Results Outline

**Full MLH1**

Basics stats

X cells, 3 subspecies (closely related) , X wild derived inbred strains sampled from across the territory

-sampling oocytes was the limiting factor –so we have additional strains with just male observations which are included in male-specific analyses. All measures of sex effects use data with sex matched strains

(estimate of replication from human counting, mention of relationship of quality and counting error

1. -female strain means are v close -male strain means are more varied!!

-Comparisons to references / literatures

-house mouse close to minimum of 1 obligate CO per chrm

2. The largest female difference is G, 1.07X the largest male difference is PWD and MSM which are 1.3 and 1.2 respectively.

(above replaces the main points for the sex specific patterns)

\*\*current issue, think about how to list the numbers in a understandable fashion.

Add strain specific glm results --- which strains are significant effects

**Evolutionary framework**

To understand the qualitative patterns

(in order to understand the variance within an evolutionary framework – we fit the mouse gwRRs to a mixed model, which specified subspecies, sex and the interaction as fixed effects. The subspecies effect – is a proxy for measuring the degree of divergence across subspecies. The strains were coded as a random effect to –simulate / approximate / the standing genetic variation across subspecies )

1.all mixed model terms were significant – Table X, So we ran post hoc fixed effect models

-glm: The G strain effect is the largest effect for gwRR. The sex\*strain interaction effects for PWD and MSM are also significant.

<the model framework // quantify – confirm significance of two general patterns (of total data)

1. G strain has large sexual dimorphism (degree of heterochiasmy) (table x)(p values for these effects)
2. MSM and PWD strain by sex effects – significantly higher just in males (pvalues from glm models)

Models in sex specific data indicate (which include more )

1. In females G and MSM have significant strain effects (table x)(p values for these effects)
2. SKIVE is a significant strain effect (intermediate),

\*\*for males; PWD, MSM and SKIVE can be classified as ‘high rec’, while rest of strains will be ‘min rec’\*\*

**===**

**WITHIN MOUSE VARIANCE in MLH1 count**

The variance and cv for within mouse MLH1 counts across cells are listed in Table 1

Variance across total sample / and within mice / the average within mouse sample

1. Se and var are larger in females, from full model

2. sex specific models, analysis (FIGURE OUT THE MAIN POINTS …. The strain effects seem very unstable across

(rank the strains by variance --- how is it affects by sample size)

1. Male – all 4 models, NO SIG strain effects’

Female :

Full.var and LEW.cv LEW sig strain

Q12.var LEW sig straing

Q12.cv PWD sig strain

2. across the models

Some strain effects, -- but inconsistent – strain effects across models

and hard to differentitae from sampling

**Sex specific model results**

**-female**: G and MSM

-male: ranges, of strain means, (ranges of mouse means)

**1.G and MSM in females, (variance)?**

**2.PWD\*male MSM\*male effects (variance)?**

PWD, MSM, and SKIVE have significant strain effects . They are grouped into the ‘High Rec’ group. PWD, MSM, and SKIVE have evolved 20, 30% and 10% higher than other means respectively.

1. Predictions for ‘uniform’ patterns across all Mus musculus strains are not met; (ie. strains within subspecies didn’t diverge uniformly)

2. Sex is a significant effect, but not in a uniform manner, the significant fixed effects are interactions (strain \* sex).

**DMC1 Variation in DSB number**

In an attempt to localize the above male specific crossover number evolution within the meiotic pathway we quantified a marker for DSBs, DMC1, in early prophase meiocytes.

To connect / the variation in CO number across – the high and low rec strain – we quantified DMC1 foci in meiocytes.

1.-basic stats, strains quantified from juvenile mice, cells selected based on stage of prophase. Means and number of cells reported in table.

1. **Lep (early Z) mean number of DMC1 foci per cell has evolved, in the direction predicted by number of COs.**

-basic stats, Number of cells, juvinille mice, 1 for each strain – this subset choosen to directly compare the differences in high and min rec strains

Boring results;

1. Early staged cells have significant more foci (DSBs) than the later stage.

**1. evolution of early number of DSBs (leptotene or early zygotene), this results in higher correlation of mean Z to CO)**

For all strains Early stages have more than later, (replicating pathway of meiosis

-only early cell stages are significantly different between the two groups

2. ratios calculated to estimate the proportion of DSBs repaired into COs (the inverse / tells can be used as a proxy for NCOs. The ratios calculated for both staged estimates – were not significantly different (t.ttes; p = p= )

3. points to the differences being established before DSB formation2. TABLE OF CO : NCO RATIOS ( DSB / MLH1 = estimated proportion of NCO. ) –**correlation with MLH1** L based ratios: WSB, G and KAZ: 7.3 and 6.5, 6.6\* PWD and MSM: 6.1 and 7.4

Z based ratios: WSB, G and KAZ: 5.4 and 5.9, 6.9 PWD and MSM: 4.8 and 5.3

+ t.tests of ratios

**Chromosome proportions, Genome wide recombination rate variation translates to the chromosome level** ( in order to decompose the cell wide rate, we decided to look at the proportion or chromosomes with different numbers of COs), gwRR evo -> chrm proportion*- This is motivation to investigate more traits at the chromosome level*

