

SNP Based Copy Number Microarrays Provide Cues to UPD, and Recessive Allele Risk Due to Inbreeding

P.R. Papenhausen, I. Gadi, H. Risheg, C. Bailey-Pridham, J. Tepperberg
LabCorp Inc., RTP, North Carolina

ABSTRACT

SNP based chromosome microarrays can provide an extremely high density whole genome analysis of copy number variation in the clinical analysis of blood from children with developmental delay. We have studied about 2,500 cases this year using both the Affymetrix 500k and 1.8 million SNP/copy number chips. The allele differentiation that the array provides allows designation of the relative distribution of homozygosity (HZ) (or loss of heterozygosity) throughout the genome, in addition to copy number. We expected that long contiguous stretches of HZ(LCSH) >15Mb in a single chromosome would correlate with UPD based on the almost 2/3 of our reported 36 cases of UPD that showed regions of both hetero and iso UPD using limited numbers of microsatellites. The added significance of HZ determinations was not fully appreciated until the following cases were noted (it is important to understand that contiguous HZ over 6 Mb is rarely observed in these studies): a 3yo showed 27 LCSH's (>10Mb) on 20 chrs, a 5yo had 8 LCSH's on 7 chrs, a 2yo had 10 LCSH on 8 chrs, a 12yo showed 12 LCSH's on 11 chrs, a 10yo had 11 LCSH's on 11 chrs and a 32yo had 9 LCSH's on 7 chrs. All cases proved to represent consanguinity with the first case a product of a brother-sister pairing. Additional probands from first and second cousin pairings showed smaller blocks with fewer chromosomes involved, as what might be expected from multi-generation recombination and dilution of consanguineous chromosomes. In all, 45 cases are presented. A separate pattern of cases showed greatly increased homozygosity, but without contiguous stretches. These cases appeared to represent geographical or ethnic isolates with limited outbreeding. It is important to determine the LCSH boundaries for levels of consanguinity and a reasonable means of reporting these findings which have implications for recessive allele risk. Patterns of homozygosity with possible threshold ranges and UPD examples are presented. Although LCSH in a single chromosome correlates with UPD, confirmation through paternal microsatellite exclusion or direct methylation specific testing is highly recommended.

INTRODUCTION

SNP based microarrays allow for the designation of regions of homozygosity through the targeting of biallelic markers at high density dispersed throughout the genome. In some individuals, long contiguous tracts of HZ can be observed, and in others shorter, but excessively numerous, runs can be found^{1,2}. The best explanation for the long stretches is identity by descent (IBD), where "autozygous" chromosome segments have been passed from a common ancestor. As these segments segregate and are cut by additional generations of recombination, they become fewer and smaller proportional to the degree of inbreeding. Those individuals showing numerous runs of shorter HZ appear to correlate with genepool restrictions based on geographical or ethnic isolation³. It is also apparently true that even outbred populations have multimegabase ancestral haplotypes that persist in various genomic regions which generally show low recombination rates. Our observations of LCSH in clinical analysis of 25000 clinical cases using the Affymetrix 6.0 genotyping chips provided confirmation of the degree of inbreeding that the microarray provides by using a simple threshold value for establishing the coefficient. In addition, we confirmed our observations of 15 years of UPD studies which suggested a UPD "signature" could be provided by high resolution SNP microarray analysis that does not require a parental trio analysis.

METHODS

The Affymetrix version 6.0 genechip (1.8 million SNP and CN targets) with GTC 2.1 software was used for most of these studies (version 5.0 in 10%). Our estimation of identity by descent was based on establishing a threshold for probable autozygosity based LCSH that was clearly above background levels. Since autosomal HZ runs rarely stretched above 7Mb in our control group of 200 consecutive patients (largely referred due to developmental delay), we assigned a 10 Mb cut-off for this simple summation [IBD = Σ LCSH(>10Mb)].

