Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

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], Andrés D. Klein[1,*]

[1] Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile. **valeria.olguin.araneda@gmail.com**, **anduranm@udd.cl**, **mlasherasp@udd.cl**, **jrubilare@udd.cl**
[2] Universidad de Santiago de Chile, Facultad de Química y Biología, Departamento de Biología, Santiago, Chile. francisco.cubillos.r@usach.cl
[3] Millennium Institute for Integrative Biology (iBio), Santiago, Chile, francisco.cubillos.r@usach.cl
[4] Program in Human Genetics, Institute of Biomedical Sciences, Biomedical Neurosciences Institute, Department of Neuroscience, Facultad de Medicina, Universidad de Chile, Santiago, Chile. patricioolguin@uchile.cl
 # equal contributors: **valeria.olguin.araneda@gmail.com**, **anduranm@udd.cl**, **mlasherasp@udd.cl**
* Correspondence: **andresklein@udd.cl**

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 1; mouse 2; Drosophila 3; Saccharomyces cerevisiae; translational research 4; genetic background: precision medicine 5; gene mapping 6.

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 as a result of 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.

1. 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 include i) availability of genetic tools for creating disease models by transgenic, knockout, and knock-in mice technologies [4–6] (https://www.jax.org/research-and-faculty/resources/mouse-mutant-resource/available-models); ii) strains are nearly isogenic, enabling study how the same genetic mutation modifies a phenotype of

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

interest in different genetic backgrounds [7–11]; iii) tissues are available for omics studies [12]. Some limitations include different evolutive pressures for mice and humans; therefore, some systems, such as the immune system, do not function in the same manner in both species [13].

Hybrid Mouse Diversity Panel (HMDP)

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 (HMDP and CC) and ii) Populations derived from several generations of pseudo-random breeding from inbred strains (DO and HS) (Figure 1).

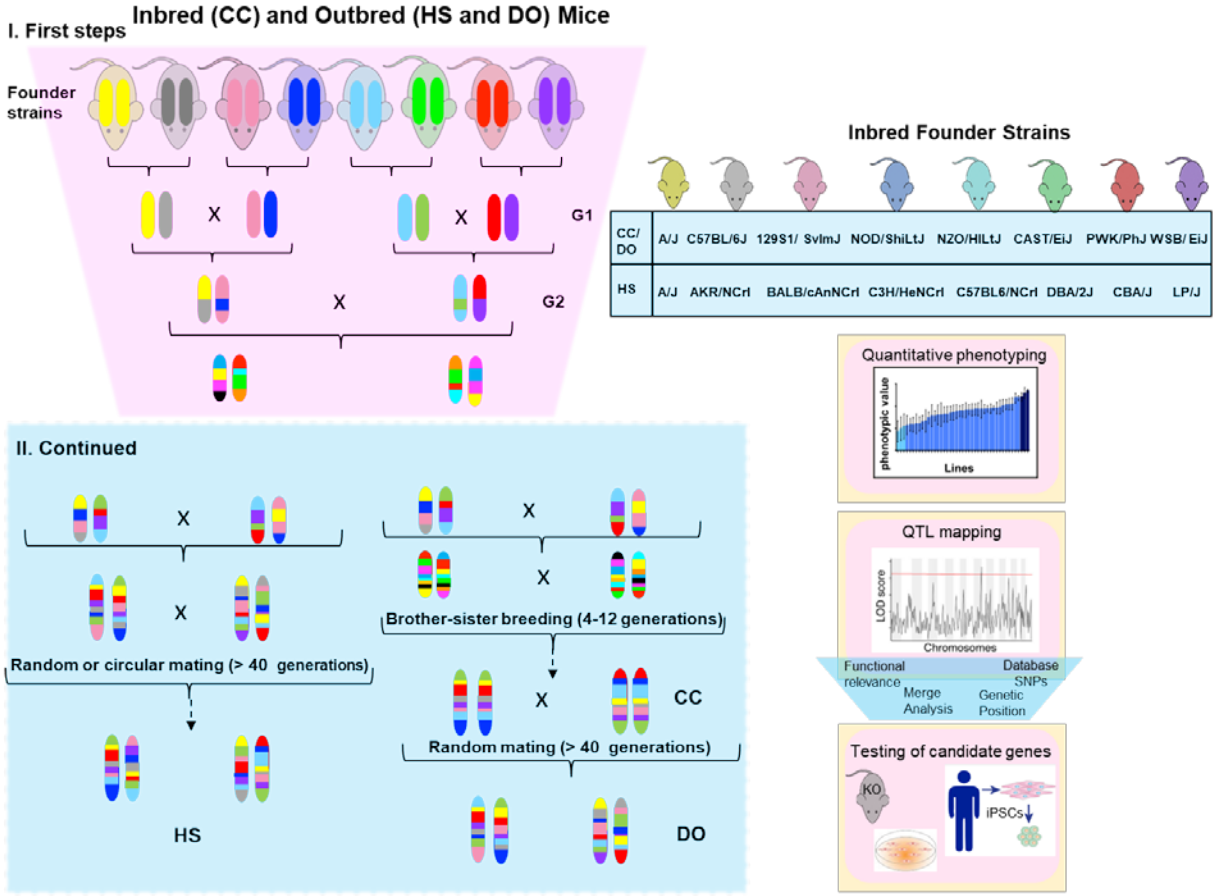

Figure 1. Breeding schemes for inbred (CC) and outbred (HS and DO) mice populations: Inbred founder strains for each panel are indicated in the right box. CC and DO populations share the same eight founder strains, five of which are standard laboratory inbred strains, and three are wild-derived strains. Different 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 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. The functional relevance of these variants can be assessed in animal models such as knock-out mice, iPSC cells derived from patients, and others.

