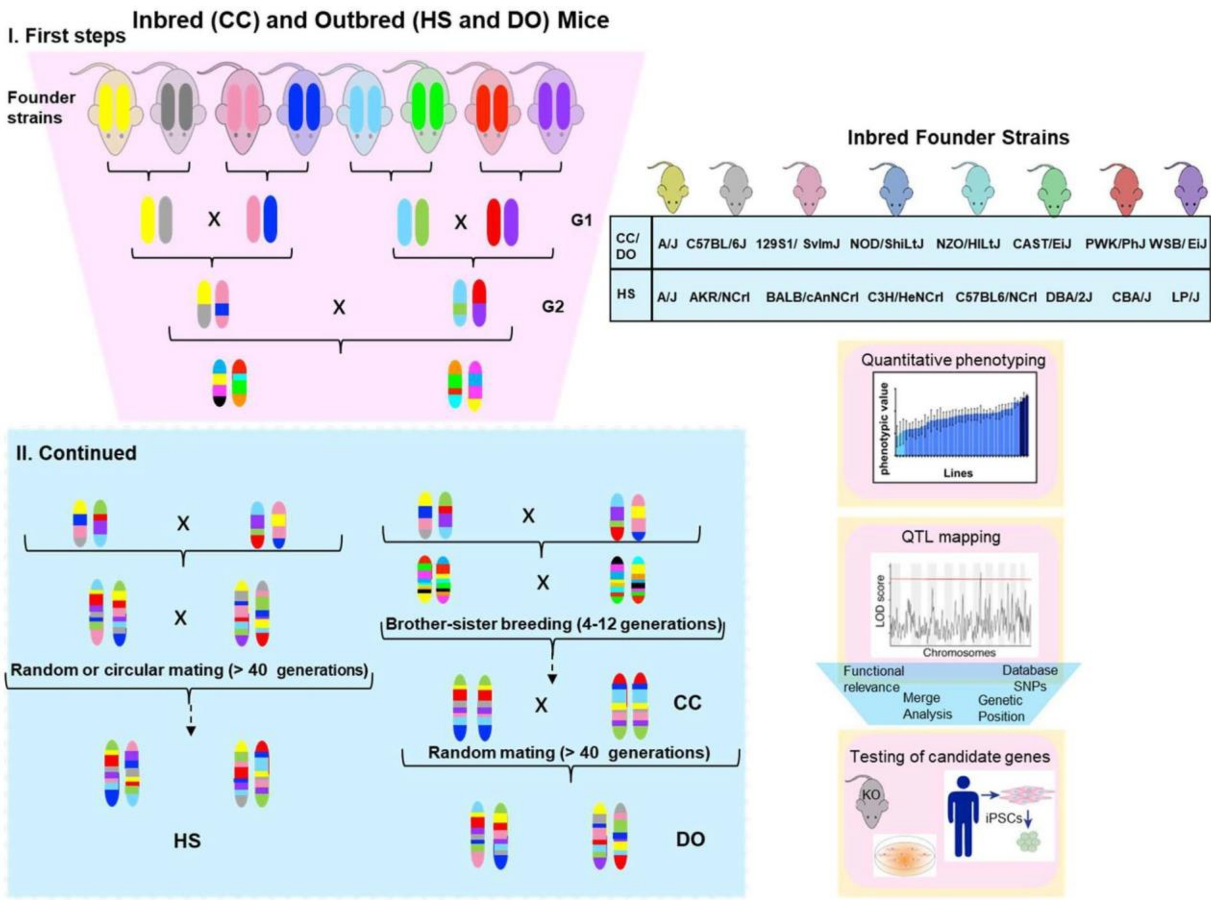


Figure 1. Breeding schemes for inbred (CC) and outbred (HS and DO) mice populations: Inbred founder strains for each panel are indicated in the right box. CC and DO populations share the same eight founder strains, five of which are standard laboratory inbred strains, while three are wild- derived strains. Colors represent the genotypes of strain chromosomes. The first steps include the combination of all eight founder genomes (outcrosses). CC is then generated as a recombinant inbred (RI) after multiple brother–sister breeding. HS and DO panels were developed as high-diversity outbred panels by over 40 generations of random outcrosses. DO was created from partially inbred Collaborative Cross (CC) mice. Quantitative phenotyping can be performed in the strains and used for gene mapping. Some signals in chromosomal locations will probably pass the threshold of significance (red line) in the LOD plot. The functional relevance of these variants can be assessed in animal models such as knockout mice and induced pluripotent stem cells (iPSC) derived from patients.

HMDP is a large panel of approximately 100 commercially available (https://www.jax.org (accessed on 22 May 2022)) and fully sequenced (www.sanger.ac.uk/science/data/mouse-genomes-project (accessed on 22 May 2022)) inbred strains: ~30 classical inbred strains and ~70 recombinant inbred (RI) strains derived mainly from crosses between C57BL/6J and DBA mice and A/J and C57BL/6J mice [14].

Advantages of using the HMDP panel are the following: (i) their genomes are known (http://mouse.cs.ucla.edu/mouseHapMap/ (accessed on 22 May 2022)); thus, it is unnecessary to spend funds performing this step; (ii) HMDP possesses ~4 million common single-nucleotide variants (SNVs), which is similar to the number present in humans [15]; (iii) high-resolution association mapping [14], which is at least an order of magnitude higher than in linkage analysis; (iv) it is possible to integrate gene mapping with other omics (transcriptomics, proteomics, and metabolomics data) [12]; (v) commercially avail- able (from The Jackson Laboratory, Harlan, and others); (vi) sufficient bioinformatics tools for data mining of complex mouse and human disease traits, such as the Systems Genetics Resource (SGR) (http://systems.genetics.ucla.edu (accessed on 22 May 2022)); (vii) servers to perform association mapping and statistical power simulation, which are also available in R to run them in house [16].

The HMDP also has limitations. For example, extensive linkage disequilibrium (LD) blocks are observed, both within and between chromosomes, probably as a result of the selection of allelic combinations conceding higher fitness during the inbreeding [17]. Consequently, regions in LD can lead to false-positive associations in GWAS analyses. Although the HMDP has a high mapping resolution, the statistical power to detect the effect of loci is small (estimated at 50% to variants explaining 10% of the trait variance) [14]. Since most loci contributing to a complex trait have an effect size below 5% [18], variants with subtle effects cannot always be detected by the HMDP. Power can be enhanced by including additional inbred and RI strains and performing meta-analyses from other panels such as the CC or traditional crosses [19].

An exciting application of the use of mouse panels in translational research comes from crossing the classical Alzheimer's disease (AD) mouse model (5XFAD) bearing mutations in APP and PSEN1 with 28 different strains of the BXD panel (AD-BXD). The F1 represents isogenic lines that were studied in a controlled environment. The AD-BXD panel mimicked several signs of the AD patients, including phenotypic variation in disease onset and severity. As in humans, the Apoe allele significantly affected spatial memory and other behavioral tests in the AD-BXD panel. Furthermore, hippocampal gene expression in the severe and mild lines agrees with transcriptomic changes observed in patients [20].

2.2. The Collaborative Cross (CC) Panel

The CC is a large panel of RI mouse strains obtained through systematically outcross- ing eight founder strains, followed by randomized breeding [21]. The founder strains of the CC include five of the widely used classical inbred laboratory strains (A/J, C57BL/6J, NOD/ShiLtJ, 129S1/SvImJ, and NZO/HILtJ), as well as three wild-derived strains descen- dent of three *M musculus* subspecies (WSB, Castaneous, and PWK) (Figure 1). These eight strains have been fully sequenced and carry ~45 million SNVs, four times more than those of classical laboratory mouse strains [22].

The genomes of the CC panel are known (http://csbio.unc.edu/CCstatus/CCGenomes (accessed on 22 May 2022)), which is helpful for genetic association studies. Haplotypes can be easily visualized or reconstructed as a mosaic of the genomes of the founders [23]. Parental strains capture approximately 90% of the genetic diversity seen in the *Mus mus- culus* species [24]. This high genetic diversity significantly reduces false candidate loci. Additionally, randomized breeding substantially increases mapping resolution by reducing population structure effects [25]. CC strains have been used to map quantitative trait loci (QTLs) to less than 5 Mb intervals [26]. Online tools are available to perform GWAS and linkage analyses [27]. Several aspects of human genetics and behavioral factors can be modeled in this system, including the heterogeneities observed in neurodevelopmental

disorders such as autistic spectrum disorders (ASDs) [28]. The CC panel allowed the discovery of novel candidate severity modifiers of ASD, e.g., *Bai3,* considered a potential target for pharmacological intervention [28].

Some considerations associated with using the CC panel are the following: (i) unique outlier phenotypes can arise in large studies, probably due to the complex genetic regu- latory networks involving multiple loci with epistatic interactions [29]; in such cases, the preferred approach for identifying causal genes is traditional F2 analysis or backcrosses [30]; (ii) because identifying loci could be time-consuming, it is suggested to perform a pilot study and expand as necessary [29]; (iii) creating a panel like the CC can generate breed- ing complications and infertility, mainly caused by genomic incompatibility introduced by the wild-derived strains. For that reason, the initial CC project aimed to produce 1000 strains but finished with only ~100 and inspired the creation of the Diversity Outbred (DO) population.

CC lines have been used for genetic association studies of many complex traits. QTL mapping for 15 metabolism- and exercise-related traits revealed five significant loci for body weight, some of which overlapped with previous human studies [31]. Gene mapping of rotarod (exercise) performance and body weight identified 45 loci, many of them related to neurological disorders and obesity in humans, suggesting a link between physical activity and neurodegeneration [32]. A study of glucose tolerance response in the CC panel identi- fied, only in female mice, a genomic region comprising 51 genes. This study highlighted sex differences in glucose response which should be considered in human studies [33]. The CC panel is also a valuable and reliable resource for studying host–pathogen interactions [29]. For example, to map genetic modifiers affecting the severity of *Pseudomonas aeruginosa* lung infections, 39 CC lines were inoculated with this pathogen. The phenotypic variability was enormous, ranging from complete resistance to lethality. It is particularly relevant to study the resistant lines since they have the biological secrets to design novel therapies for the susceptible. Genomic mapping and functional validation identified dihydropyrimidine dehydrogenase (*Dpyd*) and sphingosine-1-phosphate receptor 1 (*S1pr1*) as modifier genes. In a cohort of patients with cystic fibrosis, two SNVs in the *S1PR1* gene are associated with *Pseudomonas aeruginosa* infection [34], again indicating the translational relevance of multigenetic background studies in animal organisms.

2.3. Heterogeneous Stock and Diversity Outbred Populations

Both HS and DO are high-diversity outbred mice populations. The HS was established by breeding eight inbred strains and then outbreeding them in either a circular strategy or using random crosses (Figure 1) to minimize inbreeding [35]. After 50 or more generations, the HS-generated mice were a genetic mosaic of the founders' haplotypes [36,37]. On the other hand, the DO was established from partially inbred CC lines and is maintained indefinitely through pseudorandomized fashion non-sibling mating [38] (Figure 1). Since the DO is derived from the same eight founders as the CC, it presents the same allelic diversity as the CC strains. It can be used as a complementary tool in genetic association studies [39].

There are several advantages of using HS or DO mice compared to classical inbred mice. The outbred randomized mating increases the number of additional recombination sites compared to those of classically inbred mice; thus, each HS or DO mouse has a unique genome, which is a mosaic of the original eight founder lines, resembling human heterozygosity and allows high-resolution genetic mapping [39]. HS and DO mice have been used to finely map to intervals of 2.7 Mb [40] and less than 2 Mb [39], respectively. In addition, outbred animals are more vigorous and less prone to both early and late recessive allelic effects [41]. This genetic variability within both HS and DO populations results in a high degree of phenotypic variability; thus, outbred models enable the fine mapping of many phenotypic traits. Since the founders of CC and DO lines include wild-derived strains, unique behaviors can be observed compared to classical laboratory strains and represent a valuable tool for genetic behavior association studies [22]. A repository of

DO QTL studies can be shared between laboratories (https://dodb.jax.org (accessed on 22 May 2022)). Lastly, the founders of the HS and DO populations have been sequenced [42], reducing time and expense in locating the sequences.

Alternatively, some considerations must be made in the case of HS and DO mice. Since each outbred animal is genetically and phenotypically distinct, each HS and DO mouse requires genotyping and haplotype reconstruction to perform each QTL analysis [38]. High-resolution mapping can be achieved with these panels, but analyzing many animals is necessary for sufficient statistical power, which is not always possible [43]. Candidate modifiers of wild behaviors can be identified with outbred mice. However, it is challenging to validate in these panels because each animal has a unique genotype, in contrast to inbred lines [44].

An interesting translational study using the DO panel identified a diagnostic biomarker for human tuberculosis (TB). By applying machine learning algorithms to multidimensional data, the authors discovered CXCL1 as a putative biomarker of TB in the serum of mice. The biomarker was further validated in samples derived from human patients, discriminating active TB from latent infection and non-TB lung disease [45]. This study highlights the relevance of using population-based strategies to accelerate human biomarker discovery, validation, and testing.

3. *Drosophila melanogaster* as a Model Organism in Genetic Research: Advantages and Limitations

In addition to mouse models, *Drosophila melanogaster* has attracted many scientists. Flies are small, easy to manipulate in the laboratory, and cheap to maintain. They have a short life span (2 week generation interval) and produce many offspring. Flies show com- plex behaviors, including sleep, aggression, addiction, and social behavior [46]. Notably, about 70% of human disease-associated genes have a *Drosophila* ortholog [47]; its genome is fully sequenced and well annotated. It can be genetically modified using chemical and insertional mutagenesis, gene-specific mutations, or editions using CRISPR [47,48]. These characteristics support its use as a model system to study human diseases. As expected, the use of *Drosophila* for human research has limitations; for instance, the fly does not possess hemoglobin [49] and, thus, cannot be used for studying human pathologies related to this system.

3.1. Drosophila melanogaster Genetic Reference Panel (DGRP)

The DGRP is a collection of 205 inbred *Drosophila melanogaster* strains derived from a single natural population. Inseminated females were collected from the farmer's market in Raleigh, NC (USA), and their offspring were subjected to 20 generations of complete sibling mating [50] (Figure 2). The DGRP is a public resource available at the Bloomington Drosophila Stock Center (http://fly.bio.indiana.edu (accessed on 22 May 2022)) built for genomic association analyses. Currently, their genomes are available, and each line has minimal genetic variation [50]. Repeated measurements within each line are possible, enabling accuracy to increase the statistical power in GWA analyses. Since the DGRP is a publicly available resource, it allows different laboratories to correlate phenotypes on the same genotype and understand the pleiotropic effects of DNA variants and genes on multiple quantitative traits. Unlike the human genome, the fly genome has a structure with low LD between closely linked polymorphisms [51], which is favorable for accurate association mapping; thus, significant associated SNVs are likely causal or very near to a causal variant [52]. Lastly, experimentation in *Drosophila* has fewer ethical concerns compared to rodent models.

As with all study models, there are some limitations in DGRP that should be consid- ered. Firstly, genetic variation between the lines is a snapshot of the population from which they were derived; therefore, DGRP does not represent all the possible variations of the species. Secondly, the 205 lines usually provide enough statistical power to detect common

variants with moderate to large effects [53,54], but the statistical power is still limited for rare variants (minor allele frequency (MAF) < 0.05) [51].

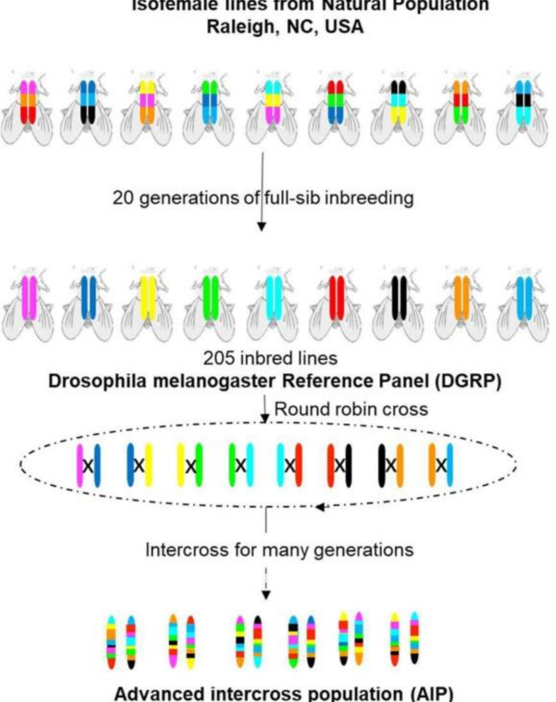


Figure 2. Generation of *Drosophila melanogaster* Genetic Reference Panel (DGRP) and Advanced Intercross Population (DGRP-AIPs). The DGRP corresponds to a sequenced panel derived from a natural fly population of Raleigh, NC (USA), and it was generated through 20 generations of full-sibling mating. The AIPs lines were derived from the DGRP by round-robin crossing and were then remapped.

3.2. DGRP for Mapping Physiological and Pathophysiological Traits

The DGRP has been used for GWA mapping of many different quantitative physi- ological traits, including food intake and sleep behavior [55,56]. Food intake is essential to animal fitness, and 25 modifiers with human orthologs were found [55]. Interestingly, diversity in mitochondrial haplotypes can directly mediate phenotypic variation in food intake [57]. Sleep has been increasingly explored in recent years with this model [56]. Flies resemble mammalian sleep and have become an important model species for identifying sleep regulation mechanisms. Analogous to human sleep studies, a DGRP GWAS high- lighted signals in the EGFR, Wnt, Hippo, and MAPK signaling pathways, suggesting that genes affecting variation in this trait are conserved [58]. DGRP studies revealed the genetic architecture of nutrient stores (glucose, glycogen, glycerol, protein, triglycerides, and wet weight) [59], developmental plasticity [60], and circadian cycle [61].

The DGRP has been used to identify candidate modifiers of retinal degeneration [62] and neurodegeneration in a Parkinson's disease (PD) model [63]. PD is a highly variable neurodegenerative disorder where variable manifestations range from cognitive distur- bances, motor alterations, and sleep and speech abnormalities to cellular pathological changes such as the formation of Lewy body inclusions and neuronal death [64]. The leucine-rich repeat kinase 2 gene G2019S mutation (LRRK2 G2019S) penetrance is incom- plete and varies among ethnic populations. In the Ashkenazy Jewish population, the low penetrance (26%) of the G2019S mutant phenotype suggests that other factors, such as the genetic background, the environment, and their interaction, act as modifiers of the variable phenotype [65,66]. In this regard, it has been reported that introducing the LRRK2 G2019S mutation in the DGRP results in considerable variability in the locomotor phenotype among backgrounds [63]. Gene mapping revealed 177 candidate modifier genes enriched in path-

ways involved in the neuronal outgrowth. The study suggests a link among LRRK2, neurite regulation, and neuronal degeneration in PD [63].

3.3. Lines Derived from DGRP and DSRP

A limitation of the DGRP is its low statistical power [51], which motivated the development of DGRP-derived advanced intercross populations (AIPs). These correspond to lines generated by crossing parentals DGRP for many generations, which were then remapped [67]. By successive crossings of a subset of parentals lines, it is possible to increase the recombination rate and, consequently, the statistical power compared to the DGRP [52]. Furthermore, the extreme QTL mapping strategy in AIPs can be used to resolve the statistical limitations of the DGRP for rare variants (MAF < 0.05). Extreme QTL map- ping refers to selecting individuals from the extremes of the phenotypic distribution for a trait (resembling a case–control study). Flies are pooled and sequenced, which is cheaper than sequencing all individuals of the initial population. This allows identifying alleles that segregate differentially among the distribution extremes (causal variant or in LD with it) [68,69]. The discovery of rare variants in DGRP will occur at higher frequencies in the AIPs after an extreme QTL mapping strategy.

A less applied strategy to increase the mapping power is to use DGRP and another panel for cross-validation, such as the Drosophila Synthetic Population Resource (DSPR). This collection of 1700 inbred lines is derived from 15 isogenic founder lines created from geographically distinct *Drosophila* populations [70]. However, some studies in both AIPs and DSPR lack overlap with candidate genes found in DGRP, probably due to the different genetic architecture or genetic variants between the panels.

4. *Saccharomyces cerevisiae* as a Model Organism in Genetic Research: Advantages and Limitations

Saccharomyces cerevisiae, the budding yeast, has gained prominence as a model organism in quantitative genetics because it has several experimental and biologically advantageous features. For example, it has a small and compact genome of approximately 12 million bp in haploids (about one two-hundredth of the human genome). It contains fewer introns and a lower proportion of intergenic sequences than higher eukaryotes [71]. Furthermore, it is easy to cultivate and maintain in large population size in the laboratory. In addition, two-thirds of all yeast genes share at least one domain of significant homology with human genes, and about 30% of known genes involved in human diseases have yeast orthologs [72].

One of the main advantages of yeast for quantitative genetics studies is its large genetic map. *S. cerevisiae* exhibits high meiotic recombination rates, with an average of about 90 crossovers per meiosis, allowing precise quantitative phenotyping [71,73,74]. The homologous recombination in yeast is highly efficient, facilitating the deletion of sequences or genes in vivo [72,75]. This efficient recombination permitted the generation of the first complete deletion mutant strain collection using gene replacement with the G418 resistance gene (KanMX) cassette in the reference *S. cerevisiae* strain [76]. Since then, similar panels have been available in different genetic backgrounds, demonstrating the high degree of genetic background dependencies for different phenotypes [77,78]. Yeasts have less genetic complexity than flies and rodents. Thus, it is easier to study the effect of a single gene because of the reduced genetic redundancy [79].

4.1. Analysis of Segregating Populations from Pairwise Crosses

QTL mapping in yeast has been the primary approach to uncovering genetic variants responsible for phenotypic differences between genetic backgrounds. Identifying QTLs has been achieved by analyzing segregating populations from pairwise crosses, mainly through linkage or bulk segregant analysis (BSA) [80,81]. Linkage mapping in yeast involves mating two or more haploid parental strains that show phenotypic variation and then phenotyping and genotyping a panel of recombinant offspring obtained from

these crosses. Recombination breaks allow causal loci to segregate with the phenotype of interest, and QTLs are identified using statistical tests [80,82]. The BSA also involves crossing two or more parental strains and subsequent phenotyping of their recombinant offspring [83]. However, the BSA method uses selective genotyping of subsets of segregants, commonly the extremes of the phenotypic distribution [84]. Typically, segregants undergo selective environmental pressure, where large pools are constructed. One expresses the trait of interest (selected pool), and others are not selected (control pool) or exhibit the opposite phenotype. After genotyping each marker, genetic regions of allelic enrichment are predicted as QTLs that contribute to the attribute of interest [85]. These approaches from pairwise crosses have been successfully applied to map yeast genetic variation responsible for nitrogen utilization [86], metabolic fluxes, ethanol tolerance [87], and high-temperature fermentation [88].

Most crosses constructed in yeast have involved the reference laboratory strain S288c or its derivatives crossed against a wild or fermentative isolate [89]. However, these strains only harbor a small fraction of the phenotypic variation of natural populations and have mosaic genomes of the founder strains [84,90]. Therefore, studies using biparental crosses provide a poor understanding of the relationship between the genetic background and the QTLs. These studies lack resolution since few generations are used; consequently, they are unable to reveal the complete architecture of polygenic traits. Moreover, laboratory strains often contain artificial auxotrophic markers that confound mapping experiments [91]. Investigators have recently established advanced-generation multi-parent populations (MPPs) in yeast to overcome these problems (Figure 3).

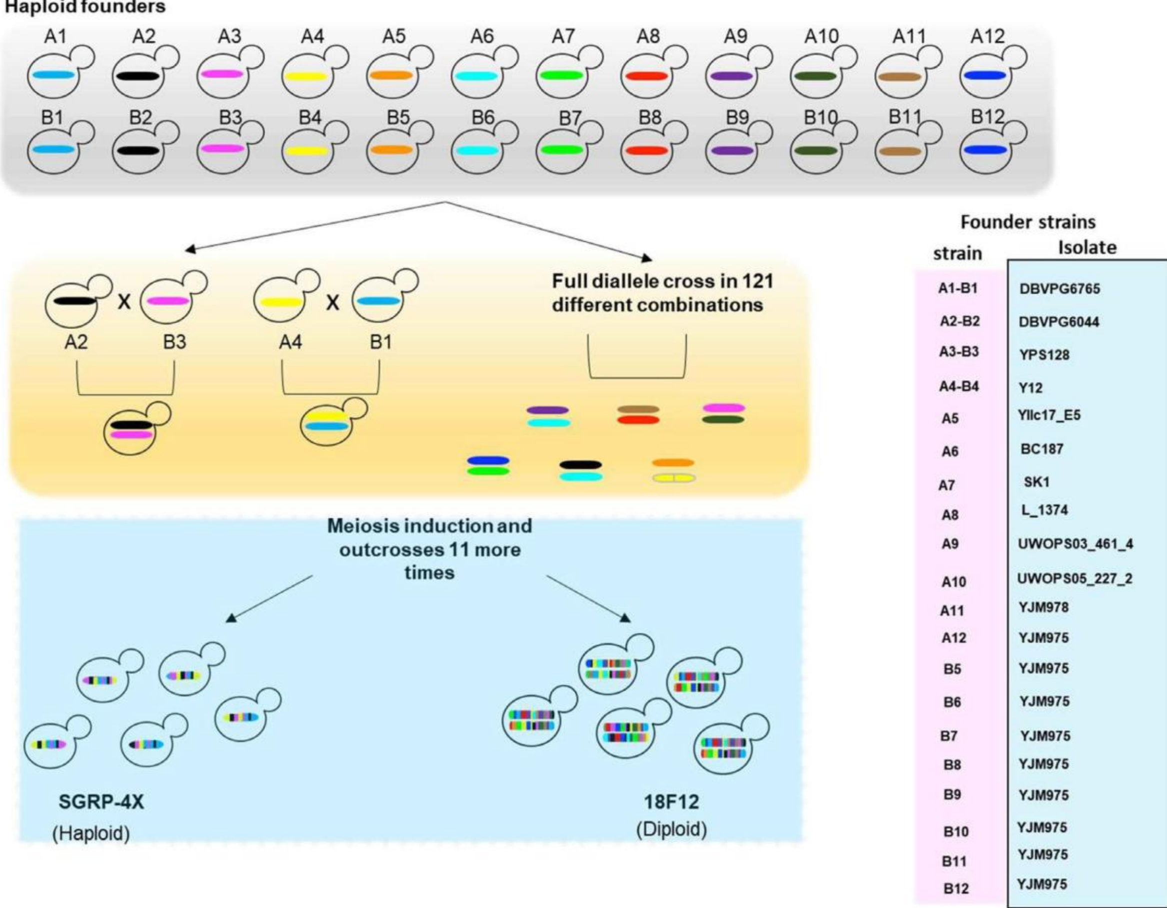


Figure 3. Cross design of SGRP-4X and 18F12 mapping populations. Haploid founder strains used for generations of these populations are indicated in the right box. Ax and Bx indicate the Mat a and Mat α haploid founder strains, respectively.

4.2. Multi-Parent Populations (MPPs)

Yeast MPPs comprise large populations with thousands to millions of individuals obtained from two main steps. Firstly, several (inbred or isogenic) founder strains from various geographical origins are crossed, and then the intercross of the resulting population is subsequentially crossed for several generations [81]. Large segregating populations are then used for mapping QTLs. The first MPP in yeast was established by Cubillos et al. [92] by crossing four strains representative of the main *S. cerevisiae* lineages (Y12 strain as representative of the SA lineage, YPS128 of the NA lineage, DBVPG6044 of the WA, and DBVPG6765 of the WE lineage) for 12 generations. The SGRP-4X contains 165 sequenced segregants, representing recombined genetic mosaics of the founder strains. Later, Linder et al. [93] extended this approach and created 18F12v1 and 18F12v2, two outbred MPPs derived from a cross of 18 genetically diverse founder strains, with each strain derived from the SGRP collection [84,92,93].

MPPs in yeast are robust mapping resources due to multiple founders and rounds of recombination in many individuals that increase both the genetic and the phenotypic diversity, as well as the linkage block resolution of the QTL mapping compared to biparental F1 or F2 populations. In fact, in yeast, it has been shown that only a few rounds of meiosis are sufficient to obtain spaced near-genic resolution [94]. Association mapping in MPPs provides more equilibrated allelic frequencies than biparental populations, increasing knowledge about the population structure [95]. Integration of this information in the QTL analysis can reduce the probability of obtaining false-positive results, thus demonstrating yeast as an accurate model system to identify dozens to hundreds of genes underlying phenotypes of interest.

4.3. Genome-Wide Association Studies (GWAS) in S. cerevisiae

GWAS utilizes the variation in large populations of unrelated individuals to provide insights into the causes of common complex traits. However, in 2012, only 36 *S. cerevisiae* genomes were available from the Saccharomyces Genome Resequencing Project, hampering GWAS studies in yeast. This situation motivated the development of a project to describe whole-genome sequence variation in numerous yeast populations (http://1002genomes.u-strasbg.fr/ (accessed on 22 May 2022)). Today, more than 2000 genomes isolated from a wide range of locations (including Australia, Europe, Russia, Vietnam, and South Africa) are available [96]. Thus, investigators can conduct GWAS in this model organism [97].

The success of GWAS in *S. cerevisiae* is a result of high diversity among natural isolates relative to humans [96], low linkage disequilibrium (extended in an average half-life of <3 kb) [98], and relatively simple quantification of phenotypes in hundreds to thousands of individuals. However, GWAS in yeast is affected by a large population structure [84,98], leading to limited statistical power and spurious associations. The increment in the number of genotyped individuals is comparable to other model organisms enabling GWAS to describe copy number variants (CNV) as having a more significant phenotypic effect than SNV in yeast and laying the foundation for GWAS in the species [99].

Many of the phenotypes addressed in yeast are directly related to the cell-autonomous features of human diseases, including neurological conditions such as Parkinson's disease [100]. Thus far, most of the disease genome-wide screenings in *S. cerevisiae* have deleted one gene at a time. To our knowledge, the genomic variability of yeast isolates is starting to be used for modeling human phenotypic variabilities. In the field of longevity and environment, a study in which 58 natural yeast strains were used led to identifying *RIM15* and *SER1* as longevity genes under caloric restrictions [101].

In the future, we expect to observe increased research using panels of organisms, where a combination of variants can be identified. This technique could be feasible in the short term for diseases that can be mimicked pharmacologically and in the medium term for disorders that can be reproduced genetically.

5. Practical Considerations and Concluding Remarks

Each of the discussed organisms and panels has advantages and disadvantages for human translational research. In addition to the already mentioned ones, researchers should consider practical factors for deciding the best model for each project. Some relevant factors are presented in Table 1.

Table 1. Practical considerations for choosing model organisms and their panels. The references are shown in brackets. When deciding the best model for a project, variables such as the percentage of homolog genes to human disease-causing genes, costs, and the possibility of automatization should be considered.

	Mus musculus	*Drosophila melanogaster*	*Saccharomyces cerevisiae*
Genome size (kb)	2,725,521 [102]	180,000 [103]	12,070 [104]
Percentage of homolog genes to human disease-causing genes	99 [105]	70 [47,106]	60 [107]
Costs to keep the panels	High	Medium	Very low
Complex behaviors	Yes	Yes	No
Discovery of cell-autonomous processes	Yes	Yes	Yes
Speed for throughput screenings and automatization of measurements	Slow	Fast	Very fast

In conclusion, the consequences of a genetic mutation can be strongly modified by the biological background in which it operates. For example, a loss-of-function mutation may be well tolerated in one genetic context and lethal in another. The most resistant individuals have the biological secrets useful for developing therapies for the most susceptible ones. Human studies are challenging; they can take a long time due to the recruitment of large cohorts, and genomic sequencing is expensive. Instead, modeling diseases in already sequenced panels of diverse model organisms followed by gene mapping and validation in smaller human cohorts can speed up translational research and precision medicine for both common and rare diseases.

Author Contributions: Conceptualization: V.O., A.D., P.O., F.A.C. and A.D.K. Writing original draft: V.O., A.D., M.L.H., J.C.R. and A.D.K. writing—review and editing: F.A.C., P.O. and A.D.K. Supervision: A.D.K. All authors have read and agreed to the published version of the manuscript.

