SUPPLEMENTAL TABLES

Supplemental Table 1: Summary of studies comparing genome-wide variant-detection accuracy from different DNA sources.

Study	Organism	Variant-detection method	Blood	Saliva	Buccal	SNPs	SNVs	Indels	CNVs	Bacterial DNA quantification	Eukaryotic DNA enrichment	Conclusion
Feigelson et al. 2007	Human	CMA	Y	N	Y	Y	N	N	N	N	N	A
Woo et al. 2007	Human	CMA	Y	N	Y	Y	N	N	N	N	N	A
Bahlo et al. 2010	Human	CMA	Y	Y	N	Y	N	N	N	N	N	B
Loomis et al. 2010	Human	CMA	Y	N	Y	Y	N	N	N	N	N	A
Yokoyama et al. 2010	Dog	CMA	Y	Y	N	Y	N	N	Y	N	N	B
Fabre et al. 2011	Human	CMA	Y	Y	N	Y	N	N	Y	N	N	C
Rincon et al. 2011	Dog	CMA	Y	N	Y	Y	N	N	Y	N	N	A
Abraham et al. 2012	Human	CMA	Y	Y	N	Y	N	N	N	N	N	B
Erickson et al. 2012	Human	CMA	Y	N	Y	Y	N	N	Y	N	N	A
Hu et al. 2012	Human	CMA	Y	Y	N	Y	N	N	N	N	N	C
Pennell et al. 2013	Human	CMA	Y	Y	Y	Y	N	N	N	N	N	D
Gudiseva et al. 2016	Human	CMA	Y	Y	N	Y	N	N	N	N	N	B
Reiner et al. 2017	Human	CMA	Y	Y	N	Y	N	N	Y	Y	N	B
Bruinsma et al. 2018	Human	CMA	Y	Y	N	Y	N	N	N	N	N	B
Wall et al. 2014	Human	WES/WGS	Y	Y	N	Y	Y	N	N	N	N	B
This study	Human	WGS	Y	Y	Y	Y	Y	Y	Y	Y	Y	-

Columns containing “Y” (yes) or “N” (no) indicate whether a study examined DNA from blood, saliva, or buccal samples; whether it examined the genotyping of SNPs, SNVs, indels, or CNVs; whether bacterial DNA was quantified; and whether eukaryotic DNA enrichment was examined. Conclusions are coded as follows: A, buccal is as good as blood; B, saliva is as good as blood; C, blood is better than saliva; D, blood and saliva are better than buccal. CMA, chromosomal microarray analysis; WES, whole-exome sequencing; WGS, whole-genome sequencing.

Supplemental Table 2: Age of participants when samples were collected.

Individual	**Age at collection of blood sample (years)**	**Age at collection of saliva and buccal samples (years)**
PGPC-0002	38	43
PGPC-0005	43	48
PGPC-0006	58	62
PGPC-0050	29	30

Supplemental Table 3: Mean and standard deviation of the values for each sample type and sequencing metric, calculated over the four study participants.

Metric	Blood	Non-enriched saliva	Enriched saliva	Non-enriched buccal	Enriched buccal
Percentage of reads aligned	99.8 (0.1)	85.3 (10.7)	97.2 (2.3)	98.4 (0.7)	99.6 (0.2)
Percentage of alignments <50 bp	0.1 (0.0)	4.8 (4.3)	0.6 (0.5)	0.4 (0.2)	0.1 (0.0)
Mean mapping quality	55.8 (0.2)	53.8 (1.7)	54.8 (0.9)	55.0 (0.3)	54.9 (0.1)
Median insert size	337.0 (10.2)	307.0 (23.4)	315.0 (8.1)	312.8 (9.5)	313.0 (4.5)
Percentage of reads that are duplicates	10.3 (0.7)	13.2 (3.6)	14.0 (3.1)	12.6 (0.9)	14.5 (3.6)
*Mean sequencing depth (genome-wide)	27.8 (0.2)	21.5 (4.2)	24.8 (1.7)	25.4 (0.6)	25.8 (1.2)
*Mean sequencing depth (mitochondrial genome only)	1746.6 (260.4)	2325.5 (640.9)	335.7 (345.0)	6726.0 (305.7)	916.6 (487.0)
Inter-quartile range of sequencing depth	10.0 (0.8)	8.5 (1.0)	10.8 (2.9)	10.0 (2.0)	15.0 (2.2)
*Percentage of positions sequenced to >20x depth	74.5 (11.5)	65.7 (12.4)	71.7 (6.4)	74.4 (3.0)	68.3 (3.7)
*Percentage of positions sequenced to >30x depth	55.6 (8.9)	25.5 (21.6)	44.2 (9.2)	44.8 (7.0)	45.5 (6.1)
*Percentage of positions sequenced to >40x depth	23.4 (17.7)	3.3 (4.2)	10.4 (3.0)	9.0 (2.8)	19.7 (3.5)

