The American Journal of Human Genetics, Volume *98*

Supplemental Information

Extensive Hidden Genomic Mosaicism Revealed in Normal Tissue

Selina Vattathil and Paul Scheet

Supplemental Figures

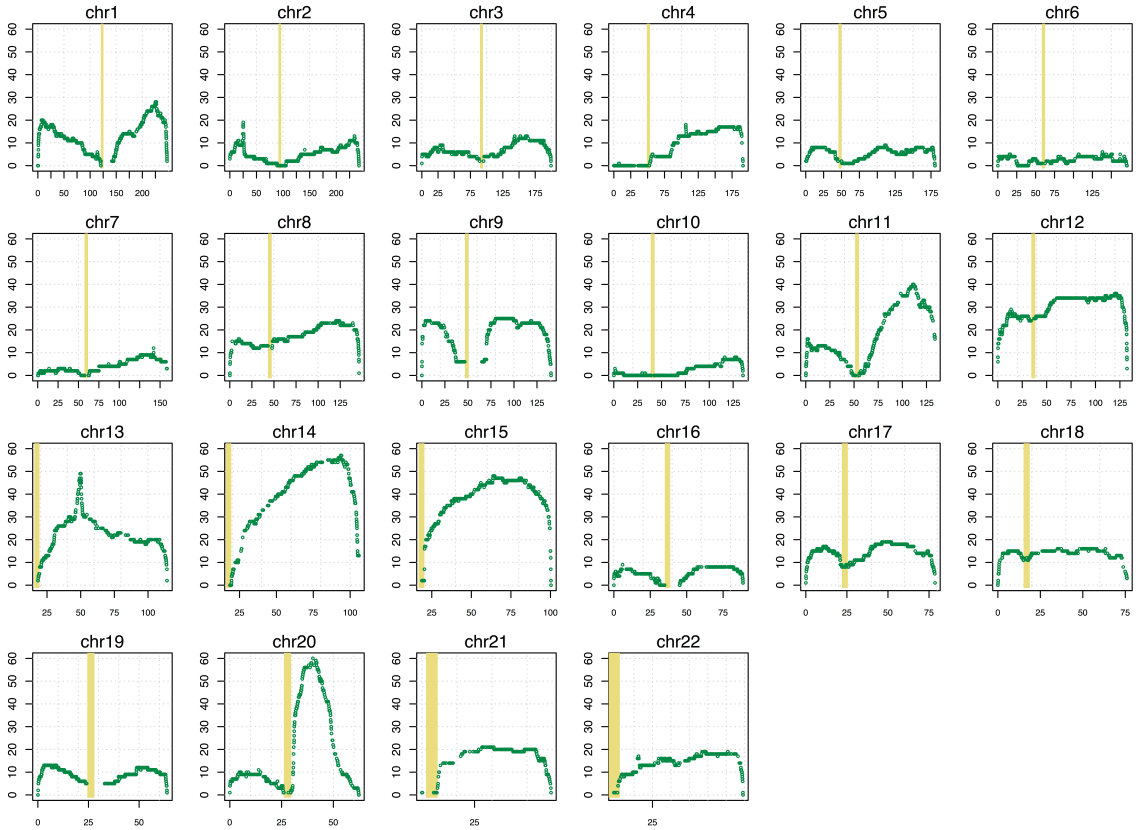

Figure S1. Number of overlapping SCNAs per gene. SCNAs discovered by hapLOH in 26,927 blood samples were considered for this tabulation. Gold bars indicate the locations of centromere regions.

Figure S2. SCNAs by chromosome. Only those events from blood samples that were unique to hapLOH, not discovered by Laurie *et al*[1]. Concordance between hapLOH and Laurie *et al.* calls was determined as described in the main text. The call set includes 698 calls. The red shading on each ideogram indicates the range of the plotted mutations. SCNAs are plotted as horizontal bars, colored by inferred copy-number: red—loss, green—CNLOH, blue—gain, gray—undetermined.

Supplemental Tables

Study name	dbGaP accession	Illumina array	DNA sources
Study of Addiction: Genetics and Environment (SAGE)	000092.v1.p1	Human1M	blood
High Density SNP Association Analysis of Melanoma	000187.v1.p1	Omni1-Quad	blood
A Genome Wide Scan of Lung Cancer and Smoking	000093.v2.p2	HumanHap550	blood
Genome-Wide Association Studies of Prematurity and its Complications	000103.v1.p1	660W-Quad	blood
The Primary Open-Angle Glaucoma Genes and Environment (GLAUGEN) Study	000308.v1.p1	660W-Quad	blood buccal
A Multi-ethnic Genome-wide scan of Prostate Cancer, with Japanese and Latino substudies	000306.v3.p1	Human1M/ 660-Quad	blood
International Consortium to Identify Genes and Interactions Controlling Oral Clefts	000094.v1.p1	610-Quad	blood
Genome-Wide Association Study of Venous Thromboembolism	000289.v2.p1	660W-Quad	blood
Genome-Wide associations of Lung Health Study (LHS)	000335.v2.p2	660W	blood

Table S1. Study names and accession numbers.

chromosome	observed peak overlap count	number of genes with peak count	gene list (if <5 genes)
20	60	1	PTPRT
14	57	17	
13	49	2	DLEU1, DLEU7
15	48	19	
11	40	28	
12	36	34	
1	28	30	
9	25	155	
8	24	16	
21	21	22	
2	19	1	DNMT3A
17	19	23	
22	19	13	
4	18	2	PPA2, ARHGEF38
18	16	52	
3	13	32	
19	13	154	
7	12	1	MTRNR2L6
5	9	13	
16	9	1	RBFOX1
10	8	20	
6	6	3	LOC100130476, TNFAIP3, PERP

Table S3. Peak per-gene SCNA overlap count per chromosome. For each gene, we counted the number of overlapping SCNAs from 26,927 blood samples. The table reports the max count per chromosome ('peak overlap count') and the number of genes with the peak count.

		Laurie *et al.* copy number		
		gain	CNLOH	loss
hapLOH copy number	gain	32	0	0
	CNLOH	0	26	0
	loss	0	0	99
	undetermined	**4**	**75**	**20**

Table S4. Copy number concordance for events with 80% reciprocal overlap. If we require 80% reciprocal overlap, 256 unique hapLOH events and 256 unique Laurie *et al.* events are called as concordant (one-to-one match). However, since we have observed that some events are being split into multiple calls in one analysis, this requirement causes some calls to be deemed discordant when they truly overlap a call in the other analysis. The copy number state matches for the 256 events with minimum 80% reciprocal overlap are presented above. For all events to which we assign a copy number, our assignment matches the classification in Laurie *et al.* We are conservative in making copy number assignments compared to Laurie *et al*.; notably, we do not assign a copy number to the majority of the events that Laurie *et al*. classifies as CNLOH. We note that they used a more sophisticated procedure for determining copy number, so we think these results reflect the conservative nature of our approach, not necessarily erroneous classifications by Laurie *et al.*

Supplemental References

1. Laurie, C.C., Laurie, C.A., Rice, K., Doheny, K.F., Zelnick, L.R., McHugh, C.P., Ling, H., Hetrick, K.N., Pugh, E.W., Amos, C., et al. (2012). Detectable clonal mosaicism from birth to old age and its relationship to cancer. Nature Genetics 44, 642-U658.