Funding: A.D.K. is funded by the Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) grant No. 1180337. P.O. is funded by the Biomedical Neuroscience Institute (BNI), Iniciativa Científica Milenio, ICN09_015, Agencia Nacional de Investigación y Desarrollo ACE210007. P.O. and A.D.K. are supported by the Pew Innovation Fund, grant No. 32422.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ashley, E.A. Towards precision medicine. *Nat. Rev. Genet.* **2016**, *17*, 507–522. [CrossRef] [PubMed]
2. Rahit, K.M.T.H.; Tarailo-Graovac, M. Genetic Modifiers and Rare Mendelian Disease. *Genes* **2020**, *11*, 239. [CrossRef] [PubMed]
3. Canales, C.P.; Walz, K. The Mouse, a Model Organism for Biomedical Research. In *Cellular and Animal Models in Human Genomics Research*; Academic Press: Cambridge, MA, USA, 2019; pp. 119–140. [CrossRef]
4. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821. [CrossRef] [PubMed]
5. Platt, R.J.; Chen, S.; Zhou, Y.; Yim, M.J.; Swiech, L.; Kempton, H.R.; Dahlman, J.E.; Parnas, O.; Eisenhaure, T.M.; Jovanovic, M.; et al. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. *Cell* **2014**, *159*, 440–455. [CrossRef] [PubMed]
6. Swiech, L.; Heidenreich, M.; Banerjee, A.; Habib, N.; Li, Y.; Trombetta, J.J.; Sur, M.; Zhang, F. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **2014**, *33*, 102–106. [CrossRef] [PubMed]
7. Calderón, J.F.; Klein, A.D. Controversies on the potential therapeutic use of rapamycin for treating a lysosomal cholesterol storage disease. *Mol. Genet. Metab. Rep.* **2018**, *15*, 135–136. [CrossRef]

8. Durán, A.; Rebolledo-Jaramillo, B.; Olguin, V.; Rojas-Herrera, M.; Heras, M.L.; Calderón, J.F.; Zanlungo, S.; Priestman, D.A.; Platt, F.M.; Klein, A.D. Identification of genetic modifiers of murine hepatic β-glucocerebrosidase activity. *Biochem. Biophys. Rep.* **2021**, *28*, 101105. [CrossRef]
9. Parra, J.; Klein, A.D.; Castro, J.; Morales, M.G.; Mosqueira, M.; Valencia, I.; Cortés, V.; Rigotti, A.; Zanlungo, S. Npc1 deficiency in the C57BL/6J genetic background enhances Niemann–Pick disease type C spleen pathology. *Biochem. Biophys. Res. Commun.* **2011**, *413*, 400–406. [CrossRef]
10. Klein, A.D.; Ferreira, N.-S.; Ben-Dor, S.; Duan, J.; Hardy, J.; Cox, T.M.; Merril, A.H., Jr.; Futerman, A.H. Identification of Modifier Genes in a Mouse Model of Gaucher Disease. *Cell Rep.* **2016**, *16*, 2546–2553. [CrossRef]
11. Rodriguez-Gil, J.L.; Watkins-Chow, D.E.; Baxter, L.L.; Elliot, G.; Harper, U.L.; Wincovitch, S.M.; Wedel, J.C.; Incao, A.A.; Huebecker, M.; Boehm, F.J.; et al. Genetic background modifies phenotypic severity and longevity in a mouse model of Niemann-Pick disease type C1. *Dis. Model. Mech.* **2020**, *13*, dmm042614. [CrossRef]
12. Klein, A.D. Modeling diseases in multiple mouse strains for precision medicine studies. *Physiol. Genom.* **2017**, *49*, 177–179. [CrossRef] [PubMed]
13. Seok, J.; Warren, H.S.; Alex, G.C.; Michael, N.M.; Henry, V.B.; Xu, W.; Richards, D.R.; McDonald-Smith, G.P.; Gao, H.; Hennessy, L.; et al. Genomic Responses in Mouse Models Poorly Mimic Human Inflammatory Diseases. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 3507–3512. [CrossRef] [PubMed]
14. Bennett, B.J.; Farber, C.R.; Orozco, L.; Kang, H.M.; Ghazalpour, A.; Siemers, N.; Neubauer, M.; Neuhaus, I.; Yordanova, R.; Guan, B.; et al. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* **2010**, *20*, 281–290. [CrossRef] [PubMed]
15. Lusis, A.J.; Seldin, M.M.; Allayee, H.; Bennett, B.J.; Civelek, M.; Davis, R.C.; Eskin, E.; Farber, C.; Hui, S.; Mehrabian, M.; et al. The Hybrid Mouse Diversity Panel: A resource for systems genetics analyses of metabolic and cardiovascular traits. *J. Lipid Res.* **2016**, *57*, 925–942. [CrossRef] [PubMed]
16. Ghazalpour, A.; Rau, C.D.; Farber, C.R.; Bennett, B.J.; Orozco, L.D.; Van Nas, A.; Pan, C.; Allayee, H.; Beaven, S.W.; Civelek, M.; et al. Hybrid mouse diversity panel: A panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm. Genome* **2012**, *23*, 680–692. [CrossRef] [PubMed]
17. Petkov, P.; Graber, J.; Churchill, G.A.; DiPetrillo, K.; King, B.; Paigen, K. Evidence of a Large-Scale Functional Organization of Mammalian Chromosomes. *PLoS Genet.* **2005**, *1*, e33. [CrossRef]
18. Flint, J.; Mott, R. Applying mouse complex-trait resources to behavioural genetics. *Nature* **2008**, *456*, 724–727. [CrossRef]
19. Kang, E.Y.; Han, B.; Furlotte, N.; Joo, J.W.J.; Shih, D.; Davis, R.C.; Lusis, A.J.; Eskin, E. Meta-Analysis Identifies Gene-by-Environment Interactions as Demonstrated in a Study of 4,965 Mice. *PLoS Genet.* **2014**, *10*, e1004022. [CrossRef]
20. Neuner, S.M.; Heuer, S.E.; Huentelman, M.J.; O'Connell, K.M.S.; Kaczorowski, C.C. Harnessing Genetic Complexity to Enhance Translatability of Alzheimer's Disease Mouse Models: A Path toward Precision Medicine. *Neuron* **2019**, *101*, 399–411. [CrossRef]
21. Srivastava, A.; Morgan, A.P.; Najarian, M.L.; Sarsani, V.K.; Sigmon, J.S.; Shorter, J.R.; Kashfeen, A.; McMullan, R.C.; Williams, L.H.; Giusti-Rodríguez, P.; et al. Genomes of the Mouse Collaborative Cross. *Genetics* **2017**, *206*, 537–556. [CrossRef]
22. Yang, H.; Wang, J.R.; Didion, J.; Buus, R.J.; Bell, T.A.; Welsh, C.E.; Bonhomme, F.; Yu, A.H.-T.; Nachman, M.W.; Piálek, J.; et al. Subspecific origin and haplotype diversity in the laboratory mouse. *Nat. Genet.* **2011**, *43*, 648–655. [CrossRef] [PubMed]
23. Aylor, D.L.; Valdar, W.; Foulds-Mathes, W.; Buus, R.J.; Verdugo, R.A.; Baric, R.S.; Ferris, M.T.; Frelinger, J.A.; Heise, M.; Frieman, M.B.; et al. Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res.* **2011**, *21*, 1213–1222. [CrossRef] [PubMed]
24. Roberts, A.; de Villena, F.P.-M.; Wang, W.; McMillan, L.; Threadgill, D.W. The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: Implications for QTL discovery and systems genetics. *Mamm. Genome* **2007**, *18*, 473–481. [CrossRef] [PubMed]
25. Saul, M.C.; Philip, V.M.; Reinholdt, L.G.; Chesler, E.J. High-Diversity Mouse Populations for Complex Traits. *Trends Genet.* **2019**, *35*, 501–514. [CrossRef] [PubMed]
26. Keele, G.; Zhang, T.; Pham, D.; Vincent, M.; Genomics, T.B.-C. *Undefined Regulation of Protein Abundance in Genetically Diverse Mouse Populations*; Elsevier: Amsterdam, The Netherlands, 2021.
27. Ram, R.; Morahan, G. Complex Trait Analyses of the Collaborative Cross: Tools and Databases. *Syst. Genet.* **2016**, *1488*, 121–129. [CrossRef]
28. Molenhuis, R.T.; Bruining, H.; Brandt, M.J.V.; Van Soldt, P.E.; Atamni, H.J.A.-T.; Burbach, J.P.H.; Iraqi, F.A.; Mott, R.F.; Kas, M.J.H. Modeling the quantitative nature of neurodevelopmental disorders using Collaborative Cross mice. *Mol. Autism* **2018**, *9*, 63. [CrossRef]
29. Noll, K.; Ferris, M.T.; Heise, M.T. The Collaborative Cross: A Systems Genetics Resource for Studying Host-Pathogen Interactions. *Cell Host Microbe* **2019**, *25*, 484–498. [CrossRef]
30. Rogala, A.R.; Morgan, A.P.; Christensen, A.M.; Gooch, T.J.; Bell, T.A.; Miller, D.R.; Godfrey, V.L.; De Villena, F.P.-M. The Collaborative Cross as a Resource for Modeling Human Disease: CC011/Unc, a New Mouse Model for Spontaneous Colitis. *Mamm. Genome* **2014**, *25*, 95–108. [CrossRef]
31. Mathes, W.F.; Aylor, D.L.; Miller, D.R.; Churchill, G.A.; Chesler, E.J.; de Villena, F.P.-M.; Threadgill, D.W.; Pomp, D. Architecture of energy balance traits in emerging lines of the Collaborative Cross. *Am. J. Physiol. Metab.* **2011**, *300*, E1124–E1134. [CrossRef]

32. Mao, J.-H.; Langley, S.A.; Huang, Y.; Hang, M.; Bouchard, K.E.; Celniker, S.E.; Brown, J.B.; Jansson, J.; Karpen, G.H.; Snijders, A.M. Identification of genetic factors that modify motor performance and body weight using Collaborative Cross mice. *Sci. Rep.* **2015**, *5*, 16247. [CrossRef]
33. Atamni, H.J.A.-T.; Ziner, Y.; Mott, R.; Wolf, L.; Iraqi, F.A. Glucose tolerance female-specific QTL mapped in collaborative cross mice. *Mamm. Genome* **2016**, *28*, 20–30. [CrossRef] [PubMed]
34. Lorè, N.I.; Sipione, B.; He, G.; Strug, L.J.; Atamni, H.J.; Dorman, A.; Mott, R.; Iraqi, F.A.; Bragonzi, A. Collaborative Cross Mice Yield Genetic Modifiers for Pseudomonas aeruginosa Infection in Human Lung Disease. *mBio* **2020**, *11*, e00097–e00117. [CrossRef]
35. Woods, L.C.S. QTL mapping in outbred populations: Successes and challenges. *Physiol. Genom.* **2014**, *46*, 81–90. [CrossRef] [PubMed]
36. Talbot, C.J.; Nicod, A.; Cherny, S.S.; Fulker, D.W.; Collins, A.C.; Flint, J. High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* **1999**, *21*, 305–308. [CrossRef]
37. Woods, L.C.S.; Mott, R. Heterogeneous Stock Populations for Analysis of Complex Traits. *Syst. Genet.* **2016**, *1488*, 31–44. [CrossRef]
38. Gatti, D.M.; Svenson, K.L.; Shabalin, A.; Wu, L.-Y.; Valdar, W.; Simecek, P.; Goodwin, N.; Cheng, R.; Pomp, D.; Palmer, A.; et al. Quantitative Trait Locus Mapping Methods for Diversity Outbred Mice. *G3 Genes|Genomes|Genet.* **2014**, *4*, 1623–1633. [CrossRef] [PubMed]
39. Logan, R.W.; Robledo, R.F.; Recla, J.M.; Philip, V.M.; Bubier, J.A.; Jay, J.J.; Harwood, C.; Wilcox, T.; Gatti, D.M.; Bult, C.J.; et al. High-precision genetic mapping of behavioral traits in the diversity outbred mouse population. *Genes Brain Behav.* **2013**, *12*, 424–437. [CrossRef]
40. Valdar, W.; Solberg, L.C.; Gauguier, D.; Burnett, S.; Klenerman, P.; Cookson, W.O.; Taylor, M.; Rawlins, J.N.P.; Mott, R.; Flint, J. Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet.* **2006**, *38*, 879–887. [CrossRef]
41. Svenson, K.L.; Gatti, D.M.; Valdar, W.; Welsh, C.E.; Cheng, R.; Chesler, E.J.; Palmer, A.A.; McMillan, L.; Churchill, G.A. High-Resolution Genetic Mapping Using the Mouse Diversity Outbred Population. *Genetics* **2012**, *190*, 437–447. [CrossRef]
42. Keane, T.M.; Goodstadt, L.; Danecek, P.; White, M.A.; Wong, K.; Yalcin, B.; Heger, A.; Agam, A.; Slater, G.; Goodson, M.; et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **2011**, *477*, 289–294. [CrossRef]
43. Parker, C.C.; Palmer, A.A. Dark Matter: Are Mice the Solution to Missing Heritability? *Front. Genet.* **2011**, *2*, 32. [CrossRef] [PubMed]
44. Chesler, E.J. Out of the bottleneck: The Diversity Outcross and Collaborative Cross mouse populations in behavioral genetics research. *Mamm. Genome* **2013**, *25*, 3–11. [CrossRef] [PubMed]
45. Koyuncu, D.; Niazi, M.K.K.; Tavolara, T.; Abeijon, C.; Ginese, M.L.; Liao, Y.; Mark, C.; Specht, A.; Gower, A.C.; Restrepo, B.I.; et al. CXCL1: A new diagnostic biomarker for human tuberculosis discovered using Diversity Outbred mice. *PLoS Pathog.* **2021**, *17*, e1009773. [CrossRef] [PubMed]
46. Kazama, H. Systems neuroscience in Drosophila: Conceptual and technical advantages. *Neuroscience* **2015**, *296*, 3–14. [CrossRef] [PubMed]
47. Yamamoto, S.; Jaiswal, M.; Charng, W.-L.; Gambin, T.; Karaca, E.; Mirzaa, G.; Wiszniewski, W.; Sandoval, H.; Haelterman, N.A.; Xiong, B.; et al. A Drosophila Genetic Resource of Mutants to Study Mechanisms Underlying Human Genetic Diseases. *Cell* **2014**, *159*, 200–214. [CrossRef]
48. Reiter, L.T.; Potocki, L.; Chien, S.; Gribskov, M.; Bier, E. A Systematic Analysis of Human Disease-Associated Gene Sequences in *Drosophila melanogaster*. *Genome Res.* **2001**, *11*, 1114–1125. [CrossRef] [PubMed]
49. Myers, E.W.; Sutton, G.G.; Delcher, A.L.; Dew, I.M.; Fasulo, D.P.; Flanigan, M.J.; Kravitz, S.A.; Mobarry, C.M.; Reinert, K.H.J.; Remington, K.A.; et al. A Whole-Genome Assembly of *Drosophila*. *Science* **2000**, *287*, 2196–2204. [CrossRef]
50. Mackay, T.F.C.; Richards, S.; Stone, E.A.; Barbadilla, A.; Ayroles, J.F.; Zhu, D.; Casillas, S.; Han, Y.; Magwire, M.M.; Cridland, J.M.; et al. The Drosophila melanogaster Genetic Reference Panel. *Nature* **2012**, *482*, 173–178. [CrossRef]
51. Huang, W.; Massouras, A.; Inoue, Y.; Peiffer, J.; Ràmia, M.; Tarone, A.M.; Turlapati, L.; Zichner, T.; Zhu, D.; Lyman, R.F.; et al. Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* **2014**, *24*, 1193–1208. [CrossRef]
52. Anholt, R.R.H.; Mackay, T.F.C. The road less traveled: From genotype to phenotype in flies and humans. *Mamm. Genome* **2017**, *29*, 5–23. [CrossRef]
53. Ober, U.; Huang, W.; Magwire, M.; Schlather, M.; Simianer, H.; Mackay, T.F.C. Correction: Accounting for Genetic Architecture Improves Sequence Based Genomic Prediction for a Drosophila Fitness Trait. *PLoS ONE* **2015**, *10*, e0132980. [CrossRef]
54. Edwards, S.M.; Sørensen, I.F.; Sarup, P.; Mackay, T.F.; Sørensen, P. Genomic Prediction for Quantitative Traits Is Improved by Mapping Variants to Gene Ontology Categories in *Drosophila melanogaster*. *Genetics* **2016**, *203*, 1871–1883. [CrossRef]
55. Garlapow, M.E.; Huang, W.; Yarboro, M.T.; Peterson, K.R.; Mackay, T.F.C. Quantitative Genetics of Food Intake in Drosophila melanogaster. *PLoS ONE* **2015**, *10*, e0138129. [CrossRef]
56. Negron, Y.L.S.; Hansen, N.F.; Harbison, S.T. The Sleep Inbred Panel, a Collection of Inbred *Drosophila melanogaster* with Extreme Long and Short Sleep Duration. *G3 Genes|Genomes|Genet.* **2018**, *8*, 2865–2873. [CrossRef] [PubMed]
57. Bevers, R.P.J.; Litovchenko, M.; Kapopoulou, A.; Braman, V.S.; Robinson, M.R.; Auwerx, J.; Hollis, B.; Deplancke, B. Mitochondrial haplotypes affect metabolic phenotypes in the Drosophila Genetic Reference Panel. *Nat. Metab.* **2019**, *1*, 1226–1242. [CrossRef]

58. Harbison, S.T.; Negron, Y.L.S.; Hansen, N.F.; Lobell, A.S. Selection for long and short sleep duration in Drosophila melanogaster reveals the complex genetic network underlying natural variation in sleep. *PLoS Genet.* **2017**, *13*, e1007098. [CrossRef] [PubMed]
59. Unckless, R.L.; Rottschaefer, S.M.; Lazzaro, B.P. A Genome-Wide Association Study for Nutritional Indices in *Drosophila*. *G3 Genes|Genomes|Genet.* **2015**, *5*, 417–425. [CrossRef]
60. Lafuente, E.; Duneau, D.; Beldade, P. Genetic basis of thermal plasticity variation in Drosophila melanogaster body size. *PLoS Genet.* **2018**, *14*, e1007686. [CrossRef] [PubMed]
61. Harbison, S.T.; Kumar, S.; Huang, W.; McCoy, L.J.; Smith, K.R.; Mackay, T.F.C. Genome-Wide Association Study of Circadian Behavior in Drosophila melanogaster. *Behav. Genet.* **2018**, *49*, 60–82. [CrossRef]
62. Chow, C.Y.; Kelsey, K.J.; Wolfner, M.F.; Clark, A.G. Candidate genetic modifiers of retinitis pigmentosa identified by exploiting natural variation in *Drosophila*. *Hum. Mol. Genet.* **2015**, *25*, 651–659. [CrossRef]
63. Lavoy, S.; Chittoor-Vinod, V.G.; Chow, C.Y.; Martin, I. Genetic Modifiers of Neurodegeneration in a *Drosophila* Model of Parkinson's Disease. *Genetics* **2018**, *209*, 1345–1356. [CrossRef]
64. Klein, A.D.; Mazzulli, J.R. Is Parkinson's disease a lysosomal disorder? *Brain* **2018**, *141*, 2255–2262. [CrossRef]
65. Marder, K.; Wang, Y.; Alcalay, R.N.; Mejia-Santana, H.; Tang, M.-X.; Lee, A.; Raymond, D.; Mirelman, A.; Saunders-Pullman, R.; Clark, L.; et al. Age-specific penetrance of *LRRK2* G2019S in the Michael J. Fox Ashkenazi Jewish LRRK2 Consortium. *Neurology* **2015**, *85*, 89–95. [CrossRef] [PubMed]
66. Olivares, G.H.; Olguín, P.; Klein, A.D. Modeling Parkinson's Disease Heterogeneity to Accelerate Precision Medicine. *Trends Mol. Med.* **2019**, *25*, 1052–1055. [CrossRef] [PubMed]
67. Mackay, T.F.; Huang, W. Charting the genotype–phenotype map: Lessons from the *Drosophila melanogaster* Genetic Reference Panel. *Wiley Interdiscip. Rev. Dev. Biol.* **2017**, *7*, e289. [CrossRef] [PubMed]
68. Swarup, S.; Huang, W.; Mackay, T.F.C.; Anholt, R.R.H. Analysis of natural variation reveals neurogenetic networks for *Drosophila* olfactory behavior. *Proc. Natl. Acad. Sci. USA* **2012**, *110*, 1017–1022. [CrossRef] [PubMed]
69. Shorter, J.; Couch, C.; Huang, W.; Carbone, M.A.; Peiffer, J.; Anholt, R.R.H.; Mackay, T.F.C. Genetic architecture of natural variation in *Drosophila melanogaster* aggressive behavior. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3555–E3563. [CrossRef] [PubMed]
70. Long, A.D.; Macdonald, S.J.; King, E.G. Dissecting complex traits using the Drosophila Synthetic Population Resource. *Trends Genet.* **2014**, *30*, 488–495. [CrossRef]
71. Duina, A.A.; Miller, M.E.; Keeney, J.B. Budding Yeast for Budding Geneticists: A Primer on the *Saccharomyces cerevisiae* Model System. *Genetics* **2014**, *197*, 33–48. [CrossRef]
72. Walberg, M.W. Applicability of Yeast Genetics to Neurologic Disease. *Arch. Neurol.* **2000**, *57*, 1129–1134. [CrossRef]
73. Mancera, E.; Bourgon, R.; Brozzi, A.; Huber, W.; Steinmetz, L.M. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature* **2008**, *454*, 479–485. [CrossRef] [PubMed]
74. Sanchez, A.; Borde, V. Methods to Map Meiotic Recombination Proteins in *Saccharomyces cerevisiae*. In *Homologous Recombination*; Humana: New York, NY, USA, 2020; pp. 295–306. [CrossRef]
75. Botstein, D.; Fink, G.R. Yeast: An Experimental Organism for Modern Biology. *Science* **1988**, *240*, 1439–1443. [CrossRef] [PubMed]
76. Giaever, G.; Nislow, C. The Yeast Deletion Collection: A Decade of Functional Genomics. *Genetics* **2014**, *197*, 451–465. [CrossRef] [PubMed]
77. Galardini, M.; Busby, B.P.; Vieitez, C.; Dunham, A.S.; Typas, A.; Beltrao, P. The impact of the genetic background on gene deletion phenotypes in *Saccharomyces cerevisiae*. *Mol. Syst. Biol.* **2019**, *15*, e8831. [CrossRef] [PubMed]
78. Parts, L.; Batté, A.; Lopes, M.; Yuen, M.W.; Laver, M.; Luis, B.S.; Yue, J.; Pons, C.; Eray, E.; Aloy, P.; et al. Natural variants suppress mutations in hundreds of essential genes. *Mol. Syst. Biol.* **2021**, *17*, e10138. [CrossRef] [PubMed]
79. Khurana, V.; Lindquist, S. Modelling neurodegeneration in Saccharomyces cerevisiae: Why cook with baker's yeast? *Nat. Rev. Neurosci.* **2010**, *11*, 436–449. [CrossRef]
80. Ehrenreich, I.M.; Magwene, P.M. Genetic Dissection of Heritable Traits in Yeast Using Bulk Segregant Analysis. *Cold Spring Harb. Protoc.* **2017**, *2017*, pdb.prot088989. [CrossRef] [PubMed]
81. Bloom, J.S.; Boocock, J.; Treusch, S.; Sadhu, M.J.; Day, L.; Oates-Barker, H.; Kruglyak, L. Rare variants contribute disproportionately to quantitative trait variation in yeast. *eLife* **2019**, *8*, e49212. [CrossRef]
82. Lander, E.S.; Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **1989**, *121*, 185–199. [CrossRef] [PubMed]
83. Michelmore, R.W.; Paran, I.; Kesseli, R.V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9828–9832. [CrossRef]
84. Liti, G.; Carter, D.M.; Moses, A.M.; Warringer, J.; Parts, L.; James, S.A.; Davey, R.P.; Roberts, I.N.; Burt, A.; Koufopanou, V.; et al. Population genomics of domestic and wild yeasts. *Nature* **2009**, *458*, 337–341. [CrossRef]
85. Swinnen, S.; Thevelein, J.M.; Nevoigt, E. Genetic mapping of quantitative phenotypic traits in Saccharomyces cerevisiae. *FEMS Yeast Res.* **2012**, *12*, 215–227. [CrossRef] [PubMed]
86. Kessi-Pérez, E.I.; Molinet, J.; Martínez, C. Disentangling the genetic bases of Saccharomyces cerevisiae nitrogen consumption and adaptation to low nitrogen environments in wine fermentation. *Biol. Res.* **2020**, *53*, 2. [CrossRef] [PubMed]
87. Haas, R.; Horev, G.; Lipkin, E.; Kesten, I.; Portnoy, M.; Buhnik-Rosenblau, K.; Soller, M.; Kashi, Y. Mapping Ethanol Tolerance in Budding Yeast Reveals High Genetic Variation in a Wild Isolate. *Front. Genet.* **2019**, *10*, 998. [CrossRef] [PubMed]

88. Wang, Z.; Qi, Q.; Lin, Y.; Guo, Y.; Liu, Y.; Wang, Q. QTL Analysis Reveals Genomic Variants Linked to High-Temperature Fermentation Performance in the Industrial Yeast. *Biotechnol. Biofuels* **2019**, *12*, 59. [CrossRef]
89. Liti, G.; Louis, E.J. Yeast Evolution and Comparative Genomics. *Annu. Rev. Microbiol.* **2005**, *59*, 135–153. [CrossRef]
90. Warringer, J.; Zörgö, E.; Cubillos, F.A.; Zia, A.; Gjuvsland, A.; Simpson, J.T.; Forsmark, A.; Durbin, R.; Omholt, S.W.; Louis, E.J.; et al. Trait Variation in Yeast Is Defined by Population History. *PLoS Genet.* **2011**, *7*, e1002111. [CrossRef]
91. Perlstein, E.; Ruderfer, D.; Roberts, D.C.; Schreiber, S.L.; Kruglyak, L. Genetic basis of individual differences in the response to small-molecule drugs in yeast. *Nat. Genet.* **2007**, *39*, 496–502. [CrossRef]
92. Cubillos, F.A.; Louis, E.J.; Liti, G. Generation of a large set of genetically tractable haploid and diploid *Saccharomyces strains*. *FEMS Yeast Res.* **2009**, *9*, 1217–1225. [CrossRef]
93. Linder, R.A.; Majumder, A.; Chakraborty, M.; Long, A. Two Synthetic 18-Way Outcrossed Populations of Diploid Budding Yeast with Utility for Complex Trait Dissection. *Genetics* **2020**, *215*, 323–342. [CrossRef]
94. Cubillos, F.; Parts, L.; Salinas, F.; Bergström, A.; Scovacricchi, E.; Zia, A.; Illingworth, C.J.R.; Mustonen, V.; Ibstedt, S.; Warringer, J.; et al. High-Resolution Mapping of Complex Traits with a Four-Parent Advanced Intercross Yeast Population. *Genetics* **2013**, *195*, 1141–1155. [CrossRef] [PubMed]
95. Diouf, I.; Pascual, L. Multiparental Population in Crops: Methods of Development and Dissection of Genetic Traits. In *Crop Breeding*; Tripodi, P., Ed.; Humana: New York, NY, USA, 2020; Volume 2264, pp. 13–32. [CrossRef]
96. Peter, J.; Schacherer, J. Population genomics of yeasts: Towards a comprehensive view across a broad evolutionary scale. *Yeast* **2015**, *33*, 73–81. [CrossRef] [PubMed]
97. Bendixsen, D.P.; Frazão, J.G.; Stelkens, R. *Saccharomyces* yeast hybrids on the rise. *Yeast* **2021**, *39*, 40–54. [CrossRef] [PubMed]
98. Schacherer, J.; Shapiro, J.A.; Ruderfer, D.M.; Kruglyak, L. Comprehensive polymorphism survey elucidates population structure of Saccharomyces cerevisiae. *Nature* **2009**, *458*, 342–345. [CrossRef]
99. Peter, J.; De Chiara, M.; Friedrich, A.; Yue, J.-X.; Pflieger, D.; Bergström, A.; Sigwalt, A.; Barre, B.; Freel, K.; Llored, A.; et al. Genome evolution across 1,011 Saccharomyces cerevisiae isolates. *Nature* **2018**, *556*, 339–344. [CrossRef]
100. Khurana, V.; Peng, J.; Chung, C.Y.; Auluck, P.K.; Fanning, S.; Tardiff, D.F.; Bartels, T.; Koeva, M.; Eichhorn, S.W.; Benyamini, H.; et al. Genome-Scale Networks Link Neurodegenerative Disease Genes to alpha-Synuclein through Specific Molecular Pathways. *Cell Syst.* **2017**, *4*, 157–170. [CrossRef]
101. Jung, P.P.; Zhang, Z.; Paczia, N.; Jaeger, C.; Ignac, T.; May, P.; Linster, C.L. Natural variation of chronological aging in the Saccharomyces cerevisiae species reveals diet-dependent mechanisms of life span control. *npj Aging Mech. Dis.* **2018**, *4*, 3. [CrossRef]
102. Breschi, A.; Gingeras, T.R.; Guigó, A.B.R. Comparative transcriptomics in human and mouse. *Nat. Rev. Genet.* **2017**, *18*, 425–440. [CrossRef]
103. Adams, M.D.; Celniker, S.E.; Holt, R.A.; Evans, C.A.; Gocayne, J.D.; Amanatides, P.G.; Scherer, S.E.; Li, P.W.; Hoskins, R.A.; Galle, R.F.; et al. The Genome Sequence of *Drosophila melanogaster*. *Science* **2000**, *287*, 2185–2195. [CrossRef]
104. Belda, I.; Ruiz, J.; Santos, A.; Van Wyk, N.; Pretorius, I.S. Saccharomyces cerevisiae. *Trends Genet.* **2019**, *35*, 956–957. [CrossRef]
105. Zhu, F.; Nair, R.R.; Fisher, E.M.C.; Cunningham, T.J. Humanising the mouse genome piece by piece. *Nat. Commun.* **2019**, *10*, 1845. [CrossRef] [PubMed]
106. Ugur, B.; Chen, K.; Bellen, H.J. Drosophila tools and assays for the study of human diseases. *Dis. Models Mech.* **2016**, *9*, 235–244. [CrossRef] [PubMed]
107. Smith, M.G.; Snyder, M. Yeast as a Model for Human Disease. *Curr. Protoc. Hum. Genet.* **2006**, *48*, 15–21. [CrossRef] [PubMed]

Biochemistry and Biophysics Reports 28 (2021) 101105

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Identification of genetic modifiers of murine hepatic β-glucocerebrosidase activity

Anyelo Durán [a,1], Boris Rebolledo-Jaramillo [a,1], Valeria Olguin [a], Marcelo Rojas-Herrera [a], Macarena Las Heras [a], Juan F. Calderón [a], Silvana Zanlungo [b], David A. Priestman [c], Frances M. Platt [c], Andrés D. Klein [a,*]

[a] Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana, Universidad del Desarrollo, Santiago, Chile
[b] Department of Gastroenterology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile
[c] Department of Pharmacology, University of Oxford, Oxford, OX1 3QT, UK

ARTICLE INFO

Keywords:
Systems genetics
Modifier genes
β-glucocerebrosidase
Inbred strains
Gaucher disease
Parkinson's disease

ABSTRACT

The acid β-glucocerebrosidase (GCase) enzyme cleaves glucosylceramide into glucose and ceramide. Loss of function variants in the gene encoding for GCase can lead to Gaucher disease and Parkinson's disease. Therapeutic strategies aimed at increasing GCase activity by targeting a modulating factor are attractive and poorly explored. To identify genetic modifiers, we measured hepatic GCase activity in 27 inbred mouse strains. A genome-wide association study (GWAS) using GCase activity as a trait identified several candidate modifier genes, including *Dmrtc2* and *Arhgef1* ($p=2.1x10^{-7}$), and *Grik5* ($p=2.1x10^{-7}$). Bayesian integration of the gene mapping with transcriptomics was used to build integrative networks. The analysis uncovered additional candidate GCase regulators, highlighting modules of the acute phase response ($p=1.01x10^{-8}$), acute inflammatory response ($p=1.01x10^{-8}$), fatty acid beta-oxidation ($p=7.43x10^{-5}$), among others. Our study revealed previously unknown candidate modulators of GCase activity, which may facilitate the design of therapies for diseases with GCase dysfunction.

1. Introduction

Hydrolytic enzymes are abundant in the lysosome; more than 60 acidic hydrolases have been described to date [1]. In addition to its digestive and recycling functions, the lysosome orchestrates metabolic adaptations to external cues [2]. The acidic lysosomal β-glucocere- brosidase (GCase) enzyme degrades glucosylceramide into glucose and ceramide [3]. GCase is encoded by the *GBA1* gene. Loss-of-function variants in this gene cause the rare lysosomal storage disorder Gaucher disease (GD) [4]. *GBA1* variants also significantly increase the risk of developing Parkinsonism and Parkinson's disease (PD) [5]. In addition to other mechanistic data, this observation highlights the role of lysosomal dysfunction as a risk factor for PD [6]. Therefore, an exogenous increase of GCase activity and other related enzymes is an attractive therapeutic strategy that has not yet been thoroughly explored.

