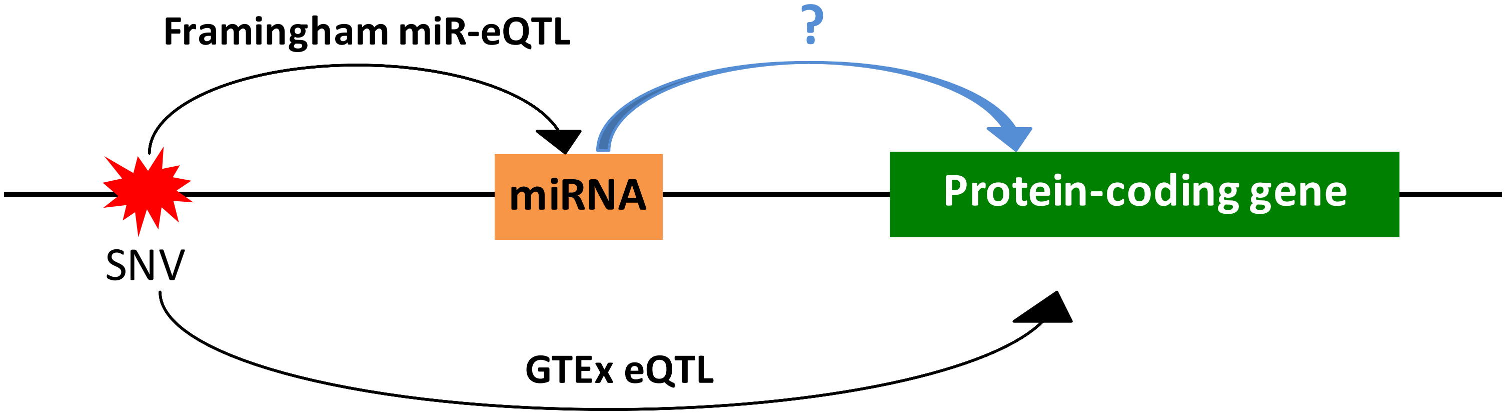
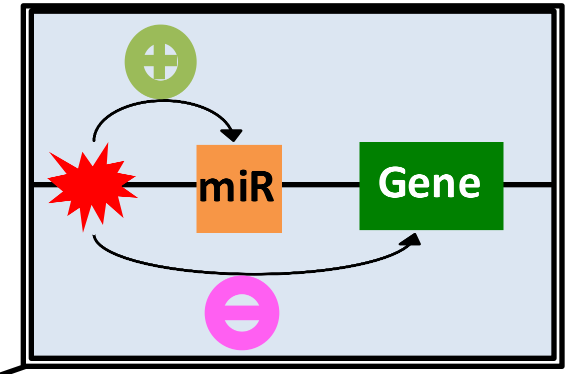
Are there detectable eQTLs in GTEx Whole Blood samples that may arise from miRNA-eQTLs (available in published Framingham dataset), wherein the "GTEx gene" is the target of the associated miRNA?

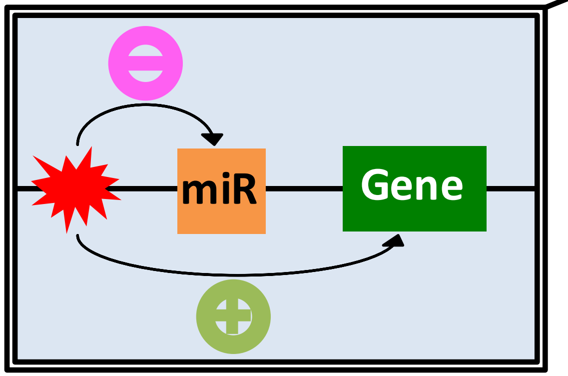


Total # w/opposite-signed effect sizes: 3052 / 4731

GTEx eQTLs (Effects)

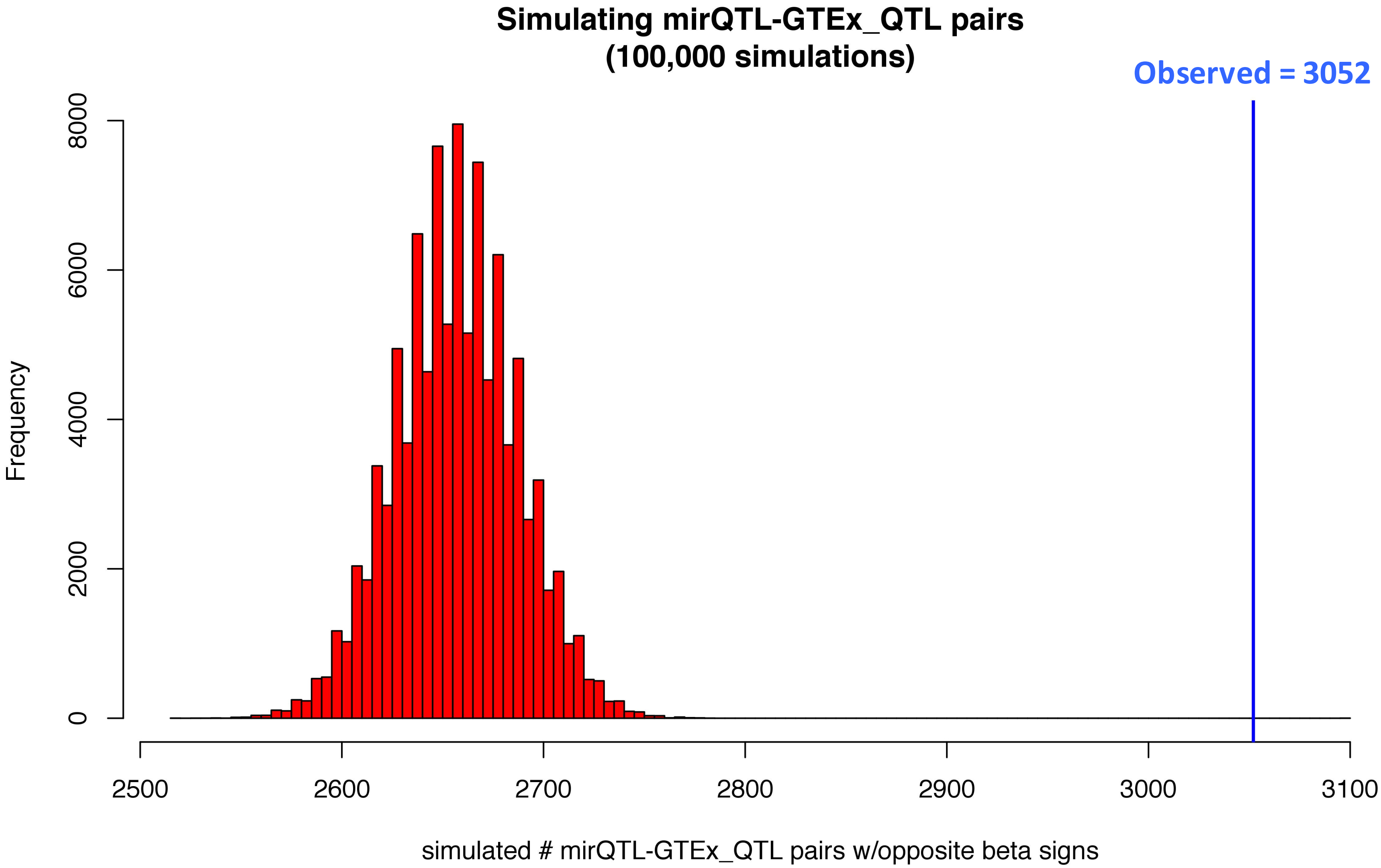
Framingham mirQTLs (Effects)	⊕	⊖	
⊕	847	2357	3204
⊖	695	832	1527
	1542	3189	





