

Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits

Anatole Ghazalpour · Christoph D. Rau · Charles R. Farber · Brian J. Bennett · Luz D. Orozco · Atila van Nas · Calvin Pan · Hooman Allayee · Simon W. Beaven · Mete Civelek · Richard C. Davis · Thomas A. Drake · Rick A. Friedman · Nick Furlotte · Simon T. Hui · J. David Jentsch · Emrah Kostem · Hyun Min Kang · Eun Yong Kang · Jong Wha Joo · Vyacheslav A. Korshunov · Rick E. Laughlin · Lisa J. Martin · Jeffrey D. Ohmen · Brian W. Parks · Matteo Pellegrini · Karen Reue · Desmond J. Smith · Sotirios Tetradis · Jessica Wang · Yibin Wang · James N. Weiss · Todd Kirchgessner · Peter S. Gargalovic · Eleazar Eskin · Aldons J. Lusis · Renée C. LeBoeuf

Received: 29 March 2012 / Accepted: 4 July 2012
© Springer Science+Business Media, LLC 2012

Abstract We have developed an association-based approach using classical inbred strains of mice in which we correct for population structure, which is very extensive in mice, using an efficient mixed-model algorithm. Our approach includes inbred parental strains as well as recombinant inbred strains in order to capture loci with effect sizes typical of complex traits in mice (in the range of 5 % of total trait variance). Over the last few years, we have typed the hybrid mouse diversity panel (HMDP)

strains for a variety of clinical traits as well as intermediate phenotypes and have shown that the HMDP has sufficient power to map genes for highly complex traits with resolution that is in most cases less than a megabase. In this essay, we review our experience with the HMDP, describe various ongoing projects, and discuss how the HMDP may fit into the larger picture of common diseases and different approaches.

Anatole Ghazalpour, Christoph D. Rau, Charles R. Farber, Brian J. Bennett, Luz D. Orozco, and Atila van Nas contributed equally to this work.

A. Ghazalpour · M. Civelek · R. C. Davis · S. T. Hui · B. W. Parks · A. J. Lusis
Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

C. D. Rau · J. N. Weiss · A. J. Lusis
Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, CA, USA

C. R. Farber
Departments of Medicine and Biochemistry and Molecular Genetics, and Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

B. J. Bennett
Department of Genetics, and Nutrition Research Institute, University of North Carolina, Chapel Hill, NC, USA

L. D. Orozco · A. van Nas · C. Pan · J. W. Joo · L. J. Martin · K. Reue · J. Wang · E. Eskin · A. J. Lusis
Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

H. Allayee
Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

S. W. Beaven
Division of Digestive Diseases, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

T. A. Drake
Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

R. A. Friedman
Department of Otolaryngology/Head and Neck Surgery, House Research Institute, Los Angeles, CA, USA

N. Furlotte · E. Kostem · E. Y. Kang · E. Eskin
Department of Computer Sciences, University of California, Los Angeles, CA, USA

J. D. Jentsch
Department of Psychology & Behavioral Neuroscience and Psychiatry & Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

H. M. Kang
Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA

Introduction

The human genome-wide association studies (GWAS) of the last several years have provided the first unbiased views of the genetics of common complex diseases, such as coronary artery disease, diabetes, and cancer. Many of the loci contain novel genes not previously connected to their respective disease, indicating that there is great potential to discover new pathways and new targets for therapeutic intervention. These GWAS do, however, have some important limitations. First, human GWAS are not well powered to study genetic interactions, such as gene-by-gene or gene-by-environment interactions (Zuk et al. 2012). Second, it will be difficult to move from locus to a disease pathway directly in humans (Altshuler et al. 2008). And third, for most diseases, GWAS have identified only a small fraction of the total genetic contributions and, thus, there is a great deal more to be discovered (Altshuler et al. 2008; Manolio et al. 2009).

To simplify genetic analysis, natural variations relevant to disease have been studied in mice and rats (Ahlqvist et al. 2011; Flint and Mackay 2009; Keane et al. 2011). This has generally involved traditional linkage mapping methods with crosses between different strains to identify quantitative trait loci (QTLs). An important problem with such analysis has been poor mapping resolution because the QTLs generally contain hundreds of genes, making the identification of the causal genes difficult.

To address these limitations, we have developed an association-based approach using classical inbred strains of mice (Bennett et al. 2010). We follow previous attempts to apply association in mice (Cervino et al. 2007; Grupe et al. 2001; Guo et al. 2007; Liao et al. 2004; Liu et al. 2007; Pletcher et al. 2004) with two differences. First, we correct for population structure, which is very extensive in mice, using an efficient mixed-model algorithm (EMMA) (Kang et al. 2008). Second, to capture loci with effect sizes typical of complex traits in mice (in the range of 5 % of total trait variance), we supplemented the population with recombinant inbred (RI) strains.

Over the last few years, we have typed the hybrid mouse diversity panel (HMDP) strains for a variety of clinical traits as well as intermediate phenotypes, and have shown that the HMDP has sufficient power to map genes for highly complex traits with resolution that is in most cases less than a megabase. In this essay, we review our experience with the HMDP, describe various ongoing projects, and discuss how the HMDP may fit into the larger picture of common diseases and different approaches.

Overview of the HMDP

The hybrid mouse diversity panel (HMDP) consists of a population of over 100 inbred mouse strains selected for usage in systematic genetic analyses of complex traits

J. W. Joo

Bioinformatics Program, David Geffen School of Medicine,
University of California, Los Angeles, CA, USA

V. A. Korshunov

Department of Medicine, Aab Cardiovascular Research Institute,
University of Rochester School of Medicine and Dentistry,
Rochester, NY, USA

R. E. Laughlin

Semel Institute for Neuroscience and Human Behavior,
University of California, Los Angeles, CA, USA

J. D. Ohmen

Department of Cell Biology and Genetics,
House Research Institute, Los Angeles, CA, USA

M. Pellegrini

Department of Molecular, Cell and Developmental Biology,
University of California, Los Angeles, CA, USA

D. J. Smith

Department of Molecular and Medical Pharmacology,
David Geffen School of Medicine, University of California,
Los Angeles, CA, USA

S. Tetradis

Molecular Biology Institute, David Geffen School of Medicine,
University of California, Los Angeles, CA, USA

S. Tetradis · A. J. Lusis

Division of Diagnostic and Surgical Science, School of
Dentistry, University of California, Los Angeles, CA, USA

J. Wang

Department of Medicine/Cardiology, David Geffen School
of Medicine, University of California, Los Angeles, CA, USA

Y. Wang

Division of Molecular Medicine, Department of Anesthesiology,
Physiology and Medicine, David Geffen School of Medicine,
University of California, Los Angeles, CA, USA

T. Kirchgessner · P. S. Gargalovic

Department of Cardiovascular Drug Discovery,
Bristol-Myers Squibb Co, Pennington, NJ, USA

R. C. LeBoeuf (✉)

Division of Metabolism, Endocrinology and Nutrition,
Department of Medicine, University of Washington,
850 Republican Street, Seattle, WA 98109-4725, USA
e-mail: leboeuf@u.washington.edu

(Table 1). Our goals in selecting the strains were to (1) increase resolution of genetic mapping, (2) have a renewable resource that is available to all investigators worldwide, and (3) provide a shared data repository that would allow the integration of data across multiple scales, including genomic, transcriptomic, metabolomic, proteomic, and clinical phenotypes. The core of our panel for association mapping (Bennett et al. 2010; Cervino et al. 2007; Grupe et al. 2001) consists of 29 classic parental inbred strains which are a subset of a group of mice commonly called the mouse diversity panel. We settled on our strains by eliminating closely related strains and removing wild-derived strains. The decision to remove wild-derived strains is based on the tradeoff between statistical power and genetic diversity. While we were sacrificing the genetic diversity by leaving out wild-derived strains, our panel increased the statistical power (assuming the same number of animals) to identify genetic variants polymorphic among the classical inbred strains which affect traits, and these variants account for a tremendous amount of phenotypic diversity among the classical inbred strains.