Higher values for the inter-quartile range of sequencing depth indicate lower read-depth uniformity. Metrics prefixed with an asterisk were corrected for the total number of reads in a given sample.

Supplemental Table 4: Statistically-significant differences in the distributions of sequencing metrics between blood samples and non-enriched saliva or buccal samples, and between enriched saliva or buccal samples and the corresponding non-enriched samples.

Metric	Blood versus non-enriched saliva	Blood versus non-enriched buccal	Enriched saliva versus non-enriched saliva	Enriched buccal versus non-enriched buccal
Percentage of reads aligned	0.000	0.005	0.062	0.062
Percentage of alignments <50 bp	0.000	0.005	0.062	0.062
Mean mapping quality	—	—	0.125	0.853
Median insert size	—	—	0.465	1.0
Percentage of reads that are duplicates	—	—	0.465	0.125
*Mean sequencing depth (genome-wide)	0.003	0.015	0.062	0.465
*Mean sequencing depth (mitochondrial genome only)	0.228	0.003	0.062	0.062
*Percentage of positions sequenced to >20x depth	—	—	0.125	0.062
*Percentage of positions sequenced to >30x depth	—	—	0.062	0.465
*Percentage of positions sequenced to >40x depth	—	—	0.062	0.062

For blood, non-enriched saliva, and non-enriched buccal samples, a Friedman repeated-measures test was first performed to determine whether there were significant differences among these three sample types. Metrics with Friedman P-value >0.05 are indicated with "—". For metrics with Friedman P-value ≤ 0.05, the P-value from the post-hoc Conover-Iman test is shown. For comparing enriched saliva or buccal samples with the corresponding non-enriched samples, the Wilcoxon signed-rank test was used, with the P-value indicated in the table. For the Wilcoxon test, a P-value of 0.062 was considered significant, as this is the smallest possible P-value when $n = 4$. Metrics prefixed with an asterisk were corrected for the total number of reads in a given sample prior to statistical testing.

Supplemental Table 5: BLAST results and mean base-quality scores for 10,000 unmapped reads from each sample.