Analytically-derived 2-sided p value for num_opposite_effects (3052) = 1.319E-30
(derived w/pen-and paper calculation & normal approx to binomial distribution)

Also: 100,000 simulations to check for significance
+ most_extreme__num_w_opposite_sign: 2784



A “cis-cis” search necessarily limits the search of eQTLs to each of the following classes:

1) only those SNVs that are close to both the miRNA and the gene:

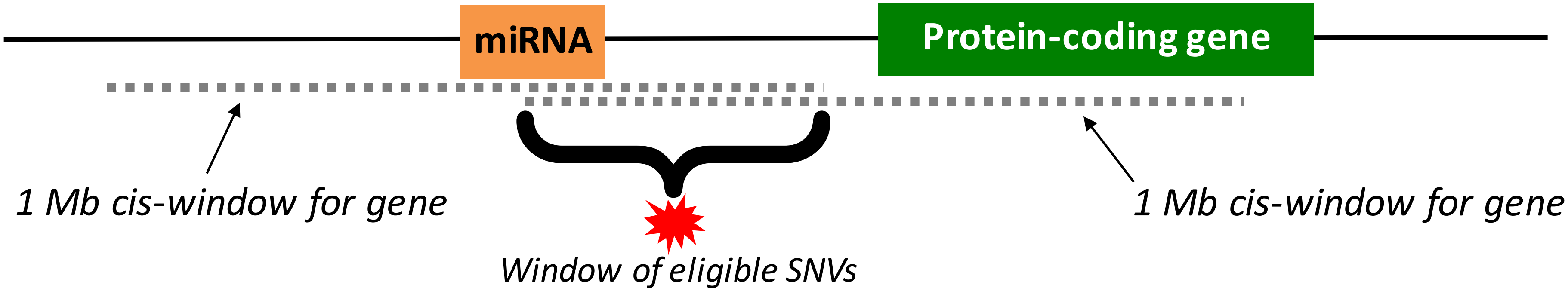
$$\mathbf{Dist_{SNV\text{-}miRNA} < 1\ Mb \quad \& \quad Dist_{SNV\text{-}Gene} < 1\ Mb}$$

2) among those SNVs, consider only those in the overlap of the restricted windows

$$\mathbf{[miRNA\_cis\_window] \cap [gene\_cis\_window]}$$

3) only miRNA-gene pairs that are close to one another

$$\mathbf{Dist_{miRNA\text{-}Gene} < 2\ Mb}$$



miRNA

Protein-coding gene

1 Mb cis-window for gene

Window of eligible SNVs

1 Mb cis-window for gene

The total number of identified eQTLs grows from ~5000 to ~10,000

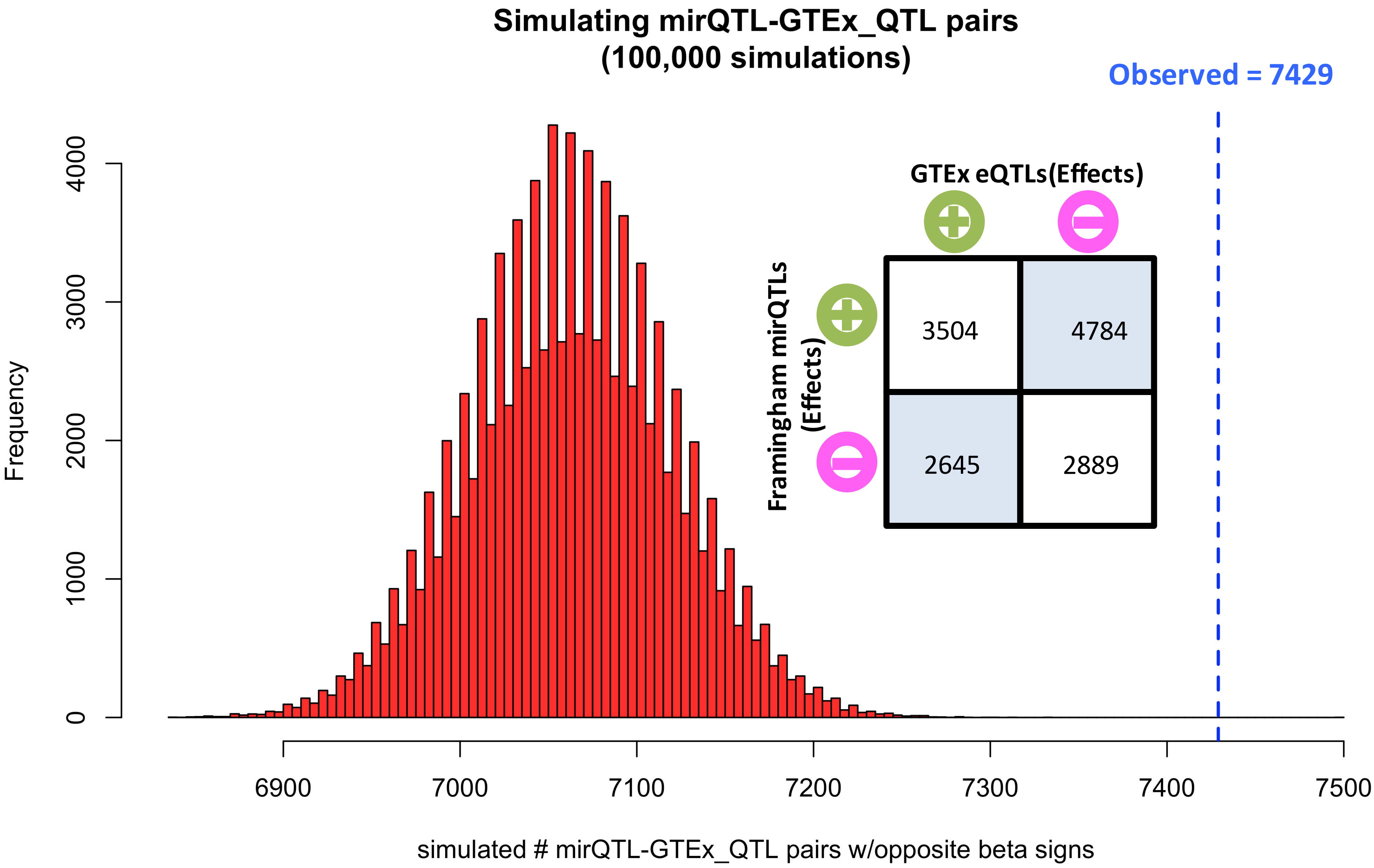
GTEx eQTLs: FDR < .10

		cis	*trans*	
Framingham mirQTLs	*cis*	13053	12	13065
	trans	655	102	757
		13708	120	

