HWE Test for X-linked Markers

It is usual to test HWE for X-linked markers using only females.

Under HWE allele frequencies for SNPs in males and females, on the X chromosome, should be the same. Should examine the difference allele frequencies when testing for HWE.

If a sample has n_m males and n_f females, and if the males have m_A, m_B alleles of types A, B, and if females have f_{AA}, f_{AB}, f_{BB} genotypes AA, AB, BB, then the probability of the data, under HWE, is

$$\frac{n_A! n_B! n_m! n_f!}{m_A! m_B! f_{AA}! f_{AB}! f_{BB}! n_t!} 2^{f_{AB}}$$

where $n_t = n_m + 2n_f$.

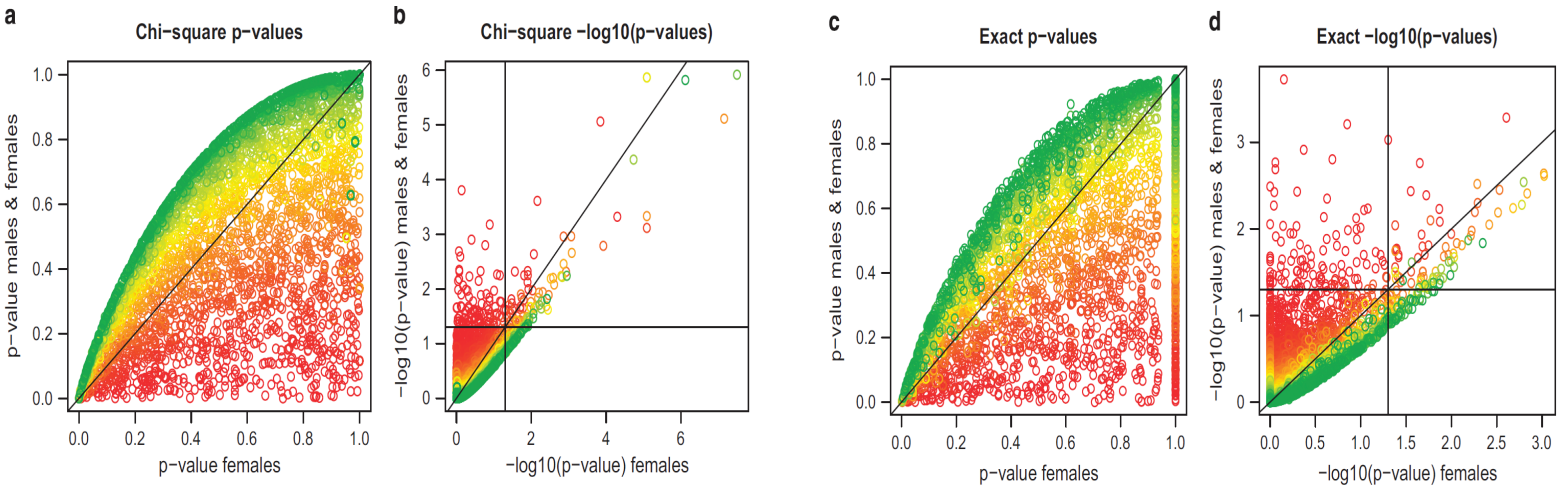
Graffelman and Weir, 2016, Heredity 116:558-568.

Example: 10 males, 10 females, 6 A alleles

If there are six A alleles in a sample that has 10 males and 10 females, there are 16 possible datasets:

Set	m_A	m_B	f_{AA}	f_{AB}	f_{BB}	Probability
1	0	10	3	0	7	0.0002
2	6	4	0	0	10	0.0004
3	0	10	0	6	4	0.0026
4	2	8	2	0	8	0.0034
5	4	6	1	0	9	0.0035
6	5	5	0	1	9	0.0085
7	0	10	2	2	6	0.0085
8	1	9	2	1	7	0.0121
9	0	10	1	4	5	0.0340
10	3	7	1	1	8	0.0364
11	4	6	0	2	8	0.0637
12	2	8	1	2	7	0.1091
13	1	9	1	3	6	0.1132
14	1	9	0	5	5	0.1358
15	3	7	0	3	7	0.1940
16	2	8	0	4	6	0.2546

