2020-04-16

Table of Contents

# Material Methods Draft v1

ToDo; Bret’s suggestion for moving thing to Mat and Methods

## 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. We use abriviated superscript notation for the remainder of the paper (e.g. *musculusWSB*).

All mice were housed UW-Madison Biotech and MSC facilities following RARC approved 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. We tested for effects on mean MLH1 counts across difference mouse rooms/facilities and found no effect (Supplemental section).

## 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). for DMC1 spreads we targeted juvinile male mice (X ages).

Because some strains have breeding issues, all mice some strains were kept in breeding pairs (before collecting, and aged out of the desired age range. We performed tests for effects on MLH1 counts and found no effects in the older male mice ages. (Also we tested other effects, including maternal age…)

The vast majority of oocyte data was collected from neonate mice between 5 to 48 hours old. This approach maintain breeding pairing and still result in prophase oocytes <(cite timeline of oocytes in 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). No effect on mlh1 count due to age effects (pooled as embryo and neonate) was found (Supplemental section).

Meiocyte spreads were made following (Peters et al. 1997). whole testes and decapsulated overaires were incubated in 3ml or 300ul of hypotonic solution for 45min respectively.

After incubation, gonads were transferred to sucrose solution for mastication. 15ul of cell slurry 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

Precursurs to COs, DSBs were quantified with the DMC1 staning antibody.

## 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 quantification, 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)

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.

(the data from these scripts differed slightly in that the total sc script returned the skeletonized SC area for the whole cell.

to quantify / charaecterize the SC (length) (synapsed chromosomes, )– two image analysis algorithms were used both which apply a ‘skeletonizing’ transformation to the synapsed chromosomes, which transforms the elongatde chromosome space into a single pixel wide ‘trace’ of the originial chromosom shape.

The DNA CrossOver script isolated single chromosomes shapes from an image, in addition to returning the SC length – also measured the location of (green and blue signal along the single SC - reflecting MLH1 foci and cetromere signal respectively).

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

Single bivalent features were measured both manual and with an automated image analysis software (Peterson, Miller, and Payseur 2019)

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

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).

### Models

**M1. Mixed Model**

**M2. Linear Model**

**M3. Linear Model**

**M4. Linear Model**

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.

## Single bivalent charaecterization

**For SC based metric**

* To account for confounding effects of sex chromosomes from pooled samples of bivalents, we considered a reduced dataset including only bivalents with SC lengths below the 2nd quartile (for SC length) in each cell. This dataset included the four or five shortest bivalents and excluded the X bivalent in oocytes. A total of 678 – 678 ‘short’ bivalents were isolated from 103 – 103 oocytes and 37 – 37 spermatocytes. Although this smaller dataset has decreased power, it offers a more comparable set of single bivalents to compare between the sexes.
* As an additional metric of chromatin compaction, we computed the total (summed) SC length of all bivalents for single cells, using the image analysis algorithm of (Wang et al. 2019). Mouse means were calculated from cell-wide total SC lengths in 2,984 – 2984 out of 3,680 – 3680 cells with MLH1 counts (Figure X).
* 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.

The single bivalent data set was curated of algorithm mistakes manually>. For charaecterizing any of the foci positions or interfocal distances any normalized refers to dividing the raw measures by the total SC for that invidiualy bivalent.

The proportions of different chromosome classes (seperated by number of foci) was tested with chi-sqaure tests.

**chromosome size effects**

Using the single bivalent dataset, the within cell distribution of bivalent SC lengths from cells with at least 17 isolated chromosome measures. The longest and shortest 4 to 5 chromosomes within each cell were identified by qualqulating the fourth and second quartille. (these comprised the short and long bivalent datasets).

* 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.

< manual measures of whole cell bivalent measures (measures of all bivalents were incorporated into this data set. ), were compiled by mixed automated bivalent measures with hand measurements. (Since the error in automated measures was so low)>

**SC Length model**

* **Linear model M1**
* **Linear model M2**
* This metric is closer to measures of interference from linkage maps which are also removed from physical scales in that they measure frequencies of crossovers. Crossover interference is stronger in male specific linkage maps compared to female maps in a variety of species (ref). **IFDnorm**

**CO Interferecne**

Our primary mode of examining interference is to compared the interfocal disatnces (IFDs) (raw and normalized by SC length.) Gamma parameter v has historically been used to charaecterize interference and was calculated using .

# 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* 123 (4). Nature Publishing Group: 442–57.

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.