Although treatments for Gaucher disease are available, they have clinical limitations [7]. Studying how GCase is modulated can allow us to i) learn about its regulation and possibly ii) develop new targeted therapies to treat these diseases which diseases be more specific. One attractive approach to identify new therapeutic targets to modulate GCase is to use the natural genetic diversity present in populations of individuals (i.e., model organisms) to identify genetic modifiers that control a given trait [8]. By integrating gene mapping with other sets of -omics, it is possible to find regulatory elements underlying the variation in a given trait [9]. This holistic population-based approach is called systems genetics [10].

Inbred mouse strains arise from crossing siblings for at least 20 generations [11]. The Hybrid Mouse Diversity Panel (HMDP) [8] corresponds to a panel of genotyped inbred strains where other -omics data are also available [8,12]. The HMDP panel has been used to perform association studies and find modifier genes for a variety of complex traits [13,14].

Here, we used a systems genetics approach to identify putative

* Corresponding author. Av. Las Condes 12.461, Postal Zip code 7590943, Santiago, Chile.
E-mail address: andresklein@udd.cl (A.D. Klein).

[1] Equal contribution.

https://doi.org/10.1016/j.bbrep.2021.101105
Received 4 June 2021; Received in revised form 11 August 2021; Accepted 12 August 2021
Available online 18 August 2021

2405-5808/© 2021 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

modifier genes/networks of GCase activity in mice. To this end, we measured hepatic GCase activity in 27 strains of mice. A genome-wide association (GWAS) analysis identified putative modifier genes. We used Mergeomics analysis to integrate GCase activity, gene mapping, and available liver transcriptomic data. Our study revealed genes, networks, and biological processes that might regulate GCase function.

2. Materials and methods

2.1. Mouse tissues

We used 8 weeks-old mice livers derived from 27 inbred mouse strains which were kindly donated by Dr. Aldons Lusis (University of California, Los Angeles). (i) 129X1/SvJ (n=5), (ii) A/J (n=5), (iii) AKR/ J (n=5), (iv) BALB/cJ (n=5), (v) BTBR T<+> tf/J (n=5), (vi) BUB/BnJ (n=3), (vii) C3H/HeJ (n=3), (viii) C57BL/6J (n=5), (ix) C58/J (n=5), (x) CAST/EiJ (n=3), (xi) CBA/J (n=5), (xii) CE/J (n=5), (xiii) DBA/2J (n=5), (xiv) FVB/NJ (n=3), (xv) KK/HlJ (n=3), (xvi) LG/J (n=4), (xvii) LP/J (n=3), (xviii) MA/MyJ (n=3), (xix) NOD/ShiLtJ (n=5), (xx) NON/ShiLtJ (n=5), (xxi) NZB/BlNJ (n=5), (xxii) NZW/LacJ (n=5), xxiii) PL/J (n=5), xxiv) RIIIS/J (n=3), xxv) SEA/GnJ (n=5), xxvi) SM/J (n=5), xxvii) SWR/J (n=5). Tissues were homogenized and adjusted to 50 mg tissue/ml in H_2O, with a Potter-Elvehjem tissue homogenizer (Omni International, USA) and then stored at −80 °C until further use. This liver was selected because of its high relevance in the generation of pathophysiological phenotypes in GD [4].

2.2. GCase activity assays

GCase activity was determined using an artificial fluorescent substrate based on 4-methylumbelliferone (4-MU) [15]. For this purpose, liver homogenates were diluted 1/10 with GCase buffer (200 mM citrate-phosphate buffer, pH 5.2, containing 0.25% Triton X-100, 1.25 mM EDTA, 4 mM 2-mercaptoethanol, all these reagents from Calbiochem, Merck KGaA, Darmstadt, Germany). Three cycles of freezing and thawing with liquid nitrogen were performed to disrupt the cell membranes. Subsequently, 10 μl of the diluted homogenates (5 μg of total protein from each sample) were mixed with and without 10 μl of 0.3 mM N-butyldeoxydeoxynojirimycin (NB-DGJ) for 30 min on ice. *N*B-DGJ is a β-glucocerebrosidase 2 (GBA2; non-lysosomal enzyme) inhibitor that does not inhibit GCase [16] (Toronto Research Chemicals, North York, Ontario, Canada). Thereafter, the tubes were placed in a 37 °C water bath, and 40 μl of the substrate 4-methylumbelliferyl-β-D-glucopyranoside (4.5 mM 4-MU-β-D-Glc in GCase buffer) (Sigma, Dorset, England) was added. The reaction was stopped after 30 min of incubation by the addition of 400 μl cold 0.5 M Na_2CO_3 at pH 10.7 (Panreac Applichem, Barcelona, Spain). Fluorescence was measured at 340 nm excitation and 460 nm emission with a gain of 40 in a semi-automated plate reader (Synergy HT, BioTek, Winooski, USA). Fluorescence values were normalized to protein content in each sample as obtained by a BSA assay (Pierce BCA Protein Assay Kit; Thermo Fisher Scientific, Illinois, USA). To calculate GCase specific activity, a 4-MU standard curve was constructed, and the final value was adjusted to 1 h of enzymatic reaction. For each biological sample, at least three technical replicates were performed.

2.3. Phylogenetic tree and GWAS using an efficient mixed model association (EMMA)

We used the average GCase activity per mouse strain as a phenotype to perform the GWAS using an Efficient Mixed Model Association Study (EMMA v.1.1.2) [17]. In addition, we included in the analysis the mouse HapMap reference panel, consisting of 4 million SNVs downloaded from http://mouse.cs.ucla.edu/mousehapmap/full.html [18]. The R package for EMMA was downloaded from http://mouse.cs.ucla.edu/emma/ [19]. P-value was recorded as the strength of the genotype-phenotype associations. To build to phylogenetic tree we used the EMMA uses a kinship matrix to run hierarchical clustering using R [19].

2.4. Gene expression array

The hepatic transcript levels of inbred mouse strains were down- loaded from the repository GSE16780 UCLA Mouse MDP Liver Affy HT M430A MDP Liver [20]. If there was more than one probe quantifying the same gene, the values were averaged.

2.5. Functional impact of gene variants

We downloaded the genomes of three strains with low and high GCase activity respectively (CBA/J, A/J, FVB/NJ, CAST/EiJ, BUB/BnJ, and C58/J) from the Mouse Phenome Database (MPD) (RRID: SCR_003212) of Jackson Laboratory (https://phenome.jax.org/) and Mouse Genomes Project of Sanger Institute [21]. The Impact of variants was assessed using SIFT (sorting intolerant from tolerant) [22] and SnpEff [23].

2.6. Integrative networks of genomic and transcriptomic data

To study how genomic and liver transcriptomic variation contributes to hepatic GCase activity variability in the 27 HMDP strains, we employed Mergeomics v1.18 [24]. To build Bayesian networks of inte- grative omics underlying GCase activity, two modules of Mergeomics are required: a) marker set enrichment analysis (MSEA) and b) weighted key driver analysis (wKDA). MSEA requires the following data inputs: 1) EMMA GWAS results: i) marker-GCase activity association (marker-- value) and ii) gene-marker mapping file (gene-marker); 2) functionally related gene sets (module-gene), which are preloaded in Mergeomics. These results are integrated through the package algorithm to find sets of genes associated with GCase activity. The parameter settings of the MSEA module included: i) type of permutation at the gene level. ii) minimum (10) and maximum (500) number of genes in the sets. iii) the minimum and maximum overlap ratio between sets of genes associated with disease/trait = 0.33 (33% overlap). iv) the number of gene or marker permutations = 2000 and finally v) the MSEA FDR cutoff was ≤25% [25], this analysis calculates the Benjamini-Hochberg FDR [26].

To identify key driver (KD) genes, which are defined as the gene hubs most significantly associated with other genes in the network, we used wKDA [24]. The wKDA module takes input data from the MSEA results generated in the previous step and a defined liver tissue Bayesian network corresponding to human and rodent expression datasets of earlier studies [27]. The parameters for running wKDA included i) Search depth of wKDA = 1, which means that we search for key-drivers whose immediate neighborhood is enriched for MSEA significant genes, ii) the edge type of wKDA = incoming and outgoing directionality, iii) the minimum overlap, is the threshold above which hubs will be designated as co-hubs, of wKDA = 0.33, and iv) the edge factor of wKDA = 0.5, which means an unweight network. This module projected sets of genes associated with liver GCase activity onto a Bayesian liver network, representing seemingly causal relationships between genes and KD genes [27]. We ran both Mergeomics modules in the R package [19].

2.7. Gene ontology enrichment

ShinyGO v0.61 [28], for gene ontology (GO) enrichment analysis for network modules, was employed. This tool has annotations for model organisms. The chromosomal gene location, metabolic pathways, gene clustering, and protein interaction networks can be plotted [28].

2.8. Statistics

Prism v9.1.0 (GraphPad software, San Diego, CA) and the R package [19] was used for statistical analysis and included a two-tailed Student's

t-test and ANOVA with Bonferroni test. Pearson for correlation analyses was employed. The significant value was considered as $p < 0.05$.

3. Results

3.1. Liver GCase activity varies among mouse strains

We measured hepatic GCase activity by fluorimetry in the liver of 27 inbred mice strains with different phylogenetic origins (Fig. 1A) using an artificial substrate 4-MU-β-D-Glc. We observe significant variability in the average enzymatic activity between the different strains (Fig. 1B). This activity was higher in the BUB/BnJ and C58/J strains ($p < 0.001$; mean ± 95%confidence interval; 39.3 ± 25.0–53.5 and 40.1 ± 29.57–50.6 respectively) compared to MA/MyJ and CBA/J (20.7 ± 10.3–31.2 and 23.8 ± 16.8–30.7 respectively).

3.2. GWAS identified putative modifier genes of GCase activity

To uncover possible modifier genes, we performed a GWAS analysis with EMMA. EMMA is a statistical test that corrects the strains' population structure and genetic relatedness. EMMA applies a mixed models for association mapping, allowing to substantially increase the compu- tational speed and reliability of the results by reducing false positives associations [17]. We found 271 significant Single Nucleotide Variants (SNVs) $p \leq 4.1 \times 10^{-6}$ that exceeded the threshold of suggestive associations previously established [17], represented in 9 non-redundant genes (Table S1 and Fig. 2A). Among all the variants identified, we found an exonic variant in the *Myo6* gene, one in the 3′UTR region of *Dmrtc2* and A*rhgef1*, among others (Table S1). Then, we organized the strains by GCase activity and plotted the genotypes of the top associated markers. The strains with low and intermediate activities present a different distribution pattern than those with high GCase activity (Fig. 2B). Strains AKR/J and PL/J showed a different distribution pattern. Perhaps in these two strains, spontaneous mutations arose in some of their generations after genotyping, as described for C3H/HeN, BL6, BALBc, and FVB [29]. A suggestive association threshold (blue line) (Fig. 2A) $p \leq 4.1 \times 10^{-6}$ was calculated in Ref. [17] and ii) a Bonfer- roni correction, which resulted in a much stricter p-value of 1.28×10^{-8}, denoted by the red line in Fig. 2A. Interestingly we observed variants at different genomic locations with a genotype-phenotype association that exceeds the empirical p-value (Fig. 2A &B).

3.3. Associations between hepatic transcripts levels and GCase enzyme activity

The significant SNVs in a GWAS can be regulators of gene expression levels. Thus, we performed correlation studies between GCase activity and liver transcripts levels. To this end, we downloaded the hepatic gene expression array data from an online repository (GSE16780 UCLA Mouse MDP Liver Affy HT M430A) [20]. The array included the probes for seven of the nine identified genes. Associations were explored (Fig. S1). No significant correlations were found. The array data that we used did not include probes for *Spag16* and *Dmrtc2* genes. Therefore, it was not possible to test correlations with these two genes. For *Erbb4* and *Zic4* we observed a trend ($p=0.07$). The signals in a Manhattan plot can be labeling coding or other no coding genomic variability. To explore this possibility, we downloaded the genomes of three low (CBA/J, A/J, and FVB/NJ) and three high (CAST/EiJ, BUB/BnJ, and C58/J) GCase activity strains from the Mouse Genomes Project of Sanger Institute [21]. We identified predicted splice and or miss sense variants in *Grik5*, *Impg1*, *Myo6*, and *Spag16* (Fig. 2C). To assess the implications of miss sense variants we used SIFT (sorting intolerant from tolerant) [22] and SnpEff [23].

3.4. Identification of modules, key drivers, and pathways associated with GCase activity

EMMA/GWAS results were used to identify modules and key driver genes within the coexpression network re-constructed for mouse/human liver, using Mergeomics v1.18. We used 20 pre-defined mouse gene sets [24] and FDR <25%. The MSEA module of Mergeomics highlighted modules which correspond to mouse liver expression data converted to human gene symbols (Table S2). The wKDA identified 4 top key drivers (*Itih4, Hsd3b5, Ocel1, Pigt*) and 18 total network hubs (Fig. 3A, Table S2). We included the significant correlations between GCase activity, and the transcripts identified in the network (Fig. 3B), and of the driver genes (Fig. S2). Gene sets obtained from the MSEA analysis were used to perform a GO term enrichment analysis. Significantly enriched path- ways included acute-phase response ($p=1.01 \times 10^{-8}$), acute inflamma- tory response ($p=1.01 \times 10^{-8}$), fatty acid beta-oxidation ($p=7.43 \times 10^{-5}$), fatty acid catabolic process ($p=8.99 \times 10^{-4}$), cellular lipid catabolic process ($p=3 \times 10^{-3}$), ion transport ($p=3 \times 10^{-2}$), cell surface receptor signaling pathway ($p=4 \times 10^{-2}$), among others associated with biological processes (Fig. 4A). Cellular component analysis highlighted blood

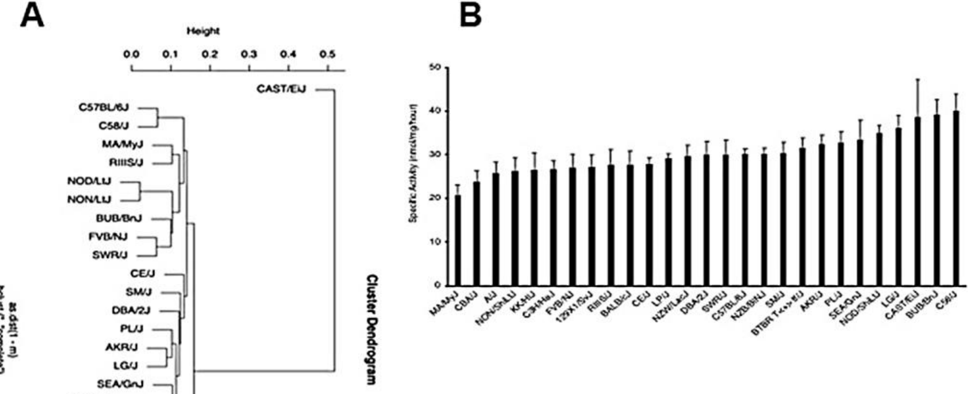
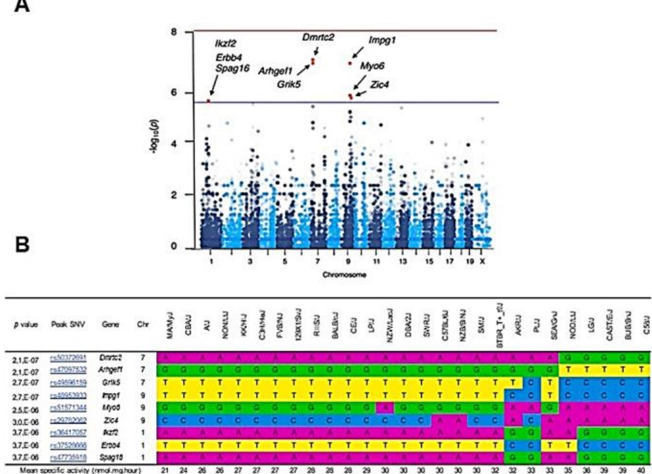
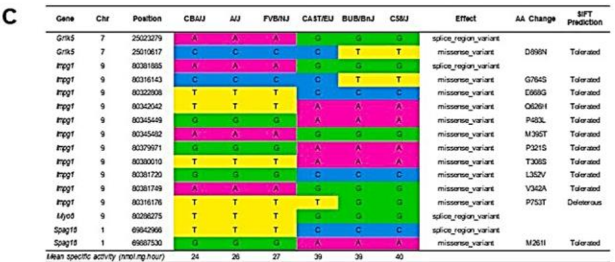


Fig. 1. Variation in the hepatic GCase activity among inbred mouse strains.
(A) Hierarchical clustering of the genetic distance among the 27 used strains, based on EMMA's kinship calculation
(B) Levels of GCase activity in the liver of 27 mouse inbred strains. Values are presented as mean ± standard error (n=5 biological sample with three technical replicates). ANOVA analysis revealed significant differences among the groups (p=0.0028).





Gene	Chr	Position	CBA/J	A/J	FVB/NJ	CAST/EiJ	BUB/BnJ	C58/J	Effect	AA Change	SIFT Prediction
Grik5	7	25023279	A	A	A	G	G	G	splice_region_variant		
Grik5	7	25010617	C	C	C	C	T	T	missense_variant	D898N	Tolerated
Impg1	9	80381885	A	A	A	G	G	G	splice_region_variant		
Impg1	9	80316143	C	C	C	C	T	T	missense_variant	G764S	Tolerated
Impg1	9	80322808	T	T	T	C	C	C	missense_variant	E668G	Tolerated
Impg1	9	80342042	T	T	T	A	A	A	missense_variant	Q526H	Tolerated
Impg1	9	80345449	G	G	G	A	A	A	missense_variant	P483L	Tolerated
Impg1	9	80345482	A	A	A	G	G	G	missense_variant	M395T	Tolerated
Impg1	9	80379971	G	G	G	A	A	A	missense_variant	P321S	Tolerated
Impg1	9	80380010	T	T	T	A	A	A	missense_variant	T308S	Tolerated
Impg1	9	80381720	G	G	G	C	C	C	missense_variant	L352V	Tolerated
Impg1	9	80381749	A	A	A	G	G	G	missense_variant	V342A	Tolerated
Impg1	9	80316176	T	T	T	T	G	G	missense_variant	P753T	Deleterous
Myo6	9	80298275	T	T	T	G	G	G	splice_region_variant		
Spag16	1	69842966	T	T	T	C	C	C	splice_region_variant		
Spag16	1	69887530	G	G	G	A	A	A	missense_variant	M261I	Tolerated
Mean specific activity (nmol.mg.hour)			24	26	27	39	39	40			

Fig. 2. GWAS identifies putative modifiers of hepatic GCase activity.

(A) Manhattan plot highlights the top associated GCaase activity genes plotted as chromosome position versus the inverse of the negative logarithm of the association p-value.

(B) Strains were organized according to GCase enzymatic activity, from lowest to highest, and the genotype of the peak associated SNV and genomic regions are shown. Adenine (A), purple; cytosine (C), blue; guanine (G), green; thymine (T), yellow; Chr, chromosome.

(C) Miss sense and coding variants in the top associated genes in three strains with low GCase activity and three high activity levels. Same color code than in (B) was used. The effect of the variant, Amino Acid (AA) change, and SIFT prediction were plotted. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

microparticle ($p=6x10^{-2}$), early endosome ($p=4x10^{-1}$), protein-lipid complex ($p=4x10^{-1}$), organelle lumen ($p=4x10^{-1}$) and others (Fig. 4B). Molecular function revealed fatty-acyl-CoA binding ($p=1x10^{-2}$), anion-sodium symporter activity ($p=1x10^{-2}$), among others (Fig. 4C).

4. Discussion

Our goal was to identify putative modifier genes/networks of hepatic GCase activity using a system genetics strategy. Identifying modulable genetic modifiers of GCase activity offers a feasible and attractive therapeutic alternative for diseases with lysosomal dysfunction, bringing us closer to a precision medicine-based approach.

Our study associated 271 SNVs (Table S1) within nine genes (*Dmrtc2, Arhgef1, Grik5, Impg1, Myo6, Zic4, Ikzf2, Erbb4, Spag16*) to liver GCase activity (Fig. 2A & B), and four key drivers (*Itih4, Hsd3b5, Ocel1, Pigt*) (Fig. 3A-D, Table S3). We found no literature linking these newly associated genes to the GCase enzyme directly. However, the identified genes are associated with several human diseases: *Grik5* to bipolar disorder [30]; *Erbb4* to schizophrenia [31]; and melanoma [32]; *Myo6* to deafness [33]; *Arhgef1* to primary Immunodeficiencies [34]; *Impg1* to

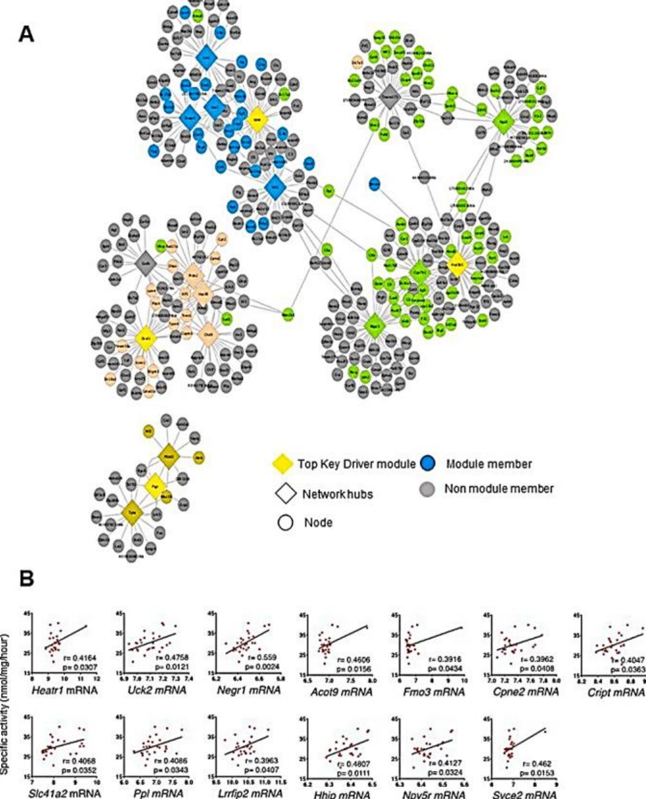


Fig. 3. Integrative network of GCase activity. The gene hubs of the network are represented with diamonds and Key Driver (KD) genes by yellow diamonds. Gene modules are indicated with a different color. Non-member genes are expressed in grey. Red edges show the multiple interactions of KD with other genes.

(A) Itih4; Inter-Alpha-Trypsin Inhibitor Heavy Chain 4.

(B) Hsd3b5; hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 5.

(C) Ocel1; Occludin/ELL Domain Containing 1.

(D) Pigt; Phosphatidylinositol Glycan Anchor Biosynthesis Class T. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

vitelliform macular dystrophies [35] and *Zic4* to Dandy-Walker malformation [36]. A connection between these disorders and GCase ac- tivity should be therefore be explored.

In addition, the integrative networks identified additional putative regulators of GCase activity. The KD genes have been linked to different functions: *Itih4* to inflammatory responses [37] and liver development and regeneration [38]. *Hsd3b5* to steroid hormones biosynthesis [39], *Ocel1* to cancer prognosis [40] and *Pigt* to glycosylphosphatidylinositol transfer (GPI) proteins [41]. Alterations in these functions have been reported in GD patients, such as i) lymphoid neoplasms [42]; ii) gam- mopathies [43]; iii) predisposition to infections [44]; iv) immune system dysregulation [45]. Widespread inflammation has been studied in depth in GD. Foamy GD macrophages, known as Gaucher cells, release in- flammatory molecules including IL1β, TNF-α, MCP-1 and IL-6, [46-48]. Our results support a role for GCase in the immune response and or this inflammatory pathway(s) can regulate GCase activity.

Our study has some limitations, i) we used public data from a liver expression array of HMDP strains instead of RNAseq. The array has limited probes to capture the transcriptomic landscape of the tissue. Thus, we cannot infer the eventual roles of other genes, different iso- forms, splice variants, and genes with a low level of expression [49]. ii) we used liver homogenates. Several cell types make up the liver, and each cell subgroup could have specific contributions to the variability of

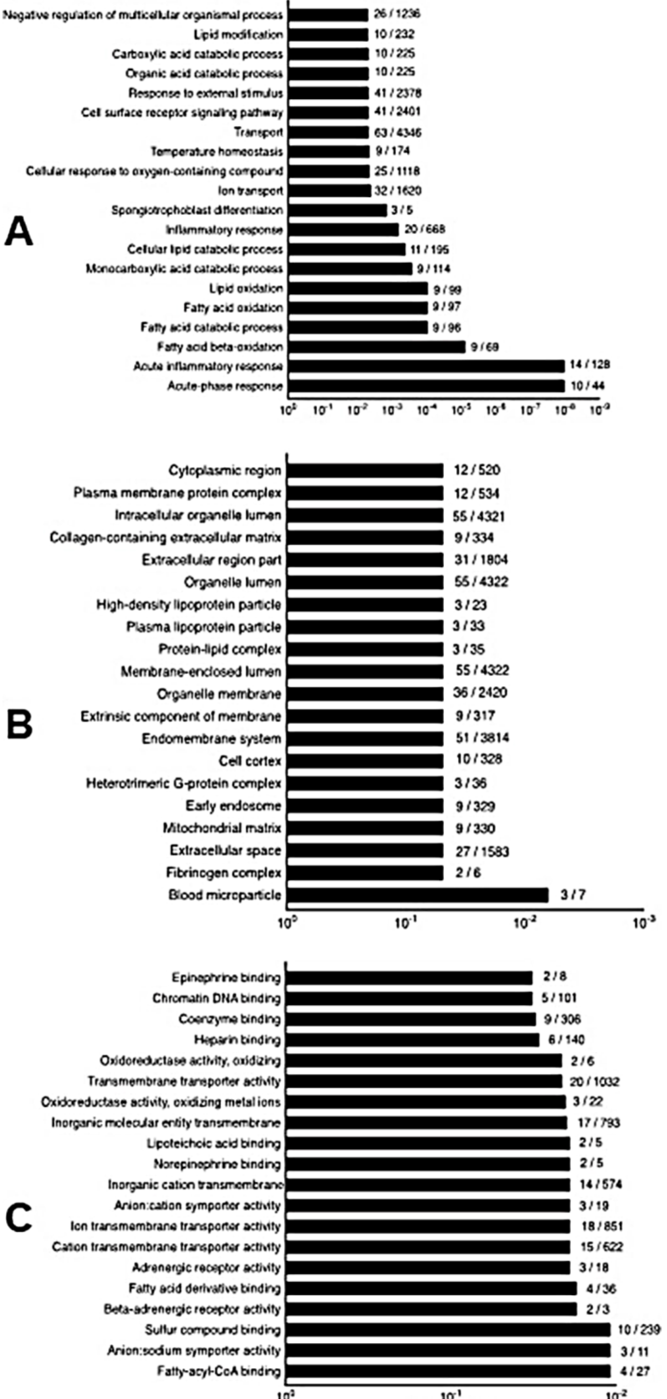


Fig. 4. Identification of enriched GO terms associated with the modules of GCase activity. (A) Biological process. **(B)** Cellular component. **(C)** Molecular function.

the phenotype studied [49–52].

In conclusion, our study has revealed candidate modulators of GCase activity. Further functional analyses are required to understand how the identified genes regulate GCase activity in hepatic cells. The newly identified targets might be relevant for designing therapies for patients with GCase dysfunction, such as Gaucher and Parkinson's disease.

Funding

This work was supported by ANID-CHILE: Fondecyt grant No 1180337 (ADK) (2018–2022) and 1190334 (2019–2023) (SZ), Lyso-Mod, funded by the European Union's Horizon 2020 research and innovation programme (RISE) under the Marie Sklodowska-Curie grant agreement No 734825 (SZ, FMP, ADK) (2017–2022), and the Mizutani Foundation for Glycoscience, grant No: 200133 (FMP, DAP, ADK) (2020). Funding for computational infrastructure was provided by FONDEQUIP EQM150093. FP is a Wolson Royal Society Merit Award Holder and a Wellcome Trust Investigator in science.

Author contributions

AD, BRJ, and DAP performed the experiments, analyzed the data, wrote the paper. VO, and JFC reviewed and edited the document and provided helpful discussions. ADK, SZ, FMP conceptualization, analyzed the data, revised the manuscript, funding acquisition.

Declaration of competing interest

None.

Acknowledgments

The authors acknowledge Dr. Aldons Lusis (University of California, Los Angeles) for kindly donating mouse livers.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2021.101105.

References

[1] T. Lübke, P. Lobel, D.E. Sleat, Proteomics of the lysosome, Biochim. Biophys. Acta Mol. Cell Res. 1793 (2009) 625–635, https://doi.org/10.1016/j.bbamcr.2008.09.018.

[2] A. Ballabio, J.S. Bonifacino, Lysosomes as dynamic regulators of cell and organismal homeostasis, Nat. Rev. Mol. Cell Biol. 21 (2020) 101–118, https://doi.org/10.1038/s41580-019-0185-4.

[3] R.O. Brady, J.N. Kanfer, R.M. Bradley, et al., Demonstration of a deficiency of glucocerebroside-cleaving enzyme in Gaucher's disease, J. Clin. Invest. 45 (1966), https://doi.org/10.1172/JCI105417.

[4] J.F. Desforges, E. Beutler, Gaucher's disease, N. Engl. J. Med. 325 (1991), https://doi.org/10.1056/NEJM199111073251906.

[5] E. Sidransky, M.A. Nalls, J.O. Aasly, et al., Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease, N. Engl. J. Med. (2009) 361, https://doi.org/10.1056/NEJMoa0901281.

[6] A.D. Klein, J.R. Mazzulli, Is Parkinson's disease a lysosomal disorder? Brain 141 (2018) 2255–2262, https://doi.org/10.1093/brain/awy147.

[7] R. Giugliani, F. Vairo, F. Kubaski, et al., Neurological manifestations of lysosomal disorders and emerging therapies targeting the CNS, Lancet Child Adolescent Health 2 (2018) 56–68, https://doi.org/10.1016/S2352-4642(17)30087-1.

[8] A. Ghazalpour, C.D. Rau, C.R. Farber, et al., Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits, Mamm. Genome 23 (2012) 680–692, https://doi.org/10.1007/s00335-012-9411-5.

[9] A.D. Klein, Modeling diseases in multiple mouse strains for precision medicine studies, Physiol. Genom. 49 (2017) 177–179, https://doi.org/10.1152/physiolgenomics.00123.

[10] M. Civelek, A.J. Lusis, Systems genetics approaches to understand complex traits, Nat. Rev. Genet. 15 (2014) 34–48, https://doi.org/10.1038/nrg3575.

[11] J. Flint, E. Eskin, Genome-wide association studies in mice, Nat. Rev. Genet. 13 (2012) 807–817, https://doi.org/10.1038/nrg3335.

[12] A.J. Lusis, M. Seldin, H. Allayee, et al., The Hybrid Mouse Diversity Panel: A Resource for Systems Genetics Analyses of Metabolic and Cardiovascular Traits, 2016. www.jlr.org.

[13] A.D. Klein, N.S. Ferreira, S. Ben-Dor, . J, et al., Identification of modifier genes in a mouse model of gaucher disease, Cell Rep. 16 (2016) 2546–2553, https://doi.org/10.1016/j.celrep.2016.07.085.

[14] R.J. Ferland, J. Smith, D. Papandrea, et al., Multidimensional genetic analysis of repeated seizures in the hybrid mouse diversity panel reveals a novel epileptogenesis susceptibility locus, G3: genes, Genomes, Genetics, 7, https://doi.org/10.1534/g3.117.042234, 2017, 2545-2558.

[15] D. Robinson, The fluorimetric determination of β-glucosidase: its occurrence in the tissues of animals, including insects, Biochem. J. 63 (1956), https://doi.org/10.1042/bj0630039.

[16] S. van Weely, M. Brandsma, A. Strijland, J.M. Tager, J.M.F.G. Aerts, Demonstration of the existence of a second, non-lysosomal glucocerebrosidase that is not deficient in Gaucher disease, Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. (1993) 1181, https://doi.org/10.1016/0925-4439(93)90090-N.

[17] M.K. Hyun, N.A. Zaitlen, C.M. Wade, A, et al., Efficient control of population structure in model organism association mapping, Genetics 178 (2008) 1709–1723, https://doi.org/10.1534/genetics.107.080101.