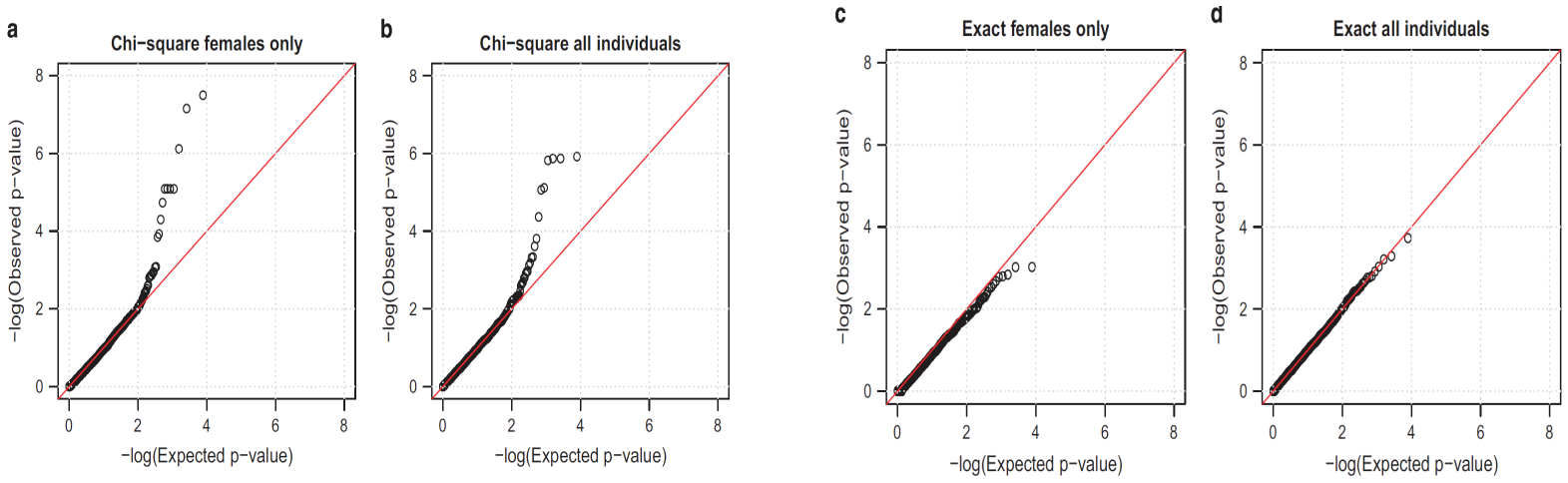
X-linked Markers: Real Data



Scatter plots of P-values in original and -log10 scale for chi-square tests (a, b) and exact tests (c, d) for HWE using females only and using both males and females for 4158 SNPs at the X chromosome of the venous thrombosis database. The horizontal and vertical black lines in (b) and (d) correspond to a significance level of 5%. Points colored according to their significance level in Fisher's test for equality of allele frequencies (range 0-1 from red to green).

Graffelman J, Weir BS. 2016. Heredity 116:558-568.