HMDP is a large panel comprised of 100 commercially available (https://www.jax.org) and fully sequenced (www.sanger.ac.uk/science/data/mouse-genomes-project) inbred strains: 70 Recombinant inbreds (RI) strains derived mainly from crosses between C57BL/6J and DBA mice (BxD RI) and A/J and C57BL/6J (AxB and BxA RI), and 30 classical inbred strains [14].

Several features are highlighted in HMDP: i) Strains have been genotyped (http://mouse.cs.ucla.edu/mouseHapMap/). Thus, it is unnecessary to spend funds performing this step; ii) HMDP possesses 4 million common Single Nucleotyde Variants

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

(SNVs), which is similar to the number of common SNVs in humans [15]; iii) High resolution with 1-2 Mb regions containing five to 20 genes or less [14], at least an order of magnitude higher than traditional linkage analysis.; iv) Because HMDP is comprised of inbred strains, they can be used for genetic association studies of an unlimited number of complex traits, they are renewable (inbred strains are permanent), and data can be integrated (transcriptomic, proteomic, and metabolomic data) [12]; v) Current resources facilitate the availability to the strains (from The Jackson Laboratory); genotyped SNVs in HMDP (http://mouse.cs.ucla.edu/mouseHapMap/); vi) integrated information from multiple layers of characterization in the inbred strains, from Systems Genetics Resource (SGR); vii) servers to perform association mapping and statistical power simulation [16].

To the contrary, HMDP, as for any different model for genetic studies presents limitations. Besides the incorporation of inbred strains, long-range linkage disequilibrium (LD) both within both blocks and between regions on separate chromosomes can still be present, probably as a result of selection for allelic combinations that concede higher fitness during the inbreeding procedure [17]. Regions in LD can lead to false-positive associations in GWAS analyses. Although HMDP has a high resolution, the statistical power to detect the effect of loci is marginal (estimated at 50% to variants explaining 10% of the trait variance) [14]. Considering that most loci contributing to complex clinical traits have effect sizes below 5% [18], variants with subtle effects cannot be detected in HMDP. Power can be enhanced by including additional inbred and RI strains and performing meta-analyses from other panels as CC or traditional crosses [19].

An interesting example comes from crossing the 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, there was a significant effect of the Apoe allele in the AD-BXD panel. Furthermore, hippocampal gene expression in the severe and mild lines agrees with transcriptomic changes observed in patients [20].

The collaborative cross (CC) panel

The CC is a large panel of recombinant inbred (RI) mouse strains obtained through systematically outcrossing eight founder strains, followed by randomized breeding [21]. The founder strains of the CC include five widely used classical inbred laboratory strains, representatives from fancy mice of the M. musculus domesticus subspecies: A/J, C57BL/6J, NOD/ShiLtJ, 129S1/ SvImJ, NZO/HILtJ, as well as three wild-derived strains descendent 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 found in classical laboratory mouse strains [22].

Hallmarks of the CC panel useful for genetic association studies include both CC founder strains (http://www.sanger.ac.uk/resources/mouse/genomes/), as well as CC inbred lines (http://csbio.unc.edu/CCstatus/CCGenomes). These mice have been genotyped or sequenced, so haplotypes can be easily visualized or reconstructed as a mosaic of genomes of founders and enabling genomic mapping by linkage [23]. Founder strains capture approximately 90% of the genetic diversity seen in the Mus musculus species [24]. This high genetic diversity significantly reduces the list of plausible candidate loci. Additionally, randomized breeding substantially increases mapping resolution by reducing population structure effects [25]. CC strains have been used to map QTLs to less than 5 Mb intervals [26]. Online tools are available to perform both GWAS and linkage style analysis in CC panel [27], and several aspects of human genetic and behavioral factors can be modeled in this system [28], which makes CC a very powerful tool in searching for causes in genetic analyses.

Some considerations associated with CC use include: i) In extensive studies with CC strains, unique outlier phenotypes can arise, probably due to complex genetic regulatory networks involving multiple loci with epistatic interactions [29]. In such cases, the analysis

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

of causal modifiers in the entire population may not present sufficient mapping power, and the preferred approach for research studies to identify causal genes has been traditional F2 or backcrosses [30]. Because the identification of loci in large panels, such as CC, could be extensively resource-intensive, initial or pilot study within the population has been suggested, starting with a small-scale approach in a minor representative set of strains and expanding as necessary [29]. The strategy in creating a CC panel generates high breeding complications and infertility rates, mainly caused by genomic incompatibility introduced by the wild-derived strains. For that reason, the initial project focused on producing 1,000 CC strains finished with only around 100 CC strains and inspired the creation of a DO population.

