Lack of evidence supporting transgenerational effects of non-transmitted paternal alleles on the murine transcriptome

Rodrigo Gularte-Mérida[1,2*], *Carole Charlier*[1], *Michel Georges*[1]

[1]Unit of Animal Genomics, GIGA Institute and Faculty of Veterinary Medicine, University of Liège, 1 Avenue de l'Hôpital, 4000 Liège, Belgium

[2]Current address: Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY

*Correspondence: gularter@mskcc.org

Abstract (217 words)

Transgenerational genetic effects are defined as the effects of untransmitted parental alleles on the phenotype of their offspring. Well-known transgenerational genetic effects, in humans and other mammals, are the effects of a parental genotype on the nurturing ability of the parents, coined "genetic nurture". However, there exist examples of transgenerational genetic effects in model organisms that are independent of nurturing effects and support the epigenetic transmission of a memory of the parental genotype possibly mediated by small RNA species. To test whether such transgenerational epigenetic effects might exist in mammals, we generated 833 isogenic C57BL/6J (B6) mice that differed only by the presence in the genome of their sire of one copy of four A/J chromosomes (MMU 15, 17, 19 or X). We measured 25 anatomical traits and performed RNA-Seq on five distinct tissues (heart, liver, pituitary, whole embryo, and placenta). There was no evidence of a significant effect from untransmitted A/J sire chromosome alleles, whether on anatomical traits or gene expression level. We observed an effect on *Mid1* expression levels in multiple tissues, but this was shown to be due to a de novo mutation that occurred in one of the sire lines. We conclude that transgenerational epigenetic memory of non-transmitted paternal alleles - if it exists - is uncommon in mice and likely other mammals.

Introduction (526 words)

Traits are typically considered to be genetic if individual phenotypes correlate with individual genotypes. In several instances, individual phenotypes also correlate with parental genotypes including with the non-transmitted alleles. These transgenerational genetic effects (TGE) include instances of "genetic nurture", i.e., situations where the parental genotype will affect the environment in which the offspring is raised. TGE aim to identify parental genotypes and differ from transgenerational epigenetics where molecular changes are the result of

38 environmental perturbations in the absence of genetic variation. This is easily understood by
39 so-called "maternal effects", which are determined by the care that is provided pre- and
40 postnatally to offspring by mothers. Similar nurture effects of non-transmitted paternal alleles
41 have been reported in humans, where the genotype of the parents may also affect the offspring's
42 environment (Mott et al. 2014, Kong et al. 2018; Cheuquemán, et al. 2021). In addition, the
43 effects of non-transmitted parental alleles on the offspring's phenotype exist that cannot be
44 explained by genetic nurture. One such example is paramutation, which involves the
45 epigenetic conversion of transmitted by non-transmitted alleles. Paramutation-like phenomena
46 have been reported in plants, *D. melanogaster*, and possibly the mouse (Nadeau 2009; Heard
47 and Martienssen 2014). A second example is the effect of gene copy number of *Sly*, *Ssty1*,
48 *Ssty2*, *Srsy*, *Rbmy,* and *Rbm31y* of MMU Y on the predisposition and severity of experimental
49 autoimmune encephalomyelitis in the next generation (Teuscher et al. 2006; Case et al. 2015).
50 In C. elegans, the longevity of $Wd5^{+/+}$ wild-type descendants of $Wd5^{-/-}$ knockout worms was
51 increased for three generations even in the absence of transmission of the $Wd5^{-}$ knockout allele,
52 when compared to worms of a fully $Wd5^{+/+}$wild-type lineage (Greer et al. 2011). Most
53 recently, promoter associated CpG DNA methylation via genetic engineering has been shown
54 to be transmitted across multiple generations and retain parental phenotypic differences
55 (Takahashi, et al. 2023). What remains unknown, however, is how common such nurture
56 independent TGE are in mammals under naturally occurring variation.

57 We took advantage of mouse consomic strains and global gene expression by RNA-Seq to
58 address this issue. Consomic mice, also referred to as chromosome substitution strains, harbour
59 one chromosome pair from a donor strain "A" in the background of a recipient strain "B". A
60 genome-wide panel of consomic mice is commercially available for which "A" alleles from
61 *A/J* mice were introgressed into the "B" background of *B6* mice (Nadeau et al. 2000). By
62 backcrossing consomic males (say with a donor chromosome 15 of "A" ancestry) for two
63 subsequent generations to the recipient "B" background strain, and selecting generation N2
64 offspring that have inherited a non-recombinant "B" chromosome 15, one obtains mice that are
65 genetically identical to the inbred "B" background strain (isogenic), except for the genotype of
66 their sire (heterozygous A/B for chromosome 15) (**Figure 1**). This design permits aggregation
67 of any potential TGE across a whole chromosome, maximizing our chances to detect such
68 effects if they exist. To study the effect of the non-transmitted paternal chromosome on many

69 phenotypes at once, we decided to study growth and body composition traits in adult mice, and
70 the transcriptome of five tissues at two developmental stages of the resulting N2 offspring by
71 RNA-Seq.

72 **Results (1194 words)**

73 We performed this experiment for four of the 21 murine chromosomes taking advantage of the
74 chromosome B6.A/J chromosome substitution strains (CSS) (MMU 15, 17, 19 and X.
75 Hereinafter mouse strains will be denoted by their standard congenic nomenclature B6.A-
76 <chr>, where B6 is the recipient strain, A is the A/J donor strain, chr refers to the specific
77 mouse chromosome. Autosomes MMU 15, 17, and 19 were selected on the basis that MMU
78 15 was found the most transcriptionally diverse chromosome relative to B6, MMU 17 is known
79 to affect body weight and body composition in other strains (Shockley and Churchill 2006,
80 Corva et al. 2001; Farber et al. 2006), MMU 19 because it has a short chromosome length, and
81 finally, MMU X was chosen for its convenience as one generation would automatically
82 generate isogenic B6/B6 males. We produced 2,456 N2 mice, resulting in 958 consomic-
83 derived *C57BL/6J* isogenic males throughout the experiment. The initial isogenic cohort was
84 composed of 123 B6.A-15 N2 $^{B6/B6}$, 132 B6.A-17 N2 $^{B6/B6}$, 136 B6.A-19 N2 $^{B6/B6}$, 220 B6.A-X
85 F1 $^{B6/B6}$, and 206 purebred *C57BL/6J* male controls. The corresponding proportions of
86 recombinant and non-recombinant mice matched expectations based on chromosome size (in
87 centimorgan), and were concordant with the observed results of Nadeau et al. (2000)
88 (**Supplemental Material, Supplemental Table S1**). 80% of N2 male mice were euthanized
89 at 60 days post-natal (dpn), dissected (recording 25 phenotypes pertaining to body weights,
90 and body composition), and major organs flash-frozen – namely heart, liver, and pituitary. The
91 other 20% of animals were collected in utero at 13.5 days post-coitum (dpc), embryo and
92 placenta severed, and flash frozen. For each of the 10-line x timepoint combinations – (B6.A-
93 15, B6.A-17, B6.A-19, B6.A-X, and B6 controls) x (60 dpn and 13 dpc) –, we extracted RNA
94 from eight randomly selected mice within 1 standard deviation (SD) of the phenotypic mean
95 for their corresponding genotypic group (60 dpn: heart, liver, and pituitary; 13 dpc: embryo,
96 placenta). We then generated two equimolar RNA pools of four individuals for the 25-line x
97 time point x tissue combinations, generating 50 libraries for next-generation RNA sequencing
98 using TrueSeq® Stranded Total RNA with RiboZero kits (Illumina), and produced an average