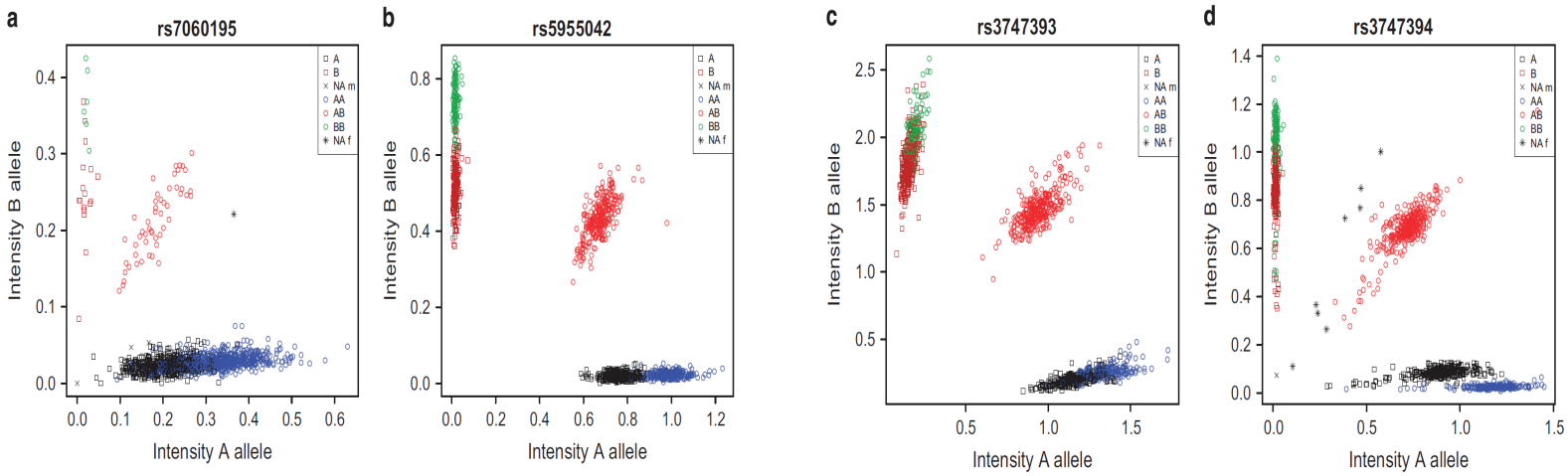
X-linked Markers: Real Data



QQ plots of - log10 transformed P-values of Chi-square and exact tests for HWE for 4158 SNPs of the venous thrombosis database. (a, c) Females only and (b, d) all individuals.

Graffelman J, Weir BS. 2016. Heredity 116:558-568.

X-linked Markers: Real Data

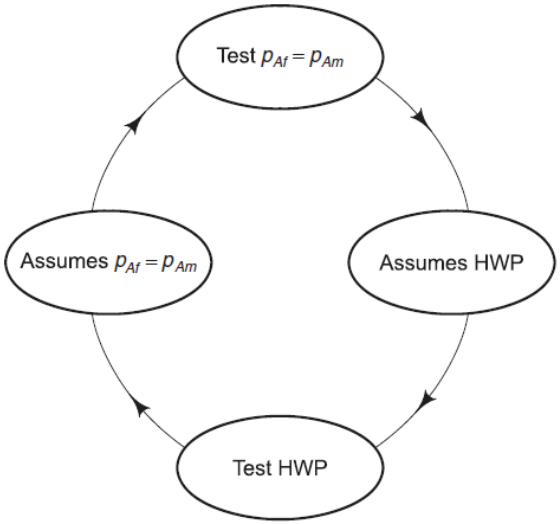


Cluster plots of allele intensities of four SNPs of the venous thrombosis database. (a and b) are significant in both the female-only (P=0.0025, P=0.0010) and all-individual test (P=0.0005, P=0.0023). (c) is non-significant in the female-only test (P=0.4261) but highly significant in the all-individual test (P=0.0012). (d) is non-significant in the female-only test (P=0.8732) and close to significant in the all-individual test (P=0.0914).

Graffelman J, Weir BS. 2016. Heredity 116:558-568.

Separate Male and Female Autosomal Counts

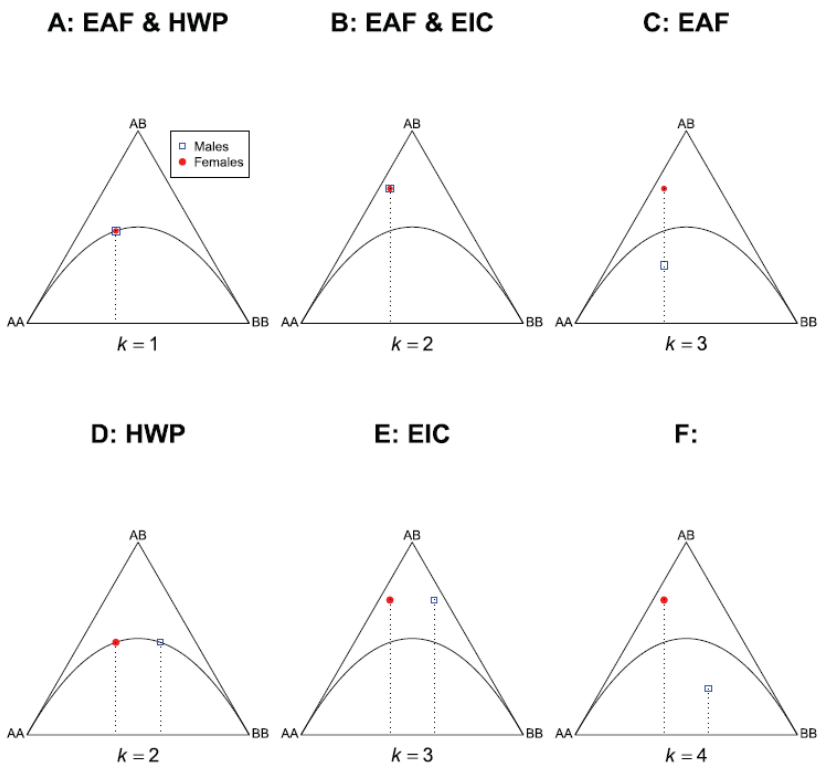
The X-linked test can be extended to autosomal markers when genotype counts are recorded separately for males and females.