In order to increase power, we included panels of RI mice, including the BXD, CXB, BXA/AXB, and BXH panels. Power calculations with the inclusion of these additional strains indicated that we have 70 % power to detect SNPs that contribute ~ 10 % of the overall variance of a complex trait (Bennett et al. 2010). We have recently shown that power can be further increased by performing meta-analysis in which data from the HMDP are combined with data from traditional crosses (Furlotte et al. 2012). Power can also be increased by typing additional commercially available RI panels, as discussed below.

A key feature of the panel is that genotyping is not necessary due to the wealth of single nucleotide polymorphism (SNP) genotypes known across the mouse strains (Keane et al. 2011; Kirby et al. 2010). The inbred strains used for the HMDP were previously genotyped by the Broad Institute and then combined with genotypes from the Wellcome Trust Center for Human Genetics (WTCHG) (Table 2). Genotypes of RI strains at the Broad Institute were inferred from WTCHG genotypes by interpolating alleles at polymorphic SNPs among parental strains, calling ambiguous genotypes missing. Of the 140,000 SNPs available, 107,145 were informative with an allele frequency >5 % and were used for GWAS in our publications (Bennett et al. 2010; Farber et al. 2011; Park et al. 2011). Additional genotyping classifications have been performed recently (Keane et al. 2011; Kirby et al. 2010), and the resulting 4 million SNP genotypes are freely available (Table 2).

In the current HMDP panel consisting of over 100 inbred and RI strains, we used $\sim 860,000$ SNPs to examine linkage disequilibrium (LD) blocks. These SNPs have

greater than 5 % minor allele frequency (mean and median minor allele frequency of 28 and 31 %) and are polymorphic between the strains. Using 0.8 as the r^2 cutoff to define LD, there are a total of 13,706 LD blocks with the median size of 42.8 kb per block (mean of 143.3 kb per block) scattered throughout the genome (Fig. 1). Since LD blocks are more likely to define the window in which a candidate gene for a locus resides, the presence of small LD in the HMDP suggests that the number of causal candidate genes will be on average fewer than five genes per locus. This is a considerable improvement over mapping resolution in traditional linkage studies and/or in outbred stock mice where, on average, the number of candidate genes for each locus ranges from 10 to 50. Large blocks (defined as LD blocks >1 Mb) are also present in the HMDP panel but not frequently. Only 1.5 % of the LD blocks (211 of 13,706 total) have a large size encompassing 12.5 % of the genome. The X chromosome is noted to contain multiple large LD blocks, suggesting that mapping resolution for this chromosome is reduced as compared to the autosomes. These large blocks reflect the regions in the genome that were inherited by all strains from the shared ancestors as described in (Frazer et al. 2007). The high-resolution mapping property of the HMDP becomes particularly important in systems genetics as one common goal is to identify causal genes that coordinately regulate network function. Such drivers have been reported in numerous studies as candidate genes for “QTL hot spots,” but the true identity of such drivers has been elusive mainly due to lack of resolution in mapping.

In addition to the excellent resolution, the HMDP has important advantages for systems genetics and for analysis of genetic interactions. The progeny from a genetic cross are unique and as such can be characterized for a limited number of phenotypes, whereas the inbred strains of the HMDP can be examined for an unlimited number of phenotypes since the data are cumulative. The same concept applies to interactions. Thus, mice of the same genotype can be examined under a variety of conditions to identify gene-by-environment interactions, and epistatic interactions can be tested using targeted perturbations on specific genetic backgrounds. This is also an important feature of other replicate mouse genetic systems such as consomic strains, collaborative cross strains, and wild-derived inbred strains.

Discoveries using the HMDP

The successful use of the HMDP to identify complex trait genes was recently highlighted in a study of bone mineral density (BMD) (Farber et al. 2011). BMD is a polygenic phenotype that is commonly investigated in human and

Table 1 The 114 strains typed within the HMDP for metabolic phenotypes

	JAX ID	JAX strain name	EMMA name	JAX URL	Type
1	691	129X1/SvJ	129X1/SvJ	http://jaxmice.jax.org/strain/000691.html	Classic
2	646	A/J	A/J	http://jaxmice.jax.org/strain/000646.html	
3	648	AKR/J	AKR/J	http://jaxmice.jax.org/strain/000648.html	
4	651	BALB/cJ	BALB/cJ	http://jaxmice.jax.org/strain/000651.html	
5	2282	BTBR T<+> tf/J	BTBRT<+>tf/J	http://jaxmice.jax.org/strain/002282.html	
6	653	BUB/BnJ	BUB/BnJ	http://jaxmice.jax.org/strain/000653.html	
7	659	C3H/HeJ	C3H/HeJ	http://jaxmice.jax.org/strain/000659.html	
8	664	C57BL/6J	C57BL/6J	http://jaxmice.jax.org/strain/000664.html	
9	662	C57BLKS/J	C57BLKS/J	http://jaxmice.jax.org/strain/000662.html	
10	668	C57L/J	C57L/J	http://jaxmice.jax.org/strain/000668.html	
11	669	C58/J	C58/J	http://jaxmice.jax.org/strain/000669.html	
12	656	CBA/J	CBA/J	http://jaxmice.jax.org/strain/000656.html	
13	657	CE/J	CE/J	http://jaxmice.jax.org/strain/000657.html	
14	671	DBA/2J	DBA/2J	http://jaxmice.jax.org/strain/000671.html	
15	1800	FVB/NJ	FVB/NJ	http://jaxmice.jax.org/strain/001800.html	
16	674	I/LnJ	I/LnJ	http://jaxmice.jax.org/strain/000674.html	
17	2106	KK/HIJ	KK/HIJ	http://jaxmice.jax.org/strain/002106.html	
18	675	LG/J	LG/J	http://jaxmice.jax.org/strain/000675.html	
19	676	LP/J	LP/J	http://jaxmice.jax.org/strain/000676.html	
20	677	MA/MyJ	MA/MyJ	http://jaxmice.jax.org/strain/000677.html	
21	1976	NOD/ShiLtJ	NOD/LtJ	http://jaxmice.jax.org/strain/001976.html	
22	2423	NON/ShiLtJ	NON/LtJ	http://jaxmice.jax.org/strain/002423.html	
23	684	NZB/BINJ	NZB/BINJ	http://jaxmice.jax.org/strain/000684.html	
24	1058	NZW/LacJ	NZW/LacJ	http://jaxmice.jax.org/strain/001058.html	
25	680	PL/J	PL/J	http://jaxmice.jax.org/strain/000680.html	
26	683	RIIS/J	RIIS/J	http://jaxmice.jax.org/strain/000683.html	
27	644	SEA/GnJ	SEA/GnJ	http://jaxmice.jax.org/strain/000644.html	
28	686	SJL/J	SJL/J	http://jaxmice.jax.org/strain/000686.html	
29	687	SM/J	SM/J	http://jaxmice.jax.org/strain/000687.html	
30	689	SWR/J	SWR/J	http://jaxmice.jax.org/strain/000689.html	
31	1673	AXB1/PgnJ	AXB-1/PgnJ	http://jaxmice.jax.org/strain/001673.html	AXB/BXA
32	1681	AXB10/PgnJ	AXB-10/PgnJ	http://jaxmice.jax.org/strain/001681.html	
33	1683	AXB12/PgnJ	AXB-12/PgnJ	http://jaxmice.jax.org/strain/001683.html	
34	1826	AXB13/PgnJ	AXB-13/PgnJ	http://jaxmice.jax.org/strain/001826.html	
35	1685	AXB15/PgnJ	AXB-15/PgnJ	http://jaxmice.jax.org/strain/001685.html	
36	1687	AXB19/PgnJ	AXB-19/PgnJ	http://jaxmice.jax.org/strain/001687.html	
37	1686	AXB19a/PgnJ	AXB-18/PgnJ	http://jaxmice.jax.org/strain/001686.html	
38	1688	AXB19b/PgnJ	AXB-20/PgnJ	http://jaxmice.jax.org/strain/001688.html	
39	1674	AXB2/PgnJ	AXB-2/PgnJ	http://jaxmice.jax.org/strain/001674.html	
40	1690	AXB23/PgnJ	AXB-23/PgnJ	http://jaxmice.jax.org/strain/001690.html	
41	1691	AXB24/PgnJ	AXB-24/PgnJ	http://jaxmice.jax.org/strain/001691.html	
42	1676	AXB4/PgnJ	AXB-4/PgnJ	http://jaxmice.jax.org/strain/001676.html	
43	1677	AXB5/PgnJ	AXB-5/PgnJ	http://jaxmice.jax.org/strain/001677.html	
44	1678	AXB6/PgnJ	AXB-6/PgnJ	http://jaxmice.jax.org/strain/001678.html	
45	1679	AXB8/PgnJ	AXB-8/PgnJ	http://jaxmice.jax.org/strain/001679.html	
46	1692	BXA1/PgnJ	BXA-1/PgnJ	http://jaxmice.jax.org/strain/001692.html	
47	1699	BXA11/PgnJ	BXA-11/PgnJ	http://jaxmice.jax.org/strain/001699.html	