RESULTS

The normal distribution of the longest stretch of homozygosity is presented in Figure 1a. Putative UPD cases which demonstrated the qualifying signature of a single stretch of contiguous HZ over 14Mb are in Table 1 and Fig 2. Additional runs of HZ (above 8Mb) associated with inbreeding, must not be present. Examples of the UPD associated LCSH are in Figure 2. Three of the molecularly confirmed UPD cases were prospectively detected and the other six were retrospectively analyzed after molecular methylation or microsatellite analysis. The stratification of the autozygosity cases are presented in Table 2 and Figures 3. Figure 4 is a representative example of a reduced gene pool.

Table 1. Confirmed and Unconfirmed UPD Cases with Single Chromosome LCSH >13.5Mb

CASE NUMBER	AGE	CHROMOSOME AND ORIGIN	LCSH INTERVAL (Mb)	LCSH LENGTH (Mb)	ASSOCIATED SYNDROME
CONFIRMED UPD CASES					
1	5yo	X.mat	COMPLETE ISO X	155.8	XX-DOUBLE FULL MUTATION
2	28yo	X.pat	COMPLETE ISO X	155.0	PREMUTATION
3	1 week	6.pat	ISO 6	176.0	TNDM
4	AF	7.mat	80.6-108.2	29.9	TRISOMYAR RISK
5	AF	12.mat	118.7-132.2	13.5	TRISOMYAR RISK
6	1.2yo	15.mat	40.2-76.1	35.9	PRADER-WELLI
7	N/A	16.mat	34.9-46.1	22.2	PRADER-WELLI
8	N/A	16.mat	80.9-91.8	22.9	PRADER-WELLI
9	4yo	19.mat	18.3-31.9 69.4-89.0 92.8-107.9 95.9-109.2	42.2 19.6	PRADER-WELLI
UNCONFIRMED UPD CASES					
10	5yo	1	70.8-84.0	24.0	TRISOMYAR RISK
11	16.3yo	2	107.0-128.3	29.5	TRISOMYAR RISK
12	5yo	3	112.2-128.1	23.5	TRISOMYAR RISK
13	5yo	4	104.6-127.4	52.8	TRISOMYAR RISK
14	4yo	5	31.7-62.8	31.1	TRISOMYAR RISK
15	4yo	6	116.4-142.9	29.5	TRISOMYAR RISK
16	4yo	12	125.2-129.2	27.8	TRISOMYAR RISK
17	4yo	13	45.1-64.4	19.3	TRISOMYAR RISK
18	15yo	13	31.6-51.9	20.3	TRISOMYAR RISK
19	4yo	13	58.7-74.8	21.1	TRISOMYAR RISK
20	3yo	16	51.3-77.7	15.9	TRISOMYAR RISK
21	5yo	11	23.8-42.9	19.7	TRISOMYAR RISK
22	4yo	21	20.5-35.4	14.9	TRISOMYAR RISK