Vicious testing circle: mutual dependency of a test for EAF in males and females and a test for HWP Notes: A allele frequencies in males and females are represented by p_{A_m} and p_{A_f}, respectively.

Graffeleman J, Weir BS. 2018. Genetic Epidemiology 42:24-48.

Separate M&F Counts: Scenarios



A) HWP and EAF. (B) Equality of inbreeding coefficients, EAF, and both sexes out of HWP. (C) Unequal inbreeding coefficients, both sexes out of equilibrium but with equal allele frequencies. (D) Both sexes in HWP but with different allele frequencies. (E) Each sex out of equilibrium with identical inbreeding coefficients and different allele frequencies. (F) Both sexes out of equilibrium, with different inbreeding coefficients and different allele frequencies.

Aside: Separate M&F Counts: Joint Exact Test

To test for both Equal Allele Frequencies (EAF) and Hardy-Weinberg Proportions (HWP):

$$\Pr(m_{AB}, f_{AB} | n, n_A, n_m) = \frac{n_A! n_B! n_m! n_f! 2^{m_{AB}+f_{AB}}}{m_{AA}! m_{AB}! m_{BB}! f_{AA}! f_{AB}! f_{BB}! (2n)!}$$

m_{AA}, m_{AB}, m_{BB}	genotype counts in males
f_{AA}, f_{AB}, f_{BB}	genotype counts in females
$n_m = m_{AA} + m_{AB} + m_{BB}$	number of males
$n_f = f_{AA} + f_{AB} + f_{BB}$	number of females
$n = n_m + n_f$	total sample size
$m_A = 2m_{AA} + m_{AB}, m_B = 2m_{BB} + m_{AB}$	numbers of A, B alleles in males
$f_A = 2f_{AA} + f_{AB}, f_B = 2f_{BB} + f_{AB}$	numbers of A, B alleles in females
$n_A = m_A + f_A, n_B = m_B + f_B$	total numbers of A, B alleles

Aside: Separate M&F Counts: HWP Exact Test

To test for HWP:

$$\Pr(n_{AB}|n, n_A) = \frac{n_A! n_B! n! 2^{n_{AB}}}{n_{AA}! n_{AB}! n_{BB}!}$$

n_{AA}, n_{AB}, n_{BB}	total genotype counts in males and females
$n = n_{AA} + n_{AB} + n_{BB}$	total sample size
$n_A = 2n_{AA} + n_{AB}, n_B = 2n_{BB} + n_{AB}$	total numbers of A, B alleles

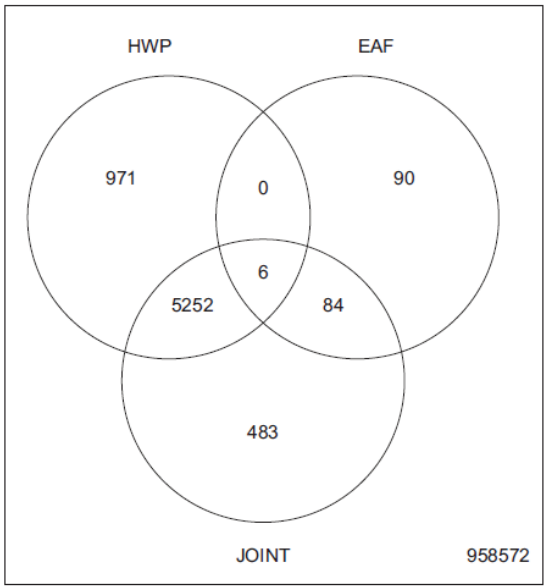
Aside: Separate M&F Counts: EAF Exact Test