[18] A. Kirby, H.M. Kang, C.M. Wade, C, et al., Fine mapping in 94 inbred mouse strains using a high-density haplotype resource, Genetics 185 (2010) 1081–1095, https://doi.org/10.1534/genetics.110.115014.

[19] R. R Core Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2020, 1–1.

[20] B.J. Bennett, C.R. Farber, L. Orozco, et al., A high-resolution association mapping panel for the dissection of complex traits in mice, Genome Res. 20 (2010) 281–290, https://doi.org/10.1101/gr.099234.109.

[21] D.J. Adams, A.G. Doran, Lilue, et al., The Mouse Genomes Project: a repository of inbred laboratory mouse strain genomes, Mamm. Genome 26 (9–10) (2015) 403–412, https://doi.org/10.1007/s00335-015-9579-6.

[22] R. Vaser, S. Adusumalli, S.N. Leng, M. Sikic, et al., SIFT missense predictions for genomes, Nat. Protoc. 11 (1) (2015) 1–9, https://doi.org/10.1038/nprot.2015.123.

[23] P. Cingolani, A. Platts, L.L. Wang, Coon, et al., A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly 6 (2) (2012) 80–92, https://doi.org/10.4161/fly.19695.

[24] D. Arneson, A. Bhattacharya, L. Shu, et al., Mergeomics: a web server for identifying pathological pathways, networks, and key regulators via multidimensional data integration, BMC Genom. 17 (2016), https://doi.org/10.1186/s12864-016-3057-8.

[25] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008), https://doi.org/10.1186/1471-2105-9-559.

[26] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. Roy. Stat. Soc. B 57 (1995) 289–300.

[27] S.S. Wang, E.E. Schadt, H. Wang, et al., Identification of pathways for atherosclerosis in mice: integration of quantitative trait locus analysis and global gene expression data, Circ. Res. 101 (2007), https://doi.org/10.1161/CIRCRESAHA.107.152975.

[28] S.X. Ge, D. Jung, R. Yao, ShinyGO: a graphical gene-set enrichment tool for animals and plants, Bioinformatics 36 (2020), https://doi.org/10.1093/bioinformatics/btz931.

[29] J. Chebib, B.C. Jackson, E. Lo´pez-Cortegano, et al., Inbred lab mice are not isogenic: genetic variation within inbred strains used to infer the mutation rate per nucleotide site, Heredity 126 (2021) 107–116, https://doi.org/10.1038/s41437-020-00361-1.

[30] M. Beneyto, L.v. Kristiansen, A. Oni-Orisan, R.E. McCullumsmith, J.H. Meador-Woodruff, Abnormal glutamate receptor expression in the medial temporal lobe in schizophrenia and mood disorders, Neuropsychopharmacology 32 (2007) 1888–1902, https://doi.org/10.1038/sj.npp.1301312.

[31] C.G. Hahn, H.Y. Wang, D.S. Cho, et al., Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia, Nat. Med. 12 (2006) 824–828, https://doi.org/10.1038/nm1418.

[32] T.D. Prickett, N.S. Agrawal, X. Wei, et al., Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4, Nat. Genet. 41 (2009), https://doi.org/10.1038/ng.438.

[33] Z.M. Ahmed, R.J. Morell, S. Riazuddin, A, et al., Mutations of MYO6 are associated with recessive deafness, DFNB37, Am. J. Hum. Genet. (2003) 72, https://doi.org/10.1086/375122.

[34] A. Bouafia, S. Lofek, J. Bruneau, L, et al., Loss of ARHGEF1 causes a human primary antibody deficiency, J. Clin. Invest. 129 (2019) 1047–1060, https://doi.org/10.1172/JCI120572.

[35] G. Manes, I. Meunier, A. Avila-Fern´andez, et al., Mutations in IMPG1 cause vitelliform macular dystrophies, Am. J. Hum. Genet. 93 (2013) 571–578, https://doi.org/10.1016/j.ajhg.2013.07.018.

[36] I. Grinberg, H. Northrup, H. Ardinger, C, et al., Heterozygous deletion of the linked genes ZIC1 and ZIC4 is involved in Dandy-Walker malformation, Nat. Genet. 36 (2004) 1053–1055, https://doi.org/10.1038/ng1420.

[37] Y. Ma, R. Li, J. Wang, et al., ITIH4, as an inflammation biomarker, mainly increases in bacterial bloodstream infection, Cytokine 138 (2021), https://doi.org/10.1016/j.cyto.2020.155377.

[38] C.D. Bhanumathy, Y. Tang, S.P.S. Monga, et al., Itih-4, a serine protease inhibitor regulated in interleukin-6-dependent liver formation: role in liver development and regeneration, Dev. Dynam. 223 (2002) 59–69, https://doi.org/10.1002/dvdy.1235.

[39] C. Park, Expression of multiple forms of 3β-hydroxysteroid dehydrogenase in the mouse liver during fetal and postnatal development, Mol. Cell. Endocrinol. 116 (1996), https://doi.org/10.1016/0303-7207(95)03707-1.

[40] M. Deng, Z. Zhang, B. Liu, et al., Low OCEL1 expression is associated with poor prognosis in human non-small cell lung cancer, Canc. Biomarkers 27 (2020) 519–524, https://doi.org/10.3233/CBM-191268.

[41] K. Ohishi, PIG-S and PIG-T, essential for GPI anchor attachment to proteins, form a complex with GAA1 and GPI8, EMBO J. 20 (2001), https://doi.org/10.1093/emboj/20.15.4088.

[42] B.E. Rosenbloom, N.J. Weinreb, A. Zimran, et al., Gaucher disease and cancer incidence: a study from the Gaucher Registry, Blood 105 (2005) 4569–4572, https://doi.org/10.1182/blood-2004-12-4672.

[43] Y. Nguyen, J. Stirnemann, F. Lautredoux, et al., Immunoglobulin abnormalities in Gaucher disease: an analysis of 278 patients included in the French gaucher disease registry, Int. J. Mol. Sci. 21 (2020), https://doi.org/10.3390/ijms21041247.

[44] L. Marodi, R. Kaposzta, J. Toth, A, et al., Impaired microbicidal capacity of mononuclear phagocytes from patients with type I Gaucher disease: partial correction by enzyme replacement therapy, Blood 86 (1995), https://doi.org/10.1182/blood.V86.12.4645.bloodjournal86124645.

[45] M.K. Pandey, T.A. Burrow, R. Rani, et al., Complement drives glucosylceramide accumulation and tissue inflammation in Gaucher disease, Nature 543 (2017) 108–112, https://doi.org/10.1038/nature21368.

[46] D.M. Mosser, J.P. Edwards, Exploring the full spectrum of macrophage activation, Nat. Rev. Immunol. 8 (2008) 958–969, https://doi.org/10.1038/nri2448.

[47] A. Mantovani, A. Sica, S. Sozzani, P, et al., The chemokine system in diverse forms of macrophage activation and polarization, Trends Immunol. 25 (2004) 677–686, https://doi.org/10.1016/j.it.2004.09.015.

[48] P. Uciechowski, W.C.M. Dempke, Interleukin-6: a masterplayer in the cytokine network, Oncology (Switzerland) 98 (2020) 131–137, https://doi.org/10.1159/000505099.

[49] S. Zhao, W.P. Fung-Leung, A. Bittner, et al., Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells, PloS One 9 (2014), https://doi.org/10.1371/journal.pone.0078644.

[50] S.A. MacParland, J.C. Liu, X.Z. Ma, et al., Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations, Nat. Commun. 9 (2018), https://doi.org/10.1038/s41467-018-06318-7.

[51] N. Aizarani, A. Saviano, Sagar, et al., A human liver cell atlas reveals heterogeneity and epithelial progenitors, Nature 572 (2019) 199–204, https://doi.org/10.1038/s41586-019-1373-2.

[52] X. Wang, L. Yang, Y.C. Wang, et al., Comparative analysis of cell lineage differentiation during hepatogenesis in humans and mice at the single-cell transcriptome level, Cell Res. 30 (2020) 1109–1126, https://doi.org/10.1038/s41422-020-0378-6.

International Journal of
Molecular Sciences

Article

A Mouse Systems Genetics Approach Reveals Common and Uncommon Genetic Modifiers of Hepatic Lysosomal Enzyme Activities and Glycosphingolipids

Anyelo Durán [1], David A. Priestman [2], Macarena Las Heras [1], Boris Rebolledo-Jaramillo [1], Valeria Olguín [1], Juan F. Calderón [1,3], Silvana Zanlungo [4], Jaime Gutiérrez [5], Frances M. Platt [2] and Andrés D. Klein [1,*]

[1] Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago 7610658, Chile
[2] Department of Pharmacology, University of Oxford, Oxford OX1 3QT, UK
[3] Research Center for the Development of Novel Therapeutic Alternatives for Alcohol Use Disorders, Santiago 7610658, Chile
[4] Department of Gastroenterology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago 8330033, Chile
[5] Cellular Signaling and Differentiation Laboratory, School of Medical Technology, Health Sciences Faculty, Universidad San Sebastian, Santiago 7510602, Chile
* Correspondence: andresklein@udd.cl

Abstract: Identification of genetic modulators of lysosomal enzyme activities and glycosphingolipids (GSLs) may facilitate the development of therapeutics for diseases in which they participate, including Lysosomal Storage Disorders (LSDs). To this end, we used a systems genetics approach: we measured 11 hepatic lysosomal enzymes and many of their natural substrates (GSLs), followed by modifier gene mapping by GWAS and transcriptomics associations in a panel of inbred strains. Unexpectedly, most GSLs showed no association between their levels and the enzyme activity that catabolizes them. Genomic mapping identified 30 shared predicted modifier genes between the enzymes and GSLs, which are clustered in three pathways and are associated with other diseases. Surprisingly, they are regulated by ten common transcription factors, and their majority by miRNA-340p. In conclusion, we have identified novel regulators of GSL metabolism, which may serve as therapeutic targets for LSDs and may suggest the involvement of GSL metabolism in other pathologies.

Keywords: metabolism; lysosomal enzymes; glycosphingolipids; systems genetics; modifier genes

Citation: Durán, A.; Priestman, D.A.; Las Heras, M.; Rebolledo-Jaramillo, B.; Olguín, V.; Calderón, J.F.; Zanlungo, S.; Gutiérrez, J.; Platt, F.M.; Klein, A.D. A Mouse Systems Genetics Approach Reveals Common and Uncommon Genetic Modifiers of Hepatic Lysosomal Enzyme Activities and Glycosphingolipids. *Int. J. Mol. Sci.* **2023**, *24*, 4915. https://doi.org/10.3390/ijms24054915

Academic Editor: Paola Pontrelli

Received: 7 January 2023
Revised: 7 February 2023
Accepted: 15 February 2023
Published: 3 March 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Hydrolytic enzymes are abundant in lysosomes [1]. In a healthy cell, the biosynthesis and catabolism of macromolecules are subject to regulatory mechanisms that maintain cellular homeostasis [2]. The degradative processes in lysosomes are controlled by their own enzymes [3,4]. Lysosomes play a central role in several biological processes, including energy metabolism, signaling, plasma membrane repair, secretion, and others [3]. Loss-of- function variants in genes encoding lysosomal proteins cause lysosomal storage disorders (LSDs), a group of diseases characterized by intracellular buildup of partially degraded material [5]. Growing evidence suggests that variants in lysosomal genes increase the risk of developing Parkinson's disease (PD) [6,7].

In the sphingolipidoses, a subset of LSDs, glycosphingolipids (GSLs) accumulate in late endocytic organelles (late endosomes/lysosomes) and participate in their pathological cascades [8]. Current treatments for LSDs include substrate reduction therapy (SRT), which aims to reduce the rate of biosynthesis of stored substrates [5,9,10], and enzyme replacement therapies (ERT) aimed at replacing a deficient enzyme [11,12]. Emerging treatments include gene and cell therapies [13–15] and chaperones for improving enzyme folding and trafficking [16]. Although there is a range of therapeutic options for LSDs,

they have limitations, such as tissue accessibility [17], antibody-mediated reaction [18], cost [19], and others. So far, therapies aimed at increasing enzyme activity or reducing lipid levels by modulating a second (modifier) gene have not been studied. In this context, a deeper understanding of the regulatory mechanisms that govern GSLs metabolism must be uncovered to fully develop this approximation.

Genome-wide association studies (GWAS) in humans and systems genetics strategies, which include gene mapping in model organisms, have identified genetic regulators of physiological and pathophysiological processes [20–22]. The Hybrid Mouse Diversity Panel (HMDP) has been a useful tool because genomes and tissue transcriptomes are freely available, allowing the combination of modifier gene mapping by GWAS and pathway analysis [23,24]. In this study, we have analyzed the activities of 11 lysosomal enzymes and several of their natural substrates in 25 strains of the HMDP panel followed by gene mapping and transcript integration. We identified a lack of correlation between most enzyme activities and their mRNA levels. Similarly, most substrates had no association between their levels and the enzyme activity that catabolizes them. Finally, we mapped pu- tative modifier genes of each lysosomal enzyme and GSL by GWAS. We found associations between the mRNA levels of many modifier genes and enzyme activities or GSL levels. We clustered the putative modifiers in pathways and identified common and uncommon genetic regulators between GSLs and lysosomal enzymes, including transcription factors that regulate them. Our discoveries may help develop novel therapeutics for diseases with altered lysosomal enzyme activities and GSLs.

2. Results

2.1. High Variability in the Hepatic Activity of Lysosomal Enzymes across Mouse Strains

We measured hepatic enzyme activity of β-hexosaminidase A and B (defective in Tay-Sachs and Sandhoff disease, respectively), α-neuraminidase (defective in Sialidosis/Mucolipidosis Type I), α-galactosidase A and B (defective in Fabry and Schindler disease), β-D-galactosidase (defective in GM1 Gangliosidosis), α-glucosidase (defective in Pompe), chitotriosidase (elevated in Gaucher disease), α-L-fucosidase (defective in fucosidosis), lysosomal acid phosphatase (elevated in patients with Gaucher), and Tartrate-resistant acid phosphatase (TRAP; altered in Gaucher disease) by fluorimetry in liver samples derived from 25 inbred mice strains using 4-methylumbelliferone (4-MU) based artificial substrates. We observed significant variability in the average enzymatic activity between the different strains (ANOVA $p \leq 0.05$) (Figure 1). We did not find changes in α-galactosidase A, lysosomal acid phosphatase, and TRAP activities across the tissues analyzed (Figure 1d,j,k). We observed unique activity distribution patterns across the strains for the other enzymes, suggesting specific modifiers for each enzyme.

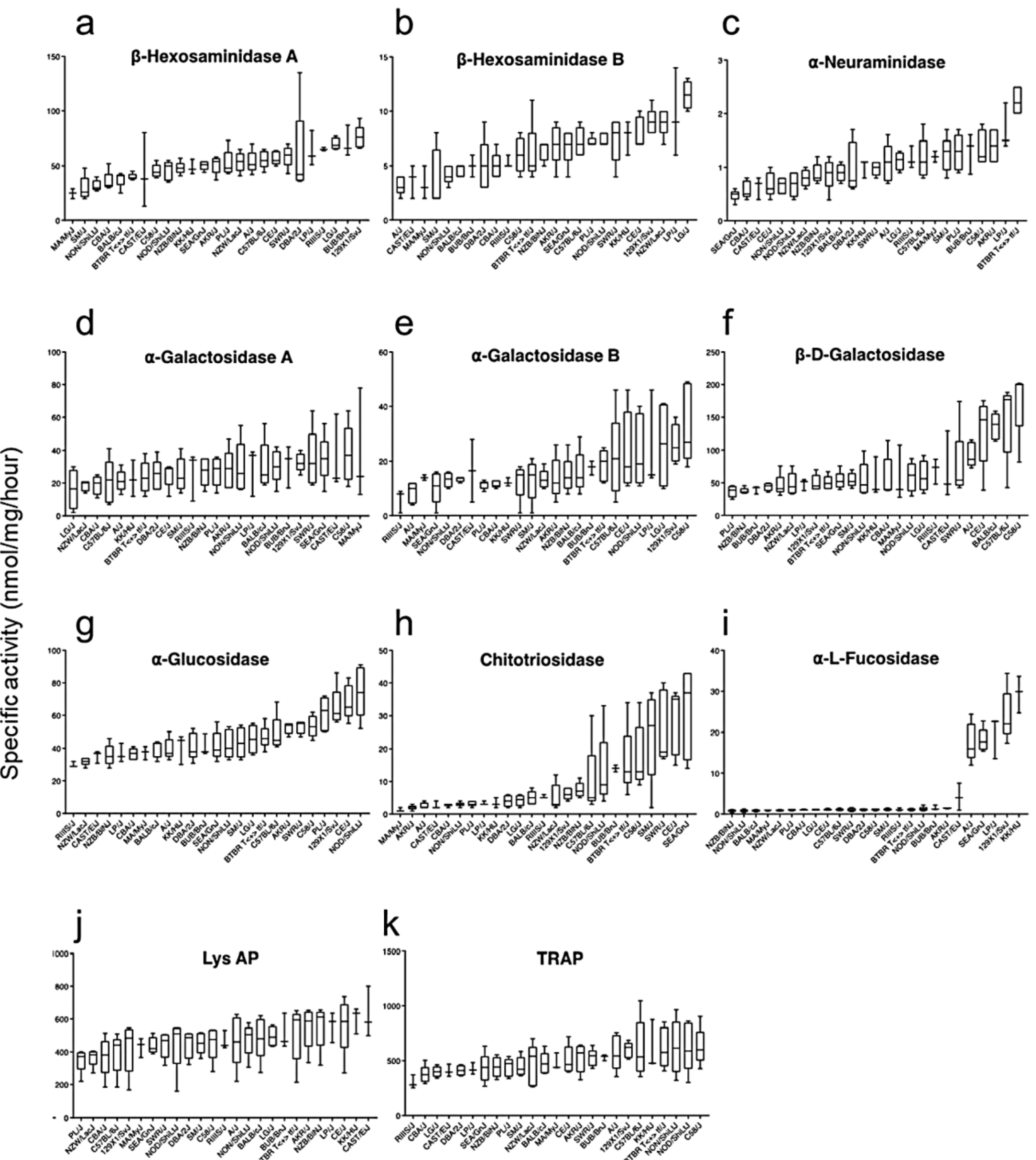


Figure 1. Hepatic variation in lysosomal enzyme activities across inbred strains. Enzyme activity are expressed as nmol/mg/hour. (**a**) Distribution of β-Hexosaminidase A. (**b**) β-Hexosaminidase B. (**c**) α-Neuraminidase. (**d**) α-galactosidase A. (**e**) α-galactosidase B. (**f**) β-D-galactosidase. (**g**) α-Glucosidase. (**h**) Chitotriosidase. (**i**) α-L-Fucosidase. (**j**) Lysosomal acid phosphatase. (**k**) Tartrate-resistant acid phosphatase, activity in the liver of 25 mouse inbred strains. Values are presented as median (n = 3–5 per strain).

2.2. Lack of Correlation between the Enzyme Activity and Its mRNA Levels

Advantages of using tissues derived from the HMDP panel of inbred mouse strains include the fact that their genomes are sequenced, and transcriptomic data are available. Thus, we analyzed potential correlations between the genes encoding lysosomal enzymes and their activities. Recently we described the natural variation of hepatic acid β-glucocerebrosidase levels across many different mouse strains and included them in this analysis [20]. We did not identify significant correlations between enzyme activity and

its transcript levels (Figure 2), with the only exception being *Glb1*, the gene encoding for β-D-galactosidase (r = 0.5775; $p \leq 0.002$) (Figure 2c). These results indicate that mRNA levels are a poor proxy for enzyme activities.

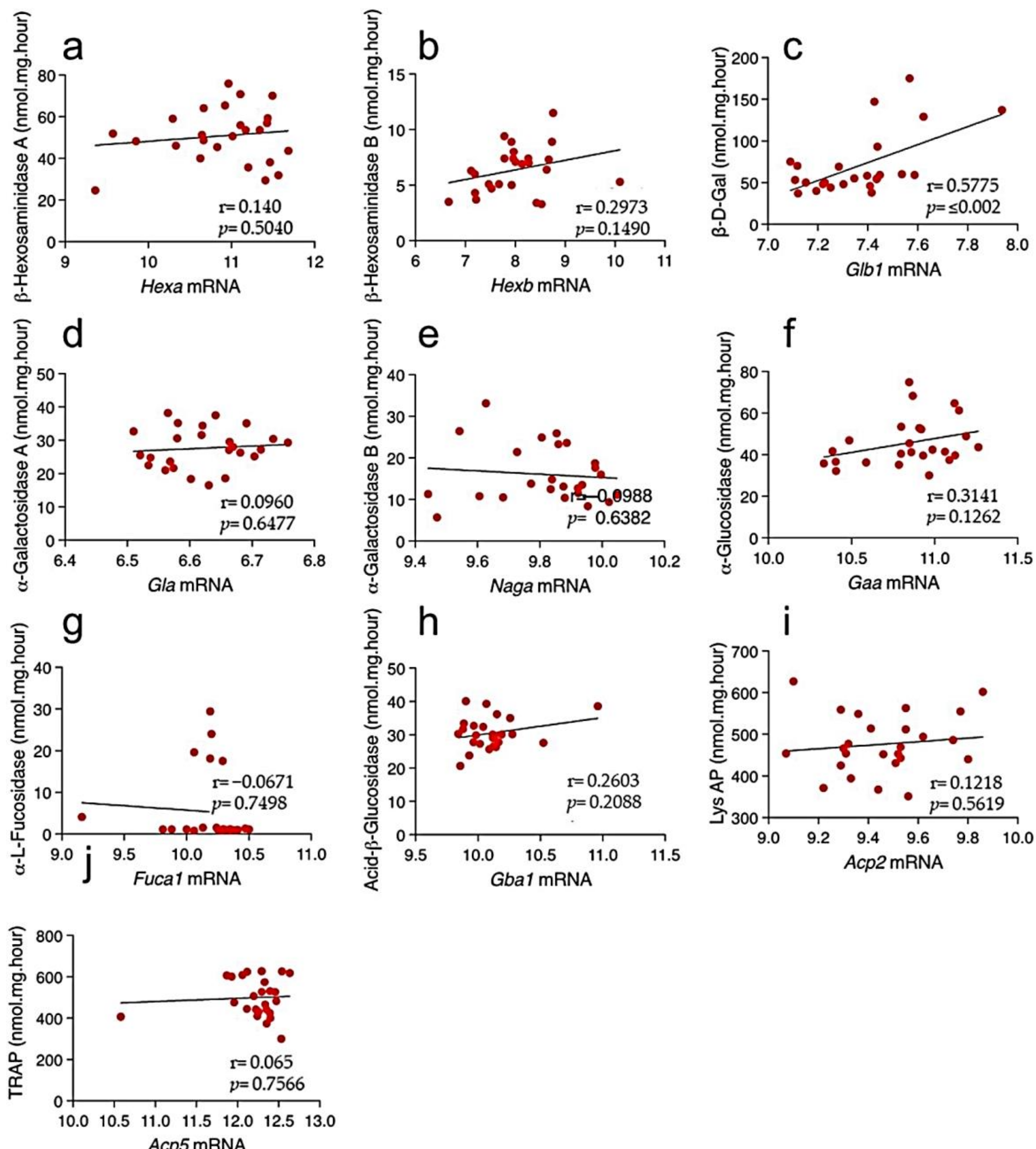


Figure 2. Correlation between expression levels and enzymatic activity in liver of mouse inbred strains. Each dot represents a mouse strain. (**a**) β-Hexosaminidase A. (**b**) β-Hexosaminidase B. (**c**) β- D-galactosidase. (**d**) α-galactosidase A. (**e**) α-galactosidase B. (**f**) α-Glucosidase. (**g**) α-L-Fucosidase. (**h**) Acid-β-glucosidase. (**i**) Lysosomal acid phosphatase. (**j**) Tartrate-resistant acid phosphatase. The Pearson's correlation was performed using 23 strains of mice. Enzyme activities are expressed as nmol/mg/hour and mRNA levels, which were downloaded from repository GSE16780 UCLA Hybrid MDP Liver Affy HT M430A [24] are expressed as log2 transformed. r, correlation; *p*, *p*-value.

2.3. High Variability in the Hepatic Glycosphingolipid Levels across Mouse Strains

Next, we measured the levels of GSLs in livers of the inbred mice strains in which we had access to enough material for three biological replicates (23/25) by Normal Phase-High-Performance Liquid Chromatography (NP-HPLC). We observed significant variability in GSLs among the strains, especially in total GSLs, GM3-Gc, GM2-Gc, GM1agc, GM3,

Gb3, GM1a, GM1b, GD1b, and GD1a (Figure 3). For example, the levels of GM3-Gc were significantly increased (ANOVA $p < 0.0001$) in NOD/ShiLtJ compared with the other samples (Figure 3b). These results indicate that GSLs levels vary across strains.

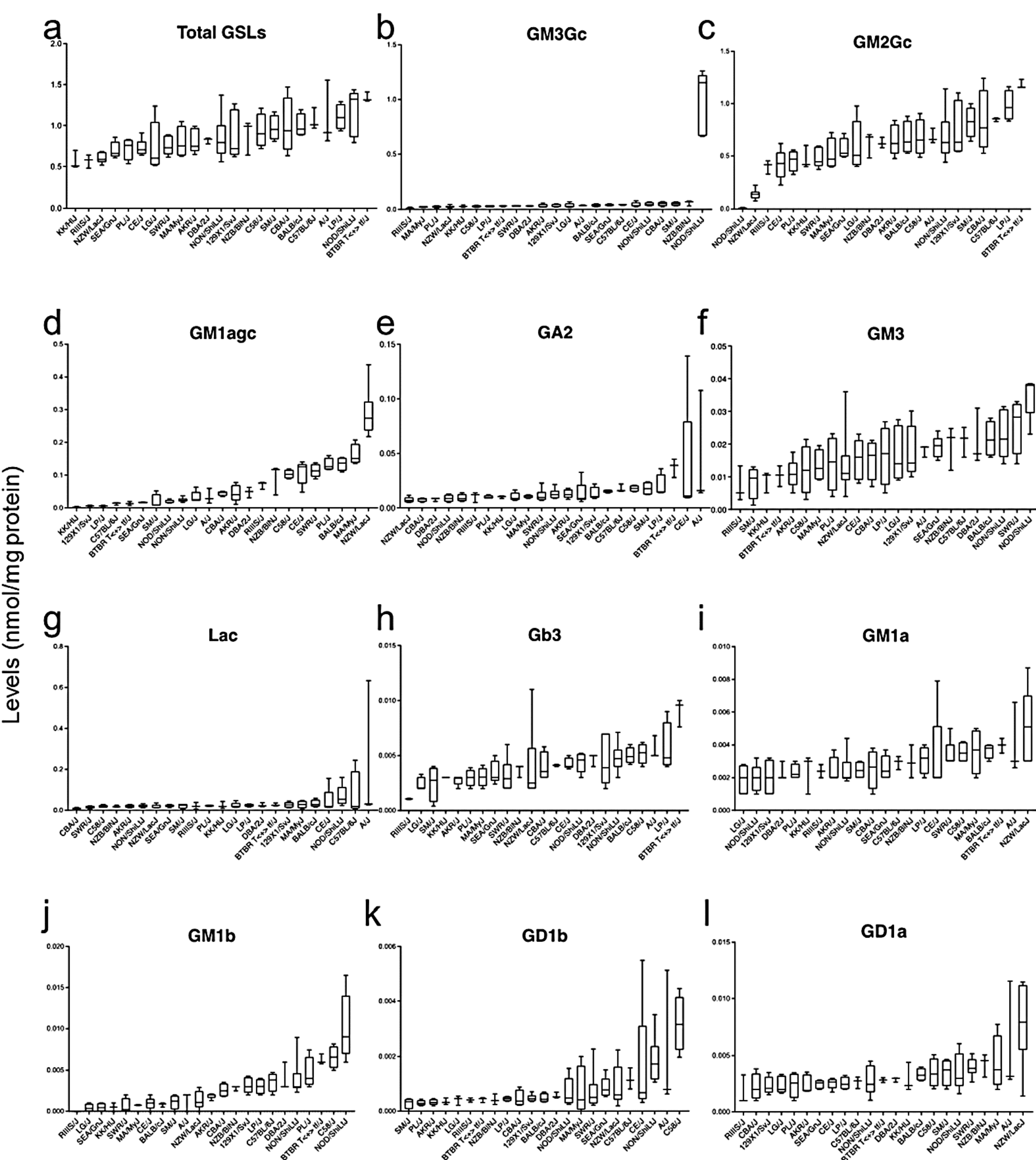


Figure 3. Variation of glycosphingolipids (GSLs) levels in liver of mouse inbred strains. Distribution of glycosphingolipids levels expressed as nmol/mg protein; (**a**) total GSLs (GM3-Gc + GM2-Gc + GM1agc + GA2 + GM3 + LacCer + Gb3 + GM1a + GM1b + GD1b + GD1a) levels. (**b**) GM3-Gc. (**c**) GM2-Gc. (**d**) GM1agc. (**e**) GA2. (**f**) GM3. (**g**) LacCer (Lac). (**h**) Gb3. (**i**) GM1a. (**j**) GM1b. (**k**) GD1b. (**l**) GD1a, in the liver of 23 mouse inbred strains. Values are presented as median (n = 3–5 per strain).

2.4. Correlations between the GSLs and the mRNA Levels of the Biosynthetic Genes

A possibility is that GSL levels could correlate with their biosynthesis rate. Since we started from frozen tissues, we could not test this directly. Instead, we utilized the transcriptomic data available from the repository GSE16780 UCLA Hybrid MDP Liver Affy HT M430A [24]. We found transcript probes for 21 mRNA of the 21 anabolic en-

zymes of the GSLs pathway and four GSL transfer proteins. The analyzed gene list of the biosynthetic pathway is presented in the Supplementary Table S1. The expression values were organized according to GSLs levels from lowest to highest and presented as a heatmap. The analysis showed significant correlations for Cgt (r = −0.4263; p = 0.042) with total GSLs (Figure 4a). For GM2-Gc with Cgt (r = −0.4582; p = 0.0279), Galgt1 (r = 0.6078; p = 0.0021), A4galt (r = 0.4903; p = 0.0176), Gltp (r = −0.454; p = 0.0296) (Figure 4b). GM3 levels correlated with Galgt1 (r = −0.579; p = 0.0038), Gltp (r = 0.4151; p = 0.0489) (Figure 4c). GM1a with Col4a3bp (r = 0.4458 p = 0.033) (Figure 4d). GM3-Gc is associated with Galgt1 (r = −0.9591, $p \leq 0.0001$) and it was the most significant correlation (Figure 4e). GM1agc levels with Slc17a2 (r = 0.4163; p = 0.0482) (Figure 5f). Gb3 with A4galt (r = 0.6011, p = 0.0024) (Figure 4g) and GM1b with Galgt1 (r = −0.5764; p = 0.004) and St8sia5 (r = −0.4194, p = 0.0046) (Figure 4h). No significant correlations were found between the majority of GSLs and biosynthetic genes (Supplementary Table S2); thus, we analyzed potential correlations between GSL levels and the enzyme activity that catabolizes them across the mouse panel.