99 of 34.8 ± 8 million 100bp paired-end reads per library on an Illumina HiSeq 2000. One may
100 expect the B6 background genome of the five lines to have diverged by the accumulation of
101 *de-novo* mutations (DNMs) since their 19-year divergence. Such DNMs may cause differences
102 in gene expression that may erroneously be attributed to TGE. To recognize such direct genetic
103 effects, we resequenced the genomes of the founder male of each strain to 10-fold depth on an
104 Illumina HiSeq2000 instrument. We identified a total of 2,794 non-A/J single nucleotide
105 polymorphisms, 1,798 small insertion-deletions, and 63 copy number variants within the B6
106 genome of CSS founder males. The B6 genome of any two founders differed on average for
107 1,464 variants (**Supplemental Material**; **Supplementary Figure S1**, **Supplementary Table**
108 **S2**).

109 RNA-Seq reads were mapped to the murine GRCm38 reference genome using the STAR
110 aligner (Dobin et al. 2012; Engström et al. 2013). Read quantification and differential
111 expression analysis between consomic-derived and *C57BL/6J* control samples (matched for
112 tissue type) were conducted using HTSeq and DESeq2, respectively (Love et al. 2014; Anders
113 et al. 2015). In a first analysis, we obtained 53 instances of significant differential expression
114 (FDR 0.05) between consomic-derived and control samples involving 47 genes (**Figure 2A;**
115 **Table 1**). Thirty-five (66 %) corresponded to increased expression in the consomic-derived
116 samples (average fold change: 1.69; range: 0.228 – 4.33), and 18 (34%) to decreased expression
117 (average fold change: -1.42; range: -0.274 – -3.68) (**Supplementary Figure S2**). The 53
118 instances were distributed across lines and sample types as shown in **Figure 2B**. Gene
119 expressions were requantified using reverse transcriptase quantitative PCR assays (RT-qPCR)
120 for the corresponding 47 genes in triplicate for all 200 samples that constituted the 50 RNA
121 pools using the BioMarkHD qPCR system, Fluidigm at IntegraGen, Évry, France. Differential
122 expression was confirmed (FDR ≤ 0.05) in at least one tissue for 10 of the 47 analysed genes.
123 *Mid1, Crem* and *Gm26448* had the strongest evidence of a transgenerational genetic effect
124 (technical replicates in **Figure 2A**; **Supplementary Table S3**).

125 To further confirm the results obtained from the 10 remaining genes, we generated an
126 independent replicate cohort with >500 male mice using the same mating design as above. The
127 independent replicate cohort yielded 244 isogenic consomic-derived N2 males (MMU 17: 38;
128 MMU 19: 19; MMU X: 98) and 89 *C57BL/6J* controls that were euthanized and dissected at

129 60 dpn. We extracted RNA from relevant tissue samples for an additional 14 to 26 animals of
130 the discovery cohort, and 14 to 28 animals of the independent replicate cohort and evaluated
131 the expression of the 10 candidate genes by RT-qPCR on the BioMarkHD using previously
132 designed assays. Initial findings were confirmed in the independent samples from the
133 discovery and replication cohort for one gene only: *Mid1* (**Figure 2A**, **Supplementary Table**
134 **S3**). *Mid1* was highly significantly overexpressed ($p < 0.0001$; Median Log_2 Fold change: 1.26
135 ± 0.4; range: 0.81 – 2.69) in all tissues of the MMU 19 consomic-derived animals
136 (**Supplementary Figure S3**).

137 The *Mid1* gene (for *Midline 1*) encodes a member of the tripartite motif (TRIM) family that
138 forms homodimers that associate with microtubules. Mutations in *Mid1* cause Opitz syndrome
139 characterized by midline defects including cleft palate, laryngeal cleft, heart defects,
140 hypospadias, and agenesis of the corpus callosum. In *Mus musculus domesticus*, *Mid1* maps
141 to the X chromosome, yet its 5' end is truncated by the pseudoautosomal boundary (PAB) on
142 the Y chromosome (**Figure 3**). *Mid1* thus spans the PAB of *Mus musculus domesticus*. We
143 examined the PAB in the five studied founder males and observed that the B6.A-19 founder
144 had a duplication at the PAR that spanned Exon 4-10 of *Mid1* (**Figure 3**). The end of the
145 duplication could not be precisely determined due to the limitations of the reference assembly
146 (Morgan et al. 2019). Furthermore, our results support the hypothesis that the duplication is
147 located on MMU Y, as all F1, N2, and N3 male offspring from this male inherited the X
148 chromosome of B6 females for three generations, and the Y chromosome from the male. Thus,
149 the differential expression of *Mid1* in the B6.A-19 consomic-derived isogenic mice when
150 compared to the four other lines including controls is most likely due to the transmission of a
151 duplicated *Mid1* transcript throughout the male lineage of the cross rather than to a TGE
152 (**Supplementary Figure S4**). *Crem,* and *Gm26448* failed to replicate their initial effect
153 through the experimental validations.

