Analysis of Genetic Responses to the Antipsychotic Medicine Haloperidol with RNA-Seq Data from Diverse Mouse Recombinant Inbred Crosses (RIX)

Vasyl Zhabotynsky[2], Paola Giusti-Rodríguez[1], James M. Holt[1], Andrew P Morgan[1], Zhaoxi Zhang[1], James Xenakis[1], James J. Crowley[1], Randal J. Nonneman[1], Daniela M. Decristo[1], Allison Ryan[1], Timothy A Bell[1], Darla R. Miller[1], Ginger S. Shaw[1], Wei Sun[4], Jennifer R Brennan[1], Leonard McMillan[3], Patrick F. Sullivan[1,5], Fernando Pardo-Manuel de Villena[1] and Fei Zou[1,2].

[1]Dept of Genetics, University of North Carolina, Chapel Hill; [2]Dept of Biostatistics, University of North Carolina, Chapel Hill; [3]Dept of Computer Science, University of North Carolina, Chapel Hill; [4]Public Health Science Division, Fred Hutchison Cancer Research Center; [5]Dept of Medical Epidemiology and Biostatistics, Karolinska Institute;

Introduction

Schizophrenia is a chronic brain disorder that affects about 1% of the population worldwide, and it is associated with substantial loss in life expectancy and personal costs.

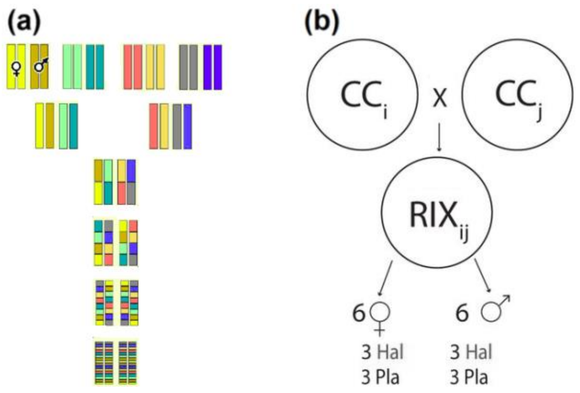
Therapeutic response to antipsychotics is known to vary a lot with a strong suggestion of genetic variation playing notable role.

Also, there is a well established literature regarding applicability of mice as a practical replacement of human subjects studying haloperidol side-effects.

More recently Kim at al. (2018) observed overlap between the genetic variation underlying the pathophysiology of schizophrenia and the molecular effects of haloperidol.

Finally, Giusti-Rodriguez at al. (2019) show a strong phenotype variation using genetically diverse mice.

Experimental design



(a) Derivation of Recombinant Inbred (RI) strain, (b) RIX cross production and haloperidol exposure.

Each RI is produced from 8 founder strains [129, A/J, B6, NOD, NZO, CAST, PWK, WSB]) ordered randomly, first going through the funnel breeding and then inbreeding process.

Experiment consists of 22 F_1 hybrids from 24 RI strains,

For each of F_1 crosses the goal was to get 6 male and 6 female mice, half treated and half untreated. In each cage there were two mice of the same sex: one treated and one untreated.

The goal was achieved for majority of the crosses, but due to mating issues could go as low as 2 treated and 2 untreated mice in one cross

Quality control filtering

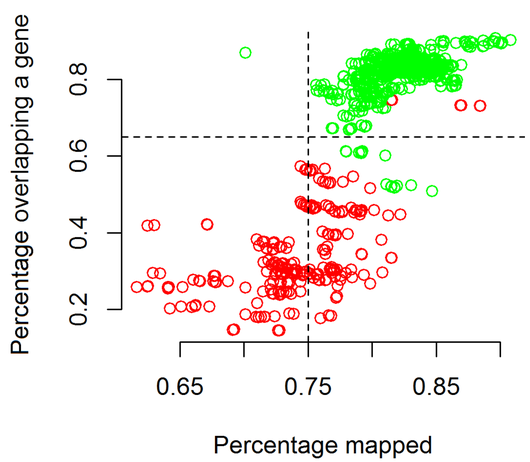
We observed a notable fraction of lanes with potential problems. After considering several QC metrics we ended up using following 3 criteria removing the sample if:

1) Duplication level was above 40% (marked red)

2) Percentage of mapped reads below 75% (x axis)

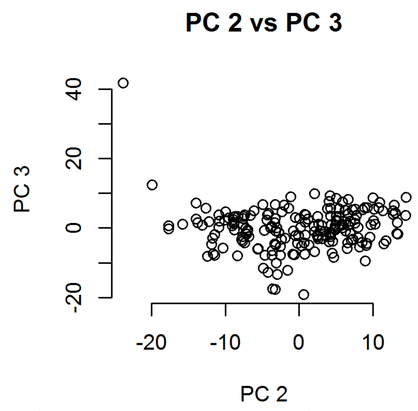
3) Percentage of mapped reads that map to a gene below 65% (y axis)

.