CC lines have been used for genetic association studies of many different complex traits, both in physiological traits as metabolism- and exercise-related traits [31], motor performance, and body weight [32], as well as in pathophysiological traits related to glucose tolerance [33], among many others. The CC panel is 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. 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 multi-genetic background studies in animal organisms.

Heterogeneous Stock and Diversity Outbred populations

Both Heterogeneous Stock (HS) and Diversity Outbred (DO) are high diversity outbred mice populations. The HS was established from breeding eight inbred strains and then outbreeding in either a circular strategy or random breeding (Figure 1) to minimize inbreeding [35]. After 50 or more generations of breeding, HS generated mice are 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 through pseudorandomized fashion, non-sibling mattings indefinitely [38] (Figure 1). Given that DO is derived from the same eight founders as 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 to using HS or DO mice compared to classical inbred mice, including outbred randomized mating that sincrease the accumulation of additional recombination sites compared to those of classically inbred mice; thus, each HS or DO mouse has a unique genome, a mosaic of the original eight founder lines, which resembles human heterozygosity and allows high-resolution genetic mapping [39]. HS and DO mice have been used to fine-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 large degree of genetic variability within both HS and DO populations results in a high degree of phenotypic variability; thus, these outbred models enable the fine-mapping of many different phenotypic traits. Furthermore, as mentioned, the founders of CC and DO lines include wild-derived strains; therefore, unique behaviors observed in these mice in comparison to classical laboratory strains represent a valuable tool in genetic behavior association studies [22]. A repository of DO QTL studies (https://dodb.jax.org) can be shared between laboratories, enabling Identification of significant loci in outbred mice that can be performed with many different genetic contexts on a mix of genetic backgrounds to minimize the role of genetic background. Finally, founders of all extant HS and DO populations have been sequenced, and these data are publicly available [42], again 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

mouse requires genotyping and haplotype reconstruction in order to perform each new QTL analysis [38]. Mapping Precision with high-resolution achieved with the use of HS or DO mice has an inverse relationship with statistical power [43], and because a large number of animals are needed for sufficient statistical power, which is not always possible. In addition, candidate modifiers identified in outbred mice lack opportunities for validation in reproducible genotypes as occurs in inbred lines. However, in the case of DO mice, this could easily be performed in CC panels that share the same founder strains [44].

An interesting translational study using the DO panel was the identification of a diagnostic biomarker for human tuberculosis (TB). By applying machine learning algorithms to multidimensional data, the authors identified 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.

2. Drosophila melanogaster as a model organism in genetic research, advantages, and limitations

In addition to the mice models, and among the model organisms, Drosophila melanogaster has gained much attraction. Flies are small, easy to manipulate in the laboratory, and cheap to keep. They have a short life span (2-week generation interval) and produce many offspring. Drosophila show complex behaviors including sleep, aggression, addiction, and social behavior [46]. Notably, about 75% of human disease-associated genes have a Drosophila ortholog; its genome is fully sequenced and well-annotated. It can be genetically modified using chemical and insertional mutagenesis and gene-specific mutations or editions using CRISPR [47, 48]. These characteristics support its use as a model system to study human genetic 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 human pathologies related to this system.

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 sib mating and then wholly sequenced [50] (Figure 2). DGRP is a public resource available at the Bloomington Drosophila Stock Center (http://fly.bio.indiana.edu). DGRP has critical features for genomic association analyses: availability of complete sequence dataminimal genetic variation within each line, thus repeated measurements enabling accuracy to increase the statistical power in GWA analyses; since the DGRP is a publicly available resource, it enables different laboratories to correlate phenotypes on the same genotype gaining an understanding of 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 the precision of association mapping; thus, significant associated SNVs are likely causal or very near to causal variant [52]; finally, there are fewer ethical issues of concern in experimental Drosophila research compared with research using rodent models.

As with all study models, there are some limitations/considerations in DGRP that should be considered. First, genetic variation between the lines is a snapshot of the population from which they were derived; therefore, DGRP does not represent all possible variations of the species. Second, the sample size of 205 lines in DGRP provide, in general, enough statistical power to detect common variants with moderately large to significant effects after correction for multiple tests [53,54], but the statistical power is still limited in the case of rare variants (MAF<0.05), which could be excluded [51].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

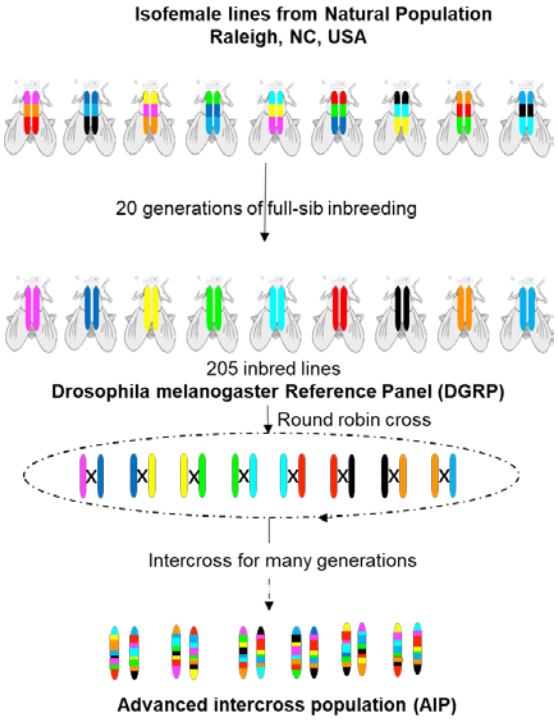

Figure 2. Generation of Drosophila Melanogaster Genetic Reference Panel (DGRP) and Advanced Intercross Population DGRP-AIPs. DGRP corresponds to a sequenced population from a natural population in Raleigh, NC (USA) through 20 generations of full-sib mating. On the other hand, AIPs lines can be generated from this base population by round-robin crossing and then remapped.