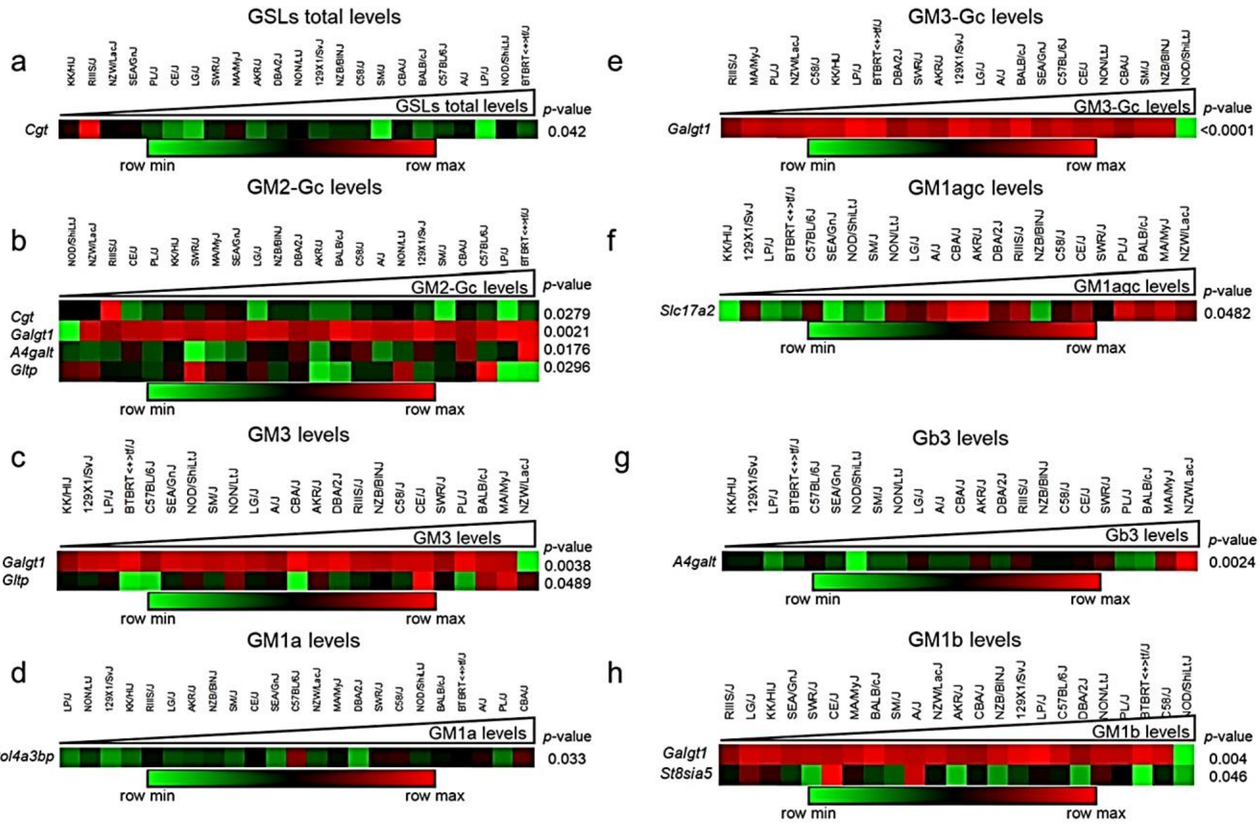


Figure 4. Correlations between GSLs and the mRNA levels of the GSL biosynthetic genes. (**a**) total GSL. (**b**) GM2-Gc. (**c**) GM3. (**d**) GM1a. (**e**) GM3-Gc. (**f**) GM1agc. (**g**) Gb3. (**h**) GM1b levels. Only the genes with significant p values with its trait using Pearson's correlations ($p \leq 0.05$) were plotted.

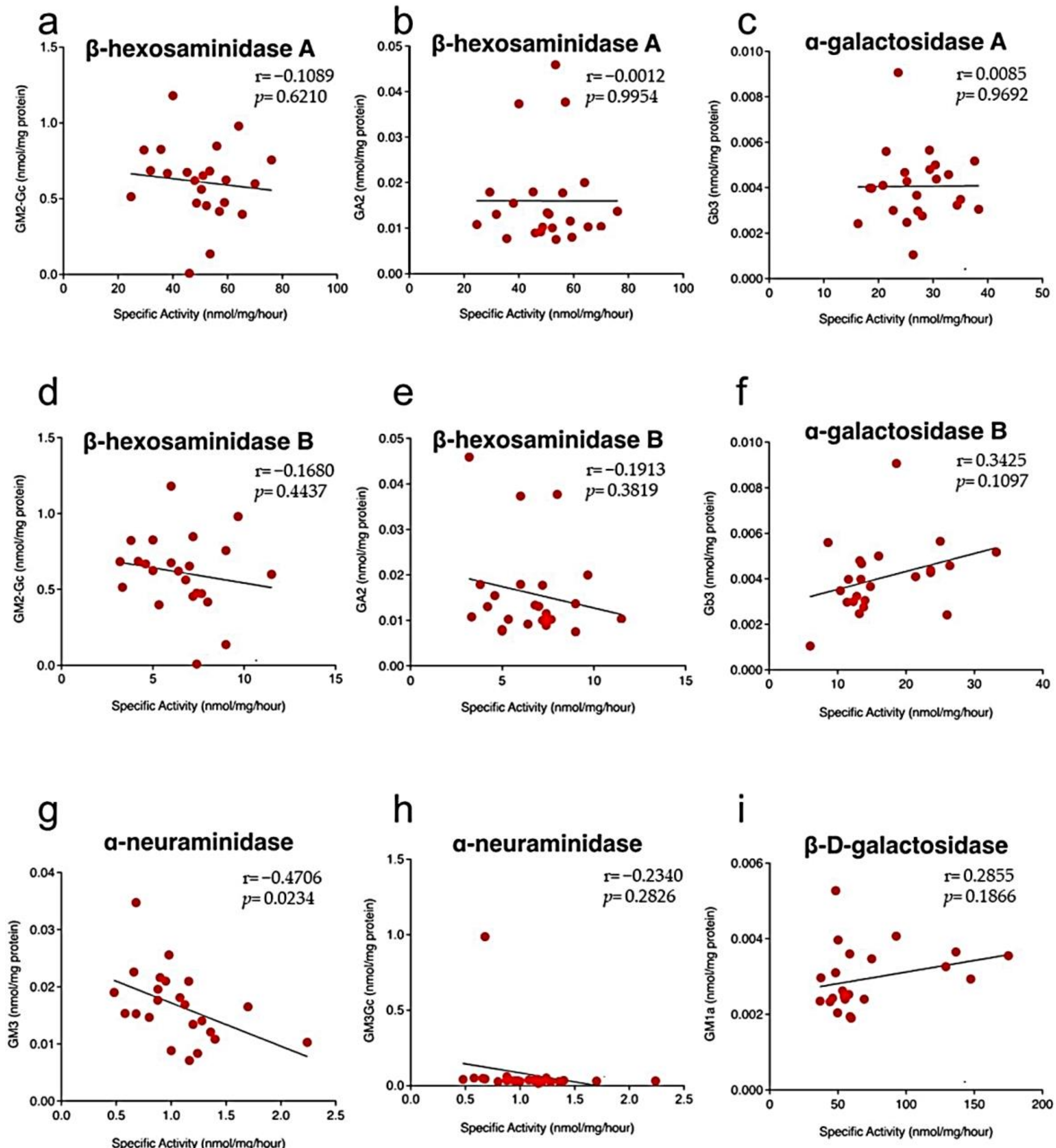


Figure 5. Correlation of hepatic lysosomal enzyme activities and specific substrates levels. Each dot represents a mouse strain. (**a**) HexA vs. GM2-Gc. (**b**) HexA vs. GA2. (**c**) α-Gal A vs. Gb3. (**d**) HexB vs. GM2-Gc. (**e**) HexB vs. GA2. (**f**) α-Gal B vs. Gb3. (**g**) Neu vs. GM3. (**h**) Neu vs. GM3-Gc. (**i**) β-D-Gal vs. GM1a; r, Pearsons' correlation; *p*, *p*-value.

2.5. *Lack of Correlation between Hepatic Lysosomal Enzyme Activity and Their Natural Substrates across Mouse Strains*

It is possible to speculate that the strains that present high activity of a particular enzyme should have reduced levels of its natural substrate because the enzyme catabolizes it. Unexpectedly, for most enzymes, we did not find significant correlations between the GSL levels and the enzyme activity that degrades it (Figure 5), except for neuraminidase and GM3-Gc (r = −0.4706; p = 0.0234) (Figure 5g). These results suggest that for most strains, the rate of biosynthesis and/or uptake of GSLs varies along with the catabolic rates which most likely are genetically regulated.

2.6. *Identification of Putative Modifier Genes of Lysosomal Enzyme Activity and Sphingolipids Levels*

To identify genetic regulators, we conducted genome-wide association studies with a quality control analysis that considered the population structure of the HMDP panel strains to reduce false associations [25,26]. We used enzyme activity levels as a trait and included the β-glucosidase activity, which we reported previously in the same and a few other strains [20]. For all the enzymes together, we identified 211 significant Single Nucleotide Variants (SNVs) that passed the empiric threshold of significance $p \leq 4.1 \times 10^{-6}$ (−log10P = 5.39), previously calculated by permutations [21,22,26], while the Bonferroni threshold was $p \leq 3.9 \times 10^{-7}$ [26]. These SNVs were located in different genomic regions (exonic, intronic, UTR3, downstream, and intergenic) (Table 1, Supplementary Table S3) in a total of 137 non-redundant genes. Similarly, we identified 3215 SNVs associated with GSLs levels (1744 non-redundant genes) whose variants are located in different genomic regions (Table 1, Supplementary Table S3). These analyses indicated that our strategy has sufficient power to map putative modifier genes.

Table 1. Summary of the putative modifier genes of enzymatic activity and GSLs levels identified by GWAS.

	Trait	Gene	Region	Chr	Position	Ref	Alt	*p*-value	SNV p<10-6	Non-redundant genes
lysosomal enzymes	α-Galactosidase A	*Stard4*	intergenic	18	33494519	C	T	2,88E-06	1	1
	α-Galactosidase B	*Barhl2*	intergenic	5	106880801	T	C	4,10E-07	1	1
	GCase	*Dmrtc2*	UTR3	7	25662483	A	G	7,46E-07	2	2
		Arhgef1	UTR3	7	25711350	G	T	7,46E-07		
	α-Glucosidase	*Tiam2*	intergenic	17	3338741	T	C	1,89E-06	3	2
		Tfb1m	intronic	17	3557483	G	T	1,89E-06		
	β-D-galactosidase	*Lyplal1*	intergenic	1	188026657	A	G	9,09E-09	88	70
		4930433B08Rik	intergenic	3	18512557	A	G	2,32E-08		
		Peak1	intronic	9	56165236	T	C	2,32E-08		
		Imp3	intergenic	9	56793621	A	G	2,32E-08		
		Scamp2	intronic	9	57409841	T	C	2,32E-08		
		Loxl1	intergenic	9	58188292	A	G	2,32E-08		
		1700072B07Rik	intergenic	9	58256079	G	A	2,32E-08		
		Arih1	intergenic	9	59348484	C	T	2,32E-08		
		Pkm	intronic	9	59506197	A	T	2,32E-08		
		Iqch	intronic	9	63413504	A	G	2,32E-08		
	Chitotriosidase 1	*Wdr89*	intergenic	12	76773815	T	C	2,45E-08	8	3
		Syne2	intronic	12	76961077	T	C	2,45E-08		
		Chchd6	intergenic	6	89566833	A	G	1,37E-06		
	α-L-Fucosidase	*Myom3*	intergenic	4	135400588	C	T	6,06E-17	103	56
		Vps45	intronic	3	95807768	C	T	1,13E-10		
		Hist2h2be	downstream	3	96027761	G	A	1,13E-10		
		Tet2	intergenic	3	133254547	A	C	1,13E-10		
		Zfp46	UTR3	4	135847850	A	G	1,13E-10		
		Hnrnpr	intergenic	4	135915162	C	T	1,13E-10		
		E2f2	UTR3	4	135750026	C	T	8,13E-09		
		Stkld1	intronic	2	26790736	A	C	6,67E-08		
		Xkr7	intergenic	2	152887679	G	A	6,67E-08		
		Ttpal	intronic	2	163432431	A	G	6,67E-08		
	TRAP	*Zfat*	intronic	15	68115989	A	C	7,70E-07	5	2
		Mir30d	intergenic	15	68244382	C	T	7,72E-07		

Table 1. *Cont.*

GSLs	GD1a	*Ctnnbl1*	intronic	2	157632357	T	A	1,46E-07	37	26
		Rap2b	intergenic	3	61765728	A	C	1,46E-07		
		Arhgef26	intronic	3	62232093	C	T	1,46E-07		
	GA2	*Dars*	intergenic	1	130350640	C	T	1,09E-11	190	103
		Abca16	intronic	7	127596308	G	A	1,09E-11		
		E130201H02Rik	intergenic	7	127763574	G	A	1,09E-11		
		Vwa3a	intronic	7	127887001	G	A	1,09E-11		
		Eef2k	intronic	7	127993389	A	G	1,09E-11		
	Lac	*Tgs1*	intergenic	4	3571870	C	G	1,38E-12	1152	723
		Dtnb	intronic	12	3586440	C	T	1,38E-12		
		Lyn	intronic	4	3673421	A	G	1,38E-12		
		Ghr	intergenic	15	3696333	T	C	1,38E-12		
		1810055G02Rik	intergenic	19	3731017	G	A	1,38E-12		
		Bambi	intergenic	18	3826103	C	A	1,38E-12		
		Hnf4g	intergenic	3	3989664	G	T	1,38E-12		
		Dnajc27	UTR3	12	4106955	G	C	1,38E-12		
		Mterf1b	intergenic	5	4503200	A	G	1,38E-12		
		Impad1	intergenic	4	4885958	C	G	1,38E-12		
	GD1b	*Ahctf1*	intergenic	1	181812047	C	A	1,03E-08	23	17
		Psen2	intronic	1	182170093	C	T	1,03E-08		
		Fhit	intronic	14	11843484	G	A	1,03E-08		
	GM3Gc	*Cdk6*	intergenic	5	3011917	T	C	1,51E-31	1811	995
		Insr	intergenic	8	3058687	C	T	1,51E-31		
		Eif4enif1	intronic	11	3143753	G	A	1,51E-31		
		Tiam2	intronic	17	3417745	T	C	1,51E-31		
		Ppp6r3	intronic	19	3539614	C	T	1,51E-31		
		Tfb1m	intronic	17	3540913	C	A	1,51E-31		
		1700102H20Rik	intergenic	17	3611518	G	A	1,51E-31		
		Pex1	intronic	5	3632859	T	G	1,51E-31		
		Ankib1	intronic	5	3711311	C	T	1,51E-31		
		Ankib1	intronic	5	3778272	C	T	1,51E-31		
	GM1b	*Hrasls*	intergenic	16	29161604	A	C	2,71E-06	2	1

2.7. Correlations between the Traits and the mRNA Levels of Putative Modifiers

To prioritize the putative modifier genes that could regulate each enzyme, we searched for correlations between the transcript levels of putative modifier genes and their traits (enzyme activity and GSL levels, respectively) (Figure 6). We found transcript probes for 67 mRNA of the 137 putative modifiers of the enzymes. The expression values were organized according to enzyme activity from lowest to highest and presented as a heatmap. The analysis showed significant correlations in Fip1l1 (r = −0.4462; p = 0.0254) with α-L-fucosidase (Figure 6a). For β-D-galactosidase with Lyplal1 (r = −0.702; p = <0.0001), Arrdc4 (r = 0.627; p = 0.0008), Pde2a (r = 0.5306; p = 0.0064), Glb1 (r = 0.5753; p = 0.0026), Bptf (r = 0.5135; p = 0.0087), Oxr1 (r = −0.447; p = 0.0251) (Figure 6b). No significant correlations were found for the other enzymes analyzed. We used SIFT to explore the impact of genetic variants on the genes identified by GWAS (benign or deleterious changes) associated with changes in enzyme activity [27], because the full genomes of the strains are known [28]. This strategy identified 308 predicted deleterious variants (Supplementary Table S4) in 43 of the 67 genes whose functions are related to organelle biogenesis (Chchd6) [29], intracellular signaling (Pde4dip) [30], and tissue development (Fam181b) [31], among others. These results suggest that amino acid substitution could affect protein function and signaling pathways leading to changes in enzyme activity.

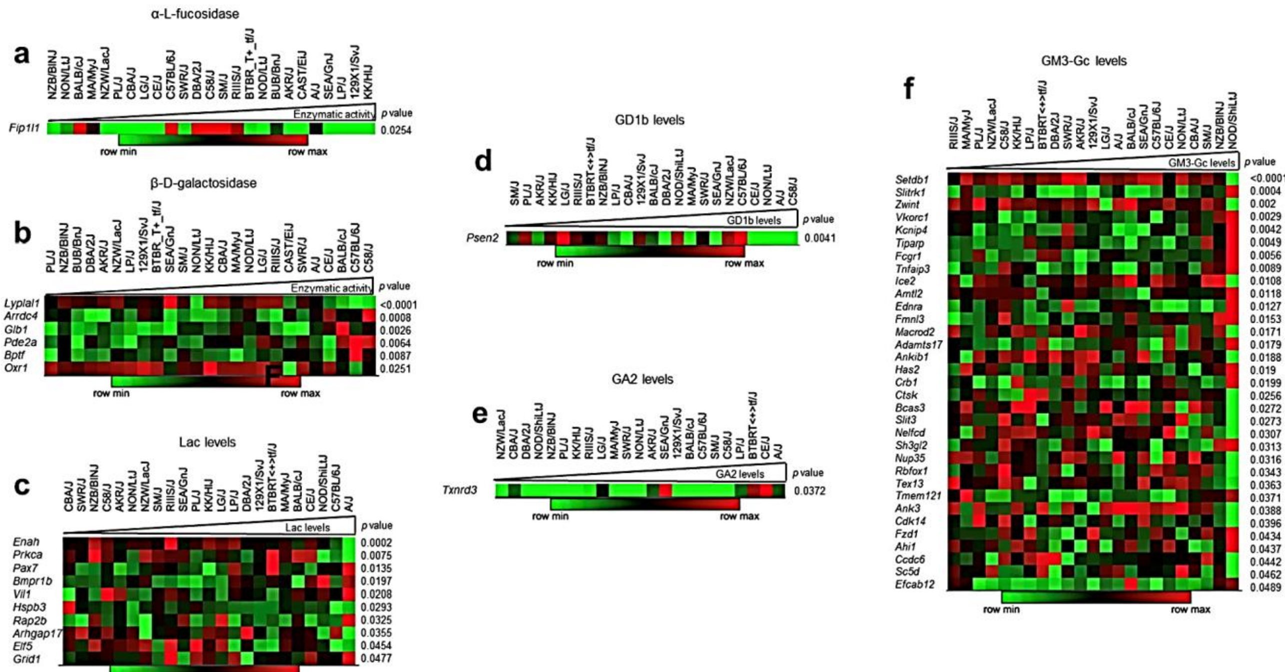


Figure 6. Correlations between the traits (enzyme activities or GSLs) and the mRNA levels of the identified modifier genes. (**a**) α-L-Fucosidase. (**b**) β-D-galactosidase. (**c**) LacCer (Lac). (**d**) GD1b. (**e**) GA2. (**f**) GM3-Gc. Only the genes with significant p values with its trait using Pearson's correlations ($p \leq 0.05$) were plotted.

The same analysis was performed to identify putative modifiers of GSL levels (Figure 6c–f). For 1744 non-redundant SNVs, we found expression values for 994 genes. The analysis identified 45 significant correlations, of which 33 were correlated with GM3-Gc levels, 10 genes with LacCer, and one gene with GD1b and GA2 (Figure 6c–f). Overall, we recorded 4.9% (52/1061) of significant correlations distributed between the two traits. We also explored the impact of genetic variants associated with changes in GSLs with SIFT [27]. This strategy identified 515 deleterious variants predicted to disrupt the protein structure (Supplementary Table S4) in 132 genes related to DNA methyltransferase activity (Setdb1) [32] and synapse (Slitrk1) [33], among others.

2.8. Enrichment Analysis and Common Modifier Genes between Glycosphingolipids Levels and Lysosomal Enzyme Activities

If there is an orchestrated regulation of GSL levels and the enzymes that degrade them, it would be expected to observe enrichment in common pathways [34]. We therefore utilized gProfiler [35] to perform enrichment analysis using the putative modifier genes lists. For the modifier of enzyme activities, we found significantly associated pathways such as cell periphery ($p = 5.9 \times 10^{-4}$), plasma membrane ($p = 2.4 \times 10^{-3}$), and integral components of the plasma membrane ($p = 2.6 \times 10^{-2}$) (Figure 7b), which could be related to endocytic processes necessary to deliver key molecules to the lysosome, including the lysosomal enzymes that can be recycled from the extracellular space. Significant biological processes analysis included regulation of cellular processes ($p = 3.9 \times 10^{-2}$) (Figure 7d) (Supplementary Table S5). We did not find significant enrichment for the molecular function category. For GSLs, we observed enrichment in terms like cytoplasm ($p = 3.5 \times 10^{-28}$), cell junction ($p = 7 \times 10^{-21}$), synapse ($p = 4.6 \times 10^{-19}$), and 70 other pathways related to cellular components (Figure 7a; Supplementary Table S5). Many of these pathways require cellular membranes, where GSLs play a structural role. Significantly enriched Gene Ontology (GO) terms included protein binding ($p = 9.1 \times 10^{-31}$), ion binding ($p = 8.8 \times 10^{-14}$), binding ($p = 2.3 \times 10^{-13}$), ATP binding ($p = 9.4 \times 10^{-13}$), carbohydrate

derivate binding (p = 1.8 × 10^{-1}), and 27 other pathways related to molecular functions (Supplementary Table S5). Biological processes terms revealed 328 pathways, including system development (p = 4.3 × 10^{-39}), anatomical structure development (5.6 × 10^{-38}), and multicellular organism development (p = 1.4 × 10^{-37}). We searched for the overlap between the cellular component domains of modifiers of enzyme activity and GSLs, which resulted in three common pathways (GO:0071944—cell periphery, GO:0005886—plasma membrane, and GO:0005887—integral component of plasma membrane) (Figure 7c) and one pathway associated with biological processes (GO:0050794; regulation of cellular process) (Supplementary Table S5).

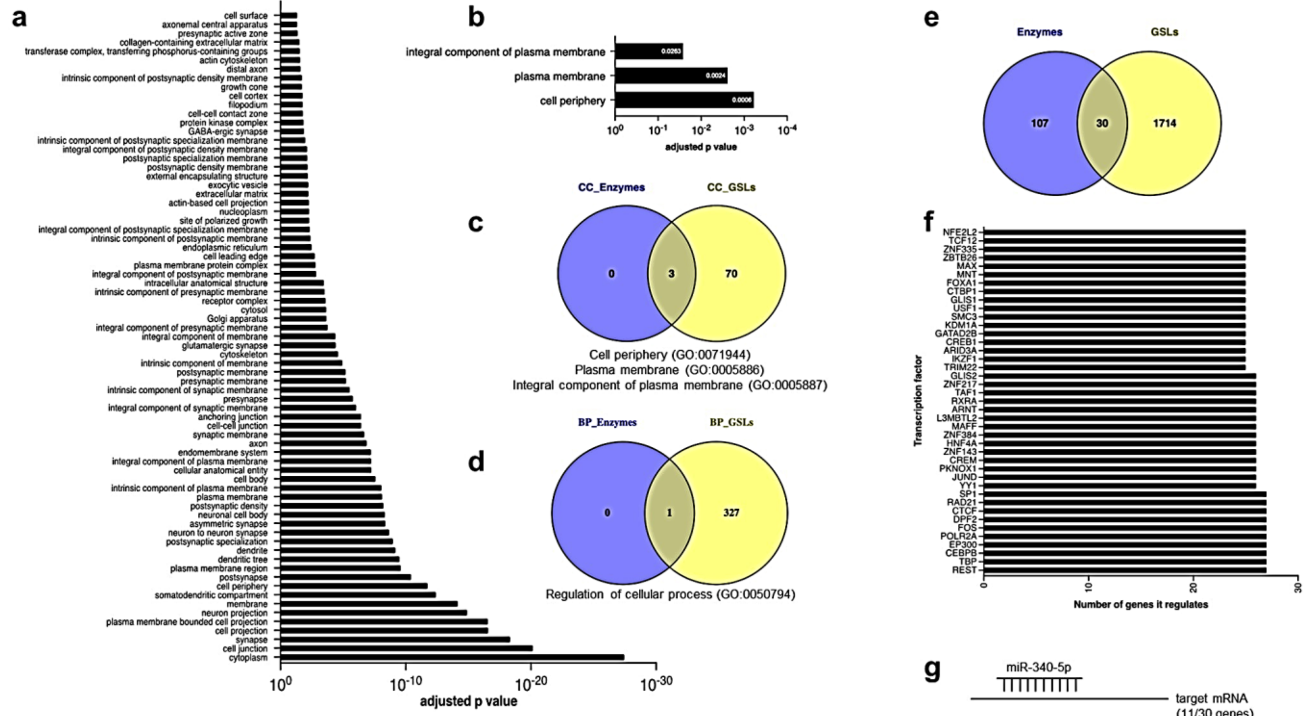


Figure 7. Enrichment analysis and common modifier genes between glycosphingolipids levels and lysosomal enzyme activities. (**a**) Cellular component functional enrichment analysis for gene sets (1744 non−redundant genes) of six substrate (GM3−Gc, GA2, Lac, GM1b, GD1b, GD1a), analyzed by g:Profiler. We found 73 GO_CC associated. (**b**) Cellular component functional enrichment analysis for gene sets (137 non−redundant genes) of eight enzymes (acid β−glucosidase, α−galactosidase A, α−galactosidase B, α−glucosidase, β−D−galactosidase, chitotriosidase, α−L−fucosidase, tartrate−resistant acid phosphatase). (**c**) Venn diagram with common GO terms cellular component between two traits. (**d**) Common GO terms biological processes between two traits. (**e**) Common genetic regulators between two traits. (**f**) Transcription factors that bind to the common genes. (**g**) Cartoon with common genetic regulators that miR−340−5p can bind.

2.9. *Common and Uncommon Modifiers between Hepatic Lysosomal Enzyme Activity and Sphingolipids Levels*

Common regulators of GSLs and enzymatic activities are relevant for understanding GSL metabolism and may be attractive therapeutic targets for LSDs. Therefore, we examined the overlap between them. We found 30 common and 1821 uncommon genes (Figure 7e). We explored their functions and identified genes involved in mitochondrial biogenesis and dynamics (Tfb1m, Timen135, Chchd6) [29,36,37], cell proliferation (Fstl5, Fzd10, Arhgap18) [38–40], platelet function (Cdh6) [41], vesicular trafficking (Vps45) [42], gene expression (Tfb1m, Zfat) [36,43], and regulating levels of the proto-oncogene MYC (Pvt1) [44]. Many of the 30 genes have been linked to diseases, such as Pvt1, Tiam2, Fstl5, Fzd10, Cdh6, Pvt1, Chchd6 in liver, colorectal, nasopharyngeal, and gastric cancer [45–50].

Others participate in neurodegenerative conditions; PD, schizophrenia, and intellectual disability (Tenm4, Pde4dip, Grid2, Arhgap18) [51,52]. These results suggest that lysosomal enzymes and GSLs may play a role in their pathophysiology and should be explored further (Table 2).

To better understand the molecular regulation of these 30 genes, we analyzed the transcription factors that bind to their promoters and/or enhancers (Figure 7f). We found no information for three of the 30 genes since they are putative (Rik) genes. The following transcription factors can bind to the 27 genes for which we have information: REST, TBP, CEBPB, EP300, POLR2A, FOS, DPF2, CTCF, RAD21, and SP1. Some of these transcription factors are broad regulators of transcription, such as TBP and POLR2A, while others are selective for specific processes, such as CTCF and RAD21. Considering all the promoters/enhancers of the 27 shared genes, we identified a total of 533 transcription factors that can bind them, although some only bind a few genes (Supplementary Table S6). We also searched for potential shared microRNA (miRNA) regulators using miRTarBase, a curated microRNA database [53]. We identified that miR-340-5p can bind to 11 of the 27 known common genes (Tusc1, Fam91a1, Zc3h12c, Adamts5, Tmem135, Tenm4, Grid2, Csnk1g3, Cdh6, Fam181b, and Pde4dip; $p = 2.2 \times 10^{-2}$) (Figure 7g). This result suggests that miRNA-340-5p regulates GSLs metabolism and may be involved in the pathogenesis of LSDs and the disorders described in Table 2.

Table 2. Common genetic modifiers associated with enzyme activity and hepatic glycosphingolipid levels in inbred mouse strains. The references related to this table are presented as supplementary material.

Gene	Description	Traits				Related functions	Associated human diseases	Previosly associated with traits	References
		Enzyme	*p*-value GWAS	GSLs	*p*-value GWAS				
Tiam2	T cell lymphoma invasion and metastasis 2	α-Glucosidase	1,89E-06	GM3-Gc	1,51E-31	neuroplasticity	liver cancer	No	[47,54]
Tfb1m	Dimethyladenosine transferase 1, mitochondrial	α-Glucosidase	1,89E-07	GM3-Gc	1,51E-31	promotion of mitochondrial biogenesis	deafness	No	[36,55]
Dok5	Insulin receptor substrate 6	β-D-galactosidase	1,43E-07	Lac	1,38E-12	promove osteoblast proliferation and differentiation / insulin and IGF-1 signaling	cancer, Alzheimer's disease	No	[56–59]
4930433b08Rik	RIKEN cDNA 4930433B08 gene	β-D-galactosidase	2,32E-08	Lac	1,38E-12	-	-	-	-
A830019l24Rik	RIKEN cDNA A830019L24 gene	β-D-galactosidase	1,43E-07	Lac	1,38E-12	-	-	-	-
Tmem135	Transmembrane protein 135	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	involved in mitochondrial dynamics	retinal diseases	No	[37,60]
Fam181b	Family with sequence similarity 181, member B	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	increased expression during mouse development	-	No	[31]
Tenm4	Teneurin transmembrane protein 4	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	cell maturation and myelination in SNC	neuropsychiatric disorders, Parkinson's	No	[51,61–63]
Plk2	Serine/Threonine-protein kinase PLK2	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	cell proliferation, alpha-synuclein phosphorylation	pulmonary fibrosis	No	[64,65]
Stk32a	Serine/Threonine kinase 32A	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	kinase activity	lung cancer	No	[66,67]
Dpysl3	Dihydropyrimidinase like 3	β-D-galactosidase	1,43E-07	GM3-Gc	1,51E-31	cell migration, cytoskeletal dynamics and inflammation	gastric cancer, amyotrophic lateral sclerosis	No	[68–70]
Prex1	PIP3 Dependent rac exchange factor 1	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	contributes to the effector activity of mouse neutrophils	prostate cancer	No	[71,72]
Fstl5	Follistatin-related protein 5	α-L-Fucosidase	6,67E-08	Lac	1,38E-12	play a role in cell proliferation	hepatocellular carcinoma	No	[38,73]
Vps45	Vacuolar protein sorting-associated protein 45	α-L-Fucosidase	1,13E-10	GM3-Gc	1,51E-31	vesicle-mediated protein trafficking from the Golgi	neutrophil disorders	No	[42,74]
Hist2h2be	Histone cluster 2 H2B family member E	α-L-Fucosidase	1,13E-10	GM3-Gc	1,51E-31	is necessary for proliferation	breast cancer	No	[75]
Pde4dip	Phosphodiesterase 4D interacting protein	α-L-Fucosidase	2,29E-07	GM3-Gc	1,51E-31	cAMP-dependent pathway to Golgi and/or centrosomes	schizophrenia	No	[52]
Tusc1	Tumor suppressor candidate 1	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	reduced cell proliferation *in vitro e in vivo*	glioblastoma	No	[76,77]
Fzd10	Frizzled class receptor 10	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	promotes cell proliferation through Wnt1	cancer	No	[39,48]
Grid2	Glutamate ionotropic receptor delta type subunit 2	α-L-Fucosidase	6,67E-08	Lac	1,38E-12	receptor for glutamate	neurodevelopmental syndrome / intellectual	No	[78]
Zc3h12c	Zinc finger CCCH-type containing 12C	α-L-Fucosidase	1,64E-06	GM3-Gc	1,51E-31	RNA stability associated with inflammatory genes	psoriasis	No	[79,80]
Arhgap18	Rho GTPase activating protein 18	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	role in migration, spreading and controls stress fiber formation	schizophrenia in Chinese population	No	[40,81]
Cdh6	Cadherin 6	α-L-Fucosidase	1,34E-06	Lac	1,38E-12	inhibit platelet aggregation	cancer	No	[41,49]
Fam91a1	Family with sequence similarity 91 member A1	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	WDR11 complex (vesicular trafficking)	adenocarcinoma	No	[82,83]
4933412e24Rik	RIKEN cDNA 4933412E24 gene	α-L-Fucosidase	6,67E-08	Lac	1,38E-12	-	-	-	-
A1bg	Alpha-1B-Glycoprotein	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	cell dynamics and acquired immune response	cervical and bladder carcinogenesis	No	[84–86]
Pvt1	Pvt1 Oncogene	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	promotes cell proliferation	cancer	No	[50,87]
Adamts5	ADAM Metallopeptidase with thrombospondin type 1 motif 5	α-L-Fucosidase	6,67E-08	GM3-Gc	1,51E-31	metalloproteinase that remoldels connective tissue	osteoarthritis	No	[88]
Csnk1g3	Casein kinase 1 gamma 3	α-L-Fucosidase	3,51E-07	Lac	1,38E-12	wnt signaling pathway	breast, brain and colon cancer	No	[89,90]
Chchd6	Coiled-coil-helix-coiled-coil-helix domain containing 6	Chitotriosidase	1,37E-06	GA2	3,27E-06	mitochondrial membrane morphology	cancer	No	[29]
Zfat	Zinc finger protein ZFAT	TRAP	7,72E-07	Lac	1,38E-12	immune response	hashimoto's disease	No	[43,91]

3. Discussion

In this study we searched for genetic modulators involved in the regulation of the lysosomal enzyme activities and the levels of substrates related to GSLs, with the idea of finding novel therapeutics targets for disorders in which they participate. By GWASs, we identified common and uncommon genetic regulators, evaluated the associations between modifier gene mRNA levels and each trait, and also clustered them in pathways. We identified 30 shared putative modifiers and described the transcription factors that are predicted to regulate them, and we noted that the miRNA340-5p can bind to 11 of these genes.