GTEx eQTLs: FDR < .05

		cis	*trans*	
Framingham mirQTLs	*cis*	10184	8	10192
	trans	583	0	583
		10767	8	

FDR < 0.10



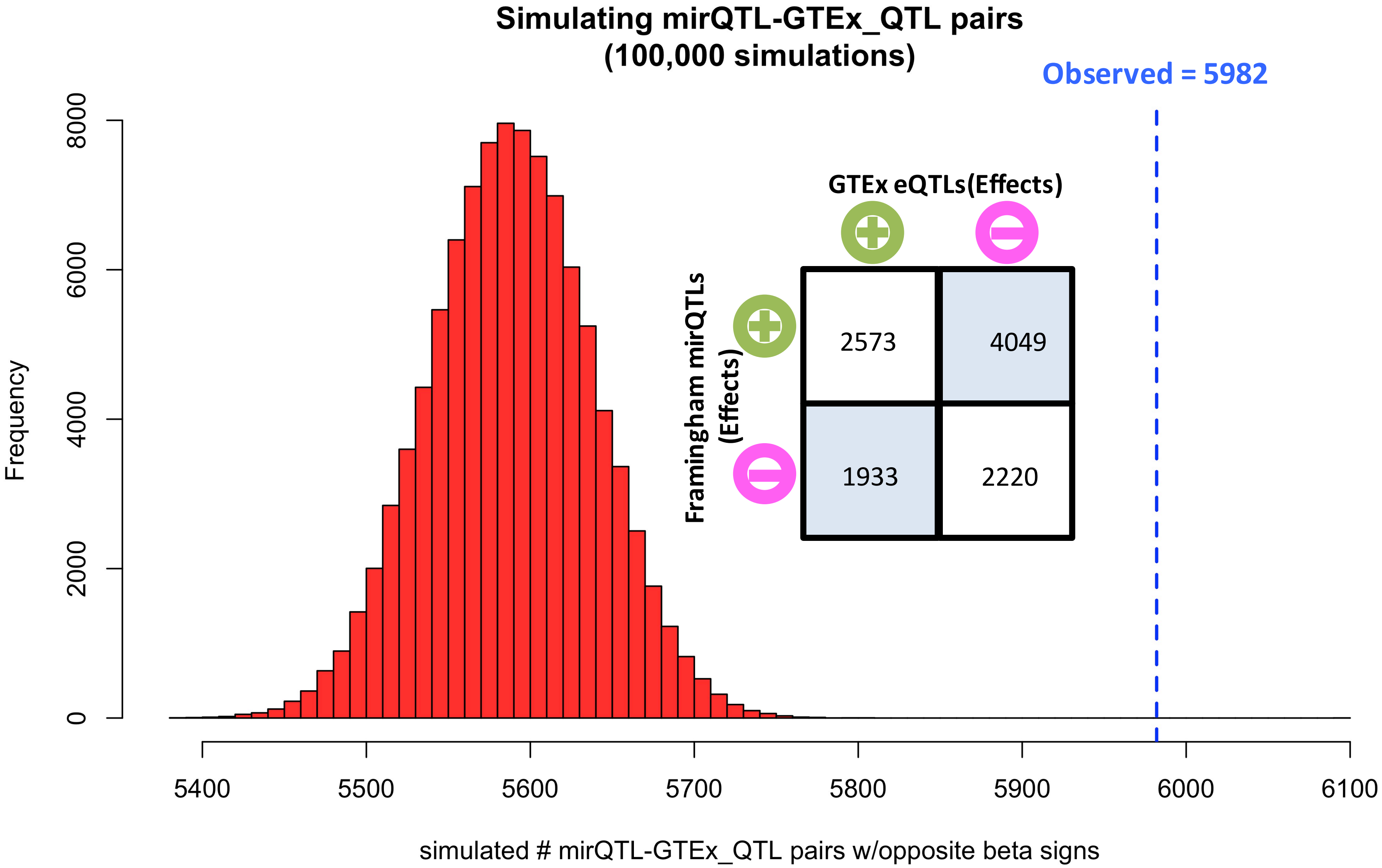
Simulating mirQTL-GTEx_QTL pairs (100,000 simulations)

Observed = 7429

GTEx eQTLs(Effects)