154 We measured 25 anatomical phenotypes on the 1,370 60 dpn N2 animals from both the
155 discovery and replication cohorts. We tested for possible TGE due to the non-transmitted
156 paternal A/J chromosomes by comparing trait means between consomic-derived and control
157 lines in two successive generations N2 and N3. There was no evidence of significant
158 differences when accounting for multiple testing ($p_{Bonf} \leq 0.05$; **Supplementary Table S4**;

159 **Supplementary Figure S5**). There was – at best – suggestive evidence (p = 0.0025) for a
160 6.7% reduction in heart weight in the B6.A-17 N2 consomic-derived line. This effect did not
161 persist to the next generation through successive backcrossing, B6.A-17 N3 (**Supplementary**
162 **Figure S6**).

163

164 **Discussion (170 words)**

165 Our work tested the hypotheses for genetically driven transgenerational effect via the paternal
166 lineage comparing 25 anatomical traits of 1,164 isogenic male mice, and transcriptome-wide
167 expression from five tissues – heart, liver, pituitary, embryo, and placenta. Our initial results
168 suggested 3 genes —*Mid1, Crem,* and *Gm26448*— as differentially expressed between an
169 isogenic-derived cohort relative to B6 purebred mice. Further validation of these genes, showed
170 *Mid1*expression was differentially expressed across all B6.A-19 N2 $^{B6/B6}$, and B6.A-19 N3$^{B6/B6}$
171 cohorts as a result of a segmental duplication at the XY-pseudoautosomal region
172 (**Supplementary Figure S4**). All other 9 tested genes (including *Crem,* and *Gm26448*) failed
173 to achieve significance in subsequent rounds of validation. We identified suggestive evidence
174 for a 6.7% reduction in heart weight in the B6.A-17 N2$^{B6/B6}$ consomic-derived line relative to
175 B6.C purebred (**Supplementary Figure S5**, **Supplementary Figure S6**). Based on these
176 results, we conclude that nurture independent TGE, if they exist at all, are extremely rare in
177 mammals under naturally occurring genetic variation and are unlikely to contribute
178 significantly to the phenotypic variance.

179 **Methods (2018 words)**

180 ***Animal husbandry***. The initial breeding population purchased from the Jackson Laboratories
181 (Bar Harbor, Maine) consisted of one male breeder and one female breeder approximately 5
182 weeks of age for each of the four chromosome substitution strains listed above, and one male
183 breeder with three female breeders for C57BL/6J. These mice were placed under quarantine
184 and then brought into the Specific Pathogen Free (SPF) facility of GIGA Research. After a
185 seven-day acclimation period, males, and females for each of the four strains were coupled to
186 maintain and expand all the strains in-house. Breeding cages were set up with one male and up
187 to three females in clear polysulfone filter top cages (Model: 1290D EURO-STANDARD

188 TYPE III, Tecniplast S.p.A, Varese- Italy). Breeding cages were checked daily to identify new-
189 born pups and their corresponding dam. Whenever possible gestating females were transferred
190 to an independent polysulfone cage for rearing pups, as SPF practices were to maintain pre-
191 weaning pups and rearing females together with their sire. The presence or absence of the sire
192 during pre-weaning did not affect any tested phenotype. At 30 ± 1 day of age males and females
193 were weaned, identified with a numbered ear tag, divided into groups of up to 5, and placed in
194 clear polysulfone filter top cages (Model: 1284L EURO-STANDARD TYPE II L, Tecniplast
195 S.p.A). During the weaning process, mice were also weighed and a 1 - 2 mm tail biopsy was
196 collected for genotyping. Breeding cages were fed a Standard SDS RM3 E diet (Cat. No.
197 801066, Special Diets Services- Essex, UK), whereas non-breeding and non-lactating cages
198 were fed a maintenance diet, Standard SDS RM1 E diet (Cat. No. 801002, Special Diets
199 Services). Enrichment was provided to all cages in the form of a half-white cardboard frites
200 box which also provides nesting material. Experimental male mice from all cohorts N2, N3
201 and B6.C were sequentially weaned and placed in cages standardised with 5 males irrespective
202 of strain. All mice were housed in accordance with the European Union Directive 2010/63/EU
203 and the FELASA Recommendations and Guidelines (Guillen 2012), and experimental
204 protocols approved by the Institutional Animal Care and Use Committee of the Université de
205 Liège.
206 **Characterization of 25 growth and body composition traits in adult and embryo male**
207 **mice.** The complete population of 966 N2 male mice from Cohort I, 125 N3 male mice from
208 Cohort III, and 501 N2 male mice from Cohort IV were weighed at 30, 40, 50, and 60 days (d)
209 of age. At 60 ± 1.5 d, mice were harvested as described in Gularte-Mérida (2015). Mice were
210 anaesthetised with Isoflurane, during which time nasal-anal and nasal-tail lengths were
211 collected. Mice were euthanised by decapitation under anaesthesia. Trunk blood was collected
212 in BD Microtainer tubes containing Dipotassium EDTA (Cat. No. 365973, BD- Belgium),
213 subsequently spun at 8000 rpm for 4 minutes, plasma was transferred to a 2 mL screw cap tube
214 and stored at -20C. The weight of femoral, gonadal, mesenteric, and retroperitoneal fat pads,
215 spleen, kidneys, liver, heart, gastrocnemius muscle, whole brain, and carcass weights were
216 recorded up to the nearest 0.001 g. Lastly, femur length was measured on the left femur from
217 the ball of the femur to the base. A collection of 11 tissues were flash frozen from each mouse
218 and stored at -80C, these were spleen, liver, kidneys, heart, muscle, gonadal fat, brain,