Table 1 continued

	JAX ID	JAX strain name	EMMA name	JAX URL	Type
48	1700	BXA12/PgnJ	BXA-12/PgnJ	http://jaxmice.jax.org/strain/001700.html	
49	1701	BXA13/PgnJ	BXA-13/PgnJ	http://jaxmice.jax.org/strain/001701.html	
50	1702	BXA14/PgnJ	BXA-14/PgnJ	http://jaxmice.jax.org/strain/001702.html	
51	1703	BXA16/PgnJ	BXA-16/PgnJ	http://jaxmice.jax.org/strain/001703.html	
52	1693	BXA2/PgnJ	BXA-2/PgnJ	http://jaxmice.jax.org/strain/001693.html	
53	1710	BXA24/PgnJ	BXA-24/PgnJ	http://jaxmice.jax.org/strain/001710.html	
54	1711	BXA25/PgnJ	BXA-25/PgnJ	http://jaxmice.jax.org/strain/001711.html	
55	1999	BXA26/PgnJ	BXA-26/PgnJ	http://jaxmice.jax.org/strain/001999.html	
56	1694	BXA4/PgnJ	BXA-4/PgnJ	http://jaxmice.jax.org/strain/001694.html	
57	1696	BXA7/PgnJ	BXA-7/PgnJ	http://jaxmice.jax.org/strain/001696.html	
58	1697	BXA8/PgnJ	BXA-17/PgnJ	http://jaxmice.jax.org/strain/001697.html	
59	1697	BXA8/PgnJ	BXA-8/PgnJ	http://jaxmice.jax.org/strain/001697.html	
60	36	BXD1/TyJ	BXD-1/TyJ	http://jaxmice.jax.org/strain/000036.html	BXD
61	12	BXD11/TyJ	BXD-11/TyJ	http://jaxmice.jax.org/strain/000012.html	
62	45	BXD12/TyJ	BXD-12/TyJ	http://jaxmice.jax.org/strain/000045.html	
63	40	BXD13/TyJ	BXD-13/TyJ	http://jaxmice.jax.org/strain/000040.html	
64	329	BXD14/TyJ	BXD-14/TyJ	http://jaxmice.jax.org/strain/000329.html	
65	95	BXD15/TyJ	BXD-15/TyJ	http://jaxmice.jax.org/strain/000095.html	
66	13	BXD16/TyJ	BXD-16/TyJ	http://jaxmice.jax.org/strain/000013.html	
67	15	BXD18/TyJ	BXD-18/TyJ	http://jaxmice.jax.org/strain/000015.html	
68	10	BXD19/TyJ	BXD-19/TyJ	http://jaxmice.jax.org/strain/000010.html	
69	75	BXD2/TyJ	BXD-2/TyJ	http://jaxmice.jax.org/strain/000075.html	
70	330	BXD20/TyJ	BXD-20/TyJ	http://jaxmice.jax.org/strain/000330.html	
71	77	BXD21/TyJ	BXD-21/TyJ	http://jaxmice.jax.org/strain/000077.html	
72	43	BXD22/TyJ	BXD-22/TyJ	http://jaxmice.jax.org/strain/000043.html	
73	31	BXD24/TyJ-Cep290<rd16>/J	BXD-24/TyJ	http://jaxmice.jax.org/strain/000031.html	
74	41	BXD27/TyJ	BXD-27/TyJ	http://jaxmice.jax.org/strain/000041.html	
75	47	BXD28/TyJ	BXD-28/TyJ	http://jaxmice.jax.org/strain/000047.html	
76	29	BXD29-Tlr4<lps-2J>/J	BXD-29/TyJ	http://jaxmice.jax.org/strain/000029.html	
77	83	BXD31/TyJ	BXD-31/TyJ	http://jaxmice.jax.org/strain/000083.html	
78	78	BXD32/TyJ	BXD-32/TyJ	http://jaxmice.jax.org/strain/000078.html	
79	3222	BXD33/TyJ	BXD-33/TyJ	http://jaxmice.jax.org/strain/003222.html	
80	3223	BXD34/TyJ	BXD-34/TyJ	http://jaxmice.jax.org/strain/003223.html	
81	3225	BXD36/TyJ	BXD-36/TyJ	http://jaxmice.jax.org/strain/003225.html	
82	3227	BXD38/TyJ	BXD-38/TyJ	http://jaxmice.jax.org/strain/003227.html	
83	3228	BXD39/TyJ	BXD-39/TyJ	http://jaxmice.jax.org/strain/003228.html	
84	3229	BXD40/TyJ	BXD-40/TyJ	http://jaxmice.jax.org/strain/003229.html	
85	3230	BXD42/TyJ	BXD-42/TyJ	http://jaxmice.jax.org/strain/003230.html	
86	37	BXD5/TyJ	BXD-5/TyJ	http://jaxmice.jax.org/strain/000037.html	
87	7	BXD6/TyJ	BXD-6/TyJ	http://jaxmice.jax.org/strain/000007.html	
88	84	BXD8/TyJ	BXD-8/TyJ	http://jaxmice.jax.org/strain/000084.html	
89	105	BXD9/TyJ	BXD-9/TyJ	http://jaxmice.jax.org/strain/000105.html	
90	3787	B6cC3-1/KccJ	BXHE1	http://jaxmice.jax.org/strain/003787.html	BXH
91	32	BXH10/TyJ	BXH-10/TyJ	http://jaxmice.jax.org/strain/000032.html	
92	9	BXH14/TyJ	BXH-14/TyJ	http://jaxmice.jax.org/strain/000009.html	
93	33	BXH19/TyJ	BXH-19/TyJ	http://jaxmice.jax.org/strain/000033.html	
94	34	BXH2/TyJ	BXH-2/TyJ	http://jaxmice.jax.org/strain/000034.html	