Framingham mirQTLs (Effects)	+	−
+	3504	4784
−	2645	2889

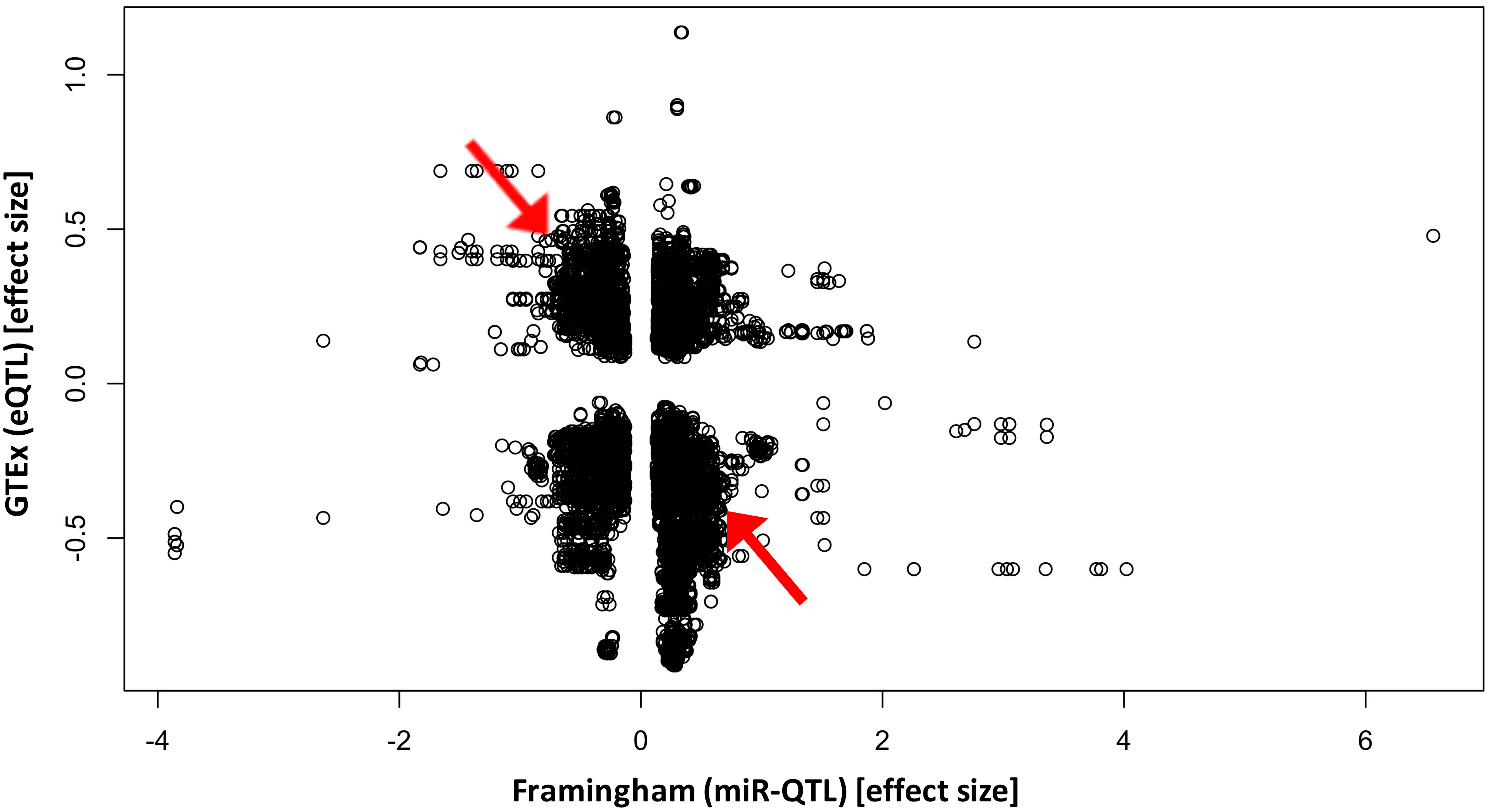
FDR < 0.05



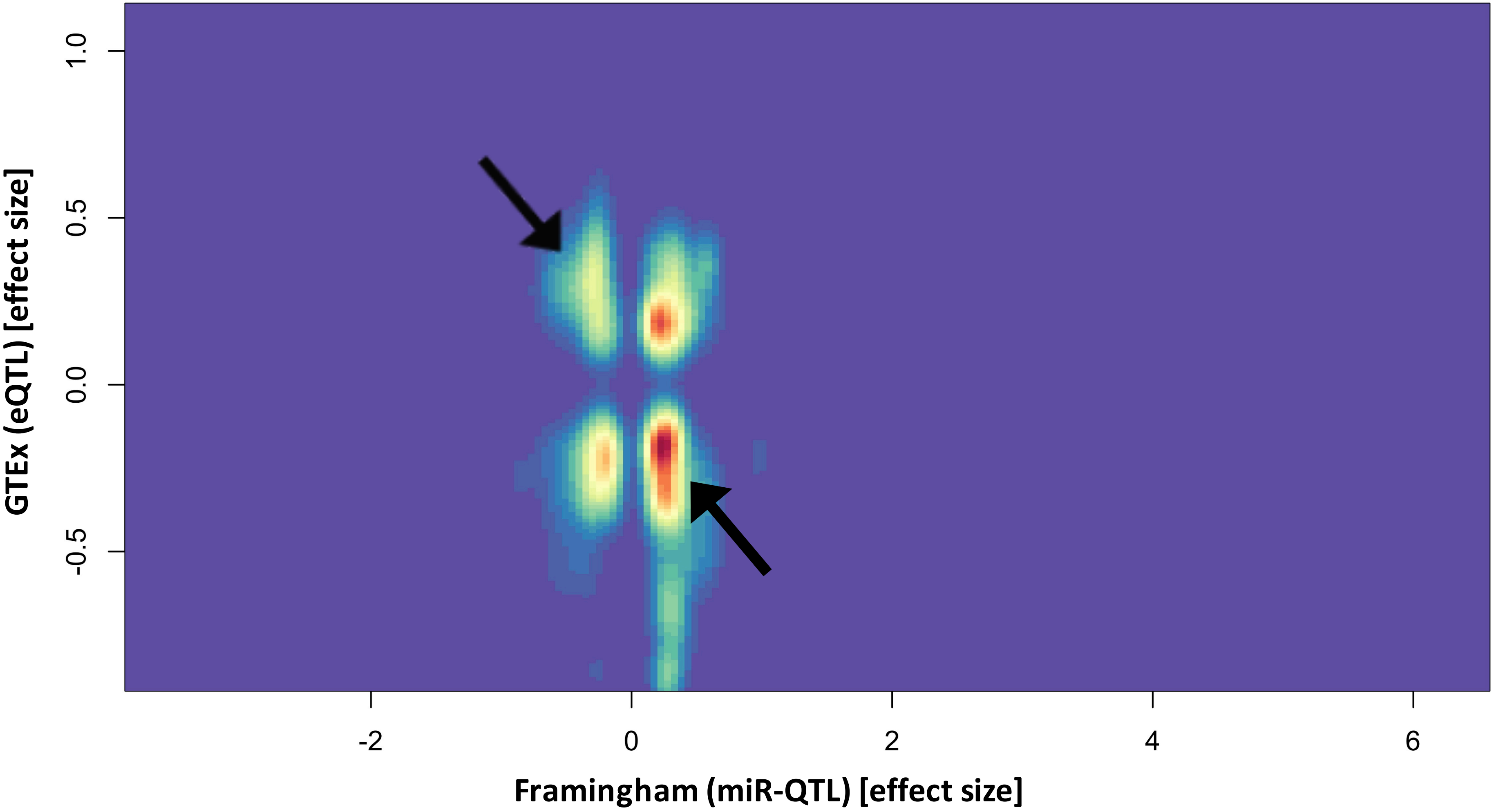
Simulating mirQTL-GTEx_QTL pairs
(100,000 simulations)