Our first unexpected finding was that most lysosomal enzyme activities do not cor- relate with their mRNA levels, nor with most of their substrate levels. Although enzyme activity can decrease with age [92], we used sex and age-matched samples; thus, the variation observed across strains was shown not to be due to any of these factors.

Another unexpected finding was that GM2-Gc levels correlate with the mRNA levels of the *Cgt* gene, which encodes for the UDP-galactose ceramide galactosyltransferase (CGT). CGT is a key enzyme for the biosynthesis of galactocerebrosides. Gangliosides, including GM2 derivates, are built from glucosylceramide and not from the galacto series [93]. However, for most of the biosynthetic genes there were no associations between the amount of lipids and the transcript levels of their anabolic pathways. Altogether, our results suggest that the GSL biosynthesis rate and uptake differ across the mouse strains, suggesting the existence of specific modifier genes for each trait.

Our third unexpected finding was that TFEB, the master transcriptional regulator of lysosomal genes [94], did not appear in the list of modifiers of lysosomal enzymes. This may be due to the fact that we screened for enzymatic activity instead of mRNA levels, and we showed a lack of correlation between transcript levels and enzyme activity under physiological conditions, at least for most enzymes. One exception was β-D-galactosidase, for which we found a positive correlation between its transcript levels and activity. Furthermore, the GWAS for this enzyme identified *Glb1*, the gene encoding for β-D-galactosidase, as aputative modifier of its activity, validating the power of discovery of our population-based strategy [95].

Our study had some limitations: First, we quantified lysosomal traits from liver homogenates that were not in living or isolated organelles, which may have diluted enzyme activity or promoted molecular interactions that might not occur in vivo because of cellular compartmentalization. Second, we could not directly measure GSLs biosyn- thesis and uptake because we started with mouse liver samples. Third, we used SNV catalogs with imputation, which may lead to false associations, though with increased mapping resolution.

Most of the enzymes we assayed are associated with LSDs [5,8]. For many LSDs, no therapies are available, and the few currently available treatments have severe limitations [5]. In this context, targeting a modifier gene could be a novel therapeutic approach. For example, lack of β-D-galactosidase activity triggers GM1 gangliosidosis, a disease with no approved therapies [96]. Our study identified the druggable *Lypla1* and *Pkm* genes as putative modifiers of β-D-galactosidase activity, which can be pharmacologically modulated [97,98]. We found other druggable genes as well for several traits, and with the current gene editing technologies virtually any gene can be targeted. The potential modifying effects of these genes and compounds can be tested in LSDs disease models.

A hallmark of the sphingolipidoses is the intracellular buildup of GSLs, so strategies aimed at reducing their levels could lead us to novel therapies [5,8]. GSLs comprise a ceramide moiety with one or more sugar residues linked to it [99]. An approved therapy for Gaucher and Niemann-Pick disease type C is Miglustat [100,101], a small molecule inhibitor of GSL biosynthesis, thus reducing their levels. Our GSLs GWAS identified more than 50 genes previously associated with sphingolipid metabolism, which served as a positive control, including *B3gnt5, Cln8, Hexb, Pnpla1, St8sia1*, and *Cgt*. *B3gnt5* regulates GSLs metabolism and lung tumorigenesis [102]. Our study also identified *Lipc* as a modifier of GM3-Gc levels, which has been previously associated with elevated serum levels of liver enzymes (alkaline phosphatase and γ-glutamyl transferase) [103], suggesting a new connection between GM3-Gc and liver damage. Variants in *LIPC, CPS1, PABPC4, CITED2, TRPS1,* and *MVK* are

associated with changes in plasma lipoprotein levels [104], connecting novel traits to GSLs metabolism.

Lysosomal leakage has been associated with Alzheimers' [105], cancer, and inflamma- tion among other conditions [106]. Recently, the phosphoinositide signaling pathway was implicated in lysosomal repair [107]. Many genes of this pathway appear in our discovery list (*Osbpl9, Osbpl6, Pde4dip, Pde2a, Pde1a, Pde7a, Pde7b, Pde4d, Pde8b, Pld5, Pik3r1, Pip4k2a, Pip5k1a, Pip5k1b, Pi4kb, Pdpk1, Atg4c, Atg10*), suggesting that integrity of the lysosomal compartment is key to the proper functioning of enzymes and/or that these enzymes and lipids participate in lysosomal repair. Furthermore, this novel lysosomal repair pathway may facilitate the development of novel therapeutics for these diseases with lysosomal leakage. Defects in the 30 shared genes are related to several pathologies, such as vision abnor- malities (*TMEM135*) [60], cancer (*CDH6* [49], *FZD10* [48], *TIAM2* [47]), neuropsychiatric disorders (*Tenm4* [51], *Pde4dip* [52], *Grid2* [78]), deafness (*TFB1M*) [55], neutrophil disorders (*VPS45*) [74] and others. Lysosomal enzymes and GSLs have been widely studied in cancer and neurodegenerative diseases [46,108–111]; however, their role in the other identified conditions should be explored.

Although not binding the complete list of shared genes, we identified some transcription factors previously known to be involved in lipid metabolism and autophagy-lysosomal functions (PPARγ, SREBF1, HNF1A, YY1, EGR1, SP1 and TFE3, E2F1, CREB1, MYC) [112–121], and many more that have not been previously linked to GSL metabolism.

We also identified miR-340-5p as a putative regulator of many common modifier genes. Changes in miR-340-5p are linked to preeclampsia, neuroinflammation [122–126], adipocyte differentiation [127], as well as obesity and diabetes [128]. GSL metabolism plays a crucial role in the two last-mentioned disorders, and inhibitors of their biosynthesis have shown promising results in animal models of these conditions, validating the relevance of our strategy [129,130].

In conclusion, we described putative regulators of hepatic lysosomal enzymes and GSLs, many of them druggable and associated with diseases where alterations in GSL metabolism have not been previously described and should be assessed. We expect our findings may facilitate the development of novel therapeutics for conditions with alterations in these traits.

4. Materials and Methods

4.1. Mouse Tissues

We used 8 weeks-old mice livers derived from 25 inbred mouse strains, which were kindly donated by Dr. Aldons Lusis (University of California, Los Angeles, CA, USA). (i) 129X1/SvJ, (ii) A/J, (iii) AKR/J, (iv) BALB/cJ, (v) BTBR T<+> tf/J, (vi) BUB/BnJ, (vii) C57BL/6J, (viii) C58/J, (ix) CAST/EiJ, (x) CBA/J, (xi) CE/J, (xii) DBA/2J, (xiii) KK/HlJ, (xiv) LG/J, (xv) LP/J, (xvi) MA/MyJ, (xvii) NOD/ShiLtJ, (xviii) NON/ShiLtJ, (xix) NZB/BlNJ, (xx) NZW/LacJ, (xxi) PL/J, (xxii) RIIIS/J, (xxiii) SEA/GnJ, (xxiv) SM/J, (xxv) SWR/J. Tissues were homogenized and adjusted to 50 mg tissue/mL in deionized water with a Potter-Elvehjem tissue homogenizer (Omni International, Kennesaw, GA, USA). Three or more livers per mouse strain were used to quantify traits (Supplementary Table S7).

4.2. Enzyme Activity Assays

Lysosomal hydrolase activities were determined using an artificial fluorescent sub- strate based on 4-methylumbelliferone (4-MU) [131]. For α-glucosidase, 1.47 mM 4-MU α- D-glucopyranoside (Sigma, Dorset, UK) in 100 mM citric acid/100 mM sodium phosphate, 0.1% TritonX-100, pH 4.0 was used as substrate [132]. The substrate for α-galactosidase A and B activities was 5 mM 4-MU α-D-galactopyranoside (Santa Cruz, CA, USA) with and without 250 mM N-acetyl-galactosamine (Sigma, Dorset, UK) in 100 mM citric acid/100 mM tri-sodium citrate, 0.1% TritonX-100, pH 4.0 [133,134]. For measuring β-hexosaminidase A and B activity, 3 mM 4-MU N-acetyl-β-D-glucosaminide (BioChemika, Dorset, UK) in 100 mM citric acid/100 mM sodium phosphate, 0.1% TritonX-100, pH 4.5 was used as sub- strate. Heat inactivation assay for β-hexosaminidase A was carried out at 50 °C for 3 h [135]. For β-galactosidase activity, 1 mM 4-MU β-D-galactose (Sigma, Dorset, UK) in 200 mM sodium acetate buffer, 100 mM NaCl, 0.1% TritonX-100, pH 4.3 was used as substrate [136]. The substrate for neuraminidase activity was 0.4 mM 4-MU α-D-N-acetylneuraminic acid (Sigma, Dorset, UK) in 0.1 M acetate buffer, 0.1% TritonX-100, pH 4.6 [137,138]. For chi- totriosidase activity, 0.013 mM 4-MU chitotrioside (Sigma, Dorset, UK) in 100 mM citric acid/200 mM

sodium phosphate, 0.1% TritonX-100, pH 5.2 was used as substrate [139,140]. For total acid phosphatase activity, 5 mM 4-MU phosphate (Sigma, Dorset, UK) with 40 mM NaCl in 200 mM citric acid/200 mM sodium phosphate, 0.1% TritonX-100, pH 4.5 was used as substrate. For tartrate-resistant acid phosphatase (TRAP) activity, 5 mM 4-MU phosphate (Sigma, Dorset, UK) with 40 mM Na Tartrate in 200 mM citric acid/200 mM sodium phosphate, 0.1% TritonX-100, pH 4.5 was used as substrate. The difference between total acid phosphatase activity and TRAP corresponded to lysosomal acid phosphatase (Lys AP) activity [141,142]. The substrate for α-L-fucosidase activity was 60 nM 4-MU α-L-fucopyranoside (Sigma, Dorset, UK) in 200 mM citric acid/200 mM sodium citrate, 0.1% TritonX-100, pH 5.0 [143,144]. We determined the acid-β-glucosidase activity in the same tissues in a previous publication [20], and further analyses were performed here based on the published activity. Liver homogenates were diluted with the buffer corre- sponding to each enzymatic determination. Three cycles of freezing (liquid nitrogen) and thawing were performed on the samples. Three biological replicates of the diluted liver extracts were incubated with the corresponding substrate at 37 °C for 30 min (or 1 h forα-neuraminidase, β-D-galactosidase, and chitotriosidase). Cold 0.5 M Na_2CO_3 (pH 10.7) was added to stop the reaction. Fluorescence intensity in samples was measured in a Synergy HT plate reader (BioTek, Winooski, VT, USA) at 360/460 nm. Protein concentration was measured using a BCA protein assay kit (Thermo Fisher Scientific, New Jersey, NJ USA). Fluorescence values were normalized to protein concentration. A 4-MU standard curve was constructed to calculate specific activity, and the final value was adjusted to one hour of enzymatic reaction.

4.3. *Glycosphingolipids Levels Quantification*

The GSLs were extracted and measured by Normal Phase-High-Performance Liquid Chromatography (NP-HPLC) following published methods [145]. Briefly, the aqueous tis- sue extract was homogenized in chloroform/methanol (C:M) (1:2 *v*/*v*) and kept overnight at 4 °C. Then, the extracts mixture was centrifuged at 3000 rpm for 10 min at room temperature. We added 0.5 mL of PBS and 0.5 mL of chloroform to the supernatant followed by a 3000- rpm centrifugation for 10 min at room temperature. The lower phase was carefully removed and dried under a stream of nitrogen gas (N2) in a heating block (42 °C), resuspended in 40 μL C:M 1:3 *v*/*v* and mixed with the upper phase. Afterwards, glycosphingolipids- derived oligosaccharides were purified from the samples using C18 columns (Telos, Kinesis, UK) previously pre-equilibrated with 1.25 mL methanol (four times) and 1.25 mL deionized water (three times). We loaded the mixed phase (lower/upper) onto a column and rinsed the sample tube with 1 × 1 mL of deionized water. Then, the C18 column was washed with 4 × 1.25 mL deionized water and eluted it with 1 × 1 mL (C:M) (98:2 *v*/*v*), 2 × 1 mL (C:M) (1:3 *v*/*v*), 1 × 1 mL methanol. The eluates were dried under N2 current and digested with a recombinant Endoglycoceramidase I (rEGCaseI) (GenScript, Oxford, UK) in buffer 50 mM sodium acetate, pH 5.0, 0.6% TritonX-100 (4 μL enzyme + 86 μL buffer) at 37 °C for 16 h. The released glycans were labeled with 310 μL of labelling mix (30 mg/mL anthranilic acid (2AA) and 45 mg/mL sodium cyanoborohydride) in 4% sodium acetate, 2% boric acid in methanol, and heated at 80 °C. Then, we cooled the samples and mixed them with 3 × 1 mL acetonitrile: deionized water (97:3) (*v*/*v*) and added them to a Discovery DPA- 6S-SPE tube (Supelco, PA, USA), pre-equilibrated with 1 × 1 mL acetonitrile, 2 × 1 mL deionized water, and 3 × 1 mL acetonitrile. The columns were cleaned with 3 × 1 mL acetonitrile: deionized water (95:5) (*v*/*v*), and the tubes were washed with 2 × 1 mL ace- tonitrile: deionized water (95:5) (*v*/*v*) and eluted in 0.6 mL deionized water. We took 60 μL from 0.6 mL sample eluted, added 140 μL acetonitrile, and injected 50 μL of this mix (deionized water: acetonitrile) (30:70) (*v*/*v*) onto NP-HPLC (Waters Alliance 2695 separations module and multi-fluorescent detector set at Ex 360/Em 425 nm). To calculate molar quantities from peaks in the chromatogram, we included a calibration standard containing 2.5 pmol 2AA-labelled chitotriose (Ludger, Oxford, UK) for each NP-HPLC run [145]. The chromatographic data were processed using Waters Empower software 3 (Waters, Milford, MA, USA). Fluorescence values by sample were normalized to protein content using a BCA Assay kit (Merck KGaA, Darmstadt, Germany).

4.4. *Genome-Wide Association Studies (GWAS)*

We used the genotype of each strain, and the enzymatic activity or substrate as trait, and its kinship matrix to perform the GWAS using The Efficient Mixed Model Association (EMMA) v.1.1.230 in the R package [26,146]. We used PLINK to remove SNVs in linkage

disequilibrium to avoid false associations [25], considering an R^2 = 0.25, leaving 127,285 independent variants out of the initial four million variants downloaded from the mouse HapMap reference panel (http://mouse.cs.ucla.edu/mousehapmap/full.html, accessed on 28 September 2020) [147].

4.5. Gene Expression Array and Heat Maps

For gene expression correlations, we obtained inbred mouse hepatic transcript data from the repository GSE16780 UCLA Hybrid MDP Liver Affy HTM430A [24]. The mRNA levels in the repository were expressed as log2 transformed and were calculated from the Affimetrix chip with the robust multiarray average (RMA) method. To plot the heatmaps, we used Morpheus software (https://software.broadinstitute.org/morpheus, accessed on 15 February 2022).

4.6. Functional Impact of Genomic Variants

The functional impact of genomic variants was assessed using the Sorting Intoler- ant From Tolerant (SIFT) software (https://sift.bii.a-star.edu.sg/www/SIFT_dbSNP.html, accessed on 12 July 2022) [27].

4.7. Enrichment Analysis

We used gProfiler [35] with the default settings to perform the pathway enrichment analyses.

4.8. Identification of Transcription Factors

We consulted the GeneHancer (GH) database, a catalogue of genome-wide enhancer- to-gene and promoter-to-gene associations, through GeneCards® (https://www.genecards.org/Guide/GeneCard, accessed on 6 September 2022) [148]. Only transcription factors with a significative GH Score were considered.

4.9. Statistics

We used Student's *t*-test, ANOVA with Bonferroni correction, and Pearson corre- lation. All tests were two-tailed. The significance was considered to be $p < 0.05$. We used an R package [146] and Prism v9.1.0 (GraphPad software, San Diego, CA, USA) for these analyses.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24054915/s1.

Author Contributions: Conceptualization, A.D.K.; methodology, A.D., D.A.P. and V.O.; bioinformatics analyses, A.D., B.R.-J., M.L.H. and J.F.C.; writing—original draft preparation, A.D. and A.D.K.; writing—review and editing, all the authors.; supervision, J.F.C., S.Z., J.G., F.M.P. and A.D.K.; funding acquisition, D.A.P., S.Z., F.M.P. and A.D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by ANID-CHILE: Fondecyt grant No 1180337 (A.D.K.) (2018– 2022), 1190334 (S.Z.) (2019–2023), EQM190110 and ACT210012 (J.F.C.), EQM150093 (B.R.-J.) and 1221362 (J.G.) (2022–2025). LysoMod, funded by the European Union's Horizon 2020 research and innovation programme (RISE) under the Marie Sklodowska-Curie grant agreement No 734825 (S.Z., F.M.P., A.D.K.) (2017–2022), and the Mizutani Foundation for Glycoscience, grant No: 200133 (F.M.P., D.A.P., A.D.K.) (2020). F.M.P. is a Wolfson Royal Society Merit Award Holder and a Wellcome Trust Investigator in science.

Institutional Review Board Statement: The animal study protocol was approved by the Institutional Review Board of Universidad del Desarrollo (protocol code number 04-2018 approved on 4 June 2018).

Informed Consent Statement: Not applicable.

Data Availability Statement: We obtained inbred mouse hepatic transcript data from the repository GSE16780 UCLA Hybrid MDP Liver Affy HTM430A reported in [24].

Acknowledgments: The authors thank Aldons J. Lusis from University of California Los Angeles (UCLA) for donating the mouse tissues and María E. Fernández-Suárez and Rodrigo Escalona-Rivano for technical support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mosen, P.; Sanner, A.; Singh, J.; Winter, D. Targeted Quantification of the Lysosomal Proteome in Complex Samples. *Proteomes* **2021**, *9*, 4. [CrossRef] [PubMed]
2. Brubaker, P.L.; Martchenko, A. Metabolic Homeostasis: It's All in the Timing. *Endocrinology* **2021**, *163*, bqab199. [CrossRef] [PubMed]
3. Settembre, C.; Fraldi, A.; Medina, D.L.; Ballabio, A. Signals from the lysosome: A control centre for cellular clearance and energy metabolism. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 283–296. [CrossRef]
4. Ballabio, A.; Bonifacino, J.S. Lysosomes as dynamic regulators of cell and organismal homeostasis. *Nat. Rev. Mol. Cell Biol.* **2019**, *21*, 101–118. [CrossRef]
5. Platt, F.M.; D'Azzo, A.; Davidson, B.; Neufeld, E.F.; Tifft, C.J. Lysosomal storage diseases. *Nat. Rev. Dis. Prim.* **2018**, *4*, 27. [CrossRef] [PubMed]
6. Sidransky, E.; Nalls, M.; Aasly, J.; Aharon-Peretz, J.; Annesi, G.; Barbosa, E.; Bar-Shira, A.; Berg, D.; Bras, J.; Brice, A.; et al. Multicenter Analysis of Glucocerebrosidase Mutations in Parkinson's Disease. *N. Engl. J. Med.* **2009**, *361*, 1651–1661. [CrossRef] [PubMed]
7. Klein, A.D.; Mazzulli, J.R. Is Parkinson's disease a lysosomal disorder? *Brain* **2018**, *141*, 2255–2262. [CrossRef]
8. Platt, F.M. Sphingolipid lysosomal storage disorders. *Nature* **2014**, *510*, 68–75. [CrossRef]
9. Jeyakumar, M.; Butters, T.D.; Cortina-Borja, M.; Hunnam, V.; Proia, R.L.; Perry, V.H.; Dwek, R.A.; Platt, F.M. Delayed symptom onset and increased life expectancy in Sandhoff disease mice treated with *N*-butyldeoxynojirimycin. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6388–6393. [CrossRef]
10. Peterschmitt, M.J.; Crawford, N.P.S.; Gaemers, S.J.M.; Ji, A.J.; Sharma, J.; Pham, T.T. Pharmacokinetics, Pharmacodynamics, Safety, and Tolerability of Oral Venglustat in Healthy Volunteers. *Clin. Pharmacol. Drug Dev.* **2020**, *10*, 86–98. [CrossRef]
11. Kan, S.-H.; Aoyagi-Scharber, M.; Le, S.Q.; Vincelette, J.; Ohmi, K.; Bullens, S.; Wendt, D.J.; Christianson, T.M.; Tiger, P.M.N.; Brown, J.R.; et al. Delivery of an enzyme-IGFII fusion protein to the mouse brain is therapeutic for mucopolysaccharidosis type IIIB. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 14870–14875. [CrossRef] [PubMed]
12. Weesner, J.A.; Annunziata, I.; Yang, T.; Acosta, W.; Gomero, E.; Hu, H.; van de Vlekkert, D.; Ayala, J.; Qiu, X.; Fremuth, L.E.; et al. Preclinical Enzyme Replacement Therapy with a Recombinant β-Galactosidase-Lectin Fusion for CNS Delivery and Treatment of GM1-Gangliosidosis. *Cells* **2022**, *11*, 2579. [CrossRef] [PubMed]
13. Sevin, C.; Deiva, K. Clinical Trials for Gene Therapy in Lysosomal Diseases With CNS Involvement. *Front. Mol. Biosci.* **2021**, *8*, 624988. [CrossRef]
14. Hobbs, J. reversal of clinical features of hurler's disease and biochemical improvement after treatment by bone-marrow transplantation. *Lancet* **1981**, *318*, 709–712. [CrossRef]
15. Dever, D.P.; Scharenberg, S.G.; Camarena, J.; Kildebeck, E.J.; Clark, J.T.; Martin, R.M.; Bak, R.O.; Tang, Y.; Dohse, M.; Birgmeier, J.A.; et al. CRISPR/Cas9 Genome Engineering in Engraftable Human Brain-Derived Neural Stem Cells. *Iscience* **2019**, *15*, 524–535. [CrossRef] [PubMed]
16. Kirkegaard, T.; Gray, J.; Priestman, D.A.; Wallom, K.-L.; Atkins, J.; Olsen, O.D.; Klein, A.; Drndarski, S.; Petersen, N.H.T.; Ingemann, L.; et al. Heat shock protein–based therapy as a potential candidate for treating the sphingolipidoses. *Sci. Transl. Med.* **2016**, *8*, 355ra118. [CrossRef]
17. Wraith, J.E. Limitations of enzyme replacement therapy: Current and future. *J. Inherit. Metab. Dis.* **2006**, *29*, 442–447. [CrossRef]
18. Broomfield, A.; Jones, S.; Hughes, S.M.; Bigger, B. The impact of the immune system on the safety and efficiency of enzyme replacement therapy in lysosomal storage disorders. *J. Inherit. Metab. Dis.* **2016**, *39*, 499–512. [CrossRef]
19. Richardson, J.S.; Kemper, A.R.; Grosse, S.D.; Lam, W.K.K.; Rose, A.M.; Ahmad, A.; Ms, A.G.; Prosser, L.A. Health and economic outcomes of newborn screening for infantile-onset Pompe disease. *Anesthesia Analg.* **2020**, *23*, 758–766. [CrossRef]
20. Durán, A.; Rebolledo-Jaramillo, B.; Olguin, V.; Rojas-Herrera, M.; Heras, M.L.; Calderón, J.F.; Zanlungo, S.; Priestman, D.A.; Platt, F.M.; Klein, A.D. Identification of genetic modifiers of murine hepatic β-glucocerebrosidase activity. *Biochem. Biophys. Rep.* **2021**, *28*, 101105. [CrossRef]
21. Ferland, R.J.; Smith, J.; Papandrea, D.; Gracias, J.; Hains, L.; Kadiyala, S.B.; O'Brien, B.; Kang, E.Y.; Beyer, B.S.; Herron, B.J. Multidimensional Genetic Analysis of Repeated Seizures in the Hybrid Mouse Diversity Panel Reveals a Novel Epileptogenesis Susceptibility Locus. *G3 Genes Genomes Genet.* **2017**, *7*, 2545–2558. [CrossRef] [PubMed]
22. Klein, A.D.; Ferreira, N.-S.; Ben-Dor, S.; Duan, J.; Hardy, J.; Cox, T.M.; Merril, A.H., Jr.; Futerman, A.H. Identification of Modifier Genes in a Mouse Model of Gaucher Disease. *Cell Rep.* **2016**, *16*, 2546–2553. [CrossRef]
23. Ghazalpour, A.; Rau, C.D.; Farber, C.R.; Bennett, B.J.; Orozco, L.D.; Van Nas, A.; Pan, C.; Allayee, H.; Beaven, S.W.; Civelek, M.; et al. Hybrid mouse diversity panel: A panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm. Genome* **2012**, *23*, 680–692. [CrossRef] [PubMed]
24. Bennett, B.J.; Farber, C.R.; Orozco, L.; Kang, H.M.; Ghazalpour, A.; Siemers, N.; Neubauer, M.; Neuhaus, I.; Yordanova, R.; Guan, B.; et al. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res.* **2010**, *20*, 281–290. [CrossRef]
25. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]

26. Kang, H.M.; Zaitlen, N.A.; Wade, C.M.; Kirby, A.; Heckerman, D.; Daly, M.J.; Eskin, E. Efficient Control of Population Structure in Model Organism Association Mapping. *Genetics* **2008**, *178*, 1709–1723. [CrossRef] [PubMed]
27. Vaser, R.; Adusumalli, S.; Leng, S.N.; Sikic, M.; Ng, P.C. SIFT missense predictions for genomes. *Nat. Protoc.* **2015**, *11*, 1–9. [CrossRef]
28. Keane, T.M.; Goodstadt, L.; Danecek, P.; White, M.A.; Wong, K.; Yalcin, B.; Heger, A.; Agam, A.; Slater, G.; Goodson, M.; et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **2011**, *477*, 289–294. [CrossRef]
29. An, J.; Shi, J.; He, Q.; Lui, K.; Liu, Y.; Huang, Y.; Sheikh, M.S. CHCM1/CHCHD6, Novel Mitochondrial Protein Linked to Regulation of Mitofilin and Mitochondrial Cristae Morphology. *J. Biol. Chem.* **2012**, *287*, 7411–7426. [CrossRef]
30. Verde, I.; Pahlke, G.; Salanova, M.; Zhang, G.; Wang, S.; Coletti, D.; Onuffer, J.; Jin, S.-L.C.; Conti, M. Myomegalin Is a Novel Protein of the Golgi/Centrosome That Interacts with a Cyclic Nucleotide Phosphodiesterase. *J. Biol. Chem.* **2001**, *276*, 11189–11198. [CrossRef]
31. Marks, M.; Pennimpede, T.; Lange, L.; Grote, P.; Herrmann, B.G.; Wittler, L. Analysis of the Fam181 gene family during mouse development reveals distinct strain-specific expression patterns, suggesting a role in nervous system development and function. *Gene* **2015**, *575*, 438–451. [CrossRef]
32. Li, H.; Rauch, T.; Chen, Z.-X.; Szabó, P.E.; Riggs, A.D.; Pfeifer, G.P. The Histone Methyltransferase SETDB1 and the DNA Methyltransferase DNMT3A Interact Directly and Localize to Promoters Silenced in Cancer Cells. *J. Biol. Chem.* **2006**, *281*, 19489–19500. [CrossRef] [PubMed]
33. Beaubien, F.; Raja, R.; Kennedy, T.E.; Fournier, A.E.; Cloutier, J.-F. Slitrk1 is localized to excitatory synapses and promotes their development. *Sci. Rep.* **2016**, *6*, 27343. [CrossRef] [PubMed]
34. Hannun, Y.A.; Obeid, L.M. Sphingolipids and their metabolism in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **2017**, *19*, 175–191. [CrossRef]
35. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J. g: Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **2019**, *47*, W191–W198. [CrossRef] [PubMed]
36. Gleyzer, N.; Vercauteren, K.; Scarpulla, R.C. Control of Mitochondrial Transcription Specificity Factors (TFB1M and TFB2M) by Nuclear Respiratory Factors (NRF-1 and NRF-2) and PGC-1 Family Coactivators. *Mol. Cell. Biol.* **2005**, *25*, 1354–1366. [CrossRef] [PubMed]
37. Lee, W.-H.; Higuchi, H.; Ikeda, S.; Macke, E.L.; Takimoto, T.; Pattnaik, B.R.; Liu, C.; Chu, L.-F.; Siepka, S.M.; Krentz, K.J.; et al. Mouse Tmem135 mutation reveals a mechanism involving mitochondrial dynamics that leads to age-dependent retinal pathologies. *eLife* **2016**, *5*, e19264. [CrossRef]
38. Zhang, D.; Ma, X.; Sun, W.; Cui, P.; Lu, Z. Down-regulated FSTL5 promotes cell proliferation and survival by affecting Wnt/β-catenin signaling in hepatocellular carcinoma. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 3386–3394.
39. Alrefaei, A.F.; Münsterberg, A.E.; Wheeler, G.N. FZD10 regulates cell proliferation and mediates Wnt1 induced neurogenesis in the developing spinal cord. *PLoS ONE* **2020**, *15*, e0219721. [CrossRef]
40. Maeda, M.; Hasegawa, H.; Hyodo, T.; Ito, S.; Asano, E.; Yuang, H.; Funasaka, K.; Shimokata, K.; Hasegawa, Y.; Hamaguchi, M.; et al. ARHGAP18, a GTPase-activating protein for RhoA, controls cell shape, spreading, and motility. *Mol. Biol. Cell* **2011**, *22*, 3840–3852. [CrossRef]
41. Dunne, E.; Spring, C.M.; Reheman, A.; Jin, W.; Berndt, M.C.; Newman, D.K.; Newman, P.J.; Ni, H.; Kenny, D. Cadherin 6 Has a Functional Role in Platelet Aggregation and Thrombus Formation. *Arter. Thromb. Vasc. Biol.* **2012**, *32*, 1724–1731. [CrossRef]
42. Frey, L.; Zie˛tara, N.; Łyszkiewicz, M.; Marquardt, B.; Mizoguchi, Y.; Linder, M.I.; Liu, Y.; Giesert, F.; Wurst, W.; Dahlhoff, M.; et al. Mammalian VPS45 orchestrates trafficking through the endosomal system. *Blood* **2021**, *137*, 1932–1944. [CrossRef]
43. Koyanagi, M.; Nakabayashi, K.; Fujimoto, T.; Gu, N.; Baba, I.; Takashima, Y.; Doi, K.; Harada, H.; Kato, N.; Sasazuki, T.; et al. ZFAT expression in B and T lymphocytes and identification of ZFAT-regulated genes. *Genomics* **2008**, *91*, 451–457. [CrossRef]
44. Jin, K.; Wang, S.; Zhang, Y.; Xia, M.; Mo, Y.; Li, X.; Li, G.; Zeng, Z.; Xiong, W.; He, Y. Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell. Mol. Life Sci.* **2019**, *76*, 4275–4289. [CrossRef] [PubMed]
45. Tang, T.; Yang, Z.-Y.; Wang, D.; Yang, X.-Y.; Wang, J.; Li, L.; Wen, Q.; Gao, L.; Bian, X.-W.; Yu, S.-C. The role of lysosomes in cancer development and progression. *Cell Biosci.* **2020**, *10*, 131. [CrossRef] [PubMed]
46. Davidson, S.M.; Vander Heiden, M.G. Critical Functions of the Lysosome in Cancer Biology. *Annu. Rev. Pharmacol. Toxicol.* **2017**, *57*, 481–507. [CrossRef]
47. Chen, J.-S.; Su, I.-J.; Leu, Y.-W.; Young, K.-C.; Sun, H.S. Expression of T-cell lymphoma invasion and metastasis 2 (TIAM2) promotes proliferation and invasion of liver cancer. *Int. J. Cancer* **2011**, *130*, 1302–1313. [CrossRef] [PubMed]
48. Terasaki, H.; Saitoh, T.; Shiokawa, K.; Katoh, M. Frizzled-10, up-regulated in primary colorectal cancer, is a positive regulator of the WNT—Beta-catenin—TCF signaling pathway. *Int. J. Mol. Med.* **2002**, *9*, 107–112. [PubMed]
49. Zhao, Z.; Li, S.; Li, S.; Wang, J.; Lin, H.; Fu, W. High expression of oncogene cadherin-6 correlates with tumor progression and a poor prognosis in gastric cancer. *Cancer Cell Int.* **2021**, *21*, 493. [CrossRef]
50. Cui, M.; Chang, Y.; Fang, Q.-G.; Du, W.; Wu, J.-F.; Wang, J.-H.; Liu, S.-T.; Luo, S.-X. Non-Coding RNA Pvt1 Promotes Cancer Stem Cell–Like Traits in Nasopharyngeal Cancer via Inhibiting miR-1207. *Pathol. Oncol. Res.* **2018**, *25*, 1411–1422. [CrossRef]