219 hypothalamus (separated after weighing), and lastly adrenal glands and pituitary gland were
220 frozen but not weighed. A tail biopsy was also collected at the end of the harvest and used for
221 DNA extraction.
222 Mouse N2 foetuses and placentas from cohort II were collected at 14 dpc and weighted and
223 flash frozen in liquid nitrogen. This was achieved by isolating each female when the coital plug
224 was identified during a morning inspection, day 0.5. Thirteen days later, females were
225 euthanised by cervical dislocation, their uterus extracted, rinsed with DPBS (Cat. No BE17-
226 512F, Lonza-Verviers), and subsequently kept on a petri-dish containing DPBS to avoid
227 dehydration while foetuses were harvested under a stereoscope. Each amniotic sac containing
228 one foetus and its respective placenta was isolated from the uterus, rinsed under a stream of
229 DPBS. The amniotic sac was transferred to a new petri dish, opened carefully to maintain the
230 yolk sac intact, and gently tearing it off from the placenta. Foetuses were excised by severing
231 the umbilical cord under the stereoscope using fine tweezers and taken out from the yolk sac,
232 rinsed in PBS, placed on a ruler, photographed, and then weighed to the nearest 0.001 g.
233 Subsequently, the right leg was gently severed from the body with fine point tweezers to be
234 used for DNA extraction. Placentas were inspected for any remaining maternal tissue, further
235 cleaning any maternal tissue if needed, and weighed to the nearest 0.001 g. Both embryo and
236 placenta were subsequently flash-frozen for later use.
237
238 ***Statistical analyses.*** Organismal phenotypes were classified into three categories, body weight
239 traits (comprising weights (g) at 14, 30, 40, 50, and 60 dpc. d), length traits (comprising Nasal-
240 Anal, Nasal-Tail, and femoral lengths (mm)), and body composition traits (comprising liver,
241 spleen, heart, kidney, brain, gastrocnemius muscle, carcass, and the four fat pad weights). To
242 test for paternal transgenerational genetic effects the specific contrasts between each "test"
243 group of B6.A-N $N2^{B6/B6}$ isogenic derived mice and B6.C were analysed using general linear
244 hypothesis testing with the R/*multcomp* package via the *glht* function set with a Tukey test
245 (Bretz et al. 2010). Confounding factors such as litter size, presence of sire in the pre-weaning
246 cage, and dam's age were tested as covariates during the model selection across all phenotypes.
247 However, none of these showed any significant effects to change the conclusions of the
248 analyses and thus were excluded from the final comparisons. A summary of power calculations
249 is presented in **Supplementary Material**.

250

251 ***Characterization of global gene expression by RNA-Seq of adult and embryonic tissues.*** To
252 characterize and analyse global gene expression we extracted RNA from 8 individuals in three
253 adult tissues — liver, heart, pituitary —, whole embryo and placenta in each of the five isogenic
254 derived sub-cohorts (B6.A-15 N2 $^{B6/B6}$, B6.A-17 N2 $^{B6/B6}$, B6.A-19 N2 $^{B6/B6}$, B6.A-X F1 $^{B6/B6}$,
255 and B6.C) for a total of 200 samples. Libraries were constructed using Illumina TrueSeq®
256 Stranded Total RNA with RiboZero kit (No. Cat. RS-122-2201 or -2202). We constructed two
257 pooled libraries for each tissue × strain combination from equimolar quantities of RNA of four
258 individuals. RNA-Seq libraries were sequenced on an Illumina HiSeq 2000 where each lane
259 contained five libraries, one per strain.
260 We obtained on average 32.5 M reads per library, a total of 162.5 M reads per lane. Sequence
261 reads were aligned to the mouse reference assembly GRCm38 with alternative loci using STAR
262 aligner (Dobin et al. 2012) using the annotated 2-pass method (annotation with Ensembl Genes
263 75) described by Engström et al. (2013), where « splice junctions » are obtained from the first
264 pass, and the second pass refines the alignment of all sequence reads. Following the alignment,
265 we analysed sequences to 1) detect differential expression in the five cohorts of isogenic mice;
266 and 2) identify if genetic variation exists among the strains that may explain the differences in
267 significantly differentially expressed genes (details in **Supplementary Material**). For the first
268 analysis, we used HTSeq to quantify the expression of each gene in the Ensembl 86 annotation
269 and we used DESeq2 to perform the differential expression analyses of all genes (Love et al.
270 2014). Comparisons were made as described in the previous section where each isogenic
271 derived sub-cohort (B6.A-15 N2 $^{B6/B6}$, B6.A-17 N2 $^{B6/B6}$, B6.A-19 N2 $^{B6/B6}$, B6.A-X F1 $^{B6/B6}$) is
272 compared to B6.C. At a Log2FoldChange of ≥ 2, power of detection using the 2 pooled RNA-
273 Seq samples ranged from ≥ 0.90 in the adult tissues(pituitary, liver, and heart), to 0.6 for the
274 embryo and placenta (**Supplementary Material**).
275

276 ***Technical and biological replications of RNA-Seq differentially expressed genes by RT-***
277 ***qPCR.*** To confirm our initial findings, we carried out a technical replication on the same 200
278 samples used to construct the RNA-Seq libraries. In these samples, we re-measured the
279 expression level of the 48 genes found differentially expressed by RNA-Seq using reverse
280 transcriptase quantitative PCR (RT-qPCR) in triplicate determinations using Fluidigm's

281 BioMarkHD qPCR system. This platform allows us to simultaneously test up to 32 assays in
282 triplicate in 96 samples (a 96 x 96 factorial design on a single reaction plate), for a total of
283 9,216 RT-qPCR reactions per plate, and each gene was measured in two or more tissues. Primer
284 assays for the RT-qPCR of the technical replicate and first validation cohorts were designed
285 by IntegraGen (Évry, France) (**Supplementary Table S5**). For the top-ranking differentially
286 expressed genes –*Mid1, Coq10b, Per3, Ide, and Crem*– we designed multiple assays targeting
287 different exon-exon junctions maximising the number of transcript variants assayed. For the
288 remaining 43 genes, one assay per gene was designed (**Supplementary Table S3**). In addition,
289 we amplified eight genes used routinely as endogenous controls, *B2m, Eif3j1, Gusb, Hprt,*
290 *Rpl37rt, Rp10, Tbp,* and *Tubb5* each with one primer assay. Complementary DNA (cDNA)
291 was synthesized from RNA using Superscript III following the manufacturer's
292 recommendations (Life Technologies, Carlsbad, CA). A reference sample was created by
293 pooling cDNA from all the samples of each tissue, from which a 1:4 serial dilution was carried
294 out to develop a four-point standard curve. All reactions were carried in two stages, where the
295 first stage cDNA is a pre-amplification using TaqMan PreAmp Master Mix (Cat No. 4384267)
296 following the manufacturer's recommendations, and the second stage is the RT-qPCR reaction
297 on the BioMarkHD system using EvaGreen(R) master mix.
298 Differential expression was analysed using relative gene expression based on the 2^{-Ct} method
299 (Livak and Schmittgen 2001). First, we perform endogenous control gene selection based on
300 the geNorm algorithm (Vandesompele et al. 2002) using qBase+ (Biogazelle, Belgium). These
301 analyses suggested the use of *Tubb5* and *Gusb* as endogenous controls in embryo, *Tbp* and
302 *Hprt* in the placenta, liver, and pituitary, and all four genes, *Tubb5*, *Gusb*, *Tbp* and *Hprt* in the
303 heart. Expression values estimated as 2^{-Ct} were calculated using the R/Bioconductor *ddCt*
304 package (Zhang et al. 2015), where the median Ct for each sample is corrected by the median
305 of the pooled reference sample, and to the median of the endogenous control genes as described
306 in Livak & Schmittgen (2001).
307 The 14 genes replicated between RNA-Seq and the RT-qPCR differential expression –*Clpx,*
308 *Srebf1, Col1a1, Crem, Gm129, Gm26448, Hpsa1a, Hpsa1b, Hspg2, Mid1, Per3, Tef,* and
309 *Thrsp*–. *These* were quantitated in an additional 800 N2 littermates from all the cohorts
310 developed throughout the course of the experiment. Roughly 20 mice for each of the four
311 B6.A-15 N2 $^{B6/B6}$, B6.A-17 N2 $^{B6/B6}$, B6.A-19 N2 $^{B6/B6}$, and B6.C isogenic littermates (Cohort