1. Basic stats – bivalent data numbers
2. Rang of co per chrm, mostly 1co 2co

Chi square tests

-low group: all Dom, KAZ, MOLF

-high PWD, MSM – have low sig p values for 1CO and 2CO tests

1. High rec strains have evolved a higher proportion of 2CO

For males Chi squared test, p = low group, p = MSM and PWD, SKIVE is intermediate)

**MSM 60% PWD 50% SKIVE 30% 20 - 10% remaining (low) strains**

1. explain why this complicates further comparisons (chromosome size effect \* chromosome class)

**Single Bivalent Level Dataset**

1. -motivation – de-construct the cell wide pattern – apply an algorithm to the MLH1 stained meiocyte images

2 - algorithm sates and performance (n bivalents (with hand foci), (Table 3 --- Ncells, Nbivs, X, proportion chrm class, )

3. average number of bivalents isolated-measures and estimates of error -- (high rate of replication) (estimates of error compare to manual measures – ref Peterson 2019, table and figure--(not all bivalents isolated average number of bivalents isolated per image, --but such large dataset – assume across all the bivalent datasets across categories ARE EQUALLY REPRESENTATIVE OF GENERAL PATTERNS.

4. <transition to driving questions>

Main motivating questions / We will use this data to address 2 questions:

Q1. Which bivalent level traits will be sexually dimorphic?

Q2. Which traits distinguish high and low recombining strains in males?

**Q1. SC – sex differences in SC length // SC area**

Our data set provides an opportunity to test if the canonical pattern of females having higher gwRR and more SC area) – have been uncoupled in the instance of rapid male specific evolution in the gwRR of PWD and MSM. A simple model might predict if SC area is the strongest predictor for gwRR / CO number, the high recombining males would have greater SC area proportional to the difference in CO number per cell.

Mouse averages for the SC based metrics ( irregaurdless of foci count ) – general comparison of chromatin compaction between sexes

Basic stats

- cells used in the final total SC data set (after filters and removing outliers)

– more total SC area (total SC)

**-t.tests for significant sex differences (almost equvilent sets of chromosomes, wouldn’t expect )**

**-Use glm / lmer for mouse average sc area – to test if there has been evolution – and interaction across the subsp and sex**

1. Total SC

Ttests, lmer, and glm – FEMALE HAVE SIG MORE SC\_AE

1. **Short bivalent data set**

Get around the XX – reduced single bivalent data set – from shortest 5 bivalents from a single cell, the XX is thought to be 3rd longest in reference genome (mb). The XY in males and distinguishable and can be is filtered out from the single bivalent data set

Short Biv

(motivation for short bivalents) – remove effects of XX

Ttests, in all but SKIVE, female mouse have significantly longer short bivalnes

Models, Lmer- and glm - Sex is the most significant effect

**1.from t.test()**

All but SKIVE have significantly long SHORT biv means -- Dom has greater sexual dimorphism compared to Musc

SKIVE is not significant (p=.11) (this might be a low number of mouse samples / maybe lower quality of the cells) … noticeable the female means are lower compared to pwd and kaz

1. <Simple model / prediction not met, suggesting a DECOUPLING of broad summaries of SC length and gwRR – female have longer SC metrics even in strains with males have more COs per cell. This is not a complete decoupling since in all strains the positive correlation of SC lengths across bivalent classes is held.

2. Longer SC-AE in females is – a consistent feature across all strains (t-tests, model’s (large sex effect)) for a cell wide summary and (a reduced single bivalent data set)

3. males from Musc and mol strains have significantly longer SC metrics than males from Dom. Suggests at chromatin compaction – between these subspecies and may (be a requirement for the rapid evolution seen in the 2 musc and mol strains)

<!-- -For almost all models, sex is the only significant effect for mouse averages of SC length.

The exception is in the Musc strains, where the SKIVE strain effect is also significant, because both male and female bivalents are shorter compared to PWD and KAZ.

-Female SC are longer than male even despite the XX. i) all bivalents are longer within cells (there isn’t a single longer bivalent), ii) shortest bivalents within cells are also longer in females. 

1. Transition, --- the SC-AE area can be thought of as the ‘area / available real-estate for COs

-- these results suggest that this area is ‘utilized’ in different ways across sexes and strains – so we next investigate the recombination landscape – (ie the relationship between the placement and number of COs along chromosomes).