Framingham mirQTLs (Effects)	GTEx eQTLs(Effects)	
	⊕	⊖
⊕	2573	4049
⊖	1933	2220

FDR < 0.05

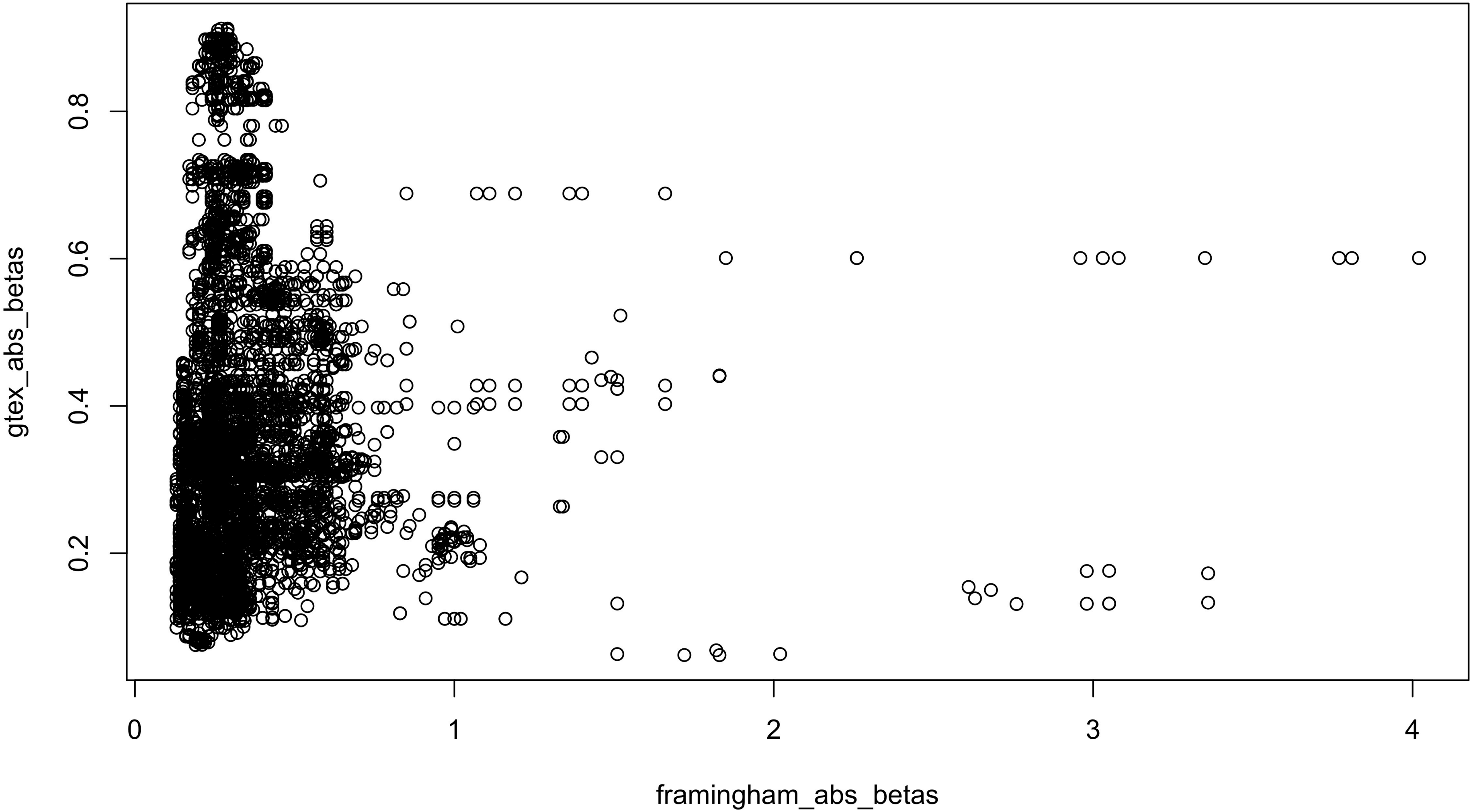

GTEx (eQTL) [effect size]
1.0
0.5
0.0
-0.5
-4
-2
0
2
4
6
Framingham (miR-QTL) [effect size]

FDR < 0.05


1.0
0.5
0.0
-0.5
GTEx (eQTL) [effect size]
-2
0
2
4
6
Framingham (miR-QTL) [effect size]

FDR < 0.05

For cases in which there are opposite effects, are the strengths of the eQTLs correlated (as measured by the abs. value of the effect sizes)?



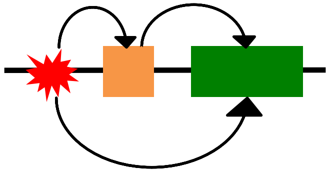
***Note very different processing for the 2 datasets though**

At FDR <= 0.05:

10775 matched eQTLs queried for primary linkages

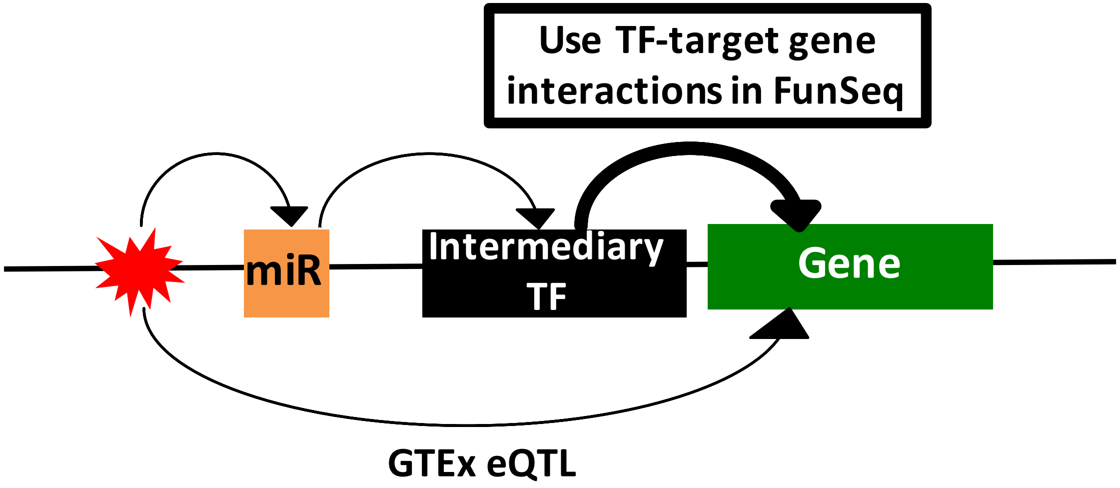
- > 206 distinct genes
- > 54 distinct micro-RNAs

- > 61 **“primary”** miRNA-target linkages identified
  - **> 8 distinct genes**
  - **> 9 distinct micro-RNAs**



Only a limited number of direct (ie, “primary”) miRNA-gene linkages identified

- Possible that the miRNA is acting on gene expression ***indirectly*** – for example, through an intermediary TF
- one caveat w/this line of reasoning: if this is the case, then why is it that the “primary target” (ie, the TF in this case) is not captured as a GTEx eQTL?



Total # of TFs in FunSeq: 115

common eQTLs between Framingham & GTEx (at FDR < 0.05): **10775**

of these 10775, # secondary (“TF-mediated”) eQTLs identified: **1074**

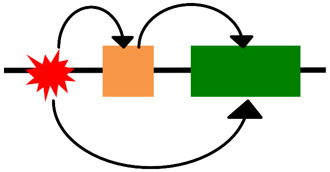
At FDR <= 0.05:

10775 matched eQTLs queried for primary linkages

- > 206 distinct genes
- > 54 distinct micro-RNAs

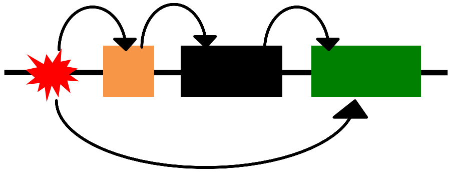
> **61** primary miRNA-target linkages identified

- **> 8 distinct genes**
- **> 9 distinct micro-RNAs**



> **1074** secondary ("TF-mediated") eQTLs identified

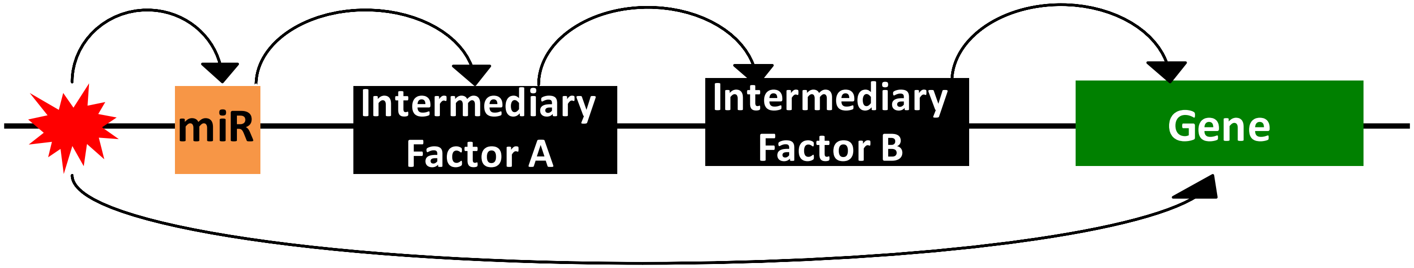
- **> 20 distinct genes**
- **> 13 distinct micro-RNAs**



Further steps

- Uniform processing using raw data from FHS (also enables use of PEER factors)
- NL2nd

- Tertiary or higher-order effects?
- combinatorial effects?
- networks/graph-based approaches involving PPIs



- TF-miRNA linkages from ENCODE-nets? (same as FunSeq?)

- more in-depth look into distinct genes identified above (master regulators?)