51. Elotan, A.; Efenckova, M.; Ebralten, J.; Ealttoa, A.; Edixson, L.; Williams, R.W.; van der Voet, M. Neuroinformatic analyses of common and distinct genetic components associated with major neuropsychiatric disorders. *Front. Neurosci.* **2014**, *8*, 331. [CrossRef]
52. Kim, S.; Cho, H.; Lee, D.; Webster, M.J. Association between SNPs and gene expression in multiple regions of the human brain. *Transl. Psychiatry* **2012**, *2*, e113. [CrossRef] [PubMed]
53. Hsu, S.-D.; Lin, F.-M.; Wu, W.-Y.; Liang, C.; Huang, W.-C.; Chan, W.-L.; Tsai, W.-T.; Chen, G.-Z.; Lee, C.-J.; Chiu, C.-M.; et al. miRTarBase: A database curates experimentally validated microRNA–target interactions. *Nucleic Acids Res.* **2010**, *39*, D163–D169. [CrossRef] [PubMed]
54. Chu, C.; Chen, J.; Chuang, P.; Su, C.; Chan, Y.-L.; Yang, Y.; Chiang, Y.; Su, Y.; Gean, P.; Sun, H.S. TIAM2S as a novel regulator for serotonin level enhances brain plasticity and locomotion behavior. *FASEB J.* **2020**, *34*, 3267–3288. [CrossRef] [PubMed]
55. Bykhovskaya, Y.; Mengesha, E.; Wang, D.; Yang, H.; Estivill, X.; Shohat, M.; Fischel-Ghodsian, N. Human mitochondrial transcription factor B1 as a modifier gene for hearing loss associated with the mitochondrial A1555G mutation. *Mol. Genet. Metab.* **2004**, *82*, 27–32. [CrossRef] [PubMed]
56. Fulzele, K.; Riddle, R.C.; DiGirolamo, D.J.; Cao, X.; Wan, C.; Chen, D.; Faugere, M.-C.; Aja, S.; Hussain, M.A.; Brüning, J.C.; et al. Insulin Receptor Signaling in Osteoblasts Regulates Postnatal Bone Acquisition and Body Composition. *Cell* **2022**, *185*, 746. [CrossRef] [PubMed]
57. Dearth, R.K.; Cui, X.; Kim, H.-J.; Kuiatse, I.; Lawrence, N.A.; Zhang, X.; Divisova, J.; Britton, O.L.; Mohsin, S.; Allred, D.C.; et al. Mammary Tumorigenesis and Metastasis Caused by Overexpression of Insulin Receptor Substrate 1 (IRS-1) or IRS-2. *Mol. Cell. Biol.* **2006**, *26*, 9302–9314. [CrossRef]
58. Ochiai, T.; Sano, T.; Nagayama, T.; Kubota, N.; Kadowaki, T.; Wakabayashi, T.; Iwatsubo, T. Differential involvement of insulin receptor substrate (IRS)-1 and IRS-2 in brain insulin signaling is associated with the effects on amyloid pathology in a mouse model of Alzheimer's disease. *Neurobiol. Dis.* **2021**, *159*, 105510. [CrossRef]
59. Xu, L.; Wu, J.; Yu, Y.; Li, H.; Sun, S.; Zhang, T.; Wang, M. Dok5 regulates proliferation and differentiation of osteoblast via ca-nonical Wnt/β-catenin signaling. *J. Musculoskelet. Neuronal Interact.* **2022**, *22*, 113–122.
60. Landowski, M.; Bhute, V.J.; Takimoto, T.; Grindel, S.; Shahi, P.K.; Pattnaik, B.R.; Ikeda, S.; Ikeda, A. A mutation in transmembrane protein 135 impairs lipid metabolism in mouse eyecups. *Sci. Rep.* **2022**, *12*, 756. [CrossRef]
61. Suzuki, N.; Fukushi, M.; Kosaki, K.; Doyle, A.D.; De Vega, S.; Yoshizaki, K.; Akazawa, C.; Arikawa-Hirasawa, E.; Yamada, Y. Teneurin-4 Is a Novel Regulator of Oligodendrocyte Differentiation and Myelination of Small-Diameter Axons in the CNS. *J. Neurosci.* **2012**, *32*, 11586–11599. [CrossRef] [PubMed]
62. Pu, J.-L.; Gao, T.; Si, X.-L.; Zheng, R.; Jin, C.-Y.; Ruan, Y.; Fang, Y.; Chen, Y.; Song, Z.; Yin, X.-Z.; et al. Parkinson's Disease in Teneurin Transmembrane Protein 4 (TENM4) Mutation Carriers. *Front. Genet.* **2020**, *11*, 598064. [CrossRef]
63. Hor, H.; Francescatto, L.; Bartesaghi, L.; Ortega-Cubero, S.; Kousi, M.; Lorenzo-Betancor, O.; Jiménez-Jiménez, F.J.; Gironell, A.; Clarimón, J.; Drechsel, O.; et al. Missense mutations inTENM4, a regulator of axon guidance and central myelination, cause essential tremor. *Hum. Mol. Genet.* **2015**, *24*, 5677–5686. [CrossRef] [PubMed]
64. Kant, T.; Newe, M.; Winter, L.; Hoffmann, M.; Kämmerer, S.; Klapproth, E.; Künzel, K.; Kühnel, M.; Neubert, L.; El-Armouche, A.; et al. Genetic Deletion of Polo-Like Kinase 2 Induces a Pro-Fibrotic Pulmonary Phenotype. *Cells* **2021**, *10*, 617. [CrossRef] [PubMed]
65. Inglis, K.J.; Chereau, D.; Brigham, E.F.; Chiou, S.-S.; Schöbel, S.; Frigon, N.L.; Yu, M.; Caccavello, R.J.; Nelson, S.; Motter, R.; et al. Polo-like Kinase 2 (PLK2) Phosphorylates α-Synuclein at Serine 129 in Central Nervous System. *J. Biol. Chem.* **2009**, *284*, 2598–2602. [CrossRef] [PubMed]
66. Dong, J.; Hu, Z.; Wu, C.; Guo, H.; Zhou, B.; Lv, J.; Lu, D.; Chen, K.; Shi, Y.; Chu, M.; et al. Association analyses identify multiple new lung cancer susceptibility loci and their interactions with smoking in the Chinese population. *Nat. Genet.* **2012**, *44*, 895–899. [CrossRef]
67. Manning, G.; Whyte, D.B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298*, 1912–1934. [CrossRef]
68. Manivannan, J.; Tay, S.; Ling, E.-A.; Dheen, S. Dihydropyrimidinase-like 3 regulates the inflammatory response of activated microglia. *Neuroscience* **2013**, *253*, 40–54. [CrossRef]
69. Kanda, M.; Nomoto, S.; Oya, H.; Shimizu, D.; Takami, H.; Hibino, S.; Hashimoto, R.; Kobayashi, D.; Tanaka, C.; Yamada, S.; et al. Dihydropyrimidinase-like 3 facilitates malignant behavior of gastric cancer. *J. Exp. Clin. Cancer Res.* **2014**, *33*, 66. [CrossRef]
70. Blasco, H.; Bernard-Marissal, N.; Vourc'H, P.; Guettard, Y.O.; Sunyach, C.; Augereau, O.; Khederchah, J.; Mouzat, K.; Antar, C.; Gordon, P.H.; et al. A Rare Motor Neuron Deleterious Missense Mutation in the*DPYSL3*(*CRMP4*) Gene is Associated with ALS. *Hum. Mutat.* **2013**, *34*, 953–960. [CrossRef]
71. Damoulakis, G.; Gambardella, L.; Rossman, K.L.; Lawson, C.; Anderson, K.; Fukui, Y.; Welch, H.C.; Der, C.; Stephens, L.; Hawkins, P.T. P-Rex1 directly activates RhoG to regulate GPCR-driven Rac signalling and actin polarity in neutrophils. *J. Cell Sci.* **2014**, *127*, 2589–2600. [CrossRef]
72. Qin, J.; Xie, Y.; Wang, B.; Hoshino, M.; Wolff, D.W.; Zhao, J.; Scofield, M.A.; Dowd, F.J.; Lin, M.-F.; Tu, Y. Upregulation of PIP3-dependent Rac exchanger 1 (P-Rex1) promotes prostate cancer metastasis. *Oncogene* **2009**, *28*, 1853–1863. [CrossRef] [PubMed]

73. Zhang, D.-Y.; Lei, J.-S.; Sun, W.-L.; Wang, D.-D.; Lu, Z. Follistatin Like 5 (FSTL5) inhibits epithelial to mesenchymal transition in hepatocellular carcinoma. *Chin. Med. J.* **2020**, *133*, 1798–1804. [CrossRef] [PubMed]
74. Vilboux, T.; Lev, A.; Malicdan, M.C.V.; Simon, A.J.; Järvinen, P.; Racek, T.; Puchalka, J.; Sood, R.; Carrington, B.; Bishop, K.; et al. A Congenital Neutrophil Defect Syndrome Associated with Mutations in *VPS45*. *N. Engl. J. Med.* **2013**, *369*, 54–65. [CrossRef] [PubMed]
75. Nayak, S.R.; Harrington, E.; Boone, D.; Hartmaier, R.; Chen, J.; Pathiraja, T.N.; Cooper, K.L.; Fine, J.L.; Sanfilippo, J.; Davidson, N.E.; et al. A Role for Histone H2B Variants in Endocrine-Resistant Breast Cancer. *Horm. Cancer* **2015**, *6*, 214–224. [CrossRef]
76. Shan, Z.; Shakoori, A.; Bodaghi, S.; Goldsmith, P.; Jin, J.; Wiest, J.S. TUSC1, a Putative Tumor Suppressor Gene, Reduces Tumor Cell Growth In Vitro and Tumor Growth In Vivo. *PLoS ONE* **2013**, *8*, e66114. [CrossRef]
77. Zhang, R.; Yu, W.; Liang, G.; Jia, Z.; Chen, Z.; Zhao, L.; Yuan, Y.; Zhou, X.; Li, D.; Shen, S.; et al. Tumor Suppressor Candidate 1 Suppresses Cell Growth and Predicts Better Survival in Glioblastoma. *Cell. Mol. Neurobiol.* **2016**, *37*, 37–42. [CrossRef]
78. Grigorenko, A.P.; Protasova, M.S.; Lisenkova, A.A.; Reshetov, D.A.; Andreeva, T.V.; Garcias, G.D.L.; Roth, M.D.G.M.; Papassotiropoulos, A.; Rogaev, E.I. Neurodevelopmental Syndrome with Intellectual Disability, Speech Impairment, and Quadrupedia Is Associated with Glutamate Receptor Delta 2 Gene Defect. *Cells* **2022**, *11*, 400. [CrossRef]
79. Matsushita, K.; Takeuchi, O.; Standley, D.M.; Kumagai, Y.; Kawagoe, T.; Miyake, T.; Satoh, T.; Kato, H.; Tsujimura, T.; Nakamura, H.; et al. Zc3h12a is an RNase essential for controlling immune responses by regulating mRNA decay. *Nature* **2009**, *458*, 1185–1190. [CrossRef]
80. Tsoi, L.C.; Spain, S.L.; Knight, J.; Ellinghaus, E.; Stuart, P.E.; Capon, F.; Ding, J.; Li, Y.; Tejasvi, T.; Gudjonsson, J.E.; et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat. Genet.* **2012**, *44*, 1341–1348. [CrossRef]
81. Guo, W.; Cai, Y.; Zhang, H.; Yang, Y.; Yang, G.; Wang, X.; Zhao, J.; Lin, J.; Zhu, J.; Li, W.; et al. Association of ARHGAP18 polymorphisms with schizophrenia in the Chinese-Han population. *PLoS ONE* **2017**, *12*, e0175209. [CrossRef]
82. Negredo, P.N.; Edgar, J.R.; Manna, P.T.; Antrobus, R.; Robinson, M.S. The WDR11 complex facilitates the tethering of AP-1-derived vesicles. *Nat. Commun.* **2018**, *9*, 596. [CrossRef]
83. Ouyang, W.-W.; Li, Q.-Y.; Yang, W.-G.; Su, S.-F.; Wu, L.-J.; Yang, Y.; Lu, B. Genetic characteristics of a patient with multiple primary cancers: A case report. *World J. Clin. Cases* **2021**, *9*, 8563–8570. [CrossRef] [PubMed]
84. Abdul-Rahman, P.S.; Lim, B.-K.; Hashim, O.H. Expression of high-abundance proteins in sera of patients with endometrial and cervical cancers: Analysis using 2-DE with silver staining and lectin detection methods. *Electrophoresis* **2007**, *28*, 1989–1996. [CrossRef] [PubMed]
85. Sun, D.; Zhao, Y.Y.; Dai, S.P.; Fang, K.; Dong, L.Y.; Ding, Q. Cloning and analysis of human alpha-1B glycoprotein precursor gene: A novel member of human immunoglobulin superfamily. *Yi Chuan Xue Bao* **2002**, *29*, 299–302. (In Japanese) [PubMed]
86. Kreunin, P.; Zhao, J.; Rosser, C.; Urquidi, V.; Lubman, D.M.; Goodison, S. Bladder Cancer Associated Glycoprotein Signatures Revealed by Urinary Proteomic Profiling. *J. Proteome Res.* **2007**, *6*, 2631–2639. [CrossRef] [PubMed]
87. Tang, J.; Li, Y.; Sang, Y.; Yu, B.; Lv, D.; Zhang, W.; Feng, H. LncRNA PVT1 regulates triple-negative breast cancer through KLF5/beta-catenin signaling. *Oncogene* **2018**, *37*, 4723–4734. [CrossRef]
88. Glasson, S.S.; Askew, R.; Sheppard, B.; Carito, B.; Blanchet, T.; Ma, H.-L.; Flannery, C.R.; Peluso, D.; Kanki, K.; Yang, Z.; et al. Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature* **2005**, *434*, 644–648. [CrossRef] [PubMed]
89. Agajanian, M.J.; Potjewyd, F.M.; Bowman, B.M.; Solomon, S.; LaPak, K.M.; Bhatt, D.P.; Smith, J.L.; Goldfarb, D.; Axtman, A.D.; Major, M.B. Protein proximity networks and functional evaluation of the casein kinase 1 gamma family reveal unique roles for CK1γ3 in WNT signaling. *J. Biol. Chem.* **2022**, *298*, 101986. [CrossRef]
90. Schittek, B.; Sinnberg, T. Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis. *Mol. Cancer* **2014**, *13*, 231. [CrossRef]
91. Inoue, N.; Watanabe, M.; Yamada, H.; Takemura, K.; Hayashi, F.; Yamakawa, N.; Akahane, M.; Shimizuishi, Y.; Hidaka, Y.; Iwatani, Y. Associations between autoimmune thyroid disease prognosis and functional polymorphisms of susceptibility genes, CTLA4, PTPN22, CD40, FCRL3, and ZFAT, previously revealed in genome-wide association studies. *J. Clin. Immunol.* **2012**, *32*, 1243–1252. [CrossRef]
92. Wallom, K.-L.; Fernández-Suárez, M.E.; Priestman, D.A.; Vruchte, D.T.; Huebecker, M.; Hallett, P.J.; Isacson, O.; Platt, F.M. Glycosphingolipid metabolism and its role in ageing and Parkinson's disease. *Glycoconj. J.* **2021**, *39*, 39–53. [CrossRef] [PubMed]
93. Ryckman, A.E.; Brockhausen, I.; Walia, J.S. Metabolism of Glycosphingolipids and Their Role in the Pathophysiology of Lysosomal Storage Disorders. *Int. J. Mol. Sci.* **2020**, *21*, 6881. [CrossRef]
94. Sardiello, M.; Palmieri, M.; Di Ronza, A.; Medina, D.L.; Valenza, M.; Gennarino, V.A.; Di Malta, C.; Donaudy, F.; Embrione, V.; Polishchuk, R.S.; et al. A Gene Network Regulating Lysosomal Biogenesis and Function. *Science* **2009**, *325*, 473–477. [CrossRef] [PubMed]
95. Olguín, V.; Durán, A.; Heras, M.L.; Rubilar, J.C.; Cubillos, F.A.; Olguín, P.; Klein, A.D. Genetic Background Matters: Population-Based Studies in Model Organisms for Translational Research. *Int. J. Mol. Sci.* **2022**, *23*, 7570. [CrossRef]
96. Rha, A.K.; Maguire, A.S.; Martin, D.R. GM1 Gangliosidosis: Mechanisms and Management. *Appl. Clin. Genet.* **2021**, *14*, 209–233. [CrossRef]

97. Adibekian, A.; Martin, B.; Chang, J.; Hsu, K.-L.; Tsuboi, K.; Bachovchin, D.; Speers, A.; Brown, S.; Spicer, T.; Fernandez-Vega, V.; et al. Characterization of a Selective, Reversible Inhibitor of Lysophospholipase 1 (LYPLA1). In *Probe Reports from the NIH Molecular Libraries Program [Internet]*; National Center for Biotechnology Information (US): Bethesda, MD, USA, 2010.
98. Wu, S.; Cao, R.; Tao, B.; Wu, P.; Peng, C.; Gao, H.; Liang, J.; Yang, W. Pyruvate Facilitates FACT-Mediated γ H2AX Loading to Chromatin and Promotes the Radiation Resistance of Glioblastoma. *Adv. Sci.* **2022**, *9*, 2104055. [CrossRef]
99. Merrill, A.H., Jr. Sphingolipid and Glycosphingolipid Metabolic Pathways in the Era of Sphingolipidomics. *Chem. Rev.* **2011**, *111*, 6387–6422. [CrossRef]
100. Lachmann, R.H.; Vruchte, D.T.; Lloyd-Evans, E.; Reinkensmeier, G.; Sillence, D.; Fernandez-Guillen, L.; Dwek, R.A.; Butters, T.D.; Cox, T.M.; Platt, F.M. Treatment with miglustat reverses the lipid-trafficking defect in Niemann–Pick disease type C. *Neurobiol. Dis.* **2004**, *16*, 654–658. [CrossRef] [PubMed]
101. Elstein, D.; Hollak, C.; Aerts, J.M.F.G.; van Weely, S.; Maas, M.; Cox, T.M.; Lachmann, R.H.; Hrebicek, M.; Platt, F.M.; Butters, T.D.; et al. Sustained therapeutic effects of oral miglustat (Zavesca,*N*-butyldeoxynojirimycin, OGT 918) in type I Gaucher disease. *J. Inherit. Metab. Dis.* **2004**, *27*, 757–766. [CrossRef] [PubMed]
102. Meng, Q.; Hu, X.; Zhao, X.; Kong, X.; Meng, Y.-M.; Chen, Y.; Su, L.; Jiang, X.; Qiu, X.; Huang, C.; et al. A circular network of coregulated sphingolipids dictates lung cancer growth and progression. *Ebiomedicine* **2021**, *66*, 103301. [CrossRef]
103. Pazoki, R.; Vujkovic, M.; Elliott, J.; Evangelou, E.; Gill, D.; Ghanbari, M.; van der Most, P.J.; Pinto, R.C.; Wielscher, M.; Farlik, M.; et al. Genetic analysis in European ancestry individuals identifies 517 loci associated with liver enzymes. *Nat. Commun.* **2021**, *12*, 2579. [CrossRef] [PubMed]
104. Willer, C.J.; Schmidt, E.M.; Sengupta, S.; Peloso, G.M.; Gustafsson, S.; Kanoni, S.; Ganna, A.; Chen, J.; Buchkovich, M.L.; Mora, S.; et al. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **2013**, *45*, 1274–1283. [CrossRef] [PubMed]
105. Lee, J.-H.; Yang, D.-S.; Goulbourne, C.N.; Im, E.; Stavrides, P.; Pensalfini, A.; Chan, H.; Bouchet-Marquis, C.; Bleiwas, C.; Berg, M.J.; et al. Faulty autolysosome acidification in Alzheimer's disease mouse models induces autophagic build-up of Aβ in neurons, yielding senile plaques. *Nat. Neurosci.* **2022**, *25*, 688–701. [CrossRef]
106. Stahl-Meyer, J.; Stahl-Meyer, K.; Jäättelä, M. Control of mitosis, inflammation, and cell motility by limited leakage of lysosomes. *Curr. Opin. Cell Biol.* **2021**, *71*, 29–37. [CrossRef] [PubMed]
107. Tan, J.X.; Finkel, T. A phosphoinositide signalling pathway mediates rapid lysosomal repair. *Nature* **2022**, *609*, 815–821. [CrossRef]
108. Alessenko, A.V.; Albi, E. Exploring Sphingolipid Implications in Neurodegeneration. *Front. Neurol.* **2020**, *11*, 437. [CrossRef]
109. Cumin, C.; Huang, Y.-L.; Rossdam, C.; Ruoff, F.; Céspedes, S.P.; Liang, C.-Y.; Lombardo, F.C.; Coelho, R.; Rimmer, N.; Konantz, M.; et al. Glycosphingolipids are mediators of cancer plasticity through independent signaling pathways. *Cell Rep.* **2022**, *40*, 111181. [CrossRef]
110. Snaebjornsson, M.T.; Janaki-Raman, S.; Schulze, A. Greasing the Wheels of the Cancer Machine: The Role of Lipid Metabolism in Cancer. *Cell Metab.* **2019**, *31*, 62–76. [CrossRef]
111. Petersen, N.H.; Olsen, O.D.; Groth-Pedersen, L.; Ellegaard, A.-M.; Bilgin, M.; Redmer, S.; Ostenfeld, M.S.; Ulanet, D.; Dovmark, T.H.; Lønborg, A.; et al. Transformation-Associated Changes in Sphingolipid Metabolism Sensitize Cells to Lysosomal Cell Death Induced by Inhibitors of Acid Sphingomyelinase. *Cancer Cell* **2013**, *24*, 379–393. [CrossRef]
112. Fatehi-Hassanabad, Z.; Chan, C.B. Transcriptional regulation of lipid metabolism by fatty acids: A key determinant of pancreatic β-cell function. *Nutr. Metab.* **2005**, *2*, 1. [CrossRef]
113. Liu, F.; Zhu, X.; Jiang, X.; Li, S.; Lv, Y. Transcriptional control by HNF-1: Emerging evidence showing its role in lipid metabolism and lipid metabolism disorders. *Genes Dis.* **2021**, *9*, 1248–1257. [CrossRef] [PubMed]
114. Pan, G.; Diamanti, K.; Cavalli, M.; Gutiérrez, A.L.; Komorowski, J.; Wadelius, C. Multifaceted regulation of hepatic lipid metabolism by YY1. *Life Sci. Alliance* **2021**, *4*, e202000928. [CrossRef] [PubMed]
115. Magee, N.; Zhang, Y. Role of early growth response 1 in liver metabolism and liver cancer. *Hepatoma Res.* **2017**, *3*, 268–277. [CrossRef] [PubMed]
116. Banerjee, A.; Mahata, B.; Dhir, A.; Mandal, T.K.; Biswas, K. Elevated histone H3 acetylation and loss of the Sp1–HDAC1 complex de-repress the GM2-synthase gene in renal cell carcinoma. *J. Biol. Chem.* **2019**, *294*, 1005–1018. [CrossRef] [PubMed]
117. Martina, J.A.; Diab, H.I.; Lishu, L.; Jeong-A, L.; Patange, S.; Raben, N.; Puertollano, R. The Nutrient-Responsive Transcription Factor TFE3 Promotes Autophagy, Lysosomal Biogenesis, and Clearance of Cellular Debris. *Sci. Signal.* **2014**, *7*, ra9. [CrossRef] [PubMed]
118. Di Malta, C.; Siciliano, D.; Calcagni, A.; Monfregola, J.; Punzi, S.; Pastore, N.; Eastes, A.N.; Davis, O.; De Cegli, R.; Zampelli, A.; et al. Transcriptional activation of RagD GTPase controls mTORC1 and promotes cancer growth. *Science* **2017**, *356*, 1188–1192. [CrossRef]
119. Polager, S.; Ofir, M.; Ginsberg, D. E2F1 regulates autophagy and the transcription of autophagy genes. *Oncogene* **2008**, *27*, 4860–4864. [CrossRef]
120. Seok, S.; Fu, T.; Choi, S.-E.; Li, Y.; Zhu, R.; Kumar, S.; Sun, X.; Yoon, G.; Kang, Y.; Zhong, W.; et al. Transcriptional regulation of autophagy by an FXR–CREB axis. *Nature* **2014**, *516*, 108–111. [CrossRef]
121. Annunziata, I.; van de Vlekkert, D.; Wolf, E.; Finkelstein, D.; Neale, G.; Machado, E.; Mosca, R.; Campos, Y.; Tillman, H.; Roussel, M.F.; et al. MYC competes with MiT/TFE in regulating lysosomal biogenesis and autophagy through an epigenetic rheostat. *Nat. Commun.* **2019**, *10*, 3623. [CrossRef]

122. Song, N.; Song, R.; Ma, P. MiR-340-5p alleviates neuroinflammation and neuronal injury via suppressing STING in subarachnoid hemorrhage. *Brain Behav.* **2022**, *12*, e2687. [CrossRef] [PubMed]
123. Guo, Y.; Gao, Y.; Liu, S. lncRNA XIST is associated with preeclampsia and mediates trophoblast cell invasion via miR-340-5p/KCNJ16 signaling pathway. *Transpl. Immunol.* **2022**, *74*, 101666. [CrossRef]
124. Ou, D.; Wu, Y.; Zhang, J.; Liu, J.; Liu, Z.; Shao, M.; Guo, X.; Cui, S. miR-340-5p affects oral squamous cell carcinoma (OSCC) cells proliferation and invasion by targeting endoplasmic reticulum stress proteins. *Eur. J. Pharmacol.* **2022**, *920*, 174820. [CrossRef] [PubMed]
125. Huang, Z.; Xu, Y.; Wan, M.; Zeng, X.; Wu, J. miR-340: A multifunctional role in human malignant diseases. *Int. J. Biol. Sci.* **2021**, *17*, 236–246. [CrossRef] [PubMed]
126. Liu, Y.; Li, X.; Zhang, Y.; Wang, H.; Rong, X.; Peng, J.; He, L.; Peng, Y. An miR-340-5p-macrophage feedback loop modulates the progression and tumor microenvironment of glioblastoma multiforme. *Oncogene* **2019**, *38*, 7399–7415. [CrossRef] [PubMed]
127. Liu, H.; Li, B.; Qiao, L.; Liu, J.; Ren, D.; Liu, W. miR-340-5p inhibits sheep adipocyte differentiation by targeting *ATF7*. *Anim. Sci. J.* **2020**, *91*, e13462. [CrossRef]
128. Zhu, Y.; Yang, X.; Zhou, J.; Chen, L.; Zuo, P.; Chen, L.; Jiang, L.; Li, T.; Wang, D.; Xu, Y.; et al. miR-340-5p Mediates Cardiomyocyte Oxidative Stress in Diabetes-Induced Cardiac Dysfunction by Targeting Mcl-1. *Oxidative Med. Cell. Longev.* **2022**, *2022*, 3182931. [CrossRef]
129. Priestman, D.A.; Van Der Spoel, A.C.; Butters, T.D.; Dwek, R.A.; Platt, F.M. N-butyldeoxynojirimycin causes weight loss as a result of appetite suppression in lean and obese mice. *Diabetes Obes. Metab.* **2007**, *10*, 159–166. [CrossRef]
130. Chatterjee, S.; Zheng, L.; Ma, S.; Bedja, D.; Bandaru, V.V.R.; Kim, G.; Rangecroft, A.; Iocco, D.; Campbell, S.A. Management of metabolic syndrome and reduction in body weight in type II diabetic mice by inhibiting glycosphingolipid synthesis. *Biochem. Biophys. Res. Commun.* **2020**, *525*, 455–461. [CrossRef]
131. Robinson, D.S. The fluorimetric determination of β-glucosidase: Its occurrence in the tissues of animals, including insects. *Biochem. J.* **1956**, *63*, 39–44. [CrossRef]
132. Shigeto, S.; Katafuchi, T.; Okada, Y.; Nakamura, K.; Endo, F.; Okuyama, T.; Takeuchi, H.; Kroos, M.A.; Verheijen, F.W.; Reuser, A.J.; et al. Improved assay for differential diagnosis between Pompe disease and acid α-glucosidase pseudodeficiency on dried blood spots. *Mol. Genet. Metab.* **2011**, *103*, 12–17. [CrossRef] [PubMed]
133. Beutler, E.; Kuhl, W. Purification and properties of human alpha-galactosidases. *J. Biol. Chem.* **1972**, *247*, 7195–7200. [CrossRef] [PubMed]
134. Mayes, J.S.; Scheerer, J.B.; Sifers, R.N.; Donaldson, M.L. Differential assay for lysosomal alpha-galactosidases in human tissues and its application to Fabry's disease. *Clin. Chim. Acta Int. J. Clin. Chem.* **1981**, *112*, 247–251. [CrossRef]
135. Padeh, B.; Navon, R. Diagnosis of Tay-Sachs disease by hexosaminidase activity in leukocytes and amniotic fluid cells. *Isr. J. Med. Sci.* **1971**, *7*, 259–263.
136. Hauser, E.C.; Kasperzyk, J.L.; D'Azzo, A.; Seyfried, T.N. Inheritance of Lysosomal Acid β-Galactosidase Activity and Gangliosides in Crosses of DBA/2J and Knockout Mice. *Biochem. Genet.* **2004**, *42*, 241–257. [CrossRef] [PubMed]
137. Potier, M.; Mameli, L.; Bélisle, M.; Dallaire, L.; Melançon, S. Fluorometric assay of neuraminidase with a sodium (4-methylumbelliferyl-α-d-N-acetylneuraminate) substrate. *Anal. Biochem.* **1979**, *94*, 287–296. [CrossRef]
138. Seyrantepe, V.; Landry, K.; Trudel, S.; Hassan, J.A.; Morales, C.R.; Pshezhetsky, A.V. Neu4, a Novel Human Lysosomal Lumen Sialidase, Confers Normal Phenotype to Sialidosis and Galactosialidosis Cells. *J. Biol. Chem.* **2004**, *279*, 37021–37029. [CrossRef] [PubMed]
139. Bussink, A.P.; Verhoek, M.; Vreede, J.; der Vlugt, K.G.-V.; Donker-Koopman, W.E.; Sprenger, R.R.; Hollak, C.E.; Aerts, J.M.F.G.; Boot, R.G. Common G102S polymorphism in chitotriosidase differentially affects activity towards 4-methylumbelliferyl substrates. *FEBS J.* **2009**, *276*, 5678–5688. [CrossRef]
140. Schoonhoven, A.; Rudensky, B.; Elstein, D.; Zimran, A.; Hollak, C.E.; Groener, J.E.; Aerts, J.M. Monitoring of Gaucher patients with a novel chitotriosidase assay. *Clin. Chim. Acta* **2007**, *381*, 136–139. [CrossRef]
141. Lau, K.H.; Onishi, T.; Wergedal, J.E.; Singer, F.R.; Baylink, D.J. Characterization and assay of tartrate-resistant acid phosphatase activity in serum: Potential use to assess bone resorption. *Clin. Chem.* **1987**, *33*, 458–462. [CrossRef]
142. Goodlad, G.A.J.; Mills, G.T. The acid phosphatases of rat liver. *Biochem. J.* **1957**, *66*, 346–354. [CrossRef]
143. di Matteo, G.; Orfeo, M.A.; Romeo, G. Human α-fucosidase: Purification and properties. *Biochim. Biophys. Acta (BBA)-Enzym.* **1976**, *429*, 527–537. [CrossRef]
144. Li, C.; Qian, J.; Lin, J.-S. Purification and characterization of α-L-fucosidase from human primary hepatocarcinoma tissue. *World J. Gastroenterol.* **2006**, *12*, 3770–3775. [CrossRef] [PubMed]
145. Neville, D.C.; Coquard, V.; Priestman, D.A.; te Vruchte, D.J.M.; Sillence, D.J.; Dwek, R.A.; Platt, F.M.; Butters, T.D. Analysis of fluorescently labeled glycosphingolipid-derived oligosaccharides following ceramide glycanase digestion and anthranilic acid labeling. *Anal. Biochem.* **2004**, *331*, 275–282. [CrossRef] [PubMed]
146. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.