Table 2. Total and Number of LCSH Intervals >10Mb

Relation	Total LOH (CN)	Mean (SD)	Number of Chromosomes with Contiguous LOH (Chromosomes)	Mean (SD)
Brother/Sister	159.3	159.3	131(1)	131(1)
Brother/Sister	488.5	488.5	131(4)	131(4)
Brother/Sister	529.9	529.9	131(4)	131(4)
Brother/Sister	656.5	656.5	131(2)	131(2)
Father/Daughter	588.1	588.1	131(2)	131(2)
1 st Cousin	133.3	133.3	131(1)	131(1)
1 st Cousin	227.8	227.8	131(1)	131(1)
1 st Cousin	258.9	258.9	131(1)	131(1)
1 st Cousin	308.9	308.9	131(1)	131(1)
1 st Cousin	351.3	351.3	131(1)	131(1)
1 st Cousin	393.8	393.8	131(1)	131(1)
1 st Cousin	436.3	436.3	131(1)	131(1)
1 st Cousin	478.8	478.8	131(1)	131(1)
1 st Cousin	521.3	521.3	131(1)	131(1)
1 st Cousin	563.8	563.8	131(1)	131(1)
1 st Cousin	606.3	606.3	131(1)	131(1)
1 st Cousin	648.8	648.8	131(1)	131(1)
1 st Cousin	691.3	691.3	131(1)	131(1)
1 st Cousin	733.8	733.8	131(1)	131(1)
1 st Cousin	776.3	776.3	131(1)	131(1)
1 st Cousin	818.8	818.8	131(1)	131(1)
1 st Cousin	861.3	861.3	131(1)	131(1)
1 st Cousin	903.8	903.8	131(1)	131(1)
1 st Cousin	946.3	946.3	131(1)	131(1)
1 st Cousin	988.8	988.8	131(1)	131(1)
1 st Cousin	1031.3	1031.3	131(1)	131(1)
1 st Cousin	1073.8	1073.8	131(1)	131(1)
1 st Cousin	1116.3	1116.3	131(1)	131(1)
1 st Cousin	1158.8	1158.8	131(1)	131(1)
1 st Cousin	1201.3	1201.3	131(1)	131(1)
1 st Cousin	1243.8	1243.8	131(1)	131(1)
1 st Cousin	1286.3	1286.3	131(1)	131(1)
1 st Cousin	1328.8	1328.8	131(1)	131(1)
1 st Cousin	1371.3	1371.3	131(1)	131(1)
1 st Cousin	1413.8	1413.8	131(1)	131(1)
1 st Cousin	1456.3	1456.3	131(1)	131(1)
1 st Cousin	1498.8	1498.8	131(1)	131(1)
1 st Cousin	1541.3	1541.3	131(1)	131(1)
1 st Cousin	1583.8	1583.8	131(1)	131(1)
1 st Cousin	1626.3	1626.3	131(1)	131(1)
1 st Cousin	1668.8	1668.8	131(1)	131(1)
1 st Cousin	1711.3	1711.3	131(1)	131(1)
1 st Cousin	1753.8	1753.8	131(1)	131(1)
1 st Cousin	1796.3	1796.3	131(1)	131(1)
1 st Cousin	1838.8	1838.8	131(1)	131(1)
1 st Cousin	1881.3	1881.3	131(1)	131(1)
1 st Cousin	1923.8	1923.8	131(1)	131(1)
1 st Cousin	1966.3	1966.3	131(1)	131(1)
1 st Cousin	2008.8	2008.8	131(1)	131(1)
1 st Cousin	2051.3	2051.3	131(1)	131(1)
1 st Cousin	2093.8	2093.8	131(1)	131(1)
1 st Cousin	2136.3	2136.3	131(1)	131(1)
1 st Cousin	2178.8	2178.8	131(1)	131(1)
1 st Cousin	2221.3	2221.3	131(1)	131(1)
1 st Cousin	2263.8	2263.8	131(1)	131(1)
1 st Cousin	2306.3	2306.3	131(1)	131(1)
1 st Cousin	2348.8	2348.8	131(1)	131(1)
1 st Cousin	2391.3	2391.3	131(1)	131(1)
1 st Cousin	2433.8	2433.8	131(1)	131(1)
1 st Cousin	2476.3	2476.3	131(1)	131(1)
1 st Cousin	2518.8	2518.8	131(1)	131(1)
1 st Cousin	2561.3	2561.3	131(1)	131(1)
1 st Cousin	2603.8	2603.8	131(1)	131(1)
1 st Cousin	2646.3	2646.3	131(1)	131(1)