- integration w/available GWAS data

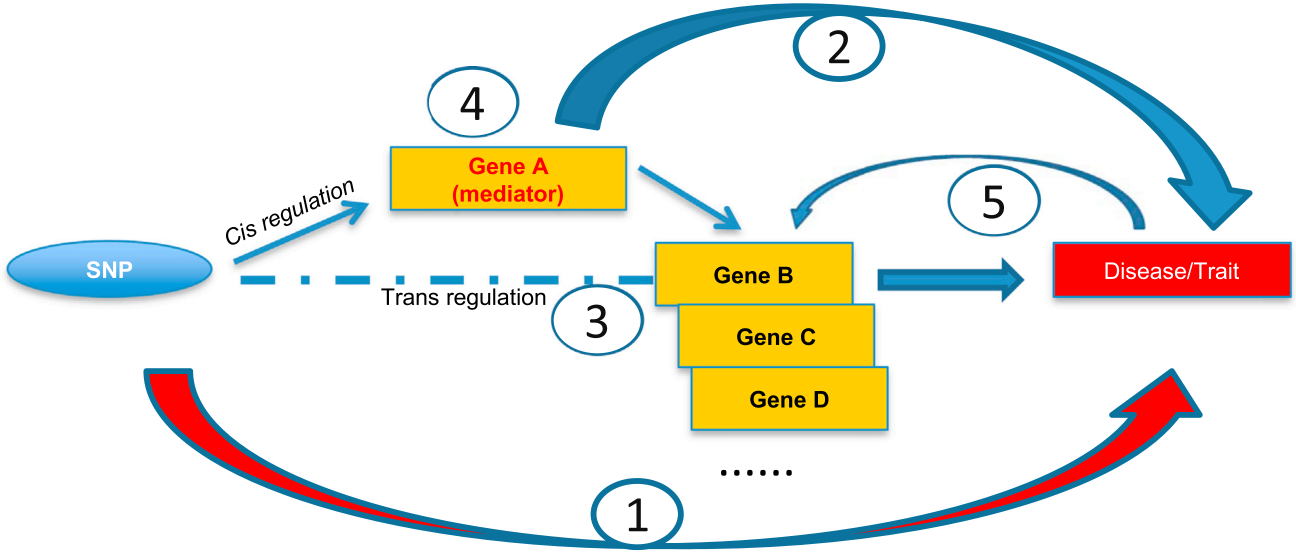
- Very recent paper in The American Journal of Human Genetics


ARTICLE

Dynamic Role of *trans* Regulation of Gene Expression in Relation to Complex Traits

Chen Yao,[1,2] Roby Joehanes,[1,2,3] Andrew D. Johnson,[1,2] Tianxiao Huan,[1,2] Chunyu Liu,[1,2] Jane E. Freedman,[4] Peter J. Munson,[5] David E. Hill,[6,7] Marc Vidal,[6,7] and Daniel Levy[1,2,*]

"Using published GWAS datasets with 39,165 single-nucleotide polymorphisms (SNPs) associated with 1,960 traits, ***we explored whole blood gene expression associations of trait-associated SNPs in 5,257 individuals from the Framingham Heart Study****. We identified 2,350 trans-eQTLs (at p < 10 7); more than 80% of them were found to have cis-associated eGenes... We hypothesized that some* ***trans-eQTLs regulate expression of distant genes by altering the expression of nearby genes*** *(cis-eGenes)."*



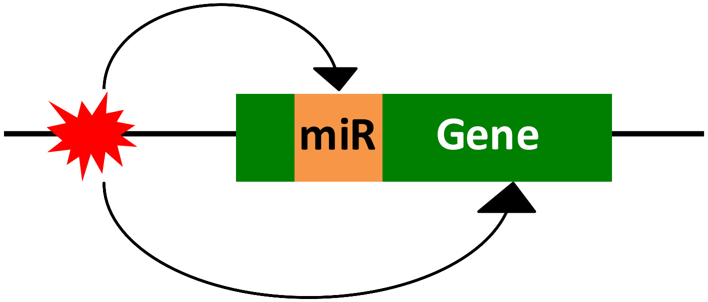
Supplementary content- incl:

- > Interior miRNAs
- > p-value calculations

miRNAs within
the genes themselves

N = 43 pairs(< 1%)

6 miRNAs, 4 genes



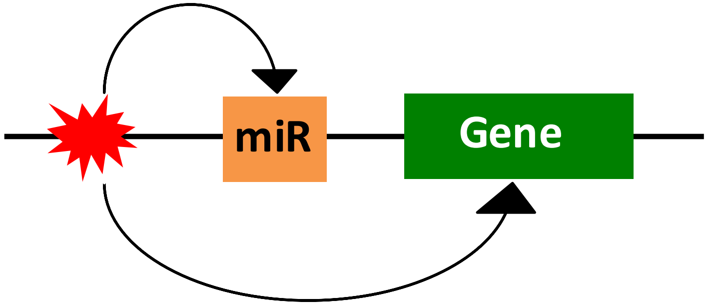
w/opposite effect sizes: 8/ 43

	⊕	⊖
⊕	21	4
⊖	4	14

miRNAs outside
the genes

N = 4688 pairs

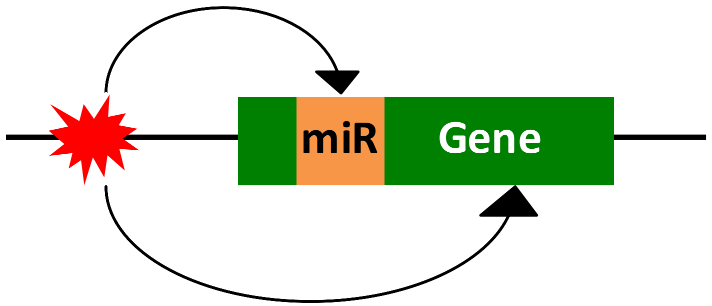
39 miRNAs, 59 genes

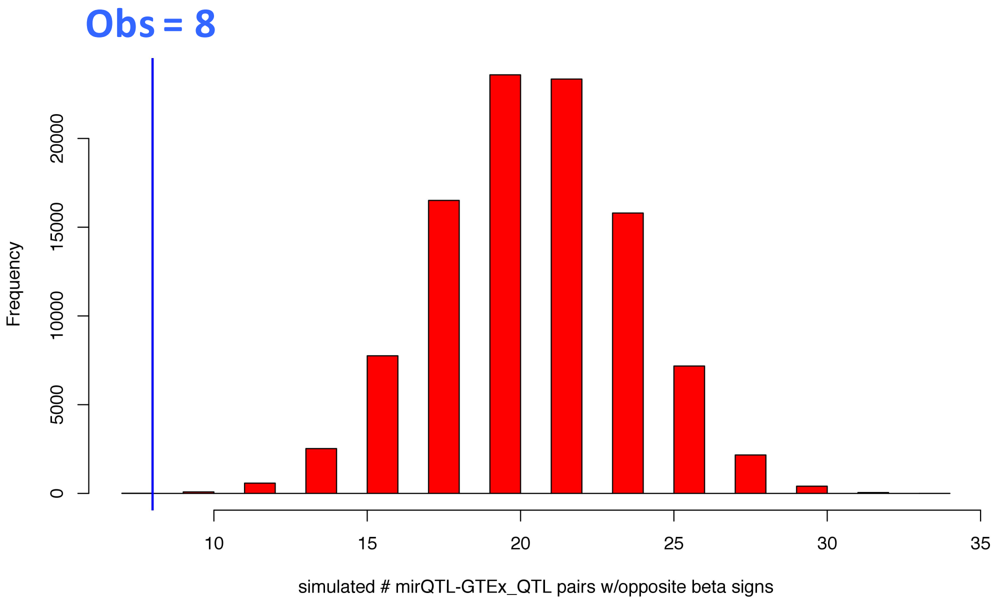


w/opposite effect sizes : 3044/ 4688

	⊕	⊖
⊕	826	2353
⊖	691	818

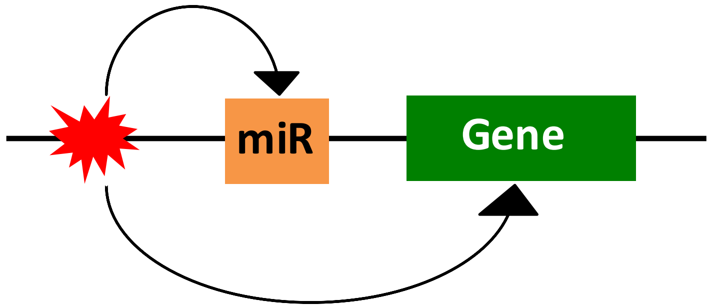
miRNAs within
the genes themselves
N = 43 pairs(< 1%)
6 miRNAs, 4 genes