147. Kirby, A.; Kang, H.M.; Wade, C.M.; Cotsapas, C.; Kostem, E.; Han, B.; Furlotte, N.; Kang, E.Y.; Rivas, M.; Bogue, M.A.; et al. Fine Mapping in 94 Inbred Mouse Strains Using a High-Density Haplotype Resource. *Genetics* **2010**, *185*, 1081–1095. [CrossRef]
148. Rebhan, M. GeneCards: Integrating information about genes, proteins and diseases. *Trends Genet.* **1997**, *13*, 163. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Chapter 3: Conclusions and future projections

This study used GWAS, pathway enrichment analyses, and integrative networks to identify potential genetic modifiers of murine hepatic lysosomal enzyme activities and GSL levels. Based on our findings, we have concluded the following:

(1) Mouse systems genetics strategies effectively identify candidate genetic regulators of lysosomal phenotypes. We have uncovered a set of putative genetic regulators for lysosomal enzyme activities and GSL levels, including transcription factors that may regulate both groups of phenotypes.

(2) Variants in many candidate modifiers are associated with human diseases where alterations in GSL metabolism have not previously been described. Therefore, the role of these lipids in the pathophysiology of these disorders should be investigated.

To validate our findings, the following prioritization criteria need to be applied:

1.- Lysosomal gene:

1a.- Exonic variant.

1b.- Variant in the UTR region.

1c.- Variant in introns/splicing.

2.- Non-lysosomal gene:

2a.- Exonic variant.

2b.- Variant in UTR region.

2c.- Variant in introns/splicing.

3.- Hits *p*-value in intergenic regions:

3a.- Regulatory function (epigenetic marks)/ncRNA-mirRNA.

3b.- Distance of the nearest gene.

3c.- Enhancers, promoters (distal or proximal), suppressors (chromatin modifiers), insulators.

4.- Common modifier of several traits

5.- Clinical implications of variants (OMIM)

6.- Higher number of variants per gene region:

With this in mind, we prioritized the following genes for future validations:

i) *Vps45*: the gene product is lysosomal (1) [69], identified as a modifier of fucosidase activity and GM3-Gc levels (4). In the fucosidase screening, appeared an intronic variant (1c) and in the GM3-Gc in the intergenic region and upstream (3c) with significant *p* values of 1,13 E-10 and 1,51 E-31, respectively; in addition, there are diseases associated with this gene like severe autosomal recessive congenital neutropenia 5, (5) (OMIM: *610035).

ii) *Pde4dip*: although the gene product is expressed in the Golgi (2) [70], it was recently shown the phosphoinositide signaling pathway, in which PDE4DIP participates, plays a central function in lysosomal repair [7]. It was uncovered as a modifier of fucosidase activity and GM3-Gc levels (4). In the fucosidase screening, an intronic variant appears (2c) and in GM3-Gc in exonic (2a) and intronic region (2c) with significant *p* values of 2.29E-7 and 1.51E-31 respectively; there are also diseases associated with this gene such as myeloproliferative disorder associated with eosinophilia (OMIM: *608117).

iii) miR-340-5p: is a putative shared regulator between eleven modifier genes (p= 2.2E-2) (4). Its dysregulation has been linked to preeclampsia [71], breast and gastric cancer [72,73].

Gain-of-function and loss-of-function studies are required to validate these candidate modifiers. Classical approaches include genetic and pharmacological tools such as CRISPR-Cas9, plasmids, anti-sense oligonucleotides, and small

molecules [74-76], which should be tested under healthy and diseased conditions. The following steps should be taken:

(a) Determine the mechanism of action, which can be done in mouse cell lines.

(b) Evaluate the potential therapeutic role, initially in cells derived from patients with fucosidosis, GM1 gangliosidosis, and other diseases with lysosomal dysfunction, such as Parkinson's. Following the cellular studies, animal models should be used to validate the results. This will determine whether the function of the modifier is specific to a particular cell type or organ or whether it has a systemic effect.

In conclusion, this study provides new insights into potential therapeutic avenues for lysosomal diseases. We believe these findings will facilitate the development of novel therapeutics for common and rare diseases with similar pathogenic cascades. This is particularly important in terms of translation and access to treatments.

Additional references

[1] C. Settembre, A. Fraldi, D.L. Medina, A. Ballabio, Signals from the lysosome: a control centre for cellular clearance and energy metabolism, Nat Rev Mol Cell Biol. 14 (2013) 283–296. https://doi.org/10.1038/nrm3565.

[2] A. Ballabio, J.S. Bonifacino, Lysosomes as dynamic regulators of cell and organismal homeostasis, Nat Rev Mol Cell Biol. 21 (2020) 101–118. https://doi.org/10.1038/s41580-019-0185-4.

[3] C. de Duve, The lysosome turns fifty, Nat Cell Biol. 7 (2005) 847–849. https://doi.org/10.1038/ncb0905-847.

[4] R. Puertollano, mTOR and lysosome regulation, F1000Prime Rep. 6 (2014). https://doi.org/10.12703/P6-52.

[5] A. González Montoro, K. Auffarth, C. Hönscher, M. Bohnert, T. Becker, B. Warscheid, F. Reggiori, M. van der Laan, F. Fröhlich, C. Ungermann, Vps39 Interacts with Tom40 to Establish One of Two Functionally Distinct Vacuole-Mitochondria Contact Sites, Dev Cell. 45 (2018) 621-636.e7. https://doi.org/10.1016/j.devcel.2018.05.011.

[6] P. Fineran, E. Lloyd-Evans, N.A. Lack, N. Platt, L.C. Davis, A.J. Morgan, D. Höglinger, R.V. V. Tatituri, S. Clark, I.M. Williams, P. Tynan, N. Al Eisa, E. Nazarova, A. Williams, A. Galione, D.S. Ory, G.S. Besra, D.G. Russell, M.B. Brenner, E. Sim, F.M. Platt, Pathogenic mycobacteria achieve cellular persistence by inhibiting the Niemann-Pick Type C disease cellular pathway, Wellcome Open Res. 1 (2017) 18. https://doi.org/10.12688/wellcomeopenres.10036.2.

[7] J.X. Tan, T. Finkel, A phosphoinositide signalling pathway mediates rapid lysosomal repair, Nature. 609 (2022) 815–821. https://doi.org/10.1038/s41586-022-05164-4.

[8] T. Iwayama, T. Okada, T. Ueda, K. Tomita, S. Matsumoto, M. Takedachi, S. Wakisaka, T. Noda, T. Ogura, T. Okano, P. Fratzl, T. Ogura, S. Murakami, Osteoblastic lysosome plays a central role in mineralization, Sci Adv. 5 (2019). https://doi.org/10.1126/sciadv.aax0672.

[9] M. Sardiello, M. Palmieri, A. di Ronza, D.L. Medina, M. Valenza, V.A. Gennarino, C. Di Malta, F. Donaudy, V. Embrione, R.S. Polishchuk, S. Banfi, G. Parenti, E. Cattaneo, A. Ballabio, A Gene Network Regulating Lysosomal Biogenesis and Function, Science (1979). 325 (2009) 473–477. https://doi.org/10.1126/science.1174447.

[10] B.-B. Chu, Y.-C. Liao, W. Qi, C. Xie, X. Du, J. Wang, H. Yang, H.-H. Miao, B.-L. Li, B.-L. Song, Cholesterol Transport through Lysosome-Peroxisome Membrane Contacts, Cell. 161 (2015) 291–306. https://doi.org/10.1016/j.cell.2015.02.019.

[11] J. Nakamura, T. Yamamoto, Y. Takabatake, T. Namba-Hamano, S. Minami, A. Takahashi, J. Matsuda, S. Sakai, H. Yonishi, S. Maeda, S. Matsui, I. Matsui, T. Hamano, M. Takahashi, M. Goto, Y. Izumi, T. Bamba, M. Sasai, M. Yamamoto, T. Matsusaka, F. Niimura, M. Yanagita, S. Nakamura, T. Yoshimori, A. Ballabio, Y. Isaka, TFEB-mediated lysosomal exocytosis alleviates high-fat diet–induced lipotoxicity in the kidney, JCI Insight. 8 (2023). https://doi.org/10.1172/jci.insight.162498.

[12] R.M. Perera, R. Zoncu, The Lysosome as a Regulatory Hub, Annu Rev Cell Dev Biol. 32 (2016) 223–253. https://doi.org/10.1146/annurev-cellbio-111315-125125.

[13] K.S. Hruska, M.E. LaMarca, C.R. Scott, E. Sidransky, Gaucher disease: mutation

and polymorphism spectrum in the glucocerebrosidase gene (GBA), Hum Mutat. 29 (2008) 567–583. https://doi.org/10.1002/humu.20676.
[14] H. HERS, α-Glucosidase deficiency in generalized glycogen-storage disease (Pompe's disease), Biochemical Journal. 86 (1963) 11–16. https://doi.org/10.1042/bj0860011.
[15] A. Vellodi, Lysosomal storage disorders, Br J Haematol. 128 (2005) 413–431. https://doi.org/10.1111/j.1365-2141.2004.05293.x.
[16] F.M. Platt, Emptying the stores: lysosomal diseases and therapeutic strategies, Nat Rev Drug Discov. 17 (2018) 133–150. https://doi.org/10.1038/nrd.2017.214.
[17] T. Kolter, K. Sandhoff, Sphingolipid metabolism diseases, Biochimica et Biophysica Acta (BBA) - Biomembranes. 1758 (2006) 2057–2079. https://doi.org/10.1016/j.bbamem.2006.05.027.
[18] K. SANDHOFF, My journey into the world of sphingolipids and sphingolipidoses, Proceedings of the Japan Academy, Series B. 88 (2012) 554–582. https://doi.org/10.2183/pjab.88.554.
[19] R. Tidhar, A.H. Futerman, The complexity of sphingolipid biosynthesis in the endoplasmic reticulum, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 1833 (2013) 2511–2518. https://doi.org/10.1016/j.bbamcr.2013.04.010.
[20] Y.A. Hannun, L.M. Obeid, Sphingolipids and their metabolism in physiology and disease, Nat Rev Mol Cell Biol. 19 (2018) 175–191. https://doi.org/10.1038/nrm.2017.107.
[21] E.B. Vitner, F.M. Platt, A.H. Futerman, Common and Uncommon Pathogenic Cascades in Lysosomal Storage Diseases, Journal of Biological Chemistry. 285 (2010) 20423–20427. https://doi.org/10.1074/jbc.R110.134452.
[22] F.M. Platt, B. Boland, A.C. van der Spoel, Lysosomal storage disorders: The cellular impact of lysosomal dysfunction, Journal of Cell Biology. 199 (2012) 723–734. https://doi.org/10.1083/jcb.201208152.
[23] L.A. Robak, I.E. Jansen, J. van Rooij, A.G. Uitterlinden, R. Kraaij, J. Jankovic, P. Heutink, J.M. Shulman, M.A. Nalls, V. Plagnol, D.G. Hernandez, M. Sharma, U.-M. Sheerin, M. Saad, J. Simón-Sánchez, C. Schulte, S. Lesage, S. Sveinbjörnsdóttir, S. Arepalli, R. Barker, Y. Ben-, H.W. Berendse, D. Berg, K. Bhatia, R.M.A. de Bie, A. Biffi, B. Bloem, Z. Bochdanovits, M. Bonin, J.M. Bras, K. Brockmann, J. Brooks, D.J. Burn, E. Majounie, G. Charlesworth, C. Lungu, H. Chen, P.F. Chinnery, S. Chong, C.E. Clarke, M.R. Cookson, J. Mark Cooper, J.C. Corvol, C. Counsell, P. Damier, J.-F. Dartigues, P. Deloukas, G. Deuschl, D.T. Dexter, K.D. van Dijk, A. Dillman, F. Durif, A. Dürr, S. Edkins, J.R. Evans, T. Foltynie, J. Dong, M. Gardner, J. Raphael Gibbs, A. Goate, E. Gray, R. Guerreiro, C. Harris, J.J. van Hilten, A. Hofman, A. Hollenbeck, J. Holton, M. Hu, X. Huang, I. Wurster, W. Mätzler, G. Hudson, S.E. Hunt, J. Huttenlocher, T. Illig, P. V Jónsson, J.-C. Lambert, C. Langford, A. Lees, P. Lichtner, P. Limousin, G. Lopez, D. Lorenz, C. Lungu, A. McNeill, C. Moorby, M. Moore, H.R. Morris, K.E. Morrison, V. Escott-Price, E. Mudanohwo, S.S. O'Sullivan, J. Pearson, J.S. Perlmutter, H. Pétursson, P. Pollak, B. Post, S. Potter, B. Ravina, T. Revesz, O. Riess, F. Rivadeneira, P. Rizzu, M. Ryten, S. Sawcer, A. Schapira, H. Scheffer, K. Shaw, I. Shoulson, J. Shulman, E. Sidransky, C. Smith, C.C.A. Spencer, H. Stefánsson, F. Bettella, J.D. Stockton, A. Strange, K. Talbot, C.M. Tanner, A. Tashakkori-Ghanbaria, F. Tison, D. Trabzuni, B.J. Traynor, A.G. Uitterlinden, D. Velseboer, M. Vidailhet, R. Walker, B. van de Warrenburg, M. Wickremaratchi, N. Williams, C.H. Williams-Gray, S. Winder-Rhodes, K. Stefánsson, M. Martinez,

N.W. Wood, J. Hardy, P. Heutink, A. Brice, T. Gasser, A.B. Singleton, Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease, Brain. 140 (2017) 3191–3203. https://doi.org/10.1093/brain/awx285.
[24] A. Fraldi, A.D. Klein, D.L. Medina, C. Settembre, Brain Disorders Due to Lysosomal Dysfunction, Annu Rev Neurosci. 39 (2016) 277–295. https://doi.org/10.1146/annurev-neuro-070815-014031.
[25] J. Magalhaes, M.E. Gegg, A. Migdalska-Richards, M.K. Doherty, P.D. Whitfield, A.H.V. Schapira, Autophagic lysosome reformation dysfunction in glucocerebrosidase deficient cells: relevance to Parkinson disease, Hum Mol Genet. 25 (2016) 3432–3445. https://doi.org/10.1093/hmg/ddw185.
[26] Z. Ou, J. Pan, S. Tang, D. Duan, D. Yu, H. Nong, Z. Wang, Global Trends in the Incidence, Prevalence, and Years Lived With Disability of Parkinson's Disease in 204 Countries/Territories From 1990 to 2019, Front Public Health. 9 (2021). https://doi.org/10.3389/fpubh.2021.776847.
[27] J. Shin, D. Kim, H.-L. Kim, M. Choi, Y. Koh, S.-S. Yoon, Oncogenic effects of germline variants in lysosomal storage disease genes, Genetics in Medicine. 21 (2019) 2695–2705. https://doi.org/10.1038/s41436-019-0588-9.
[28] G.A. Baselli, O. Jamialahmadi, S. Pelusi, E. Ciociola, F. Malvestiti, M. Saracino, L. Santoro, A. Cherubini, P. Dongiovanni, M. Maggioni, C. Bianco, F. Tavaglione, A. Cespiati, R.M. Mancina, R. D'Ambrosio, V. Vaira, S. Petta, L. Miele, U. Vespasiani-Gentilucci, A. Federico, J. Pihlajamaki, E. Bugianesi, A.L. Fracanzani, H.L. Reeves, G. Soardo, D. Prati, S. Romeo, L.VC. Valenti, V. Borroni, A. Liguori, L. Ronzoni, A. Cherubini, L. Santoro, M. Tomasi, A. Lombardi, M. Ostadreza, E. Casirati, I. Marini, S. Maier, C. Rosso, G.S. Baroni, C. Santaniello, M. Dallio, Rare ATG7 genetic variants predispose patients to severe fatty liver disease, J Hepatol. 77 (2022) 596–606. https://doi.org/10.1016/j.jhep.2022.03.031.
[29] F.D. Bulut, D. Kor, S. Kılavuz, B. Şeker Yılmaz, İ. Kaplan, F. Ekinci, E. Burgaç, İ. Varol, B. Köşeci, S. Tuğ Bozdoğan, E. Kara, F. Demir, A. Deniz, F. Temiz, N. Önenli Mungan, Expanding the phenotypic landscape of Gaucher disease type 3c with a novel entity - Transient neonatal cholestasis, Eur J Med Genet. 66 (2023) 104764. https://doi.org/10.1016/j.ejmg.2023.104764.
[30] G.E. Morris, P.S. Braund, J.S. Moore, N.J. Samani, V. Codd, T.R. Webb, Coronary Artery Disease–Associated *LIPA* Coding Variant rs1051338 Reduces Lysosomal Acid Lipase Levels and Activity in Lysosomes, Arterioscler Thromb Vasc Biol. 37 (2017) 1050–1057. https://doi.org/10.1161/ATVBAHA.116.308734.
[31] J. Fan, V.L. Hale, L.T. Lelieveld, L.J. Whitworth, E.M. Busch-Nentwich, M. Troll, P.H. Edelstein, T.M. Cox, F.J. Roca, J.M.F.G. Aerts, L. Ramakrishnan, Gaucher disease protects against tuberculosis, Proceedings of the National Academy of Sciences. 120 (2023). https://doi.org/10.1073/pnas.2217673120.
[32] K.G. Lassen, C.I. McKenzie, M. Mari, T. Murano, J. Begun, L.A. Baxt, G. Goel, E.J. Villablanca, S.-Y. Kuo, H. Huang, L. Macia, A.K. Bhan, M. Batten, M.J. Daly, F. Reggiori, C.R. Mackay, R.J. Xavier, Genetic Coding Variant in GPR65 Alters Lysosomal pH and Links Lysosomal Dysfunction with Colitis Risk, Immunity. 44 (2016) 1392–1405. https://doi.org/10.1016/j.immuni.2016.05.007.
[33] N.J. Weinreb, J. Charrow, H.C. Andersson, P. Kaplan, E.H. Kolodny, P. Mistry, G. Pastores, B.E. Rosenbloom, C.R. Scott, R.S. Wappner, A. Zimran, Effectiveness of enzyme replacement therapy in 1028 patients with type 1 Gaucher disease after 2 to 5 years of treatment: a report from the Gaucher Registry, Am J Med. 113 (2002) 112–119. https://doi.org/10.1016/S0002-

9343(02)01150-6.
[34] J.A. Weesner, I. Annunziata, T. Yang, W. Acosta, E. Gomero, H. Hu, D. van de Vlekkert, J. Ayala, X. Qiu, L.E. Fremuth, D.N. Radin, C.L. Cramer, A. d'Azzo, Preclinical Enzyme Replacement Therapy with a Recombinant β-Galactosidase-Lectin Fusion for CNS Delivery and Treatment of GM1-Gangliosidosis, Cells. 11 (2022) 2579. https://doi.org/10.3390/cells11162579.
[35] F.M. Platt, G.R. Neises, R.A. Dwek, T.D. Butters, N-butyldeoxynojirimycin is a novel inhibitor of glycolipid biosynthesis., J Biol Chem. 269 (1994) 8362–5.
[36] P.B. Deegan, O. Goker-Alpan, T. Geberhiwot, R.J. Hopkin, E. Lukina, A. Tylki-Szymanska, A. Zaher, C. Sensinger, S.J.M. Gaemers, V. Modur, B.L. Thurberg, J. Sharma, B. Najafian, M. Mauer, P. DasMahapatra, W.R. Wilcox, D.P. Germain, Venglustat, an orally administered glucosylceramide synthase inhibitor: Assessment over 3 years in adult males with classic Fabry disease in an open-label phase 2 study and its extension study, Mol Genet Metab. 138 (2023) 106963. https://doi.org/10.1016/j.ymgme.2022.11.002.
[37] J. Hobbs, Reversal of clinical features of hurler's disease and biochemical improvement after treatment by bone-marrow transplantation, The Lancet. 318 (1981) 709–712. https://doi.org/10.1016/S0140-6736(81)91046-1.
[38] H. Allewelt, M. Taskindoust, J. Troy, K. Page, S. Wood, S. Parikh, V.K. Prasad, J. Kurtzberg, Long-Term Functional Outcomes after Hematopoietic Stem Cell Transplant for Early Infantile Krabbe Disease, Biology of Blood and Marrow Transplantation. 24 (2018) 2233–2238. https://doi.org/10.1016/j.bbmt.2018.06.020.
[39] S. Worgall, D. Sondhi, N.R. Hackett, B. Kosofsky, M. V. Kekatpure, N. Neyzi, J.P. Dyke, D. Ballon, L. Heier, B.M. Greenwald, P. Christos, M. Mazumdar, M.M. Souweidane, M.G. Kaplitt, R.G. Crystal, Treatment of Late Infantile Neuronal Ceroid Lipofuscinosis by CNS Administration of a Serotype 2 Adeno-Associated Virus Expressing CLN2 cDNA, Hum Gene Ther. 19 (2008) 463–474. https://doi.org/10.1089/hum.2008.022.
[40] J. Kido, K. Sugawara, K. Nakamura, Gene therapy for lysosomal storage diseases: Current clinical trial prospects, Front Genet. 14 (2023). https://doi.org/10.3389/fgene.2023.1064924.
[41] T. Kirkegaard, J. Gray, D.A. Priestman, K.-L. Wallom, J. Atkins, O.D. Olsen, A. Klein, S. Drndarski, N.H.T. Petersen, L. Ingemann, D.A. Smith, L. Morris, C. Bornæs, S.H. Jørgensen, I. Williams, A. Hinsby, C. Arenz, D. Begley, M. Jäättelä, F.M. Platt, Heat shock protein–based therapy as a potential candidate for treating the sphingolipidoses, Sci Transl Med. 8 (2016). https://doi.org/10.1126/scitranslmed.aad9823.
[42] J. Gray, M.E. Fernández-Suárez, M. Falah, D. Smith, C. Smith, E. Kaya, A.M. Palmer, C.K. Fog, T. Kirkegaard, F.M. Platt, Heat shock protein amplification improves cerebellar myelination in the Npc1nih mouse model, EBioMedicine. 86 (2022) 104374. https://doi.org/10.1016/j.ebiom.2022.104374.
[43] K. Subramanian, N. Rauniyar, M. Lavalleé-Adam, J.R. Yates, W.E. Balch, Quantitative Analysis of the Proteome Response to the Histone Deacetylase Inhibitor (HDACi) Vorinostat in Niemann-Pick Type C1 disease, Molecular & Cellular Proteomics. 16 (2017) 1938–1957. https://doi.org/10.1074/mcp.M116.064949.
[44] E. Goina, P. Peruzzo, B. Bembi, A. Dardis, E. Buratti, Glycogen Reduction in Myotubes of Late-Onset Pompe Disease Patients Using Antisense Technology, Molecular Therapy. 25 (2017) 2117–2128.

https://doi.org/10.1016/j.ymthe.2017.05.019.
[45] M. Schlander, M. Beck, Expensive drugs for rare disorders: to treat or not to treat? The case of enzyme replacement therapy for mucopolysaccharidosis VI, Curr Med Res Opin. 25 (2009) 1285–1293. https://doi.org/10.1185/03007990902892633.
[46] J.S. Richardson, A.R. Kemper, S.D. Grosse, W.K.K. Lam, A.M. Rose, A. Ahmad, A. Gebremariam, L.A. Prosser, Health and economic outcomes of newborn screening for infantile-onset Pompe disease, Genetics in Medicine. 23 (2021) 758–766. https://doi.org/10.1038/s41436-020-01038-0.
[47] A. Broomfield, S.A. Jones, S.M. Hughes, B.W. Bigger, The impact of the immune system on the safety and efficiency of enzyme replacement therapy in lysosomal storage disorders, J Inherit Metab Dis. 39 (2016) 499–512. https://doi.org/10.1007/s10545-016-9917-1.
[48] J.E. Wraith, Limitations of enzyme replacement therapy: Current and future, J Inherit Metab Dis. 29 (2006) 442–447. https://doi.org/10.1007/s10545-006-0239-6.
[49] M. Nuijten, Pricing Zolgensma – the world's most expensive drug, J Mark Access Health Policy. 10 (2022). https://doi.org/10.1080/20016689.2021.2022353.
[50] M. Cao, X. Luo, K. Wu, X. He, Targeting lysosomes in human disease: from basic research to clinical applications, Signal Transduct Target Ther. 6 (2021) 379. https://doi.org/10.1038/s41392-021-00778-y.
[51] E.H. Schuchman, M.D. Ledesma, C.M. Simonaro, New paradigms for the treatment of lysosomal storage diseases: targeting the endocannabinoid system as a therapeutic strategy, Orphanet J Rare Dis. 16 (2021) 151. https://doi.org/10.1186/s13023-021-01779-4.
[52] A.C. Puhl, S. Ekins, Advancing the Research and Development of Enzyme Replacement Therapies for Lysosomal Storage Diseases, GEN Biotechnology. 1 (2022) 156–162. https://doi.org/10.1089/genbio.2021.0013.
[53] J. Su, X. Jin, K. She, Y. Liu, L. Song, Q. Zhao, J. Xiao, R. Li, H. Deng, F. Lu, Y. Yang, In vivo adenine base editing corrects newborn murine model of Hurler syndrome, Molecular Biomedicine. 4 (2023) 6. https://doi.org/10.1186/s43556-023-00120-8.
[54] O. Goker-Alpan, Divergent phenotypes in Gaucher disease implicate the role of modifiers, J Med Genet. 42 (2005) e37–e37. https://doi.org/10.1136/jmg.2004.028019.
[55] D. Chang, M.A. Nalls, I.B. Hallgrímsdóttir, J. Hunkapiller, M. van der Brug, F. Cai, G.A. Kerchner, G. Ayalon, B. Bingol, M. Sheng, D. Hinds, T.W. Behrens, A.B. Singleton, T.R. Bhangale, R.R. Graham, A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci, Nat Genet. 49 (2017) 1511–1516. https://doi.org/10.1038/ng.3955.
[56] J.D. Riordan, J.H. Nadeau, From Peas to Disease: Modifier Genes, Network Resilience, and the Genetics of Health, The American Journal of Human Genetics. 101 (2017) 177–191. https://doi.org/10.1016/j.ajhg.2017.06.004.
[57] N.A. Rosenberg, L. Huang, E.M. Jewett, Z.A. Szpiech, I. Jankovic, M. Boehnke, Genome-wide association studies in diverse populations, Nat Rev Genet. 11 (2010) 356–366. https://doi.org/10.1038/nrg2760.
[58] M. Slatkin, Linkage disequilibrium — understanding the evolutionary past and mapping the medical future, Nat Rev Genet. 9 (2008) 477–485. https://doi.org/10.1038/nrg2361.
[59] M. Civelek, A.J. Lusis, Systems genetics approaches to understand complex

traits, Nat Rev Genet. 15 (2014) 34–48. https://doi.org/10.1038/nrg3575.
[60] J.H. Sul, L.S. Martin, E. Eskin, Population structure in genetic studies: Confounding factors and mixed models, PLoS Genet. 14 (2018) e1007309. https://doi.org/10.1371/journal.pgen.1007309.
[61] J. Nsengimana, D.T. Bishop, Design Considerations for Genetic Linkage and Association Studies, in: 2012: pp. 237–262. https://doi.org/10.1007/978-1-61779-555-8_13.
[62] K.M.T.H. Rahit, M. Tarailo-Graovac, Genetic Modifiers and Rare Mendelian Disease, Genes (Basel). 11 (2020) 239. https://doi.org/10.3390/genes11030239.
[63] A. Breschi, T.R. Gingeras, R. Guigó, Comparative transcriptomics in human and mouse, Nat Rev Genet. 18 (2017) 425–440. https://doi.org/10.1038/nrg.2017.19.
[64] Mouse Genome Sequencing Consortium 1; Robert H Waterston, Initial sequencing and comparative analysis of the mouse genome, Nature. 420 (2002) 520–562. https://doi.org/10.1038/nature01262.
[65] L. Suner, F. Delhommeau, Gaucher's Disease, New England Journal of Medicine. 386 (2022) 1932–1932. https://doi.org/10.1056/NEJMicm2116167.
[66] A.D. Klein, Modeling diseases in multiple mouse strains for precision medicine studies, Physiol Genomics. 49 (2017) 177–179. https://doi.org/10.1152/physiolgenomics.00123.2016.
[67] B.J. Bennett, C.R. Farber, L. Orozco, H. Min Kang, A. Ghazalpour, N. Siemers, M. Neubauer, I. Neuhaus, R. Yordanova, B. Guan, A. Truong, W. Yang, A. He, P. Kayne, P. Gargalovic, T. Kirchgessner, C. Pan, L.W. Castellani, E. Kostem, N. Furlotte, T.A. Drake, E. Eskin, A.J. Lusis, A high-resolution association mapping panel for the dissection of complex traits in mice, Genome Res. 20 (2010) 281–290. https://doi.org/10.1101/gr.099234.109.
[68] A. Ghazalpour, C.D. Rau, C.R. Farber, B.J. Bennett, L.D. Orozco, A. van Nas, C. Pan, H. Allayee, S.W. Beaven, M. Civelek, R.C. Davis, T.A. Drake, R.A. Friedman, N. Furlotte, S.T. Hui, J.D. Jentsch, E. Kostem, H.M. Kang, E.Y. Kang, J.W. Joo, V.A. Korshunov, R.E. Laughlin, L.J. Martin, J.D. Ohmen, B.W. Parks, M. Pellegrini, K. Reue, D.J. Smith, S. Tetradis, J. Wang, Y. Wang, J.N. Weiss, T. Kirchgessner, P.S. Gargalovic, E. Eskin, A.J. Lusis, R.C. LeBoeuf, Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits, Mammalian Genome. 23 (2012) 680–692. https://doi.org/10.1007/s00335-012-9411-5.
[69] Pevsner J, Hsu SC, Hyde PS, Scheller RH. Mammalian homologues of yeast vacuolar protein sorting (vps) genes implicated in Golgi-to-lysosome trafficking. Gene. (1996) Dec 12;183(1-2):7-14. doi: 10.1016/s0378-1119(96)00367-8. PMID: 8996080.
[70] Verde I, Pahlke G, Salanova M, Zhang G, Wang S, Coletti D, Onuffer J, Jin SL, Conti M. Myomegalin is a novel protein of the golgi/centrosome that interacts with a cyclic nucleotide phosphodiesterase. J Biol Chem. (2001) Apr 6;276(14):11189-98. doi: 10.1074/jbc.M006546200. Epub 2000 Dec 27. PMID: 11134006.
[71] Guo Y, Gao Y, Liu S. lncRNA XIST is associated with preeclampsia and mediates trophoblast cell invasion via miR-340-5p/KCNJ16 signaling pathway. Transpl Immunol. 2022 Oct;74:101666. doi: 10.1016/j.trim.2022.101666. Epub 2022 Jul 6. PMID: 35809813.
[72] Hou LK, Yu Y, Xie YG, Wang J, Mao JF, Zhang B, Wang X, Cao XC. miR-340 and ZEB1 negative feedback loop regulates TGF-β- mediated breast cancer progression. Oncotarget. (2016) May 3;7(18):26016-26. doi: 10.18632/oncotarget.8421. PMID: 27036021; PMCID: PMC5041961.
[73] Yin G, Zhou H, Xue Y, Yao B, Zhao W. MicroRNA-340 promotes the tumor growth of human gastric cancer by inhibiting cyclin G2. Oncol Rep. (2016) Aug;36(2):1111-8. doi: 10.3892/or.2016.4876. Epub 2016 Jun 15. PMID:

27374211.
[74] Li H, Yang Y, Hong W, Huang M, Wu M, Zhao X. Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. Signal Transduct Target Ther. (2020) Jan 3;5(1):1. doi: 10.1038/s41392-019-0089-y. PMID: 32296011; PMCID: PMC6946647.
[75] Roberts TC, Langer R, Wood MJA. Advances in oligonucleotide drug delivery. Nat Rev Drug Discov. (2020) Oct;19(10):673-694. doi: 10.1038/s41573-020-0075-7. Epub 2020 Aug 11. PMID: 32782413; PMCID: PMC7419031.
[76] Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, Galver L, Kelley R, Karlsson A, Santos R, Overington JP, Hingorani AD, Casas JP. The druggable genome and support for target identification and validation in drug development. Sci Transl Med. (2017) Mar 29;9(383):eaag1166. doi: 10.1126/scitranslmed.aag1166. PMID: 28356508; PMCID: PMC6321762.