	Percentage of BLAST hits matching…				Mean base-quality score
Sample	**Bacteria**	**Eukaryota**	**Other**	**(No BLAST hit)**	
PGPC-0002 (blood)	0.3%	83.4%	0.4%	15.9%	25.8
PGPC-0002 (non-enriched saliva)	51.9%	2.0%	0.5%	45.6%	36.8
PGPC-0002 (enriched saliva)	18.3%	2.7%	3.2%	75.8%	27.3
PGPC-0002 (non-enriched buccal)	47.0%	8.1%	0.3%	44.5%	33.8
PGPC-0002 (enriched buccal)	14.4%	21.1%	2.8%	61.7%	26.5
PGPC-0005 (blood)	0.1%	79.2%	1.9%	18.8%	25.6
PGPC-0005 (non-enriched saliva)	44.5%	0.7%	0.1%	54.7%	37.5
PGPC-0005 (enriched saliva)	30.8%	0.5%	0.7%	68.0%	34.9
PGPC-0005 (non-enriched buccal)	72.5%	3.4%	0.7%	23.4%	36.1
PGPC-0005 (enriched buccal)	43.4%	16.6%	1.1%	38.9%	31.4
PGPC-0006 (blood)	0.2%	82.4%	0.2%	17.2%	25.7
PGPC-0006 (non-enriched saliva)	60.5%	1.0%	0.3%	38.2%	37.3
PGPC-0006 (enriched saliva)	31.1%	4.1%	0.6%	64.2%	30.6
PGPC-0006 (non-enriched buccal)	78.7%	2.9%	1.5%	17.0%	37.0
PGPC-0006 (enriched buccal)	30.6%	7.9%	1.5%	60.0%	28.4
PGPC-0050 (blood)	0.2%	72.9%	5.8%	21.1%	26.3
PGPC-0050 (non-enriched saliva)	60.8%	0.7%	0.2%	38.2%	37.7
PGPC-0050 (enriched saliva)	53.8%	2.0%	0.4%	43.9%	35.5
PGPC-0050 (non-enriched buccal)	43.9%	7.5%	1.6%	47.0%	30.3
PGPC-0050 (enriched buccal)	44.9%	22.3%	3.0%	29.9%	31.9

Supplemental Table 6: Number of known and novel SNVs and indels detected in each sample both before and after filtering.

	Before filtering		After filtering	
Sample	**Known variants**	**Novel variants**	**Known variants**	**Novel variants**
SNVs				
HuRef (blood 1)	3735426	51311	3604318	30513
HuRef (blood 2)	3738089	51960	3530091	20072
PGPC-0002 (blood)	3758148	52072	3631494	32066
PGPC-0002 (non-enriched saliva)	3751290	52137	3619413	30333
PGPC-0002 (enriched saliva)	3751555	51776	3633447	34612
PGPC-0002 (non-enriched buccal)	3749332	65094	3625655	46581
PGPC-0002 (enriched buccal)	3747049	52235	3558352	26128
PGPC-0005 (blood)	3763060	51068	3637246	32770
PGPC-0005 (non-enriched saliva)	3755741	64249	3638593	45634
PGPC-0005 (enriched saliva)	3761199	51887	3638540	34045
PGPC-0005 (non-enriched buccal)	3759176	52507	3606760	26486
PGPC-0005 (enriched buccal)	3760032	51326	3632810	34504
PGPC-0006 (blood)	3791306	47362	3663978	29276
PGPC-0006 (non-enriched saliva)	3789490	49466	3666675	30617
PGPC-0006 (enriched saliva)	3787068	46846	3674442	29909
PGPC-0006 (non-enriched buccal)	3790856	65890	3668556	46690
PGPC-0006 (enriched buccal)	3787712	46318	3662284	32383
PGPC-0050 (blood)	3744699	61662	3612079	40774
PGPC-0050 (non-enriched saliva)	3742340	75658	3638365	57616
PGPC-0050 (enriched saliva)	3732870	62149	3614610	47013
PGPC-0050 (non-enriched buccal)	3742896	76092	3618312	58060
PGPC-0050 (enriched buccal)	3753009	62334	3632516	45149
Indels				
HuRef (blood 1)	784131	59780	221427	3140
HuRef (blood 2)	792645	63920	223383	3380
PGPC-0002 (blood)	783941	58499	220715	3242
PGPC-0002 (non-enriched saliva)	776354	57939	224607	3424
PGPC-0002 (enriched saliva)	784950	58074	223559	3479
PGPC-0002 (non-enriched buccal)	784814	59170	221124	3427
PGPC-0002 (enriched buccal)	793303	62579	219572	3330
PGPC-0005 (blood)	786065	59261	222524	3339
PGPC-0005 (non-enriched saliva)	778716	114745	223409	3256
PGPC-0005 (enriched saliva)	792151	65465	224568	3436
PGPC-0005 (non-enriched buccal)	784059	59917	223916	3303
PGPC-0005 (enriched buccal)	797013	62695	223633	3407
PGPC-0006 (blood)	790455	57424	220864	2950
PGPC-0006 (non-enriched saliva)	792593	68544	226099	3269
PGPC-0006 (enriched saliva)	794109	58801	225231	3181
PGPC-0006 (non-enriched buccal)	796330	62686	225128	3256
PGPC-0006 (enriched buccal)	808146	64177	223346	3225
PGPC-0050 (blood)	784458	60909	221209	3860
PGPC-0050 (non-enriched saliva)	790086	114347	222866	4000
PGPC-0050 (enriched saliva)	785422	68412	218964	4055
PGPC-0050 (non-enriched buccal)	786786	63034	223073	4092
PGPC-0050 (enriched buccal)	807396	68010	223414	4223

