GENETICS

AUTOSOMAL INHERITANCE OF SEX-LINKED MARKER

27.03.2007

While optimizing the analysis strategy for a 500,000 SNP Affymetrix array set, I found 6 autosomal SNPs that show highly significant sex-dependent allele differences: rs2809868, rs4862188, rs2880301, rs3883011, rs3883013 and rs3883014.

Table : SNP genotyping results for 6 (putative autosomal) SNPs in 270 hapmap individuals that were significant for sex differences at $p<1x10^{-20}$.

Affx ID	dbSNP ID	Chromosome	Chromosomal position	Alignment bases tested /match	Repeat	TF	Genotypes male	Genotypes female
1 SNP_A-2180412	rs2809868	1 Y	239,284,269 12,312,998	201/100% 201/98.6%	LINE L1ME e	none	AA 11 AG 127 GG 0	AA 9 AG 51 GG 68
4 SNP_A-2306557	rs4862188	4 Y#	184,730,519 219,08,325	201/100% 201/95.1%	LTR LTR12C	AML-1	CC 0 CT 138 TT 0	CC 128 CT 0 TT 0
13 SNP_A-2257575	rs2880301	13 Y	18,998,534 26,968,489	201/100% 201/97.0%	LINE L1PA15	PAX6, SP-1	CC 0 CT 137 TT 0	CC 128 CT 0 TT 0
15.1 SNP_A-1941964	rs3883011	15 Y	82,889,398 26,014,938	201/100% 201/98.6%	LTR MLT1A1	none	CC 6 CG 124 GG 0	CC 128 CG 0 GG 0
15.2 SNP_A-2088494	rs3883013	15 Y	82,889,661 26,015,201	201/100% 201/98.6%	as above	HNF-1C	TT 0 CT 142 CC 0	TT 126 CT 0 CC 0
15.3 SNP_A-4252837	rs3883014	15* Y*	82,889,733 26,015,273	201/100% 201/98.1%	as above	STAT5B Pbx1	GG 11 CG 111 CC 0	GG 127 CG 0 CC 0

Y chromosome is not the second best hit
* several more sequence matches on chr15 and Y are not shown here

Sure, there could be autosomal marker that influences male/female outcome but there is a more likely explanation: All SNPs have paralogue sequence stretches on the Y chromosome that are co-amplified during PCR. From the initial genotyping results it is most likely that only the Y chromosomal stretch is being mutated in SNP 4, 13 and 15.2.

These SNPs are perfect sex marker, as they include an autosomal control allele (in comparison to pure Y markers like SNPs in SRY). They are always unambiguous (in contrast to pure X marker where only heterocygotes are informative).

They even offer advantage to commercial STR kits of the Amelogenin/Amely gene situated

(in the Y parautosomal region) as they would not be affected by excess homologous X chromosomal material as often found in forensic situations. In addition, they might overcome some other weakness of the Amelogenin test where a second assay is usually recommended.

If you will ever see a case-control study that is highlighting any of these SNPs, you can be sure that this study had a distorted male-female ratio between case and controls.