2. We focus on two metrics/ aspects i) the placement of single foci along a bivalent (just 1CO) and ii) the placement of two foci on the same bivalent / the interfocal distance of 2CO bivalents (reflects crossover interference)

**Q1. 1CO foci placement / foci position**

Motivation:

1. -established sex difference, General summary of Rec landscape –

Investigating evolution of ‘typical recombination landscape’ (from Sardell Kirkpatrick) Has the typical rec landscape evolved in house mouse? (in terms of the sex specific patterns)

2. - biological significance – sex differnces in centromere suppression, position bias

Are centromeres suppressing CO? and do telomeres promote clustering of COs at the telomeres? –beyond the scope of this paper

Framework notes

-Normalized positions to control for differences in SC length (across sexes)

-(stick to full biv data set) – if results not sig, try smaller data sets

-basic stats – table X – number of 1CO bivalents per category

1. **Sex is the most significant factor influencing the normalized placement of single foci along a bivalent (by t-tests, mixed and linear models).**

**(--Dom strains more significant sex differences compared to musc?)**

**1. Sex is the most significant factor influencing the normalized placement of single foci along a bivalent (by t-tests, mixed and linear models effects).**

**2. (Dom and musc strains more significant sex differences compared to Molossinus (– with the dom strains being more telomeric in males compared to mol )**

**-**The sex effect is highly significant for glms of the normalized foci 1 position. In models for the full dataset, LEW, MOLF, and KAZ had significant strain effects, which MOLF being the largest. The LEW and the MOLF strain effects are replicated in models sub setting the subspecies.

-Males have more telomeric positions of single foci while female single foci are closer positioned to 50% of the total length.

-Some strains MOLF and LEW, have significant strain effects, indicating both male and female have positions closer to the middle of the bivalent than the total average.

- *SC length and the normalized 1CO position are conserved chromosome level traits of heterochiasmy in house mouse (and other species).*

-caveats / complications – chromosome size effects

(positions / landscapes differ between chromsomes

Small chromosome (Mb physical) have more uniform landscape – either due to i) bias positions in absolute scale (thus more of small chrms) (Keeney) or ii) same mimimum requirement and less physical area for obligate CO to occur (--or there’s less space for the position to)

the minimum number of

**Q1. IFD and Interference**

**Motivation backgound**

What are the points for the IFD metrics?

Raw: test of mechanical force mediated through SC

Normalized: (closer to general desription of rec landscape?)

( comparison between sexes which have inherent SC length differences, -- general pattern of how much of each bivalent separates 2 foci? )

Results

- Female normalized IFDs are an average of 45 to 50% of the length of the SC-AE across strains. Male normalized IFDs are on average 51 to 60% the length of the SC-AE across strains.

- Generally all females have enrichment of short normalized IFD (except KAZ). This enrichment is most pronounced in strains with high rec males, (PWD, MSM, SKIVE) there is a cut off of low normalized IFDs ~30%. In the remaining strains the normalized IFD ranges overlap between males and females.

- *Females have weaker interference as indicated by normalized IFD.*

*- High rec males have stronger interference, in terms of raw and normalized IFD measures. They also have a lower threshold for IFDs, 30%, this could indicate stricter control over the REC landscape to enrich for more 2CO bivalents*

Basic stats – table X

X number of

**Results**

NOT THAT INTERESTING – confirms pattern from genetic maps

Q2 motivation and framework

Q2 interfernce

**High rec strains have longer interference**

**Q2 SC LEGNTH**

-the positive correlation of SC length and chromosome class is kept (but because of prev outlined issues with comparing across strains with different proportions – this complicated to compare High vs Low using the single bivalent data set

**T.TEST, SC based metric – distinguishing between high and low strains (t.test, log regression)**

Total.SC (mouse mean), t.test – Significant p = 0.01

Long.biv (mouse mean), t.test - NS

Short.Biv (mouse mean). t.test - NS

All bivalents (mouse means?)

**Glms** **what are the patterns of variation**

Total.SC (mouse mean)

glm\_M1 – musc, Mol, G, LEW, MSM, SKIVE, KAZ have sig effects

glm\_M2 -- all strains have significant effects

Long.biv (mouse mean)

glm\_M1 – Musc, and MSM have significant effects

glm\_M2 -- MSM, SKIVE, CZECH have significant effects

Short.Biv (mouse mean)

glm\_M1: Musc, MSM, G have significant effects

glm\_M2: all (but LEW)

more strain effects are significant when subsp removed.

More effects are significant for the total sc metric

**Q2 Normalized Foci pos**

(in order to test another aspect of the rec landscape ) --

In prev Section above confirmed males have a rec landscape biased to the telomere,

In this section we ask if

Metric – normalized 1CO pos (1CO)

Physical size effects – can’t be disentangled from chromosome specific (identity effects) in this data

Long and short bins to account for chromosome size effects

t.test – is there difference between high and low rec groups?

mouse means all:

mouse mean long:

mouse mean short:

NONE OF THESE TESTS INDICATE THAT there’s a significant difference between groups – looking at normalized 1CO pos. (this holds when SKIVE is a low – but that is different than EXCLUDING SKIVE

glms – has there been any evolution of this trait in house mouse?

mouse means all:

mouse mean long:

mouse mean short:

NS values – chromosome specific effects

The bins of 1COs are made of different chromosomes across these two groups,

Maybe issue with resolution / power – too hard to detect differences