Supplemental Table 7: Descriptive statistics for the allele fraction distributions in each sample type (blood, saliva without enrichment, saliva with enrichment, buccal without enrichment, and buccal with enrichment) for heterozygous SNVs, homozygous SNVs, heterozygous indels, and homozygous indels.

Variant type	**Blood**	**Non-enriched saliva**	**Enriched saliva**	**Non-enriched buccal**	**Enriched buccal**
Per-sample means					
Heterozygous SNVs	0.495 (0.001)	0.493 (0.002)	0.494 (0.000)	0.493 (0.003)	0.495 (0.002)
Homozygous SNVs	0.998 (0.001)	0.999 (0.000)	0.999 (0.000)	0.999 (0.000)	0.999 (0.000)
Heterozygous indels	0.513 (0.002)	0.508 (0.002)	0.508 (0.001)	0.508 (0.001)	0.507 (0.001)
Homozygous indels	0.974 (0.001)	0.975 (0.002)	0.975 (0.001)	0.974 (0.001)	0.973 (0.001)
Per-sample standard deviations					
Heterozygous SNVs	0.122 (0.002)	0.123 (0.003)	0.123 (0.002)	0.124 (0.004)	0.123 (0.002)
Homozygous SNVs	0.012 (0.003)	0.008 (0.000)	0.009 (0.000)	0.009 (0.000)	0.009 (0.001)
Heterozygous indels	0.151 (0.003)	0.149 (0.001)	0.150 (0.000)	0.150 (0.001)	0.149 (0.001)
Homozygous indels	0.105 (0.002)	0.105 (0.002)	0.105 (0.001)	0.107 (0.002)	0.110 (0.002)

The mean and standard deviation of the allele fractions were calculated for each variant type in each sample. Means and standard deviations of the per-sample means and standard deviations were then computed for each sample type (over the four study participants). For heterozygous SNVs detected in blood, for instance, the mean allele fractions for individuals PGPC-0002, PGPC-0005, PGPC-0006, and PGPC-0050 were 0.494, 0.495, 0.495, and 0.495, respectively; the mean and standard deviations of these values are 0.495 and 0.001, respectively. Similarly, the standard deviations of the heterozygous SNV allele fractions for individuals PGPC-0002, PGPC-0005, PGPC-0006, and PGPC-0050 were 0.123, 0.122, 0.124, and 0.119, respectively; the mean and standard deviation of these values are 0.122 and 0.002.

Supplemental Table 8: SNV- and indel-detection concordance between blood samples and non-enriched saliva or buccal samples, and between enriched saliva or buccal samples and the corresponding non-enriched samples, for filtered, known SNVs and indels (those present in gnomAD) in coding exons, all exons, introns, and intergenic regions.