To test for EAF:

$$\Pr(n_A|n, m_A) = \frac{n_A!n_B!n_m!n_f!}{m_A!m_B!f_A!f_B!}$$

m_A, m_B	numbers of A, B alleles in males
f_A, f_B	numbers of A, B alleles in females
$n_m = m_A + m_B$	total number of male alleles
$n_f = f_A + f_B$	total number of female alleles
$n_A = m_A + f_A, n_B = m_B + f_B$	total numbers of A, B alleles
$n = n_m + n_f = n_A + n_B$	total number of alleles in males and females

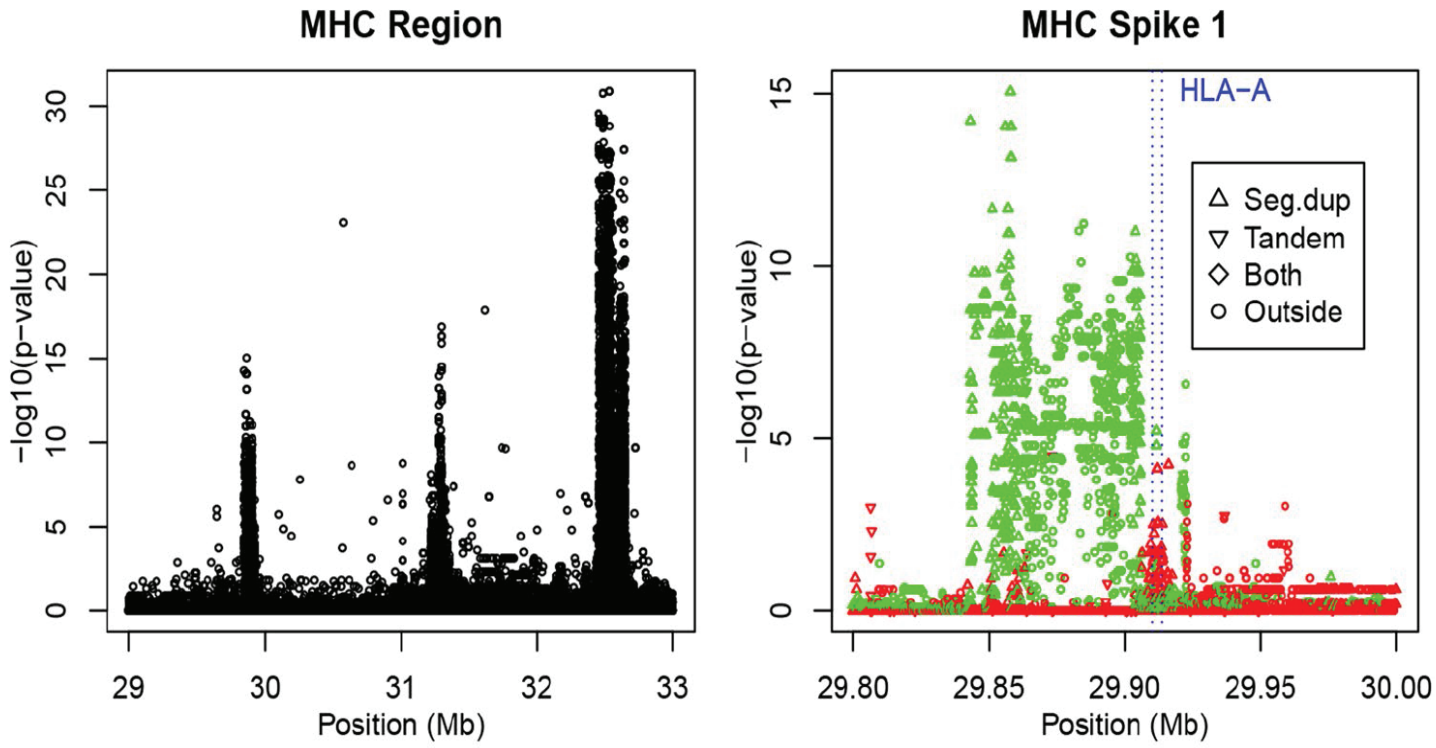
Separate M&F Counts: 1000 Genomes Result



Venn diagrams of HWP, EAF, and joint exact test results for all nonmonomorphic complete SNPs on chromosome 1 of the JPT sample Notes: Circles enclose the number of significant SNPs (at $\alpha = 0.001$) for the different tests.

Graffeleman J, Weir BS. 2018. Genetic Epidemiology 42:24-48.

MHC Region HWE Tests



Green: heterozygote deficiency. Red: heterozygote excess.

Graffelman J, Jain D, Weir B. 2017. Human Genetics 136:727-741.