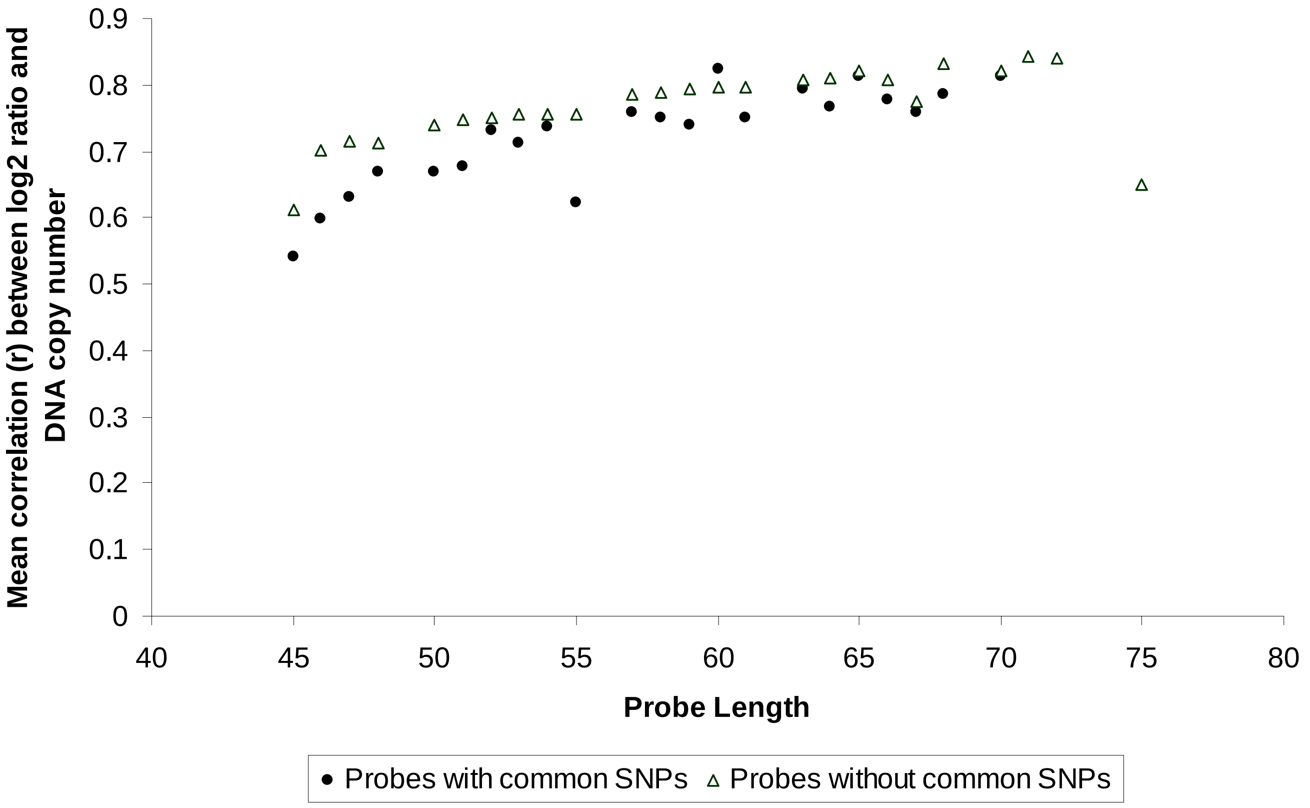
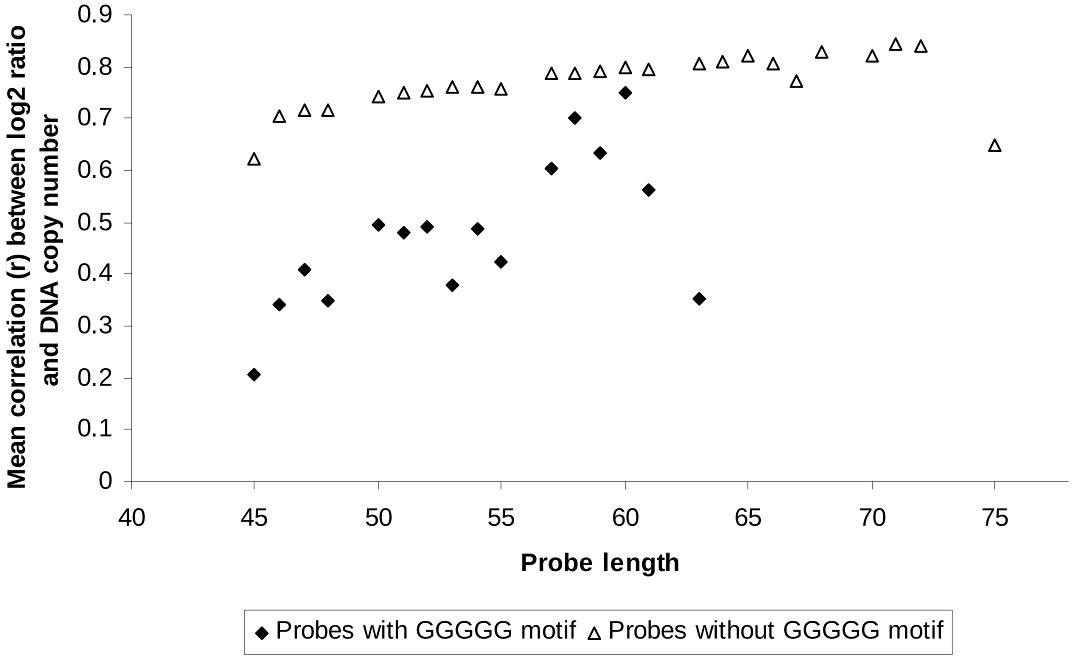
Legends to Supplementary Figures