312 I and II), 20 mice from each of B6.A-17 N3 $^{B6/B6}$, B6.A-19 N3 $^{B6/B6}$, and B6.C N3 controls
313 (Cohort III), and 20 mice from each of B6.A-17 N2 $^{B6/B6}$, B6.A-19 N2 $^{B6/B6}$, and B6.C
314 independent replicate backcross (Cohort IV). The cDNA synthesis reactions, RT-qPCR
315 technical replicate, and first validation results were performed by IntegraGen (Évry, France),
316 using *Tbp, Hprt, Tubb5,* and *Gusb* as endogenous control genes. Expression values for this first
317 validation were estimated as previously described and analysed by comparing the expression
318 values from the different B6.A-<N> N2 $^{B6/B6}$ to B6.C as described in the statistical analysis
319 section.
320 Based on the gene expression results from the previous RT-qPCR expression analyses, a
321 second of validation was performed on *Crem* and *Gm26448* in the pituitary. These two genes
322 and corresponding endogenous controls were measured in 32 B6.A-15 N2 $^{B6/B6}$, and 17 B6.A-
323 17 N2 $^{B6/B6}$ and 23 B6.C new male littermates from Cohort I. Also, included were an additional
324 14 B6.A-17 N2 $^{B6/B6}$ and 27 B6.C new littermates from Cohort IV. Complementary DNA from
325 these samples was synthesised using SuperScript IV following manufacturer's
326 recommendations. Where *Crem*, *Tbp* and *Hprt* had pre-designed assays from IDT
327 PrimeTime(R) probes and these were chosen to measure their expression, whereas *Gm26448*
328 was measured with a custom IDT PrimeTime(R) assay (**Supplementary Table S3**). Real-Time
329 qPCR reactions were carried out according to manufacturer's recommendations. Briefly, each
330 reaction was composed of 1X PrimeTime(R) Universal Master Mix (Cat No. 1055772, 0.2 µM
331 PrimeTime(R), and 10ng of cDNA.
332
333

334 **Data availability**

335 Sequence data for the mouse genomes and isogenic transcriptomes will be available through
336 the European Nucleotide Archive project (PRJEB70518). Genotype and phenotype data will
337 be publicly accessible at http://www.github.com/RodrigoGM/tgv_2023/

338

339 **Acknowledgments**

340 This project was funded by a grant from the Walloon Ministry of Agriculture (DGARNE). We
341 are grateful to all the personnel of the GIGA Mouse Housing and Transgenics Platform, in

342 particular Benoit Remy, Audrey Tromme, and Fabien Ectors. We also give special thanks to
343 the members of the GIGA Genomics Platform Nadine Cambisano, Naïma Ahariz, and Wouter
344 Coppieters for their assistance in sample processing.

345

346 **TABLES**

347 **Table 1.** Number of significant differentially expressed genes in discovery and validation data
348 sets

Discovery Population (N2.I)	**Technical Replicate (N2.I)**	**Biological Replicate (N2.I)**	**Independent / Replicate Backcross (N2.III)**	**Persistence Backcross (N3)**
RNA-Seq 2 pooled libs (N = 4 / lib; 50 libs) 53 DE genes FDR ≤ 0.05	qPCR 8 per group (N = 200) 12 genes p ≤ 0.05 1 gene p < 0.0001	qPCR ~20 per group (N = 332) 2 genes p ≤ 0.05 1 gene p < 0.0001	qPCR ~20 per group (N = 121) 2 genes p ≤ 0.1 1 gene p < 0.0001	qPCR ~20 per group (N = 185) 0 genes p ≤ 0.05 1 gene p < 0.0001

349

350 **Figure Legend**

351 **<u>Figure 1</u>. A)** Single-locus two-allele model illustration of the mating design used to perform
352 a test of hypothesis for paternal transgenerational genetic effects in isogenic mice. Where
353 males heterozygous AA' are mated to AA females. Resulting offspring are genotyped and
354 only those with an AA genotype are compared to a control offspring on the right. **B)**
355 Schematic representation of the cross performed for each B6.A/J–<N> CSS. Cross begins at
356 F0 where a single male homozygous for each of chromosomes 15, 17, 19, and X, was mated
357 to a single B6 female, and a single B6 control male is mated in parallel. The resulting males
358 from the F1 were subsequently backcrossed with multiple B6 females, and N2 offspring
359 genotyped for 236 SNP polymorphic between A/J and B6 to identify isogenic mice that
360 inherit a complete B6 chromosome. Isogenic B6.A-<N> N2 $^{B6/B6}$ are then compared to the
361 control contemporary group of B6/B6 purebred mice. A persistence N3 cross was performed
362 by selecting an isogenic B6.A-<N> N2 $^{B6/B6}$ male mice and mated to multiple B6/B6 females

363 and compared to a contemporary group of B6/B6 males. The only exception was B6.A-X in
364 which B6.A-X F1 $^{B6/B6}$ isogenic mice are obtained within a single generation.