1 st Cousin	2688.8	2688.8	131(1)	131(1)
1 st Cousin	2731.3	2731.3	131(1)	131(1)
1 st Cousin	2773.8	2773.8	131(1)	131(1)
1 st Cousin	2816.3	2816.3	131(1)	131(1)
1 st Cousin	2858.8	2858.8	131(1)	131(1)
1 st Cousin	2901.3	2901.3	131(1)	131(1)
1 st Cousin	2943.8	2943.8	131(1)	131(1)
1 st Cousin	2986.3	2986.3	131(1)	131(1)
1 st Cousin	3028.8	3028.8	131(1)	131(1)
1 st Cousin	3071.3	3071.3	131(1)	131(1)
1 st Cousin	3113.8	3113.8	131(1)	131(1)
1 st Cousin	3156.3	3156.3	131(1)	131(1)
1 st Cousin	3198.8	3198.8	131(1)	131(1)
1 st Cousin	3241.3	3241.3	131(1)	131(1)
1 st Cousin	3283.8	3283.8	131(1)	131(1)
1 st Cousin	3326.3	3326.3	131(1)	131(1)
1 st Cousin	3368.8	3368.8	131(1)	131(1)
1 st Cousin	3411.3	3411.3	131(1)	131(1)
1 st Cousin	3453.8	3453.8	131(1)	131(1)
1 st Cousin	3496.3	3496.3	131(1)	131(1)
1 st Cousin	3538.8	3538.8	131(1)	131(1)
1 st Cousin	3581.3	3581.3	131(1)	131(1)
1 st Cousin	3623.8	3623.8	131(1)	131(1)
1 st Cousin	3666.3	3666.3	131(1)	131(1)
1 st Cousin	3708.8	3708.8	131(1)	131(1)
1 st Cousin	3751.3	3751.3	131(1)	131(1)
1 st Cousin	3793.8	3793.8	131(1)	131(1)
1 st Cousin	3836.3	3836.3	131(1)	131(1)
1 st Cousin	3878.8	3878.8	131(1)	131(1)
1 st Cousin	3921.3	3921.3	131(1)	131(1)
1 st Cousin	3963.8	3963.8	131(1)	131(1)
1 st Cousin	4006.3	4006.3	131(1)	131(1)
1 st Cousin	4048.8	4048.8	131(1)	131(1)
1 st Cousin	4091.3	4091.3	131(1)	131(1)
1 st Cousin	4133.8	4133.8	131(1)	131(1)
1 st Cousin	4176.3	4176.3	131(1)	131(1)
1 st Cousin	4218.8	4218.8	131(1)	131(1)
1 st Cousin	4261.3	4261.3	131(1)	131(1)
1 st Cousin	4303.8	4303.8	131(1)	131(1)
1 st Cousin	4346.3	4346.3	131(1)	131(1)
1 st Cousin	4388.8	4388.8	131(1)	131(1)
1 st Cousin	4431.3	4431.3	131(1)	131(1)
1 st Cousin	4473.8	4473.8	131(1)	131(1)
1 st Cousin	4516.3	4516.3	131(1)	131(1)
1 st Cousin	4558.8	4558.8	131(1)	131(1)
1 st Cousin	4601.3	4601.3	131(1)	131(1)
1 st Cousin	4643.8	4643.8	131(1)	131(1)
1 st Cousin	4686.3	4686.3	131(1)	131(1)
1 st Cousin	4728.8	4728.8	131(1)	131(1)
1 st Cousin	4771.3	4771.3	131(1)	131(1)
1 st Cousin	4813.8	4813.8	131(1)	131(1)
1 st Cousin	4856.3	4856.3	131(1)	131(1)
1 st Cousin	4898.8	4898.8	131(1)	131(1)
1 st Cousin	4941.3	4941.3	131(1)	131(1)
1 st Cousin	4983.8	4983.8	131(1)	131(1)
1 st Cousin	5026.3	5026.3	131(1)	131(1)
1 st Cousin	5068.8	5068.8	131(1)	131(1)
1 st Cousin	5111.3	5111.3	131(1)	131(1)
1 st Cousin	5153.8	5153.8	131(1)	131(1)
1 st Cousin	5196.3	5196.3	131(1)	131(1)
1 st Cousin	5238.8	5238.8	131(1)	131(1)
1 st Cousin	5281.3	5281.3	131(1)	131(1)
1 st Cousin	5323.8	5323.8	131(1)	131(1)
1 st Cousin	5366.3	5366.3	131(1)	131(1)
1 st Cousin	5408.8	5408.8	131(1)	131(1)
1 st Cousin	5451.3	5451.3	131(1)	131(1)
1 st Cousin	5493.8	5493.8		