Table 1 continued

	JAX ID	JAX strain name	EMMA name	JAX URL	Type
95	3784	BXH20/KccJ	BXHA1	http://jaxmice.jax.org/strain/003784.html	
96	3786	BXH22/KccJ	BXHB2	http://jaxmice.jax.org/strain/003786.html	
97	11	BXH4/TyJ	BXH-4/TyJ	http://jaxmice.jax.org/strain/000011.html	
98	38	BXH6/TyJ	BXH-6/TyJ	http://jaxmice.jax.org/strain/000038.html	
99	14	BXH7/TyJ	BXH-7/TyJ	http://jaxmice.jax.org/strain/000014.html	
100	76	BXH8/TyJ	BXH-8/TyJ	http://jaxmice.jax.org/strain/000076.html	
101	8	BXH9/TyJ	BXH-9/TyJ	http://jaxmice.jax.org/strain/000008.html	
102	351	CXB1/ByJ	CXB-1/ByJ	http://jaxmice.jax.org/strain/000351.html	CxB
103	1631	CXB10/HiAJ	CXB-10/HiAJ	http://jaxmice.jax.org/strain/001631.html	
104	1632	CXB11/HiAJ	CXB-11/HiAJ	http://jaxmice.jax.org/strain/001632.html	
105	1633	CXB12/HiAJ	CXB-12/HiAJ	http://jaxmice.jax.org/strain/001633.html	
106	1634	CXB13/HiAJ	CXB-13/HiAJ	http://jaxmice.jax.org/strain/001634.html	
107	352	CXB2/ByJ	CXBE	http://jaxmice.jax.org/strain/000352.html	
108	353	CXB3/ByJ	CXB-3/ByJ	http://jaxmice.jax.org/strain/000353.html	
109	354	CXB4/ByJ	CXBH	http://jaxmice.jax.org/strain/000354.html	
110	355	CXB5/ByJ	CXB-5/ByJ	http://jaxmice.jax.org/strain/000355.html	
111	356	CXB6/ByJ	CXB-6/ByJ	http://jaxmice.jax.org/strain/000356.html	
112	357	CXB7/ByJ	CXB-7/ByJ	http://jaxmice.jax.org/strain/000357.html	
113	1629	CXB8/HiAJ	CXB-8/HiAJ	http://jaxmice.jax.org/strain/001629.html	
114	1630	CXB9/HiAJ	CXB-9/HiAJ	http://jaxmice.jax.org/strain/001630.html	

These strains are commercially available from The Jackson Laboratory (JAX) (Bar Harbor, ME). Additional strains will soon be entered into the HMDP data base (see Table 2). For entry of phenotypic measures into the HMDP database and for calculating associations using EMMA programming (see Table 2), the JAX strain names need to be converted to the EMMA strain names that we have provided here

Table 2 URL sites useful for the HMDP as well as sites used to develop the HMDP, including sites at The Jackson Laboratory (JAX) (Bar Harbor, ME)

Name	Description	URL
EMMA	Obtain EMMA in an R package	http://mouse.cs.ucla.edu/emma/
EMMA server	Input data and obtain analysis	http://mouse.cs.ucla.edu/emmaserver/
EMMA power simulator	R package allowing statistical power experiment via in silico mapping	http://mouse.cs.ucla.edu/power/
EMMA study design Webservice	Provides power simulations with various background genetic effects and threshold estimations	http://mouse.cs.ucla.edu/emmaserver/powerSimulation/
Mouse HapMap	Latest genetic maps that are regularly updated	http://mouse.cs.ucla.edu/mousehapmap/
Mouse genome informatics (JAX)	Search tools for gene names and chromosomal positions, phenotypes, expression, orthology, and other features	http://www.informatics.jax.org/
SGR HMDP database	Database for HMDP and several F2 studies with cQTL, eQTL, heatmap and correlation analyses	http://systems.genetics.ucla.edu
Genenetwork	Study relationships among phenotypes and genotypes	http://www.genenetwork.org/webqtl/main.py
Mouse phenome database (JAX)	Phenotypes across hundreds of strains contributed by over 200 investigators	http://phenome.jax.org/
Wellcome trust center for human genetics	Center to understand the genetic foundations of human variation and disease	http://www.well.ox.ac.uk/home
Perlegen	Perlegen mouse SNP browser covering 12 classical inbred strains and 4 wild-derived strains	http://mouse.cs.ucla.edu/perlegen/
Mouse genomes project	Nucleotide sequence database for many key mouse strains	http://sanger.ac.uk/resources/mouse/genomes/

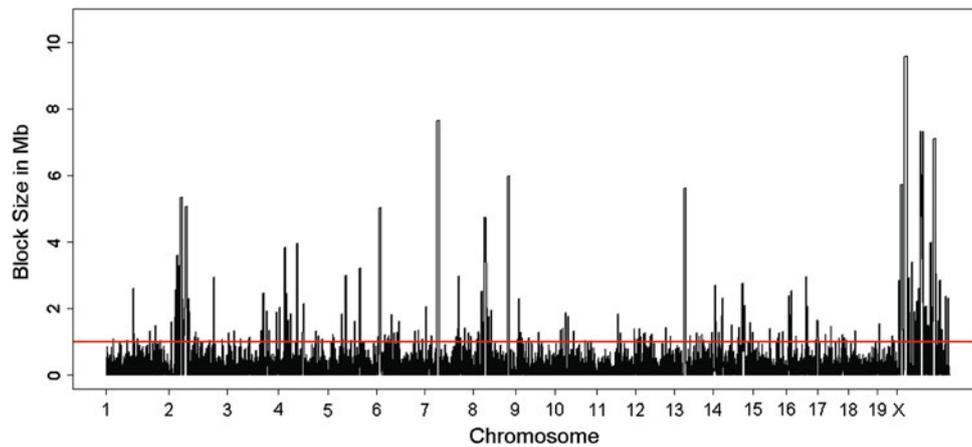


Fig. 1 Map of linkage disequilibrium (LD) blocks along the genome in the hybrid mouse diversity panel (HMDP) population. The LD blocks were determined using $\sim 860,000$ SNPs that have $>5\%$ minor allele frequency (mean and median minor allele frequency of 28 and 31 %) and are polymorphic between the strains. Using 0.8 as the r^2 cutoff to define LD, there are a total of 13,706 LD blocks with a

median size of 42.8 kb per block (mean of 143.3 kb per block) scattered throughout the genome. Only 1.5 % of the LD blocks (211 of 13,706 total) have a large size encompassing 12.5 % of the genome. The location of the large blocks in the genome can be identified by lines that cross the red bar. See text for more details

rodent genetic studies and is the single strongest predictor of osteoporotic fracture (Cummings et al. 2002; Farber and Rosen 2010). For this study, total body, spine, and femur areal BMD data were generated on 16 week-old male mice from 96 HMDP strains. The whole bone transcriptome was also profiled using Illumina gene expression microarrays. The authors used EMMA to perform genome-wide association for the three BMD measures. A total of four genome-wide significant associations were identified on Chromosomes (Chrs) 7, 11, 12, and 17, each affecting BMD at one or more sites. The Chr 12 association for total-body BMD was chosen for further analysis since the 3 Mb window surrounding the association contained only 14 candidate genes. Interestingly, the most significant association was with a nonsynonymous SNP (rs29131970) in the additional sex-combs like 2 (*Asxl2*) gene (Fig. 2). This polymorphism was predicted to have deleterious effects on ASXL2 protein function. To gain further support for *Asxl2* being the causal gene, existing human genome-wide association data (generated in $\sim 6,000$ Icelandic subjects) was used to evaluate SNPs within the human syntenic region for association with BMD (Styrkarsdottir et al. 2008). One SNP (rs7563012) was significant after Bonferroni correction and was located in intron 3 of the human *ASXL2* gene (Fig. 2). Together these data suggested that *Asxl2* influenced BMD in both humans and mice. Consistent with this hypothesis, BMD was found to be lower in *Asxl2*^{-/-} knockout mice.

Network analysis in the HMDP is another powerful approach for investigating complex traits from a systems-level perspective. Coexpression networks can be used to annotate genes of unknown function based on the known

functions of the genes to which they are most closely connected. This “guilt by association” approach has been shown to be a robust gene annotation tool (Wolfe et al. 2005) and was used to determine the mechanism through which *Asxl2* influenced BMD. Weighted Gene Co-expression Network Analysis (WGNCA) was first used to generate a coexpression network using the bone microarray data, which identified *Asxl2* as being connected to genes involved in myeloid cell differentiation. In bone, osteoclasts are bone-resorbing cells of myeloid origin (Teitelbaum and Ross 2003). Additionally, in a human protein-protein interaction network, ASXL2 interacts with TRAF6, a key component of the major signaling pathway regulating osteoclastogenesis (Teitelbaum and Ross 2003). Thus, based on network inferences, *Asxl2* was predicted to be involved in the differentiation of osteoclasts. To test this prediction, expression of *Asxl2* was knocked down in osteoclast precursors. An $\sim 50\%$ reduction in *Asxl2* transcript levels inhibited the formation of TRAP⁺ (a marker of mature osteoclasts) multinuclear cells. These data suggested that *Asxl2* influences BMD, at least in part, through its regulation of osteoclastogenesis. This work highlights the ability of using the HMDP and systems genetics to move from association to gene to mechanism in a single step.