DGRP in physiological and pathophysiological traits

The DGRP has been used for GWA mapping of many different quantitative physiological traits, including food intake and sleep behavior [55,56]. Food intake is an essential component of animal fitness, and 25 modifiers with human orthologs of food consumption have been 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 in many aspects and have become an important model species for identifying sleep regulation mechanisms. A GWAS highlighted signals in the EGFR, Wnt, Hippo and MAPK signaling pathways [58]. Several candidate genes have human homologs previously identified in human sleep studies, suggesting that genes affecting variation in this trait could be conserved. DGRP has been extended as a tool in genetic association studies to other traits such as nutrient stores [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 PD model [63]. PD is a highly variable neurodegenerative disorder where variable manifestations range from cognitive disturbances, motor alterations, sleep and speech abnormalities, to cellular pathological changes as the formation of Lewy body inclusions and neuronal death [64]. It has been shown that the penetrance of the leucine-rich repeat kinase 2 gene G2019S mutation (LRRK2 G2019S) is incomplete and varies among different 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 the introduction of the LRRK2 G2019S mutation in the fly panel DGRP results in considerable variability in the locomotor phenotype in different DGRP backgrounds [63]. Furthermore, using the natural genetic variation of these model organisms, putative modifiers of the disease-related phenotypes have been identified and functionally validated in the panel [63].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

Lines derived from DGRP and DSRP

As mentioned above, DGRP has some limitations in terms of statistical power (51),(51),(51), which has motivated the development of DGRP-derived advanced intercross populations (AIPs). These in turn correspond to lines generated by crossing parentals DGRP for many generations, which are then remapped [67]. By crossing a subset of parentals lines in a large number, it is possible to increase the recombination rate and consequently the statistical power compared to using single-line DGRP [52]. Also, the "extreme QTL mapping" strategy in AIPs can be used as a tool to resolve the statistical limitations of DGRP in the case of rare variants (MAF<0.05), which may have critical phenotypic effects. Extreme QTL mapping refers to selecting individuals from the extremes of the phenotypic distribution for a trait (resembling a case-control study), pooling and sequencing them (which is cheaper than sequencing all individuals in the initial population), and then identifying alleles that segregate differentially among the extremes of the distribution (variant causal or in LD with this) [68,69]. Rare variants in DGRP will accur at higher frequencies in AIPs after an extreme QTL mapping strategy and can then be analyzed.

Another minor applied strategy proposed to increase the mapping power is to use DGRP and another developed panel for cross-validation as the Drosophila Synthetic Population Resource (DSPR). Their 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 show a lack of overlap with candidate genes found in DGRP, which could be due to different genetic architecture or genetic variants between the panels.

3. 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 the quantitative genetics field 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 KanMX cassette in the reference S. cerevisiae strain (76). S(76). S(76). Since then, similar collections have been available in different genetic backgrounds, demonstrating the high degree of genetic background dependencies for different phenotypes [77, 78]. One of these research outcomes highlights the ability of knockouts of closely related gene products' dreading masking effect to be less like in yeast than in other complex systems, given the reduced genetic redundancy in this model organism [79].

Analysis of segregating populations from pairwise crosses

Quantitative trait loci mapping in yeast has been the primary approach used to identify 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 [80,81]. Linkage mapping in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

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 the link with genetic markers, allowing causal loci that segregate along with the closest genetic marker; thus, QTLs are identified using statistical tests for association between markers and a trait of interest [80, 82]. The bulk segregant analysis (BSA) also involves crossing two or more parental strains and subsequent phenotyping of their recombinant offspring [83]. However, this method involves selective genotyping of subsets of segregants, commonly the extremes of the phenotypic distribution [84]. Typically, segregants undergo selective environmental pressure, where pools are constructed that are too large. One expresses the trait of interest (selected pool) and others arenot selected (control pool) or exhibit the opposite phenotype. After genotyping each marker, genetic regions of allelic enrichment, are predicted as QTL that contributes to the attribute of interest [85]. These approaches from pairwise crosses have been successfully applied to map the genetic variation responsible for quantitative traits in Saccharomyces cerevisiae for several traits, including 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 in natural populations and have mosaic genomes of the founder strains [84,90]. Therefore, studies using these biparental crosses: i) provide a poor understanding of the relationship between the genetic background and the QTLs, ii) lack resolution since a small number of generations is used, and consequently iii) are unable to reveal the complete architecture of polygenic traits. Moreover, iv) laboratory strains often contain artificial auxotrophic markers that confound mapping experiments [91]. Investigators have recently established advanced-generation MultiParent Populations (MPPs) in yeast to overcome these problems.