		Coding exons			All exons			Introns			Intergenic regions		
Sample Type 1	**Sample Type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**
SNVs													
HuRef Blood 1	HuRef Blood 2	94.7	4.2	1.0	95.4	3.6	1.1	95.6	3.0	1.4	94.2	4.0	1.7
Blood	Non-enriched saliva	96.2	1.7	2.1	96.7	1.5	1.8	96.8	1.6	1.6	96.1	1.8	2.0
Blood	Non-enriched buccal	96.1	2.0	1.9	96.5	1.7	1.7	96.6	1.8	1.6	95.7	2.2	2.0
Enriched saliva	Non-enriched saliva	96.3	1.8	1.9	96.9	1.5	1.6	97.2	1.4	1.4	96.5	1.7	1.8
Enriched buccal	Non-enriched buccal	96.2	1.7	2.1	96.5	1.6	1.9	96.7	1.6	1.7	95.7	2.0	2.3
Indels													
HuRef Blood 1	HuRef Blood 2	85.7	6.9	7.4	89.3	5.3	5.4	87.6	5.9	6.5	87.3	5.9	6.9
Blood	Non-enriched saliva	80.2	10.1	9.7	87.4	5.9	6.7	87.3	5.9	6.9	86.7	5.9	7.4
Blood	Non-enriched buccal	81.4	7.7	10.9	87.0	5.9	7.0	86.9	6.2	7.0	86.0	6.6	7.4
Enriched saliva	Non-enriched saliva	81.9	10.4	7.7	87.8	6.4	5.8	87.6	6.1	6.3	86.8	6.2	7.0
Enriched buccal	Non-enriched buccal	82.5	8.2	9.4	87.5	6.4	6.0	87.2	6.5	6.4	85.8	6.8	7.5

For further details, see Table 1.

Supplemental Table 9: SNV- and indel-detection concordance between blood samples and non-enriched saliva or buccal samples, and between enriched saliva or buccal samples and the corresponding non-enriched samples, for filtered, novel SNVs and indels (those not present in gnomAD) in coding exons, all exons, introns, and intergenic regions.

		Coding exons			All exons			Introns			Intergenic regions		
Sample Type 1	**Sample Type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**	**Concordant**	**Unique to sample type 1**	**Unique to sample type 2**
SNVs													
HuRef Blood 1	HuRef Blood 2	84.8	9.1	6.1	65.5	29.5	5.0	57.4	33.9	8.7	50.2	41.9	7.8
Blood	Non-enriched saliva	17.5	2.3	80.2	44.4	7.2	48.3	56.3	11.6	32.0	56.9	15.2	27.8
Blood	Non-enriched buccal	57.6	8.4	34.0	64.7	12.5	22.8	50.7	12.6	36.7	49.1	15.9	35.0
Enriched saliva	Non-enriched saliva	17.8	5.4	76.8	43.5	11.5	45.0	56.4	13.9	29.7	57.1	18.4	24.6
Enriched buccal	Non-enriched buccal	51.3	19.0	29.7	61.9	16.3	21.9	49.7	14.0	36.3	48.3	17.2	34.5
Indels													
HuRef Blood 1	HuRef Blood 2	85.7	0.0	14.3	70.8	11.1	18.1	67.4	14.3	18.3	64.1	14.3	21.7
Blood	Non-enriched saliva	70.3	10.8	18.9	70.0	12.4	17.6	65.0	17.0	18.0	62.2	16.5	21.3
Blood	Non-enriched buccal	67.6	13.5	18.9	69.7	12.7	17.6	64.5	16.9	18.6	62.6	15.9	21.5
Enriched saliva	Non-enriched saliva	81.1	10.8	8.1	72.8	13.8	13.5	64.8	18.3	16.9	62.2	19.4	18.4
Enriched buccal	Non-enriched buccal	70.0	20.0	10.0	71.1	14.9	14.0	65.9	18.1	16.0	63.5	18.1	18.4

Note that there were very few novel indels detected in coding exons – the raw counts for the number of concordant novel indels in coding exons was 6, 26, 25, 30, and 28 for HuRef Blood 1 versus HuRef Blood 2, blood versus non-enriched saliva, blood versus non-enriched buccal, enriched saliva versus non-enriched saliva, and enriched buccal versus non-enriched buccal, respectively. For further details, see Table 1.

Supplemental Table 10: Accuracy of discordant SNVs and indels.