In another study, (Park et al. 2011) utilized the HMDP resource to report on gene networks associated with conditional fear. In this study, the authors combined behavioral phenotypes with gene expression data in two regions of the brain (striatum and hippocampus) to identify groups of genes that coordinately regulate the behavior of the animals. Overall, they observed significant overlap between

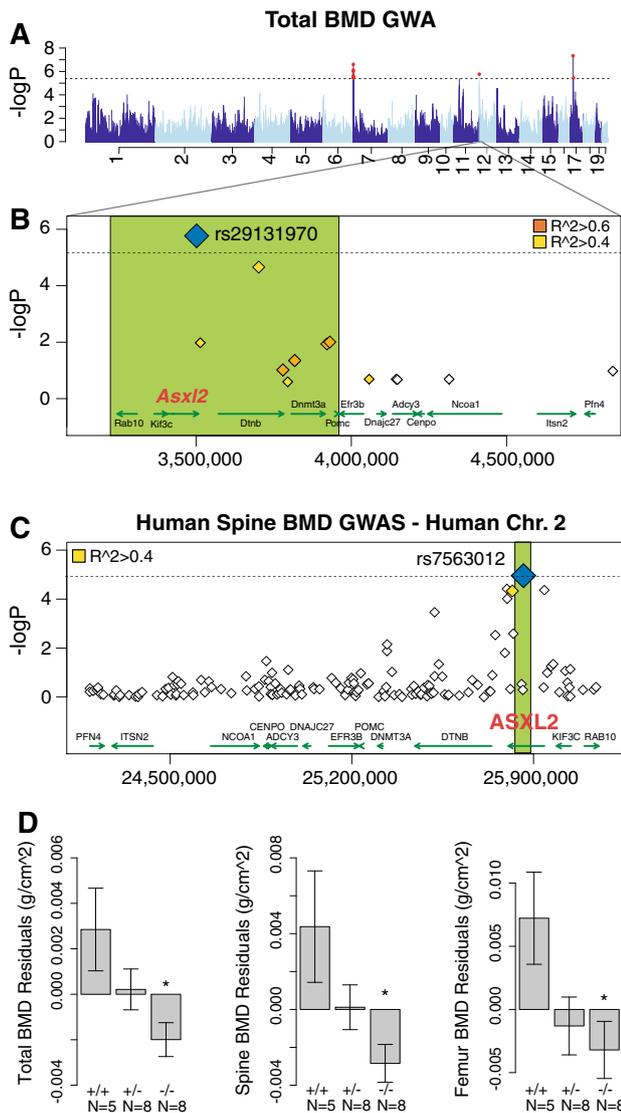


Fig. 2 Variation in *Asxl2* in mice and humans is associated with bone mineral density (BMD). **a** Genome-wide association in the HMDP for total BMD identifies an association on chromosome (Chr) 12. **b** A nonsynonymous SNP (rs29131970) in *Asxl2* that was predicted to alter protein function was the most significantly associated Chr 12 SNP in the HMDP. **c** Human SNPs within *ASXL2* were also associated with BMD in ~6,000 Icelandic individuals. **d** Male mice deficient in *Asxl2* ($-/-$) display significant decreases relative to wild-type controls ($+/+$) in total BMD, spine BMD, and femur BMD residuals after adjustments for age and body weight. Data shown in **d** are residual mean \pm SEM, $*P < 0.05$

the local QTLs and QTL hotspots in the two tissues, as well as module conservation and preservation of highly connected genes (also known as “hubs”) in the striatum and hippocampus networks. The authors were also able to identify tissue-specific network modules between the striatum and the hippocampus, and after performing functional enrichment analysis of the modules, they arrived at pathways likely to contribute to the differences in hippocampus

and striatum function. Finally, using modules as functional units, the authors were able to demonstrate correlations with behavioral traits, thus helping to prioritize candidate genes and pathways for behavioral traits.

In addition to identifying cellular mechanisms and genes underlying physiological traits, the HMDP has been used to investigate the relationships across various biological scales at the global level (Ghazalpour et al. 2011). For example, among the intermediate phenotypes that have been examined in liver are transcript levels (in triplicate, using the Affymetrix platform) and a set of peptides corresponding to about 1,000 proteins (using quantitative mass spectrometry analysis) (Ghazalpour et al. 2011). The correlation between protein and transcript levels was quite weak, with a correlation coefficient of less than 0.4 in most cases, similar to what has been observed in previous studies with yeast and worms. More surprising was the finding that transcript levels were much more strongly correlated with clinical traits (primarily metabolic) than were protein levels. One possible explanation is that transcript levels may be reactive rather than causal with respect to physiologic traits (Ghazalpour et al. 2011).

Overall, the HMDP is being used to develop a multi-scale understanding of a number of complex traits, including a recent report on elevated heart rate (Smolock et al. 2012). There are already over 70 traditional clinical traits reported and there are ongoing studies related to diet-induced obesity, hearing loss, heart failure, atherosclerosis, lipoprotein metabolism, bone metabolism, vascular injury, hematopoietic stem cells, air pollution, gut flora, addictive behavior, hepatotoxicity, and diabetic complications. In addition, gene expression microarrays have been used to quantify mRNA levels in liver, bone, adipose, brain, peritoneal macrophages, aorta, and heart, and proteomic and metabolomic profiling has been performed in liver.

Integration of the HMDP with other resources

The HMDP is just one of several recently proposed approaches to improve the resolution of mouse genetic studies. Other approaches include the Collaborative Cross (CC) (Churchill et al. 2004), outbred designs (Valdar et al. 2006; Yalcin et al. 2010), and the use of consomic strains (Gregorova et al. 2008; Singer et al. 2004; Takada et al. 2008). Each approach has advantages and disadvantages relative to the HMDP. The CC is a recently developed panel of RI strains that are descendants from eight founder strains. A key difference between the HMDP and the CC is that three of the CC founders are wild-derived strains. Wild-derived strains introduce a significantly larger amount of genetic variation and corresponding phenotypic variation compared to the HMDP. For this reason, it is

likely that more genes in the CC will have effects on traits than in the HMDP. However, the additional variation may make it relatively more difficult to map quantitative loci polymorphic in both panels since the increased variation of the CC and outbred panels will reduce the relative effect size of the same variant compared to the HMDP (Kang et al. 2008). The CC will ultimately contain approximately 300 strains compared to the current 100 strains of the HMDP. However, the HMDP may be enlarged to 260 strains (see Future directions and conclusions section) (Collaborative Cross Consortium 2012). Both the CC and the HMDP use inbred strains so they share the advantage of accumulation of data on each strain over time as more and more studies are performed. Utilizing inbred strains also facilitates performing studies with perturbations because identical animals can be phenotyped both with and without a perturbation.

An advantage of outbred designs is that they have higher resolution than the HMDP (up to 100 kb). On the other hand, a disadvantage in outbred designs, either utilizing a specially designed Heterogeneous Stock (Valdar et al. 2006) or commercially available outbred mice (Yalcin et al. 2010), is that each animal is unique. Overall, several mapping strategies are, or will be, available to tackle the genetics of complex diseases.

Consomic strains are also proving valuable for reducing genetic intervals containing candidate genes (Hoover-Plow et al. 2006; Prows et al. 2008) and for identifying causal genes (Burrage et al. 2010). However, consomic strains have limited uses for initial genomic studies, as only two alleles are sampled and consomic strains by definition consist of large areas of LD which were captured from the donor strain during breeding. Nonetheless, consomic strains have enriched the ability to test for epistasis and to reduce genetic intervals by the generation of congenic strains.