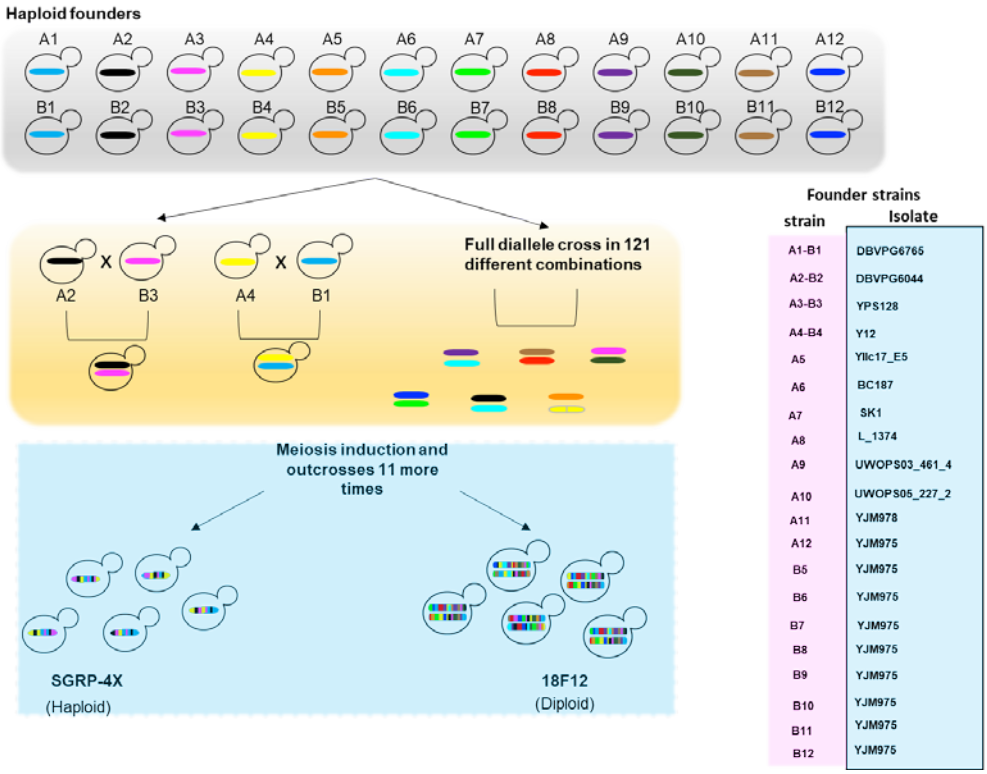

Figure 3. Cross design of SGRP-4X and 18F12 mapping populations. Haploid founder strains used for generations of these populations are indicated in the right box. Ax and Bx indicate Mat a and Mat α haploid founder strain, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

MPPS

Yeast (MPPs) comprise large populations with thousands to millions of individuals obtained from two main steps. First, several (inbred or isogenic) founder strains from various geographical origins are crosses, and then the intercross of the resulting population is crossed for several subsequent generations [81]. Large segregating populations are then used for mapping QTLs. The first MPP in yeast was established by Cubillos et al. [92], 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. SGRP-4X contains 165 sequenced segregants, representing recombined genetic mosaics of the original 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, each strain derived from the SGRP collection [84,92,93].

Multiparental populations in yeast are robust mapping resources as a resykt if multiple founders and rounds of recombination in a large number of individuals that increase both the genetic and phenotypic diversity and 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] to gain a low linkage disequilibrium. And, concerning association mapping panels, MPP design provides more equilibrated allelic frequencies, which increases the 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.

GWAS in S. cerevisiae

Genome-wide association studies (GWAS) utilize the variation in large populations of unrelated individuals to provide insight 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 a project to describe whole-genome sequence variation in more numerous yeast populations (http://1002genomes.u-strasbg.fr/). Today, there are more than two thousand genomes isolated from a wide range of locations (which include Australia, Europe, Russia, Vietnam, and South Africa) and sources (such as wine, dairy products, trees, insects, flowers, fruit, and dairy products) available [96], thus increasing the interest and ability of the investigators to conduct GWAS in this model organism [97].

The success of GWAS in S. cerevisiae occurs as a result of high diversity among natural isolates relative to humans [96], low linkage disequilibrium (LD) (extended in an average half-life of < 3 kb) [98] and relatively simple quantification of phenotypes in hundred to thousands of individuals. However, GWAS in yeast were 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 into the species [99].

Many of the phenotypes addressed in yeast are directly related to the study 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 S. cerevisiae 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].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

We expect to observe increased research using this kind of approach soon, where a combination of variants can be identified. This technique could be feasible for diseases that can be mimicked pharmacologically.

Concluding remarks

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 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, using already sequenced panels of diverse model organisms followed by validation in smaller human cohorts can speed up translational research and precision medicine for both common and rare diseases.

Funding: ADK is funded by Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) grant No 1180337. PO is funded by Biomedical Neuroscience Institute (BNI) grant No P09-015-F, Iniciativa Científica Milenio. PO and ADK are supported by Pew Innovation Fund, grant No 32422.