365 **Figure 2.** Differential expression analysis of RNA-Seq across five tissues comparing B6.A-
366 <N> $^{B6/B6}$ and B6 controls. **A)** QQ-plots differential expression -$\log_{10}$ (*p*-values) of all genes
367 compared across three cohorts, the Original N2 backcross (2012), a second independent
368 Replicate N2 backcross (2013), and a persistence backcross N3 (2013). Within the N2
369 backcross, differential expression was measured three times. RNA-Seq of 50 libraries each
370 composed of 4 mouse tissues, qRT-PCR of the same 200 mice for 53 genes with FDR $\leq$ 0.05
371 genes from RNA-Seq, and qRT-PCR of technically replicating genes on isogenic littermates
372 from the same cohort. The latter set of technically replicating genes was measured in the
373 Replicate N2, and Persistence N3 backcrosses. Colours represent the sire lineage. *P*-values
374 were calculated by comparing each B6.A-<N> N2 $^{B6/B6}$ isogenic lineage to B6 Controls.
375 Except for *Mid1,* from left to right our p-values show a loss of replication based on the
376 significance and the directionality of the effects across all differentially expressed genes. **B)**
377 Number of genes significant per B6.A–<N> N2 $^{B6/B6}$ vs B6.C comparison across the five
378 tested tissues for each of the four tested CSS. B6.A-15 N2 $^{B6/B6}$ vs B6.C had the highest
379 amount of differentially expressed genes, particularly in the liver. B6.A-19 N2 $^{B6/B6}$ vs B6.C
380 had the second, totalling 14 genes, of which *Mid1* was consistently differentially expressed
381 across all 5 tissues.

382 **Figure 3. A)** Segmented copy number profiles from the sex chromosomes of the B6.A-19 F0
383 founder male. At the distal end of mouse chromosome (MMU) X – left panel–, results show
384 a duplication on the XY pseudo-autosomal region (PAR) highlighted with a circle. However,
385 no duplication is observed in the MMU Y PAR highlighted in a circle (right panel). **B)** A
386 closer view of the distal ends of MMU X and Y pseudo-autosomal boundaries (PAB), reveals
387 the duplication event spans the *Mid1* gene, explaining the significant differential over-
388 expression of B6.A-19 N2 $^{B6/B6}$ isogenic mice when compared to their B6 contemporary
389 group.

390 **Supplemental Information**

391 **Supplemental Material.** Extended methods for generating the backcross population,
392 genotyping, selection of isogenic mice, power calculations, and sequence analysis of founder
393 males

394 **Supplementary Table S1.** Differential expression analysis of isogenic B6/B6 derived mice
395 vs B6 Controls.

396 **Supplementary Table S2.** Means ± SE for body weights, growth, body length, and body
397 composition traits measured in isogenic B6/B6 N2 mice derived from sires with different
398 genotypes.

399 **Supplementary Table S3.** Gene expression assay design to validate differentially expressed
400 genes from RNA-Seq between B6.A-<N> and B6.C.

401 **Supplementary Table S4.** Summary of genotype and allelic frequencies across all 11
402 backcrosses.

403 **Supplementary Table S5.** DNA sequence polymorphisms and de-novo mutations in B6.A-
404 15, B6.A-17, B6.A-19 and B6.A-X.

405 **Supplemental Figure S1.** The mutational and copy number landscapes of de-novo and A/J
406 alleles across the B6.A-<N> founder males of the tested CSS.

407 **Supplemental Figure S2.** Volcano plots of differential expression across all derived isogenic
408 N2 mice from the four CSS.

409 **Supplemental Figure S3.** Global expression and *Mid1* effects comparisons across replicates.

410 **Supplemental Figure S4.** Inheritance model of the *Mid1* segmental duplication across the
411 B6-A-19 male lineage throughout the discovery, validation, and persistence experimental
412 crosses.

413 **Supplemental Figure S5.** QQ-plots of $-\log_{10}$ *p*-values of all 25 anatomical traits across all
414 three analyzed cohorts.

415 **Supplemental Figure S6.** Heart, and body weights weight across the Discovery (N2),
416 Replicate (N2), and Persistence (N3) backcrosses in B6.A-17 N2 $^{B6/B6}$, B6.A-X F1 $^{B6/B6}$, and
417 B6 Control mice.

418

419 **References**

420 Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-
421 throughput sequencing data. *Bioinformatics* **31**: 166-169.
422 Case LK, Wall EH, Osmanski EE, Dragon JA, Saligrama N, Zachary JF, Lemos B,
423 Blankenhorn EP, Teuscher C. 2015. Copy number variation in Y chromosome
424 multicopy genes is linked to a paternal parent-of-origin effect on CNS autoimmune
425 disease in female offspring. *Genome Biology* **16**: 28-14.
426 Cheuquemán C, Maldonado R. 2021. Non-coding RNAs and chromatin: key epigenetic
427 factors from spermatogenesis to transgenerational inheritance. Review. Biol Res.
428 2021 Dec 20;54(1):41.
429 Corva PM, Horvat S, Medrano JF. 2001. Quantitative trait loci affecting growth in high
430 growth (hg) mice. *Mammalian Genome* **12**: 284-290.
431 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,
432 Gingeras TR. 2012. STAR: ultrafast universal RNA-Seq aligner. *Bioinformatics* **29**:
433 15-21.
434 Engström PG, Steijger T, Sipos B, Grant GR, Kahles A, Rätsch G, Goldman N, Hubbard TJ,
435 Harrow J, Guigo R et al. 2013. Systematic evaluation of spliced alignment programs
436 for RNA-seq data. *nature methods* **10**: 1185-1191.
437 Farber CR, Corva PM, Medrano JF. 2006. Genome-wide isolation of growth and obesity
438 QTL using mouse speed congenic strains. *BMC Genomics* **7**: 102.
439 Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska EA, Mansuy
440 IM. 2014. Implication of sperm RNAs in transgenerational inheritance of the effects
441 of early trauma in mice. *Nature Neuroscience* doi:10.1038/nn.3695.
442 Greer EL, Maures TJ, Ucar D, Hauswirth AG, Mancini E, Lim JP, Benayoun BA, Shi Y,
443 Brunet A. 2011. Transgenerational epigenetic inheritance of longevity in
444 Caenorhabditis elegans. *Nature* **479**: 365-371.
445 Heard E, Martienssen RA. 2014. Transgenerational epigenetic inheritance: myths and
446 mechanisms. *Cell* **157**: 95-109.
447 Kong A, Thorleifsson G, Frigge ML, Vilhjalmsson BJ, Young AI, Thorgeirsson TE,
448 Benonisdottir S, Oddsson A, Halldorsson BV, Masson G et al. 2018. The nature of
449 nurture: Effects of parental genotypes. *Science* **359**: 424-428.
450 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time
451 quantitative PCR and the 2(-Delta C(T)) Method. *Methods* **25**: 402-408.
452 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
453 RNA-Seq data with DESeq2. In *biorxivorg*, doi:10.1101/002832.
454 Morgan AP, Bell TA, Crowley JJ, Pardo-Manuel de Villena F. 2019. Instability of the
455 Pseudoautosomal Boundary in House Mice. *Genetics* **212**: 469-487.
456 Mott R, Yuan W, Kaisaki P, Gan X, Cleak J, Edwards A, Baud A, Flint J. 2014. The
457 Architecture of Parent-of-Origin Effects in Mice. *Cell* **156**: 332-342.
458 Nadeau JH. 2009. Transgenerational genetic effects on phenotypic variation and disease risk.
459 *Human Molecular Genetics* **18**: R202.
460 Nadeau JH, Singer JB, Matin A, Lander ES. 2000. Analysing complex genetic traits with
461 chromosome substitution strains. *Nature Genetics* **24**: 221-225.
462 Shockley KR, Churchill GA. 2006. Gene expression analysis of mouse chromosome
463 substitution strains. *Mammalian Genome* **17**: 598-614.
464 Takahashi Y, Morales Valencia M, Yu Y, Ouchi Y, Takahashi K, Shokhirev MN, Lande K,
465 Williams AE, Fresia C, Kurita M, Hishida T, Shojima K, Hatanaka F, Nuñez-
466 Delicado E, Rodriguez Esteban C, Izpisua Belmonte JC. 2023. Transgenerational