Resources for design and analysis of HMDP

UCLA maintains several resources for the design and analysis of HMDP studies (Table 2). The two most relevant to investigators are the EMMA association webserver and EMMA design webserver. Investigators utilizing the HMDP can upload their collected phenotypes to the association webserver which will perform association mapping using EMMA and return population structure-corrected *P* values for each SNP. The analysis is performed on a high-performance computing infrastructure at UCLA, eliminating the need for investigators applying the HMDP to invest in computational resources to perform the analysis. The EMMA design webserver allows an investigator

to estimate the power of a proposed study design through simulations (Kirby et al. 2010) that can help guide an investigator in the design of HMDP studies. In addition, curated genotypes of the HMDP strains are also available.

A systems genetics database

Because the HMDP mice are inbred, with fixed genotypes, the data generated from their study are cumulative. To facilitate the integration and analysis of such data, we also developed a database called the Systems Genetics Resource (SGR). The database comprises mouse genomic, transcriptomic, metabolomic, proteomic, and clinical trait data from the HMDP as well as selected traditional mouse crosses and several human studies. The data are accompanied by detailed descriptions of how the data were acquired, with protocols and links to related published papers. A summary of current data sets contained in the database is presented in Table 3.

We developed a web-based interface where data can be queried for information on specific gene and trait correlations, gene expression in various tissues, or quantitative trait loci, as well as be downloaded for other types of analyses. Such information can be used, for example, to prioritize candidate genes in genetic studies and *trans*-acting loci can be used to generate hypotheses about regulatory pathways (Fig. 3). The intermediate phenotypes can also be used to model gene networks and causal interactions. The power of the SGR is expected to expand as more data are added. Some of the data is also available in the Genenetwork and the Mouse Genome Informatics databases (Table 2).

The SGR is a resource that can be used to answer specific questions and to understand the relationships among different genes. For example, data generated from primary macrophages of the HMDP helped us to determine that a gene of interest, the interferon inducible helicase 1 (*Ifih1*), shows gene-by-environment interactions and that it is under the control of the inflammatory stimulus bacterial lipopolysaccharide (LPS) (Fig. 3a). Using eQTL, we also found that the expression of *Ifih1* is controlled by three loci on Chrs 5, 8, and 13 (Fig. 3b). We can also compare the expression patterns of *Ifih1* among the different tissues available on the database, which include adipose, aorta, heart, liver, and macrophages treated in three different conditions (Fig. 3c). Similarly, we found that the three *trans*-eQTL on Chrs 5, 8, and 13 are specific to macrophages treated with LPS, but there is also a strong *cis*-eQTL in the liver and other regulatory loci in Chr 2. Such information can allow us to identify candidate regulators, to examine gene-by-environment interactions and tissue specificities, and to prioritize candidate genes for clinical QTL.

Table 3 Data sets currently on line on the systems genetics resource (SGR) database

- (A) Hybrid mouse diversity panel: original panel analysis: chow fed, males, 16 weeks
 (B) Hybrid mouse diversity panel: chow fed, males, 16 weeks: second set
 (C) Hybrid mouse diversity panel: isoproterenol-induced hypertrophy and heart failure
 (D) Hybrid mouse diversity panel: Akita Diabetes Study, pilot
 (E) Hybrid mouse diversity panel: diet-induced obesity
 (F) Hybrid mouse diversity panel: human ApoB100 transgenics
 (G) Hybrid mouse diversity panel: peritoneal macrophages, inflammatory responses
 (H) Hybrid mouse diversity panel: atherosclerosis using ApoE Leiden, CETP transgenics
 (I) Hybrid mouse diversity panel: hearing phenotypes
 (J) Hybrid mouse diversity panel: mandible morphology
 (K) Hybrid mouse diversity panel: attention-related behavioral traits
 (L) Hybrid mouse diversity panel: parental strain survey
 (M) Hybrid mouse diversity panel: vascular injury response
 (N) Hybrid mouse diversity panel: F1 hybrids
 (O) Metabolic syndrome in men (Metsim) study
 (P) Human aortic endothelial cells (EC) culture
 (Q) Human aortic smooth muscle cell (SMC) culture
 (R) (BALB/cBy.Ldlr^{-/-} × C57BL/6JLdlr^{-/-}) F2
 (S) (C57BL/6J × DBA/2J)F2 on db/db background
 (T) (CeH/HeJ.ApoE^{-/-} × C57BL/6J.ApoE^{-/-}) F2
 (U) (C3H/HeJ × C57BL/6J) F2
 (V) (C57BL/6J × CAST/Ei) F2
 (W) Genome tagged mice, global congenic strains: DBA/2J on C57BL/6J background and CAST/Ei on C57BL/6J background
 (X) Recombinant inbred congenic strains, Demant, C3H × C57BL/6
 (Y) (C57BL/6J × DBA/2J) F2

This resource is updated regularly with submissions from the HMDP user groups

Future directions and conclusions

While the resolution of the HMDP is excellent at most loci, the power is marginal. As judged by QTL studies, few loci contributing to complex clinical traits have effect sizes as large as 10 % and most are below 5 % (Flint and Mott 2008). Thus, using the panel of 100 strains (Bennett et al. 2010), only a subset of the loci contributing to complex traits are likely to be identified. As mentioned above, the power can be enhanced by integrating the results from traditional crosses or by expanding the number of inbred and recombinant inbred strains. Recently, two panels of advanced intercross RI lines have become available from The Jackson Laboratory: The LXS RI panel includes 62 strains and 50 more BXD strains have recently been

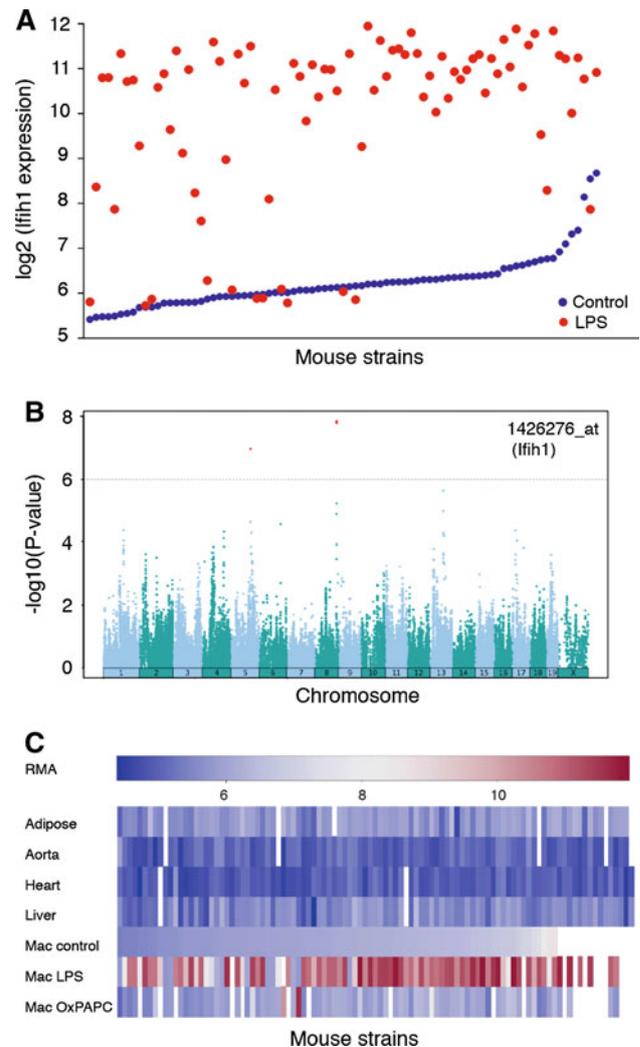


Fig. 3 Database plots for interferon-inducible helicase 1 (*Ifih1*). Sample plots for a given gene of interest that can be obtained from our online database. **a** Lipopolysaccharide (LPS) response of *Ifih1* in macrophages of the HMDP. **b** Genome-wide association for the expression of *Ifih1* in LPS-treated macrophages. **c** Relative expression levels among mouse strains of the HMDP in adipose, aorta, heart, and liver, and for macrophages treated with control, LPS or oxidation products of 1-palmitoyl-2-arachidonyl-*sn*-glycero-3-phosphocholine (OxPAPC) media. Robust microarray average (RMA) refers to an algorithm for gene expression microarray background corrections. We used the Affymetrix GCOS RMA algorithm

developed. Also available from The Jackson Laboratory are several sets of cryopreserved strains: 19 strains from the AKXD RI panel, 13 from the AKXL panel, and 15 from the NXSM panel. Thus, it is possible to increase the size of the HMDP to over 260 strains. Recombinant Inbred Congenic Lines, Chromosome Substitution Strains, and Genome Tagged Mice (Peters et al. 2007) could also be employed to increase both power and resolution. A particularly useful complement to the HMDP will be the CC strain set now being generated (Casici 2012; Threadgill and