		Unique to sample type 1	Unique to sample type 2	Unique to sample type 1	Unique to sample type 1
Sample Type 1	**Sample Type 2**	**Known**		**Novel**	
SNVs					
Blood	Non-enriched saliva	6/13	5/12	0/12	2/13
Blood	Non-enriched buccal	8/15	5/11	2/10	2/14
Enriched saliva	Non-enriched saliva	5/15	1/10	2/10	3/15
Enriched buccal	Non-enriched buccal	4/8	5/10	6/17	3/15
Indels					
Blood	Non-enriched saliva	11/13	11/13	8/12	7/12
Blood	Non-enriched buccal	12/15	8/13	6/10	11/12
Enriched saliva	Non-enriched saliva	9/12	10/14	9/13	3/11
Enriched buccal	Non-enriched buccal	11/13	9/11	7/12	9/14

A total of 200 SNVs and 200 indels detected in one sample type but not a second sample type were visually inspected in Integrative Genomics Viewer to determine their correctness. The "unique to sample type 1" columns are of the form X/Y, where X is the number of variants detected in sample type 1 but not sample type 2 that were deemed correct by visual confirmation and Y is the total number inspected (and analogously for the "unique to sample type 2" columns). After correcting for multiple hypothesis testing (Bonferroni), none of the comparisons was statistically significant in terms of having a greater proportion of variants unique to sample type 1 deemed correct than sample type 2 or vice versa (χ^2 test). A small number of variants were difficult to verify and thus are omitted from the table, which is why the total of the denominators is less than 400.

Supplemental Table 11: Number of reads supporting the alternate and reference alleles for the variant chrMT.12684G>A, which is one of the SNVs comprising a SNV cluster found in the enriched saliva samples from PGPC-0002 and PGPC-0005 but not any of the other sample types from these individuals.

Sample	**Number of reads supporting alternate allele**	**Number of reads supporting reference allele**
PGPC-0002 (blood)	36	3083
PGPC-0002 (non-enriched saliva)	18	2326
PGPC-0002 (enriched saliva)	19	63
PGPC-0002 (non-enriched buccal)	22	10269
PGPC-0002 (enriched buccal)	23	1113
PGPC-0005 (blood)	24	1486
PGPC-0005 (non-enriched saliva)	26	2568
PGPC-0005 (enriched saliva)	34	132
PGPC-0005 (non-enriched buccal)	25	9387
PGPC-0005 (enriched buccal)	50	378
PGPC-0006 (blood)	25	2029
PGPC-0006 (non-enriched saliva)	13	2579
PGPC-0006 (enriched saliva)	16	420
PGPC-0006 (non-enriched buccal)	13	9178
PGPC-0006 (enriched buccal)	12	1977
PGPC-0050 (blood)	11	1866
PGPC-0050 (non-enriched saliva)	20	4432
PGPC-0050 (enriched saliva)	26	1000
PGPC-0050 (non-enriched buccal)	27	9486
PGPC-0050 (enriched buccal)	29	1218

Highlighted cells indicate samples in which the variant was detected by the Genome Analysis Toolkit; these include the enriched buccal sample from PGPC-0005, even though most other SNVs comprising the clusters were not detected in this sample.

Supplemental Table 12: Number of common and rare CNVs detected in each sample using the ERDS and CNVnator-based workflow.