34 samples were dropped at this step.

One more sample that was on the boundary in the defined above rules was also suspicious in top principal components, so it also was removed.



Additionally, we had to dropp 4 more samples for which we could not confidently recover cross information, leaving us with 193 mice.

Model

The results obtained from fitting each strain separately were prone to outliers.

As a practical way to model 24 strains (and potentially even more with RIX design) we propose to incorporate founder status for each strain. Due to the way RIX are produced, founder status can be inferred for each RIX for most of the genes.

Observed distribution of number of founders in our population is ranging from 3 to 8,with 98% of the genes having at least 6 founders.

founders	3	4	5	6	7	8
genes	1	24	210	2439	6157	4692

Given F - number of founders for the gene (up to 8), we denote the genotype of each gene from a RIX as A_iA_j where A_i, A_j - particular founder allele (i, j = 1...F) - with the first allele (A_i) coming from mother and the second allele (A_j) coming from father.

We would code this $fnd_{f,s}$ information as 1 if A_i is a founder of the given strain. Note, this implies that exactly 2 out of F covariates would be 1 and the rest would be 0.

Along with other covariates such as library depth (k_s), sex (sex_s), treatment (trt_s) and principal components ($PC_{1,s}$ … $PC_{P,s}$) we modeled total expression y_s with Negative Binomial distribution in the following manner:

$$y_s \sim f_{NB}(y_s; \mu_s, \phi), \text{ for } s = 1, 2, ..., N,$$

$$\log(\mu_s) = \beta_0 + \beta_\kappa \times \kappa_s + \beta_{sex} \times sex_s + \beta_{trt} \times trt_s + \sum_{f=1}^{F-1} \beta_f \times fnd_{f,s} + \sum_{k=1}^{P} \beta_{PCk,s} \times PC_{k,s}$$

We tested treatment, sex and additive effects with likelihood ratio test.

$$H_0 : \beta_{trt} = 0, \; H_0 : \beta_{sex} = 0 \text{ or } H_0 : \beta_1 = ...\beta_{F-1} = 0,$$

Chromosome X

To model X chromosome we modify the above autosomal model by accounting for the fact that in male mice only maternal chromosome is present.
Female mice are treated similarly to autosomal subsection model and male mice, any cross with founders A_iA_j are treated as A_iA_i setting $fnd_{i,s}$ to 2 and leaving all the other founder variables to be 0.

Results

Comparisons to previous results

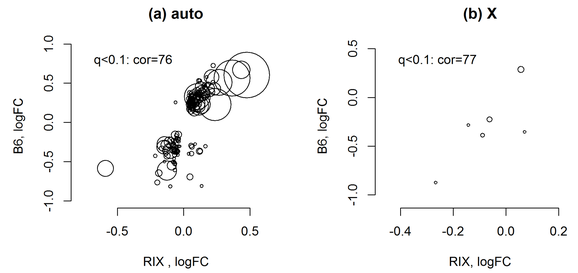
Considering multiple approaches of principle component selection, we ended up using the most consensual result - 27 PCs.

Overall number of additive effects was comparable to Crowley at al (2015) with 7,542 significant genes at q-value 0.10. For treatment effect we compared the results with Kim at al. (2018)

The overlap of the genes tested in both datasets is higher than one would expect to observe by chance.

q-val	Ref.	RIX	Both	Exc.
0.05	56	492	17	14.8
0.1	1264	774	164	86.3
0.15	3947	1013	395	77.4

Direction of significant effects



RIX dataset agrees with Kim at al. (2018) dataset regarding direction of effects

Excess of up-regulated genes

Using RIX at q-value 0.05 we discovered 316 up-regulated genes and 198 down-regulated, which proportion is significantly different from 0.50 (p-value 2e-7).

Pathway results

We observed consistent message about stronger up-regulation during pathway analysis.

GO Process results suggest that haloperidol is altering synaptic plasticity and cell signaling, via alterations in channel expression, localization, or modulation

GO Component results suggest that haloperidol may be altering neuronal morphology or density

Summary

Utilizing RIX dataset requires more extensive modeling,

It confirms multiple previously observed results in a more diverse population.

It discovers more genes and allows to observe more detailed information regarding found genes (preferential up-regulation),

It also produces consistent results in pathway analysis and provides insights regarding mechanisms of haloperidol work.

Citations

1 Kim, Yunjung, et al. "Comparative genomic evidence for the involvement of schizophrenia risk genes in antipsychotic effects." *Molecular psychiatry* 23.3 (2018): 708-712.

2 Giusti-Rodríguez, Paola, et al. "Antipsychotic behavioral phenotypes in the mouse Collaborative Cross recombinant inbred inter-crosses (RIX)." *BioRxiv* (2019): 761353.

3 Crowley, James J., et al. "Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance." *Nature genetics* 47.4 (2015): 353-360.