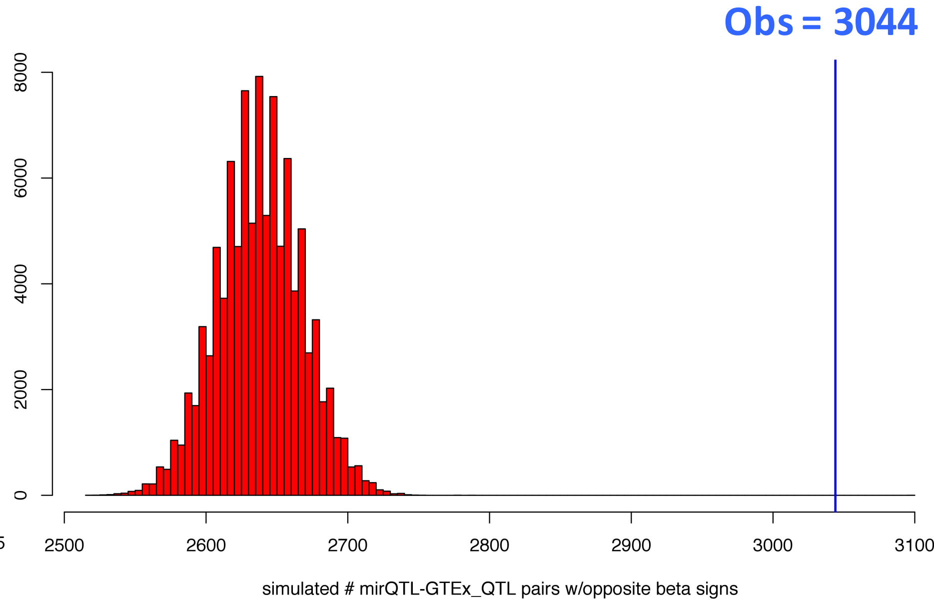




* P-val ≈ 9.2 E-16

miRNAs outside
the genes
N = 4688 pairs
39 miRNAs, 59 genes

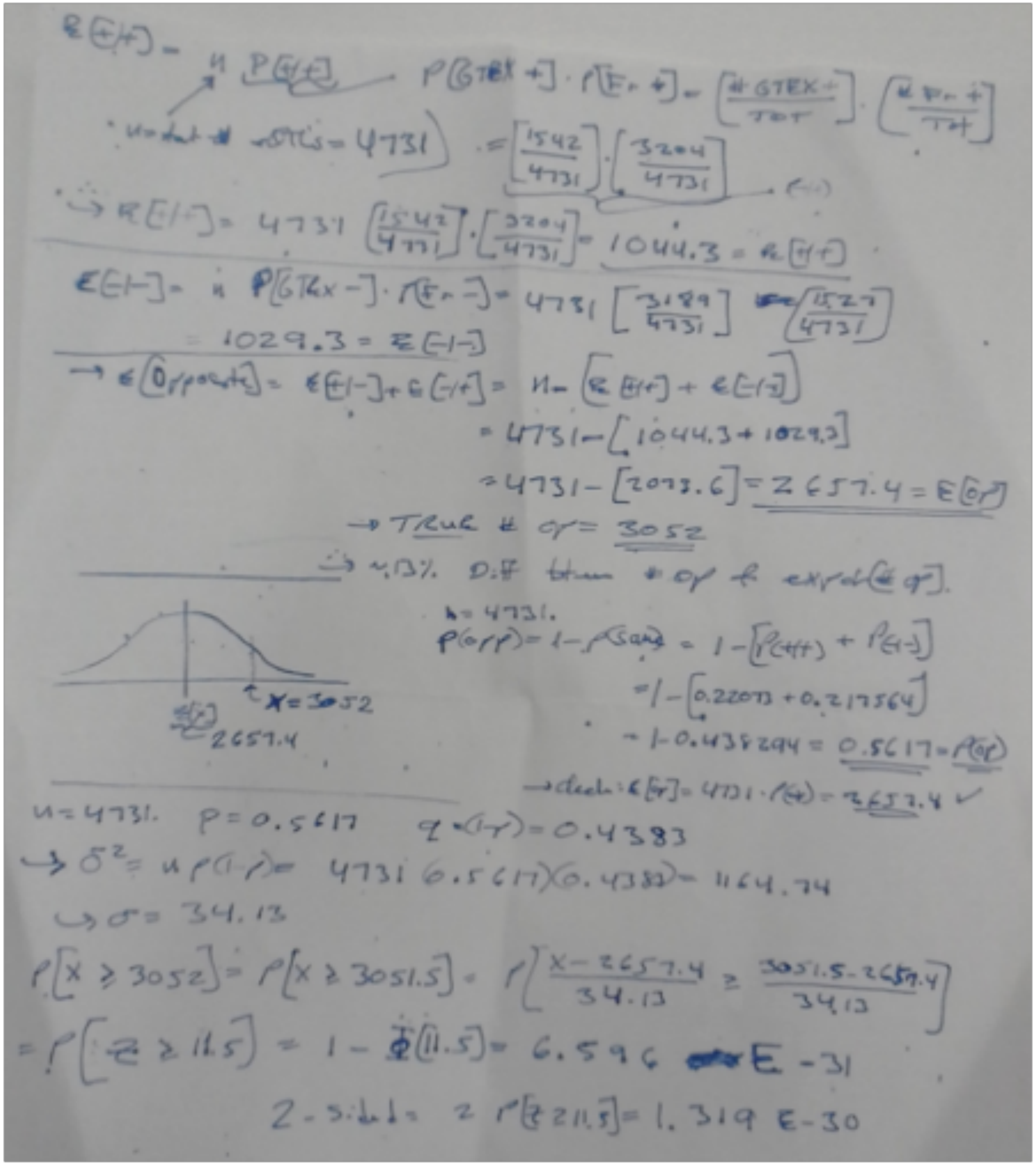




* P-val ≈ 1.3 E-30

P-val derivations in supplementary slides

p_value_derivation_all_miRNAs



$E[+/+] = n\,P[+/+]$; $P[\text{GTEX}\ +]\cdot P[\text{Fr}\ +] = \left[\frac{\#\text{GTEX}+}{\text{TOT}}\right]\cdot\left[\frac{\#\text{Fr}+}{\text{Tot}}\right]$

n = total # miRs = 4731

$= \left[\frac{1542}{4731}\right]\cdot\left[\frac{3204}{4731}\right] \rightarrow (+/+)$

$\rightarrow E[+/+] = 4731\left[\frac{1542}{4731}\right]\cdot\left[\frac{3204}{4731}\right] = 1044.3 = E[+/+]$

$E[-/-] = n\,P[\text{GTEx}\ -]\cdot P[\text{Fr}\ -] = 4731\left[\frac{3189}{4731}\right]\left[\frac{1527}{4731}\right]$

$= 1029.3 = E[-/-]$

$\rightarrow E[\text{Opposite}] = E[+/-] + E[-/+] = n - (E[+/+] + E[-/-])$

$= 4731 - [1044.3 + 1029.3]$

$= 4731 - [2073.6] = 2657.4 = E[\text{Op}]$

→ TRUE # op = 3052

→ ~13% Diff btwn # op & expected (# op).

n = 4731.

