Material Methods

2020-03-03

Table of Contents

Todo

Add the single Bivalent predictions to one of the sections.

# Mouse Husbandry

To access natural genetic variation for *Mus musculus*, wild derived inbred strains were used. The Mus musculus strains PWD/PhJ, PERC/Eij, WSB/EiJ, LEWE/EiJ, MSM/MsJ, MOLF/EiJ, CAST/EiJ, CZECHII/EiJ were purchased from Jackson labs (Maine, USA). The strains of KAZ/TUA, TOM/TUA, AST/TUA, and HMI/TUA were generated from frozen embryos purchased from Biological Resource Center (BRC) at Riken (Ibaraki, Japan). (<https://en.brc.riken.jp>) and rederived by the UW-Biotech centre (Madison UW). The related murid species *Mus caroli* CAROLI/EiJ and *Mus spretus* SPRET/EiJ were purchased from Jackson labs (<https://www.jax.org>) and *Mus spicilegus* SPIC/Eij from Riken.

All mice were housed UW-Madison Biotech and MSC facilities - following the protocols. A breeding colony of wild derived *Mus musculus* mice sampled from Gough Island (GI)is maintained at UW Veterinary school facilities. Mice were fed on dry standard breeder chow. Some strains sunflower seeds, nestlets and larger cages were added to improve fertility and litter survival. Adult mice were euthanized by CO asphyxiation. Neonate and embryonic mice were euthanized by decapitation following the protocols approved by the Institutional Animal Care and Use Committee at the University of Wisconsin-Madison.

Over the course of data collection some breeding colonies mice were moved from facilities. Additionally the GI strain was kept at a separate facility. (evidence for environmental (room) effect on (MLH1 / SC lengths / a variety of meiotic traits was tested for. Effects on average MLH1 counts were compared – for strains which were kept at mulitple institutions. ((We observed no effect, Supplemental figure))

# Tissue Collection and Immunohistochemistry

Spermatocyte spreads were collected and perpared as described in (Peterson, Miller, and Payseur 2019) . The majority of mice used were between 5 and 12 weeks, (supplemental table). However some strains had problems breeding, so testes from older mice were collected. We found a (small/no) age effect on MLH1 counts for male mice (Supplemental Figure)).

The only difference between preparation of spermatocyte and oocyte spreads was the volumns of the hypotonic buffer was 300 u l instead of 3ml.

The vast majority of oocyte data was collected from neonate litters/mice between 5 to 48 hours old. This approach maintain breeding pairing and still result in prophase oocytes (cite timeline of oocytes in neonates). (differences were assessed between embryo samples and neonates?).

Precise staging of embryos via copulary plugs was not feasible in many of these wild derived strains due to their behavior. Embryonic samples were collected when pregnancy noted in females, embryos were staged based on (X table markers).

Meiocyte spreads were made following (Peters et al. 1997). For testes the tunica was removed and whole testes was incubated in 3ml of hypotonic solution for 45min. For ovaries after dissection, the pair of ovaries were decapsulated in cold PBS and both were incubated in 300ul of hypotonic solution for 45 min.

After incubation, gonads were transferred to sucrose solution -for mastication to make a cell slurry which was transferred to 80ul of 2% PFA solution. Cells were fixed in this solution and dried in a humid chamber at room temperature overnight. The next morning, slides were treated with a photoflow (Kodak, cite) wash.

Staining / Immunohistochemistry

SYCP3 was used to mark visualize the Axial element of the synaptonemal complex. Sicne all mouse chromosomes are telocentric/acrocentric, CREST was used to indicate the centromere – indicate the side of the bivalent with the centromere.

the genome wide recombination rate was charaecterized / estimated with staining of MLH1 antibody, The same staining protocol was applied to spermatocyte and oocyte spreads. The staining protocol was based on X with some adaptations, and same as previously described in (Peterson, Miller, and Payseur 2019).

Precursurs to COs, DSBs were quantified with the DMC1 staning antibody. Due to resource constraints – only a subset of strains were stained for this marker. Only spermatocyte spreads produced enough good cells for DMC1 quantification.

# Image Processing

Images of cells were capture using a Zeiss Axioplan 2 microscope with AxioLab camera and Axio Vision software (Zeiss, Cambridge, UK).

Preprocessing, including cropping, noise reduction, and histogram adjustments, was performed using Photoshop (v13.0).

For MLH1 pachytene characterization, cells with a full karyotype (19 acrocentric bivalents and XY for spermatocytes or 20 acrocentric bivalents for oocytes), distinct foci, and intact bivalents were included for quantification. Image file names were anonymized before manual scoring of MLH1 or DMC1 foci.

Repeatability was assessed by X (comparing mulitple counted images)

Hand measures of meiotic chromosomes generated by using ImageJ/Fiji (v1.52) (Schindelin et al. 2012).

The total SC per cell and single bivalent measures were generated using the approaches discribed in (Wang et al. 2019) and (Peterson, Miller, and Payseur 2019) respectively.

Foci counts/numbers were manually scored.

Single bivalent features were measured both manual and with an automated image analysis software (**???**)

DMC1 – leptotene and sygotene cells were identified by SC morphology (by degree of synapsis) (cite another paper?). foci counted manually, error ccalulated from a subset of cells counted twice (supplement)

total number of DMC1 foci per cell were reported.