467 inheritance of acquired epigenetic signatures at CpG islands in mice. Cell 186, 715-
468 731.
469 Teuscher C, Noubade R, Spach K, McElvany B, Bunn JY, Fillmore PD, Zachary JF,
470 Blankenhorn EP. 2006. Evidence that the Y chromosome influences autoimmune
471 disease in male and female mice. *Proceedings of the National Academy of Sciences of*
472 *the United States of America* **103**: 8024-8029.
473 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.
474 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric
475 averaging of multiple internal control genes. *Genome Biology* **3**: RESEARCH0034.
476

Figure 1

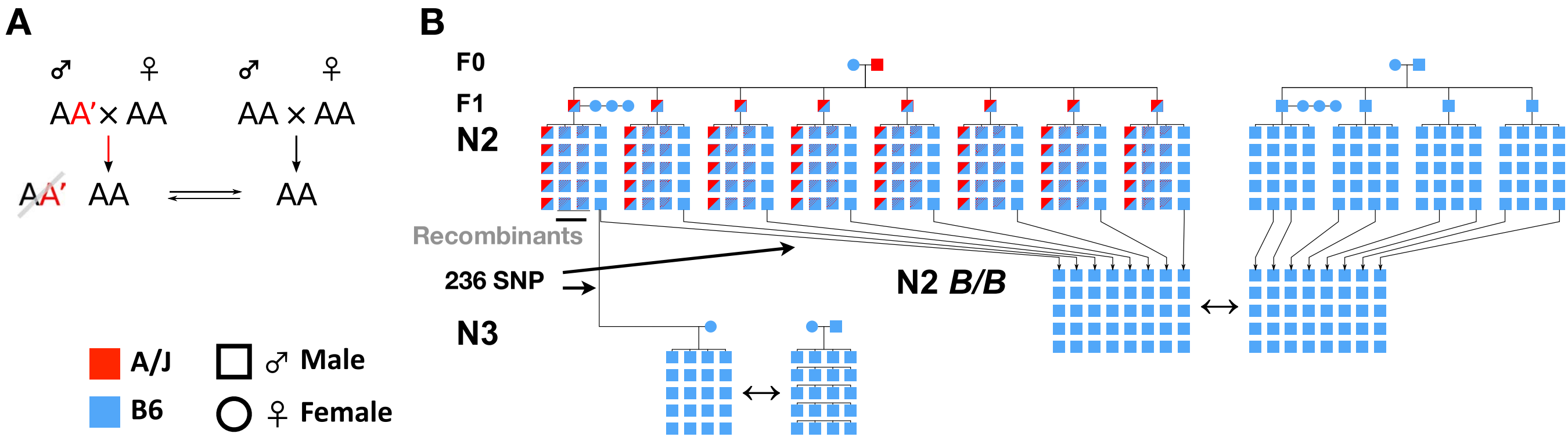

A
♂ ♀
AA′ × AA
AA′ AA
♂ ♀
AA × AA
AA
A/J
B6
♂ Male
♀ Female
B
F0
F1
N2
Recombinants
236 SNP
N2 B/B
N3