Churchill 2012). Preliminary analyses of the partially inbred CC lines have been promising and substantial resources for their characterization, including complete genomic sequencing, have been planned (Collaborative Cross Consortium 2012).

As discussed above, human genetic studies of complex traits are limited in several respects, and the HMDP provides a partial solution to some of these limitations. First, human studies are poorly powered to identify genetic interactions, and these can be addressed more effectively in mice. As mentioned above, the HMDP is very convenient for genetic analysis of environmental interactions because mice of the same genotype can be examined under different conditions. Also, epistasis is likely to complicate studies of diseases such as diabetic complications and atherosclerosis, where one set of genes contributes to a predisposing factor (diabetes and elevated cholesterol, respectively), and another set of genes affects the response to these factors. In the HMDP, sensitizing genes can be introduced by breeding dominant mutations onto each of the HMDP strains and examining the F1 progeny. For example, we have bred a dominant hyperlipidemia-inducing gene, *APOE-Leiden*, onto a number of the HMDP strains and find atherosclerosis to be concordant with previous studies in which recessive mutations were transferred onto different backgrounds (B. Bennett and A. J. Lusis unpublished). Second, it will be difficult to identify, directly in humans, the pathways perturbed by novel GWAS genes. For example, the striking relationship between an allele of *APOE* and Alzheimer's has been known for nearly 20 years and yet the mechanism remains uncertain. Clearly, studies in mice, where access to tissues and environmental conditions can be standardized, will simplify such analyses. Moreover, studying the genes in the context of natural variation, as opposed to transgenic or gene-targeted mice, may well offer important advantages. Third, for most common disease traits, human GWAS have been able to identify only a small fraction of the total heritability. There are undoubtedly many explanations but, clearly, much remains to be discovered. Studies in mice will most likely identify different, although overlapping, gene sets and perhaps different pathways. Traits that are substantially influenced by environmental factors will be addressed with greater power in mice because such factors can be controlled.

The HMDP panel should be useful as a tool for investigation of basic biological processes as well as complex clinical traits. An example is the study by (Ghazalpour et al. 2011) that examines the relationship between transcript levels, protein levels, and metabolic traits. By providing many thousands of genetic perturbations in various combinations, the HMDP enables global dissection of the relationships between biological scales such as DNA methylation, transcription factor binding, histone modification, and transcription.

Acknowledgments Funding for this work was supported in part by the U.S. National Institutes of Health (NIH) grants NIAMS RO1 AR057759 (CRF) and NHLBI K99 HL102223 (BJB); the Ruth L. Kirschstein National Research Service Awards T32 HL69766 (CDR), T32-HL007895 (JW), NINDS RO1 NS050148 and NIGMS RO1 GM098273 (DJS), NHLBI RO1 HL098227 (RCL), NIDDK DP3 DK094311 (AJL, RCD, RCL), NIH RO1 GM095656 (MP), NIDDK RO1 DK072206 (TAD), RO1 DK080339 (STH), 5RO1 DC010856-02 (JDO, RAF), PO1 HL028481 (KR), and PO1 HL090553 (KR). EE was supported by National Science Foundation grants 0513612, 0731455, 0729049, 0916676, and 1065276, and NIH grants K25-HL080079, U01 DA024417, P01-HL30568, and PO1-HL28481. LDO was supported by USPHS National Research Service Award GM07104 and USPHS National Research Service Award T32-HG002536. This research was supported in part by the University of California, Los Angeles, CA, subcontract of contract N01-ES-45530 from the National Toxicology Program and National Institute of Environmental Health Sciences to Perlegen Sciences (EE) and from the Russian Program #2011-1.5-501-006-024 (VAK) and NIH RO1 HL105623 (VAK).

References

- Ahlqvist E, Ekman D, Lindvall T, Popovic M, Forster M, Hultqvist M, Klaczowska D, Teneva I, Johannesson M, Flint J, Valdar W, Nandakumar KS, Holmdahl R (2011) High-resolution mapping of a complex disease, a model for rheumatoid arthritis, using heterogeneous stock mice. *Hum Mol Genet* 20:3031–3041
- Altshuler D, Daly MJ, Lander ES (2008) Genetic mapping in human disease. *Science* 322:881–888
- Bennett BJ, Farber CR, Orozco L, Kang HM, Ghazalpour A, Siemers N, Neubauer M, Neuhaus I, Yordanova R, Guan B, Truong A, Yang WP, He A, Kaye P, Gargalovic P, Kirchgessner T, Pan C, Castellani LW, Kostem E, Furlotte N, Drake TA, Eskin E, Lusis AJ (2010) A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Res* 20:281–290
- Burrage LC, Baskin-Hill AE, Sinasac DS, Singer JB, Croniger CM, Kirby A, Kulbokas EJ, Daly MJ, Lander ES, Broman KW, Nadeau JH (2010) Genetic resistance to diet-induced obesity in chromosome substitution strains of mice. *Mamm Genome* 21:115–129
- Casici T (2012) Mouse genetics: fruits of the Collaborative Cross. *Nat Rev Genet* 13:223
- Cervino AC, Darvasi A, Fallahi M, Mader CC, Tsinoemas NF (2007) An integrated in silico gene mapping strategy in inbred mice. *Genetics* 175:321–333
- Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J, Beavis WD, Belknap JK, Bennett B, Berrettini W, Bleich A, Bogue M, Broman KW, Buck KJ, Buckler E, Burmeister M, Chesler EJ, Cheverud JM, Clapcote S, Cook MN, Cox RD, Crabbe JC, Crusio WE, Darvasi A, Deschepper CF, Doerge RW, Farber CR, Forejt J, Gaile D, Garlow SJ, Geiger H, Gershenfeld H, Gordon T, Gu J, Gu W, de Haan G, Hayes NL, Heller C, Himmelbauer H, Hitzemann R, Hunter K, Hsu HC, Iraqi FA, Ivandic B, Jacob HJ, Jansen RC, Jepsen KJ, Johnson DK, Johnson TE, Kempermann G, Kendziorski C, Kotb M, Kooy RF, Llamas B, Lammert F, Lassalle JM, Lowenstein PR, Lu L, Lusis A, Manly KF, Marcucio R, Matthews D, Medrano JF, Miller DR, Mittleman G, Mock BA, Mogil JS, Montagutelli X, Morahan G, Morris DG, Mott R, Nadeau JH, Nagase H, Nowakowski RS, O'Hara BF, Osadchuk AV, Page GP, Paigen B, Paigen K, Palmer AA, Pan HJ, Peltonen-Palotie L, Peirce J, Pomp D, Pravenec M, Prows DR, Qi Z, Reeves RH, Roder J, Rosen GD, Schadt EE, Schalkwyk LC, Seltzer Z, Shimomura K, Shou S, Sillanpaa MJ, Siracusa LD, Snoeck HW, Spearow JL, Svenson