References

1. Ashley, E.A. Towards Precision Medicine. Nat Rev Genet 2016, 17, 507–522, doi:10.1038/NRG.2016.86.
2. Rahit, K.M.T.H.; Tarailo-Graovac, M. Genetic Modifiers and Rare Mendelian Disease. Genes (Basel) 2020, 11, doi:10.3390/GENES11030239.
3. Canales, C.P.; Walz, K. The Mouse, a Model Organism for Biomedical Research. Cellular and Animal Models in Human Genomics Research 2019, 119–140, doi:10.1016/B978-0-12-816573-7.00006-7.
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, doi:10.1126/SCIENCE.1225829.
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, doi:10.1016/J.CELL.2014.09.014.
6. Swiech, L.; Heidenreich, M.; Banerjee, A.; Habib, N.; Li, Y.; Trombetta, J.; Sur, M.; Zhang, F. In Vivo Interrogation of Gene Function in the Mammalian Brain Using CRISPR-Cas9. Nat Biotechnol 2015, 33, 102–106, doi:10.1038/NBT.3055.
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, doi:10.1016/J.YMGMR.2018.05.001.
8. Durán, A.; Rebolledo-Jaramillo, B.; Olguin, V.; Rojas-Herrera, M.; las Heras, M.; 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, doi:10.1016/J.BBREP.2021.101105.
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, doi:10.1016/J.BBRC.2011.08.096.
10. Klein, A.D.; Ferreira, N.S.; Ben-Dor, S.; Duan, J.; Hardy, J.; Cox, T.M.; Merrill, A.H.; Futerman, A.H. Identification of Modifier Genes in a Mouse Model of Gaucher Disease. Cell Rep 2016, 16, 2546–2553, doi:10.1016/J.CELREP.2016.07.085.
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, doi:10.1242/DMM.042614.
12. Klein, A.D. Modeling Diseases in Multiple Mouse Strains for Precision Medicine Studies. Physiol Genomics 2017, 49, 177–179, doi:10.1152/PHYSIOLGENOMICS.00123.2016.
13. Junhee Seok; H. Shaw Warren; 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 U S A 2013, 110, 3507–3512, doi:10.1073/PNAS.1222878110/-/DCSUPPLEMENTAL/SAPP.PDF.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

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, doi:10.1101/GR.099234.109.

15. Lusis, A.J.; Seldin, M.M.; Allayee, H.; Bennett, B.J.; Civelek, M.; Davis, R.C.; Eskin, E.; Farber, C.R.; 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, doi:10.1194/JLR.R066944.

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, doi:10.1007/S00335-012-9411-5.

17. Pstkov, P.M.; Graber, J.H.; Churchill, G.A.; Dipetrillo, K.; King, B.L.; Paigen, K. Evidence of a Large-Scale Functional Organization of Mammalian Chromosomes. PLoS Genet 2005, 1, 312–322, doi:10.1371/JOURNAL.PGEN.0010033.

18. Flint, J.; Mott, R. Applying Mouse Complex-Trait Resources to Behavioural Genetics. Nature 2008, 456, 724–727, doi:10.1038/NATURE07630.

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, doi:10.1371/JOURNAL.PGEN.1004022.

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.e5, doi:10.1016/J.NEURON.2018.11.040.

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, doi:10.1534/GENETICS.116.198838.

22. Yang, H.; Wang, J.R.; Didion, J.P.; Buus, R.J.; Bell, T.A.; Welsh, C.E.; Bonhomme, F.; Yu, A.H.T.; Nachman, M.W.; Pialek, J.; et al. Subspecific Origin and Haplotype Diversity in the Laboratory Mouse. Nat Genet 2011, 43, 648–655, doi:10.1038/NG.847.

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, doi:10.1101/GR.111310.110.

24. Roberts, A.; Pardo-Manuel De Villena, F.; 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, doi:10.1007/S00335-007-9045-1.

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, doi:10.1016/J.TIG.2019.04.003.

26. Keele, G.; Zhang, T.; Pham, D.; Vincent, M.; Genomics, T.B.-C.; 2021, undefined Regulation of Protein Abundance in Genetically Diverse Mouse Populations. Elsevier.

27. Ram, R.; Morahan, G. Complex Trait Analyses of the Collaborative Cross: Tools and Databases. Methods Mol Biol 2017, 1488, 121–129, doi:10.1007/978-1-4939-6427-7_5.

28. Molenhuis, R.T.; Bruining, H.; Brandt, M.J.V.; van Soldt, P.E.; Abu-Toamih Atamni, H.J.; 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, doi:10.1186/S13229-018-0252-2.

29. Noll, K.E.; 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, doi:10.1016/J.CHOM.2019.03.009.

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, doi:10.1007/S00335-013-9499-2.

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 Endocrinol Metab 2011, 300, doi:10.1152/AJPENDO.00707.2010.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

32. Mao, J.H.; Langley, S.A.; Huang, Y.; Hang, M.; Bouchard, K.E.; Celniker, S.E.; Brown, J.B.; Jansson, J.K.; 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, doi:10.1038/SREP16247.