Figure 2

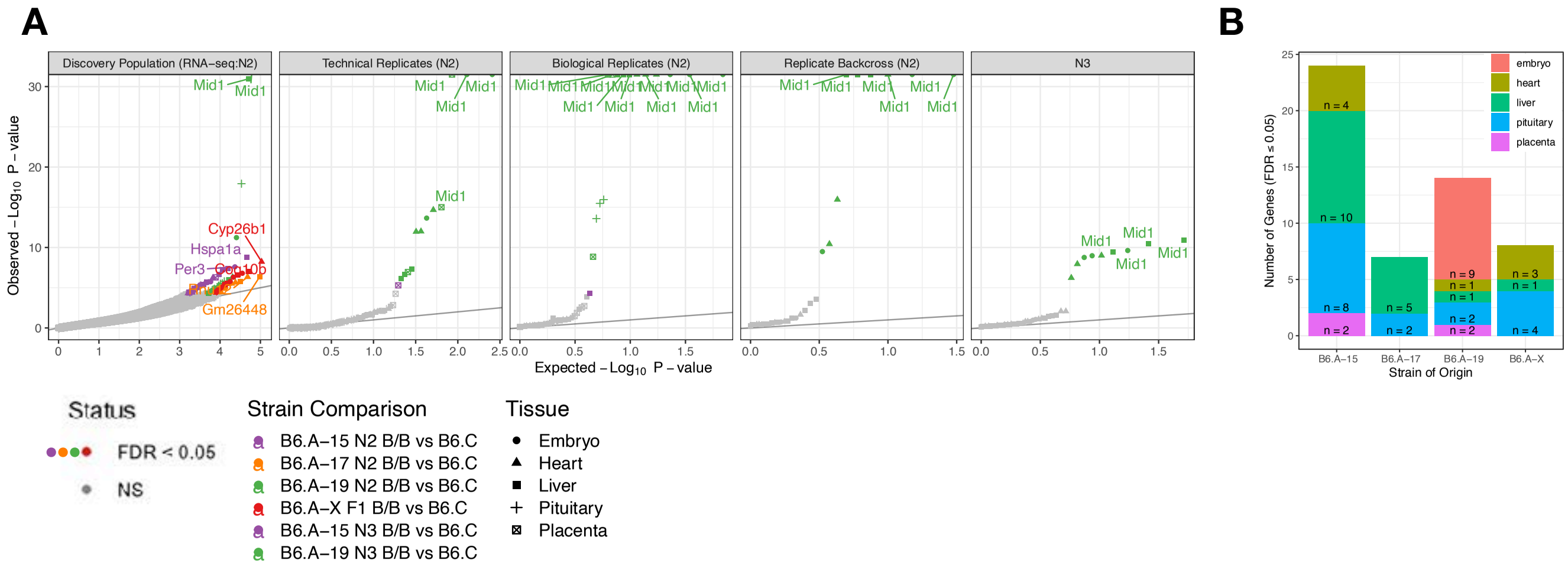
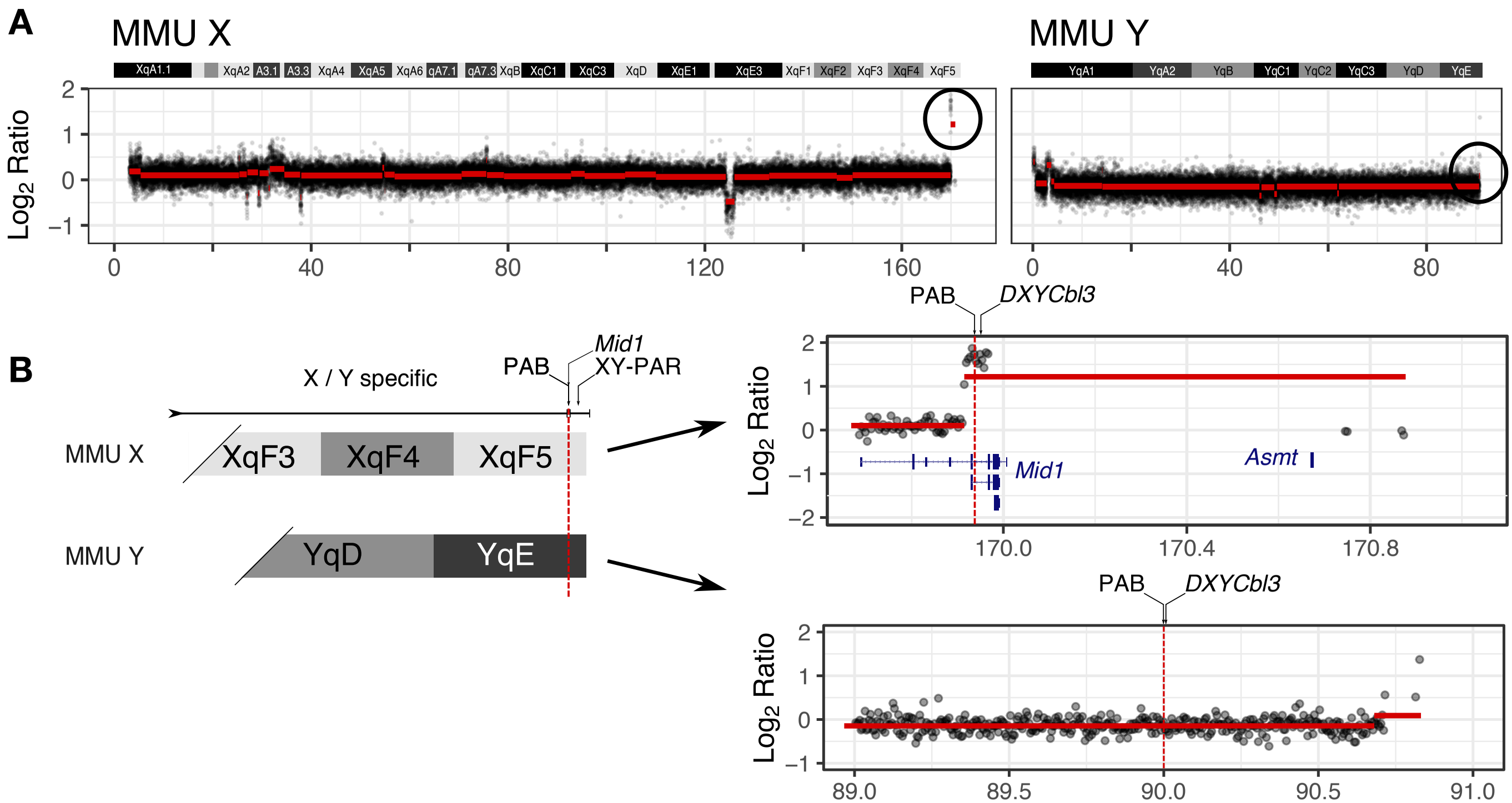

A
B
Discovery Population (RNA–seq:N2)
Technical Replicates (N2)
Biological Replicates (N2)
Replicate Backcross (N2)
N3
Observed – Log10 P – value
Expected – Log10 P – value
Mid1
Cyp26b1
Hspa1a
Per3
Gm26448
Status
FDR < 0.05
NS
Strain Comparison
B6.A–15 N2 B/B vs B6.C
B6.A–17 N2 B/B vs B6.C
B6.A–19 N2 B/B vs B6.C
B6.A–X F1 B/B vs B6.C
B6.A–15 N3 B/B vs B6.C
B6.A–19 N3 B/B vs B6.C
Tissue
Embryo
Heart
Liver
Pituitary
Placenta
Number of Genes (FDR ≤ 0.05)
Strain of Origin
B6.A–15
B6.A–17
B6.A–19
B6.A–X
embryo
heart
liver
pituitary
placenta
n = 4
n = 10
n = 8
n = 2
n = 5
n = 2
n = 9
n = 1
n = 1
n = 2
n = 2
n = 3
n = 1
n = 4

Figure 3


A
MMU X
MMU Y
XqA1.1
XqA2
A3.1
A3.3
XqA4
XqA5
XqA6
qA7.1
qA7.3
XqB
XqC1
XqC3
XqD
XqE1
XqE3
XqF1
XqF2
XqF3
XqF4
XqF5
YqA1
YqA2
YqB
YqC1
YqC2
YqC3
YqD
YqE
Log2 Ratio
B
X / Y specific
PAB
Mid1
XY-PAR
MMU X
XqF3
XqF4
XqF5
MMU Y
YqD
YqE
PAB
DXYCbl3
Mid1
Asmt
PAB
DXYCbl3