# Statisical Analysis

DMC1 – Ttest

MLH1 counts per cell are good /decent proxies for genome wide recombination rates estimates (cite).

The distributions of CO counts per cell was assessed for normality (supplemental figure). Mouse mean MLH1 count was used to get around the bad discreet nature of count data. (normality was confirmed with X distribution, supplemental figure).

The mean, variance and coefficient of variance for MLH1 counts per cell were calculated for each mouse (in R). these mouse level summary statistics were analyzed – within our model frameworks.

We chose to use a **mixed model framework** to interpret the patterns of variation, built using the lmer4 package (cite) in R (v3.5.2)(Team 2015). (lme)

The use of wild derived inbred strains enabled us to design an utilize a mixed model for our analytical framework. We built our model to predict the strain average MLH1 count per cell. The main effects and their interaction of subspecies and sex were coded as fixed. While, strain (or genetic background) was coded as a random effect to reflect a random sample of natural genetic variation from each subspecies. A strains to have a specific sex effect was also included.

This model allows us to estimate and test the fixed effects (sex and subsp), and variation in mean MLH1 counts due to strain effect. We use the subspecies term as a way to quantify between group variation (or divergence) and the random strain effect to quantify within group variation (or polymorphism). (in a sex specific manner). From the mixed model the estimates of heterochiasmy (sexual dimorphism) are a combination of subspecies and strain specific sex effects.

We can also compare the size of the sex effects between (subspecies) and within (strain) mouse groups. (ie is the male or female strain effect larger? is there more variance due to random strain effect in male or females?) (we used this frame work as proxy for ‘polymorphism’ and ‘divergence’ in MLH1 counts across house mouse and closely related species.) I may have to do more research/reading on testing the random effects.

### Single bivalet quant

Single bivalent chromosomes were isolated from cytology cell images as discribed in (X ). the single bivalents were ‘curated’ (removed errors in the segmentation algorithm)

to account for effects of CO number and physical chromosome size – the data was also analyzed after spliting up by chromosome class and in long and short chromosome bins.

* Split up ‘classes’ of bivalents depending on the number of foci – to tests/ comparisons more comprable.
* The performance of the segmentaion algorithm was assessed by counting the number of objects segmented per single cell image (expected 19 or 20 in males and females respectively). All automated bivalent data was manually curated before analysis.
* whole cell bivalent measures (measures of all bivalents ), were compiled by mixed automated bivalent measures with hand measurements. (Since the error in automated measures was so low)

**SC length**

The axial element (AE) of the SC – for single bivalents (synapsesd homolog pairs), was quantified as the skeletonized (single pixel width line) –

-comparison to human measures previously in permyscus species – and in musculuc (shown in the supplement)

* The long bivalent and short bivalent data sets were compiled by calculating the 1st and 4th quartille of SC lengths for each cell with at least 16 bivalent measures. Bivalents from each cell were assigned into the long and short bins based on these SC Length cut offs.

**CO interference**

* how we choose to measure interference, the average distance between2 foci on same bivalent.
* downside to using the gamma distribution – not the primary meteric for CO interference.
* Also pooling intervals for categories and fitting a gamma distribution (in supplement)

**rec landscape** **1CO normalized position**

* how why measure the recombination landscape – coohse to charaecterize the rec landscaoe bby isloating the 1CO bivalents; since there are known sexual dimorphic features of this and since bivalents with more than 1 foci will be primarily influenced by CO interference. We look at both the raw and

the distance to the telomere and centromere end of the SC shape were measured – for all / pooled bivalents

-specific R packages -non-parametric (tests and measures for dealing with count data) -hypothesis tested (size, DSB number)

-Mixed models and glm models were built

-tests for fixed effects -tests of random effects in mixed models was LTR of exactRLRT function in X package.

single bivalent analysis

# Numbers of mice cells ect for each method?

above belongs more in results

**outline**

1. framework for analyzing the MLH1 evolutionary patterns across mus species complex. (what is the main question we were tackling when designing this framework?)
2. write equation, describe terms (proxies for within and between polymorphism) logic / choices for fixed and random effects
3. what are the 2 general / basic questions this framework can be applied to.

$X mouse average MLH1 ~ sex + subspecies + random(strain) + $

1. how can DMC1 measures be added to the mixed model
2. descriptions of mlh1 outside of framework (missing female data)

# References

Peters, Antoine HFM, Annemieke W. Plug, Martine J. van Vugt, and Peter De Boer. 1997. “SHORT COMMUNICATIONS A Drying-down Technique for the Spreading of Mammalian Meiocytes from the Male and Female Germline.” *Chromosome Research* 5 (1): 66–68.

Peterson, April L, Nathan D Miller, and Bret A Payseur. 2019. “Conservation of the Genome-Wide Recombination Rate in White-Footed Mice.” *Heredity*. Nature Publishing Group, 1–16.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, and Benjamin Schmid. 2012. “Fiji: An Open-Source Platform for Biological-Image Analysis.” *Nature Methods* 9 (7): 676.

Team, RStudio. 2015. “RStudio: Integrated Development Environment for R.” Boston, MA. <http://www.rstudio.com>.

Wang, RJ, BL Dumont, P Jing, and BA Payseur. 2019. “A First Genetic Portrait of Synaptonemal Complex Variation.” *PLoS Genetics* 15 (8): e1008337–e1008337.