33. Abu-Toamih Atamni, H.J.; Ziner, Y.; Mott, R.; Wolf, L.; Iraqi, F.A. Glucose Tolerance Female-Specific QTL Mapped in Collaborative Cross Mice. Mamm Genome 2017, 28, 20–30, doi:10.1007/S00335-016-9667-2.

34. Lorè, N.I.; Sipione, B.; He, G.; Strug, L.J.; Atamni, H.J.; Dorman, A.; Mott, R.; Iraqi, F.A.; Bragonzia, A. Collaborative Cross Mice Yield Genetic Modifiers for Pseudomonas Aeruginosa Infection in Human Lung Disease. mBio 2020, 11, doi:10.1128/MBIO.00097-20.

35. Solberg Woods, L.C. QTL Mapping in Outbred Populations: Successes and Challenges. Physiol Genomics 2014, 46, 81–90, doi:10.1152/PHYSIOLGENOMICS.00127.2013.

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, doi:10.1038/6825.

37. Woods, L.C.S.; Mott, R. Heterogeneous Stock Populations for Analysis of Complex Traits. Methods Mol Biol 2017, 1488, 31–44, doi:10.1007/978-1-4939-6427-7_2.

38. Gatti, D.M.; Svenson, K.L.; Shabalin, A.; Wu, L.Y.; Valdar, W.W.; Simecek, P.; Goodwin, N.; Cheng, R.; Pomp, D.; Palmer, A.; et al. Quantitative Trait Locus Mapping Methods for Diversity Outbred Mice. G3 (Bethesda) 2014, 4, 1623–1633, doi:10.1534/G3.114.013748.

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, doi:10.1111/GBB.12029.

40. Valdar, W.; Solberg, L.C.; Gauguier, D.; Burnett, S.; Klenerman, P.; Cookson, W.O.; Taylor, M.S.; 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, doi:10.1038/NG1840.

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, doi:10.1534/GENETICS.111.132597.

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, doi:10.1038/NATURE10413.

43. Parker, C.C.; Palmer, A.A. Dark Matter: Are Mice the Solution to Missing Heritability? Front Genet 2011, 2, doi:10.3389/FGENE.2011.00032.

44. Chesler, E.J. Out of the Bottleneck: The Diversity Outcross and Collaborative Cross Mouse Populations in Behavioral Genetics Research. Mamm Genome 2014, 25, 3–11, doi:10.1007/S00335-013-9492-9.

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, doi:10.1371/JOURNAL.PPAT.1009773.

46. Kazama, H. Systems Neuroscience in Drosophila: Conceptual and Technical Advantages. Neuroscience 2015, 296, 3–14, doi:10.1016/J.NEUROSCIENCE.2014.06.035.

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, doi:10.1016/J.CELL.2014.09.002.

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, doi:10.1101/GR.169101.

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, doi:10.1126/SCIENCE.287.5461.2196.

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, doi:10.1038/NATURE10811.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

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, doi:10.1101/GR.171546.113.

52. Anholt, R.R.H.; Mackay, T.F.C. The Road Less Traveled: From Genotype to Phenotype in Flies and Humans. Mamm Genome 2018, 29, 5–23, doi:10.1007/S00335-017-9722-7.

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, doi:10.1371/JOURNAL.PONE.0132980.

54. Edwards, S.M.; Sørensen, I.F.; Sarup, P.; Mackay, T.F.C.; Sørensen, P. Genomic Prediction for Quantitative Traits Is Improved by Mapping Variants to Gene Ontology Categories in Drosophila Melanogaster. Genetics 2016, 203, 1871–1883, doi:10.1534/GENETICS.116.187161.

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, doi:10.1371/JOURNAL.PONE.0138129.

56. Serrano Negron, Y.L.; Hansen, N.F.; Harbison, S.T. The Sleep Inbred Panel, a Collection of Inbred Drosophila Melanogaster with Extreme Long and Short Sleep Duration. G3 (Bethesda) 2018, 8, 2865–2873, doi:10.1534/G3.118.200503.

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, doi:10.1038/S42255-019-0147-3.

58. Harbison, S.T.; Serrano Negron, Y.L.; 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, doi:10.1371/JOURNAL.PGEN.1007098.

59. Unckless, R.L.; Rottschaefer, S.M.; Lazzaro, B.P. A Genome-Wide Association Study for Nutritional Indices in Drosophila. G3 (Bethesda) 2015, 5, 417–425, doi:10.1534/G3.114.016477.

60. Lafuente, E.; Duneau, D.; Beldade, P. Genetic Basis of Thermal Plasticity Variation in Drosophila Melanogaster Body Size. PLoS Genet 2018, 14, doi:10.1371/JOURNAL.PGEN.1007686.

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 2019, 49, 60–82, doi:10.1007/S10519-018-9932-0.

62. Chow, C.Y.; Kelsey, K.J.P.; Wolfner, M.F.; Clark, A.G. Candidate Genetic Modifiers of Retinitis Pigmentosa Identified by Exploiting Natural Variation in Drosophila. Hum Mol Genet 2016, 25, 651–659, doi:10.1093/HMG/DDV502.

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, doi:10.1534/GENETICS.118.301119.

64. Klein, A.D.; Mazzulli, J.R. Is Parkinson's Disease a Lysosomal Disorder? Brain 2018, 141, 2255–2262, doi:10.1093/BRAIN/AWY147.