$P(\text{opp}) = 1 - P(\text{Same}) = 1 - [P(+/+) + P(-/-)]$

$= 1 - [0.22073 + 0.217564]$

$= 1 - 0.438294 = 0.5617 = P(\text{op})$

→ check: $E[\text{op}] = 4731\cdot P(\text{op}) = 2657.4$ ✓

n = 4731. $p = 0.5617$ $q = (1-p) = 0.4383$

$\rightarrow \sigma^2 = np(1-p) = 4731\,(0.5617)(0.4383) = 1164.74$

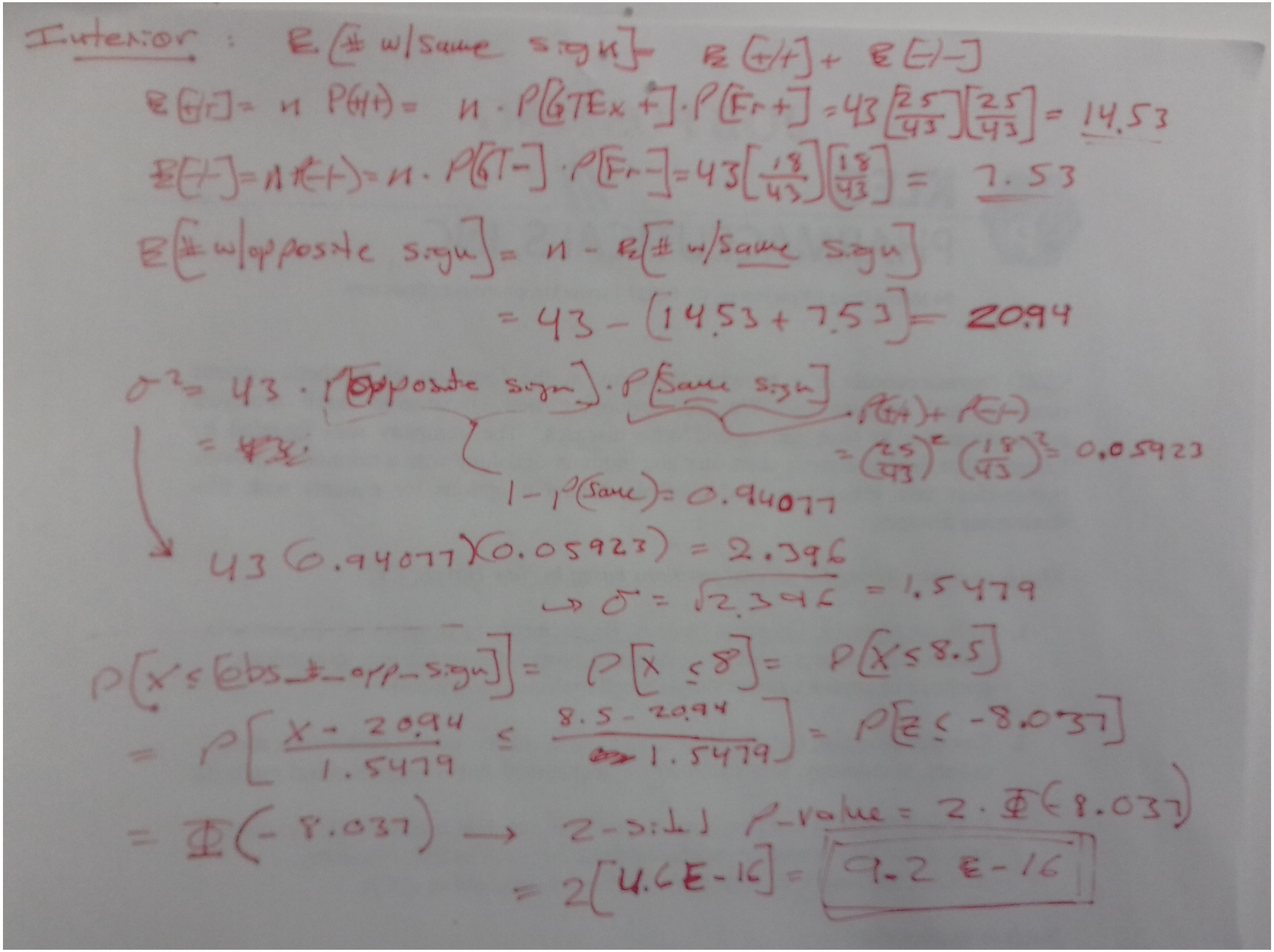
$\rightarrow \sigma = 34.13$

$P[X \geq 3052] = P[X \geq 3051.5] = P\left[\frac{X - 2657.4}{34.13} \geq \frac{3051.5 - 2657.4}{34.13}\right]$

$= P[Z \geq 11.5] = 1 - \Phi(11.5) = 6.596\ \text{E}-31$

2-sided = $2\,P[Z \geq 11.5] = 1.319\ \text{E}-30$

p_value_derivation_interior_miRNAs



Interior : $E[\#\text{ w/ same sign}] = E[+/+] + E[-/-]$

$E[+/+] = n\,P(+/+) = n \cdot P[\text{GTEx } +] \cdot P[\text{Fr } +] = 43\left[\frac{25}{43}\right]\left[\frac{25}{43}\right] = \underline{14.53}$

$E[-/-] = n\,P(-/-) = n \cdot P[\text{GT } -] \cdot P[\text{Fr } -] = 43\left[\frac{18}{43}\right]\left[\frac{18}{43}\right] = \underline{7.53}$

$E[\#\text{ w/ opposite sign}] = n - E[\#\text{ w/ same sign}]$

$= 43 - [14.53 + 7.53] = 20.94$

$\sigma^2 = 43 \cdot \underbrace{P[\text{opposite sign}]}_{1 - P(\text{Same}) = 0.94077} \cdot \underbrace{P[\text{Same sign}]}_{P(+/+) + P(-/-) = \left(\frac{25}{43}\right)^2 \left(\frac{18}{43}\right)^2 = 0.05923}$

$= \cancel{43}$

$\rightarrow 43(0.94077)(0.05923) = 2.396$

$\rightarrow \sigma = \sqrt{2.396} = 1.5479$

$P[X \le \text{obs\_\#\_opp\_sign}] = P[X \le 8] = P[X \le 8.5]$

$= P\left[\frac{X - 20.94}{1.5479} \le \frac{8.5 - 20.94}{1.5479}\right] = P[Z \le -8.037]$

$= \Phi(-8.037) \longrightarrow$ 2-sided p-value $= 2 \cdot \Phi(-8.037)$

$= 2[4.6E-16] = \boxed{9.2\ E-16}$