Supplementary Figure 1. Increased probe length attenuates the effect of common SNPs on probe performance. Plot of average correlation coefficient (r) versus probe length for probes with and without common SNPs. There are significant differences between mean correlation coefficient at probe lengths 45, 46, 50, 51, and 55 bp (2 sample t-test with Bonferroni correction). At probe lengths greater than 55 bp, significant differences in mean correlation coefficient no longer exist between probes with common SNPs and those without suggesting that the effect of common SNPs on probe performance is attenuated by increased probe length.

Supplementary Figure 2. Presence of *GGGGG* motif in probe sequence continues to affect probe performance after stratification by probe length. Differences in means are significant for lengths 45-47, 50-55 and 57 (2 sample t-test with Bonferroni correction). All other probe lengths are likely not significant due to small sample size ($N < 29$).



Supplementary Figure 1



Supplementary Figure 2

SUPPLEMENTARY TABLES

Supplementary Table 1. Summary of hybridization results in seven patients with known copy number of 15q12.2-q13.1

	UC4, Angelman syndrome deletion	**NA19240, Normal control**	**UC1, chr15 duplication**	**UC7, Marker chr15**	**UC5, chr15 triplication**	**UC6, Accessory duplicated chr15**	**UC2, Accessory triplicated chr15**
Copy number of variant region	1	2	3	3	4	4	6
Mean $\log_2$ ratio of 91069 unique probes in variant region	-0.363	0.048	0.098	0.168	0.280	0.335	0.528
Mean $\log_2$ ratio of top 10% of probes by r in variant region	-0.468	-0.064	0.108	0.164	0.405	0.414	0.833
St Dev of 91069 probes	0.300	0.235	0.199	0.273	0.290	0.289	0.322
Theoretically predicted $\log_2$ ratio	-1	0	0.585	0.585	1	1	1.585

Supplementary Table 2. Repeat content of probes in 15q12.2-q13.1.

Repeat Name	Number of probes overlapping repeat in upper quartile	Second quartile - enrichment vs upper quartile	Third quartile - enrichment vs upper quartile	Fourth quartile - enrichment vs upper quartile
ERVK	25	1.00	2.64	7.16
Alu	165	1.52	2.72	5.36
Low complexity	32	1.22	1.06	2.06
ERV1	1437	1.22	1.56	2.04
SINE	712	1.08	1.33	1.78
LTR	3526	1.14	1.34	1.61
MaLR	976	1.16	1.41	1.55
Simple	100	1.87	1.65	1.50
DNA	656	1.09	1.11	1.22
L1	2973	1.10	1.17	1.10
MER1	275	0.97	0.90	1.10
LINE	3749	1.08	1.09	0.98
ERVL	1088	1.00	1.01	0.95
MIR	547	0.95	0.90	0.70
L2	720	0.97	0.77	0.51

Overlap of probes with repeats was assessed from the map position of each probe in hg17 using RepeatMasker. Repeats are ranked according to their relative enrichment in the lower versus upper quartile.

Supplementary Table 3. Correlation between predictors of probe performance.

	Contains Common SNPs	Contains repeats	Contains GGGGG motif	GC content	Probe Length
Contains common SNPs					
Contains repeats	NS				
Contains GGGGG motif	NS	NS			
GC content	-	-	+*		
Probe length	+	+	-*	-*	

NS, not statistically significant. +, positive correlation. -, negative correlation. * indicates possible correlations that are not independent.

Supplementary Table 4. Repeat content of probes in 17q12.

Repeat Name	Number of probes overlapping repeat in upper quartile	Second quartile - enrichment vs upper quartile	Third quartile - enrichment vs upper quartile	Fourth quartile - enrichment vs upper quartile
ERVK	0	0	0	0
Alu	78	1.71	3.00	6.06
Low complexity	47	0.77	0.70	0.49
ERV1	39	1.90	2.95	8.54
SINE	755	1.08	1.23	1.35
LTR	199	1.31	1.48	2.98
MaLR	146	1.16	1.09	1.53
Simple	159	0.43	0.35	0.27
DNA	147	1.16	1.50	2.20
L1	337	0.96	0.99	0.75
MER1	83	1.23	1.76	2.18
LINE	1002	0.87	0.79	0.52
ERVL	14	1.21	1.43	2.71
MIR	677	1.01	1.03	0.80
L2	603	0.80	0.66	0.39

Overlap of probes with repeats was assessed from the map position of each probe in hg17 using RepeatMasker. Repeats are ranked according to the observed relative enrichment in the lower versus upper quartile for the array designed against 15q12.2-q13.1, as shown in Supplementary Table 2.