- K, Tarantino LM, Threadgill D, Toth LA, Valdar W, de Villena FP, Warden C, Whatley S, Williams RW, Wiltshire T, Yi N, Zhang D, Zhang M, Zou F (2004) The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet* 36:1133–1137
- Collaborative Cross Consortium (2012) The genome architecture of the collaborative cross mouse genetic reference population. *Genetics* 190:389–401
- Cummings SR, Bates D, Black DM (2002) Clinical use of bone densitometry: scientific review. *JAMA* 288:1889–1897
- Farber CR, Rosen C (2010) Genetics of osteoporosis. In: Robertson RP (ed) *Translational endocrinology & metabolism*. The Endocrine Society, Chevy Chase, pp 87–116
- Farber CR, Bennett BJ, Orozco L, Zou W, Lira A, Kostem E, Kang HM, Furlotte N, Berberyan A, Ghazalpour A, Suwanwela J, Drake TA, Eskin E, Wang QT, Teitelbaum SL, Lusis AJ (2011) Mouse genome-wide association and systems genetics identify *Asx12* as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genet* 7:e1002038
- Flint J, Mackay TF (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res* 19:723–733
- Flint J, Mott R (2008) Applying mouse complex-trait resources to behavioural genetics. *Nature* 456:724–727
- Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J, Morensoni MM, Nilsen GB, Pethiyagoda CL, Stuve LL, Johnson FM, Daly MJ, Wade CM, Cox DR (2007) A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 448:1050–1053
- Furlotte NA, Kang EY, Van Nas A, Farber CR, Lusis AJ, Eskin E (2012) Increasing association mapping power and resolution in mouse genetic studies through the use of meta-analysis for structured populations. *Genetics* 191(3):959–967
- Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Weitz K, Camp DG 2nd, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, Gargalovic P, Eskin E, Kirchgessner T, Smith DJ, Smith RD, Lusis AJ (2011) Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet* 7:e1001393
- Gregorova S, Divina P, Storchova R, Trachtulec Z, Fotopulosova V, Svenson KL, Donahue LR, Paigen B, Forejt J (2008) Mouse consomic strains: exploiting genetic divergence between *Mus m. musculus* and *Mus m. domesticus* subspecies. *Genome Res* 18:509–515
- Grupe A, Germer S, Usuka J, Aud D, Belknap JK, Klein RF, Ahluwalia MK, Higuchi R, Peltz G (2001) In silico mapping of complex disease-related traits in mice. *Science* 292:1915–1918
- Guo Y, Lu P, Farrell E, Zhang X, Weller P, Monshouwer M, Wang J, Liao G, Zhang Z, Hu S, Allard J, Shafer S, Usuka J, Peltz G (2007) In silico and in vitro pharmacogenetic analysis in mice. *Proc Natl Acad Sci USA* 104:17735–17740
- Hoover-Plow J, Shchurin A, Hart E, Sha J, Hill AE, Singer JB, Nadeau JH (2006) Genetic background determines response to hemostasis and thrombosis. *BMC Blood Disord* 6:6
- Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. *Genetics* 178:1709–1723
- Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M, Furlotte NA, Eskin E, Nellaker C, Whitley H, Cleak J, Janowitz D, Hernandez-Pliego P, Edwards A, Belgard TG, Oliver PL, McIntyre RE, Bhomra A, Nicod J, Gan X, Yuan W, van der Weyden L, Steward CA, Bala S, Stalker J, Mott R, Durbin R, Jackson IJ, Czechanski A, Guerra-Assuncao JA, Donahue LR, Reinholdt LG, Payseur BA, Ponting CP, Birney E, Flint J, Adams DJ (2011) Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* 477:289–294
- Kirby A, Kang HM, Wade CM, Cotsapas C, Kostem E, Han B, Furlotte N, Kang EY, Rivas M, Bogue MA, Frazer KA, Johnson FM, Beilharz EJ, Cox DR, Eskin E, Daly MJ (2010) Fine mapping in 94 inbred mouse strains using a high-density haplotype resource. *Genetics* 185:1081–1095
- Liao G, Wang J, Guo J, Allard J, Cheng J, Ng A, Shafer S, Puech A, McPherson JD, Foerzler D, Peltz G, Usuka J (2004) In silico genetics: identification of a functional element regulating H2-Ealpha gene expression. *Science* 306:690–695
- Liu P, Vikis H, Lu Y, Wang D, You M (2007) Large-scale in silico mapping of complex quantitative traits in inbred mice. *PLoS One* 2:e651
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753
- Park CC, Gale GD, de Jong S, Ghazalpour A, Bennett BJ, Farber CR, Langfelder P, Lin A, Khan AH, Eskin E, Horvath S, Lusis AJ, Ophoff RA, Smith DJ (2011) Gene networks associated with conditional fear in mice identified using a systems genetics approach. *BMC Syst Biol* 5:43
- Peters LL, Robledo RF, Bult CJ, Churchill GA, Paigen BJ, Svenson KL (2007) The mouse as a model for human biology: a resource guide for complex trait analysis. *Nat Rev Genet* 8:58–69
- Pletcher MT, McClurg P, Batalov S, Su AI, Barnes SW, Lagler E, Korstanje R, Wang X, Nusskern D, Bogue MA, Mural RJ, Paigen B, Wiltshire T (2004) Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. *PLoS Biol* 2:e393
- Prows DR, Hafertepen AP, Winterberg AV, Gibbons WJ Jr, Wesselkamper SC, Singer JB, Hill AE, Nadeau JH, Leikauf GD (2008) Reciprocal congenic lines of mice capture the *aliq1* effect on acute lung injury survival time. *Am J Respir Cell Mol Biol* 38:68–77
- Singer JB, Hill AE, Burrage LC, Olszens KR, Song J, Justice M, O'Brien WE, Conti DV, Witte JS, Lander ES, Nadeau JH (2004) Genetic dissection of complex traits with chromosome substitution strains of mice. *Science* 304:445–448
- Smolock EM, Ilyushkina IA, Ghazalpour A, Gerloff J, Murashev AN, Lusis AJ, Korshunov VA (2012) A genetic locus on mouse chromosome 7 controls elevated heart rate. *Physiol Genomics* 44(13):689–698
- Styrkarsdottir U, Halldorsson BV, Gretarsdottir S, Gudbjartsson DF, Walters GB, Ingvarsson T, Jonsdottir T, Saemundsdottir J, Center JR, Nguyen TV, Bagger Y, Gulcher JR, Eisman JA, Christiansen C, Sigurdsson G, Kong A, Thorsteinsdottir U, Stefansson K (2008) Multiple genetic loci for bone mineral density and fractures. *N Engl J Med* 358:2355–2365
- Takada T, Mita A, Maeno A, Sakai T, Shitara H, Kikkawa Y, Moriwaki K, Yonekawa H, Shiroishi T (2008) Mouse inter-subspecific consomic strains for genetic dissection of quantitative complex traits. *Genome Res* 18:500–508
- Teitelbaum SL, Ross FP (2003) Genetic regulation of osteoclast development and function. *Nat Rev Genet* 4:638–649
- Threadgill DW, Churchill GA (2012) Ten years of the collaborative cross. *G3 (Bethesda)* 2:153–156
- Valdar W, Solberg LC, Gauguier D, Burnett S, Klenerman P, Cookson WO, Taylor MS, Rawlins JN, Mott R, Flint J (2006) Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet* 38:879–887

- Wolfe CJ, Kohane IS, Butte AJ (2005) Systematic survey reveals general applicability of “guilt-by-association” within gene coexpression networks. *BMC Bioinformatics* 6:227
- Yalcin B, Nicod J, Bhomra A, Davidson S, Cleak J, Farinelli L, Osteras M, Whitley A, Yuan W, Gan X, Goodson M, Klenerman P, Satpathy A, Mathis D, Benoist C, Adams DJ, Mott R, Flint J (2010) Commercially available outbred mice for genome-wide association studies. *PLoS Genet* 6(9):e1001085
- Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 109:1193–1198