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, doi:10.1212/WNL.0000000000001708.

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, doi:10.1016/J.MOLMED.2019.09.004.

67. Mackay, T.F.C.; Huang, W. Charting the Genotype-Phenotype Map: Lessons from the Drosophila Melanogaster Genetic Reference Panel. Wiley Interdiscip Rev Dev Biol 2018, 7, doi:10.1002/WDEV.289.

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 U S A 2013, 110, 1017–1022, doi:10.1073/PNAS.1220168110.

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 U S A 2015, 112, E3555–E3563, doi:10.1073/PNAS.1510104112.

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, doi:10.1016/J.TIG.2014.07.009.

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, doi:10.1534/GENETICS.114.163188.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

72. Walberg, M.W. Applicability of Yeast Genetics to Neurologic Disease. Arch Neurol 2000, 57, 1129–1134, doi:10.1001/ARCHNEUR.57.8.1129.

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, doi:10.1038/NATURE07135.

74. Sanchez, A.; Borde, V. Methods to Map Meiotic Recombination Proteins in Saccharomyces Cerevisiae. Methods Mol Biol 2021, 2153, 295–306, doi:10.1007/978-1-0716-0644-5_21.

75. Botstein, D.; Fink, G.R. Yeast: An Experimental Organism for Modern Biology. Science 1988, 240, 1439–1443, doi:10.1126/SCIENCE.3287619.

76. Giaever, G.; Nislow, C. The Yeast Deletion Collection: A Decade of Functional Genomics. Genetics 2014, 197, 451–465, doi:10.1534/GENETICS.114.161620.

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, doi:10.15252/MSB.20198831.

78. Parts, L.; Batté, A.; Lopes, M.; Yuen, M.W.; Laver, M.; Luis, B.-J.S.; Yue, J.-X.; Pons, C.; Eray, E.; Aloy, P.; et al. Natural Variants Suppress Mutations in Hundreds of Essential Genes. Molecular Systems Biology 2021, 17, e10138, doi:10.15252/MSB.202010138.

79. Khurana, V.; Lindquist, S. Modelling Neurodegeneration in Saccharomyces Cerevisiae: Why Cook with Baker's Yeast? Nat Rev Neurosci 2010, 11, 436–449, doi:10.1038/NRN2809.

80. Ehrenreich, I.M.; Magwene, P.M. Genetic Dissection of Heritable Traits in Yeast Using Bulk Segregant Analysis. Cold Spring Harb Protoc 2017, 2017, 480–484, doi:10.1101/PDB.PROT088989.

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, doi:10.7554/ELIFE.49212.

82. Lander, E.S.; Botstein, S. Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. Genetics 1989, 121, 185, doi:10.1093/GENETICS/121.1.185.

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 U S A 1991, 88, 9828–9832, doi:10.1073/PNAS.88.21.9828.

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, doi:10.1038/NATURE07743.

85. Swinnen, S.; Thevelein, J.M.; Nevoigt, E. Genetic Mapping of Quantitative Phenotypic Traits in Saccharomyces Cerevisiae. FEMS Yeast Res 2012, 12, 215–227, doi:10.1111/J.1567-1364.2011.00777.X.

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, doi:10.1186/S40659-019-0270-3.

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. Frontiers in Genetics 2019, 10, 998, doi:10.3389/FGENE.2019.00998/BIBTEX.

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. Biotechnology for Biofuels 2019, 12, 1–18, doi:10.1186/S13068-019-1398-7/FIGURES/7.

89. Liti, G.; Louis, E.J. Yeast Evolution and Comparative Genomics. Annu Rev Microbiol 2005, 59, 135–153, doi:10.1146/ANNUREV.MICRO.59.030804.121400.

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, doi:10.1371/JOURNAL.PGEN.1002111.

91. Perlstein, E.O.; Ruderfer, D.M.; 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, doi:10.1038/NG1991.

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, doi:10.1111/J.1567-1364.2009.00583.X.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 May 2022 doi:10.20944/preprints202205.0314.v1

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, doi:10.1534/GENETICS.120.303202.

94. Cubillos, F.A.; 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, doi:10.1534/GENETICS.113.155515.

95. Diouf, I.; Pascual, L. Multiparental Population in Crops: Methods of Development and Dissection of Genetic Traits. Methods Mol Biol 2021, 2264, 13–32, doi:10.1007/978-1-0716-1201-9_2.

96. Peter, J.; Schacherer, J. Population Genomics of Yeasts: Towards a Comprehensive View across a Broad Evolutionary Scale. Yeast 2016, 33, 73–81, doi:10.1002/YEA.3142.

97. Bendixsen, D.P.; Frazão, J.G.; Stelkens, R. Saccharomyces Yeast Hybrids on the Rise. Yeast 2021, doi:10.1002/YEA.3684.

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, doi:10.1038/NATURE07670.

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, doi:10.1038/S41586-018-0030-5.

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 α-Synuclein through Specific Molecular Pathways. Cell Syst 2017, 4, 157-170.e14, doi:10.1016/J.CELS.2016.12.011.

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, doi:10.1038/S41514-018-0022-6.