Sample	Number of common CNVs	Number of rare CNVs
Deletions		
HuRef (blood 1)	379	7
HuRef (blood 2)	362	7
PGPC-0002 (blood)	312	13
PGPC-0002 (non-enriched saliva)	295	12
PGPC-0002 (enriched saliva)	276	11
PGPC-0002 (non-enriched buccal)	208	7
PGPC-0002 (enriched buccal)	68	3
PGPC-0005 (blood)	309	10
PGPC-0005 (non-enriched saliva)	210	7
PGPC-0005 (enriched saliva)	250	7
PGPC-0005 (non-enriched buccal)	282	10
PGPC-0005 (enriched buccal)	139	6
PGPC-0006 (blood)	264	11
PGPC-0006 (non-enriched saliva)	269	12
PGPC-0006 (enriched saliva)	236	13
PGPC-0006 (non-enriched buccal)	249	11
PGPC-0006 (enriched buccal)	70	2
PGPC-0050 (blood)	270	15
PGPC-0050 (non-enriched saliva)	193	12
PGPC-0050 (enriched saliva)	54	2
PGPC-0050 (non-enriched buccal)	247	15
PGPC-0050 (enriched buccal)	144	10
Duplications		
HuRef (blood 1)	109	4
HuRef (blood 2)	117	4
PGPC-0002 (blood)	90	4
PGPC-0002 (non-enriched saliva)	83	4
PGPC-0002 (enriched saliva)	87	4
PGPC-0002 (non-enriched buccal)	79	4
PGPC-0002 (enriched buccal)	36	4
PGPC-0005 (blood)	101	2
PGPC-0005 (non-enriched saliva)	99	2
PGPC-0005 (enriched saliva)	91	4
PGPC-0005 (non-enriched buccal)	101	3
PGPC-0005 (enriched buccal)	75	3
PGPC-0006 (blood)	83	3
PGPC-0006 (non-enriched saliva)	83	0
PGPC-0006 (enriched saliva)	80	1
PGPC-0006 (non-enriched buccal)	89	3
PGPC-0006 (enriched buccal)	45	2
PGPC-0050 (blood)	88	7
PGPC-0050 (non-enriched saliva)	75	7
PGPC-0050 (enriched saliva)	28	4
PGPC-0050 (non-enriched buccal)	82	9
PGPC-0050 (enriched buccal)	62	5

Supplemental Table 13: Concordance between blood samples and non-enriched saliva or buccal samples, and between enriched saliva or buccal samples and the corresponding non-enriched samples, for rare CNVs (those with <1% frequency in MSSNG parents).

		Concordant	Unique to sample type 1	Unique to sample type 2	Concordant	Unique to sample type 1	Unique to sample type 2	Concordant	Unique to sample type 1	Unique to sample type 2
Sample type 1	**Sample type 2**	**[1 kb,5 kb)**			**[5 kb,10 kb)**			**[10 kb,…)**		
Deletions										
HuRef Blood 1	HuRef Blood 2	1	1/1	0/0	1	0/0	0/0	3	0/0	0/0
Blood	Non-enriched saliva	10	12/12	7/8	14	2/2	0/0	11	0/0	0/0
Blood	Non-enriched buccal	13	9/9	3/3	14	2/2	3/3	10	1/1	0/0
Enriched saliva	Non-enriched saliva	6	6/6	11/12	9	2/2	5/5	8	0/2	3/3
Enriched buccal	Non-enriched buccal	3	0/0	13/13	5	0/1	11/12	2	0/10	8/8
Duplications										
HuRef Blood 1	HuRef Blood 2	0	0/0	0/0	0	0/0	0/0	3	0/0	0/0
Blood	Non-enriched saliva	2	1/1	0/0	1	2/2	0/1	9	1/1	0/0
Blood	Non-enriched buccal	2	1/1	0/7	2	1/1	0/0	8	2/2	0/0
Enriched saliva	Non-enriched saliva	2	0/1	0/0	0	0/2	1/2	5	0/3	4/4
Enriched buccal	Non-enriched buccal	1	0/0	2/9	0	0/0	1/1	7	0/6	1/1

For each rare CNV that was discordant between two sample types, we used visual inspection of alignments to classify it as true or false. The "concordant" columns contain the number of CNVs detected in both sample type 1 and sample type 2. The "unique to sample type 1" columns are of the form X/Y, where X is the number of CNVs detected in sample type 1 but not sample type 2 that were verified as correct and Y is the total number of such CNVs (and analogously for the "unique to sample type 2" column). For example, 13 out of the 13 rare deletions between 1 and 5 kb detected in non-enriched buccal samples but not in enriched buccal samples were deemed to be correct by visual inspection of alignments. Larger values of X indicate greater sensitivity, while larger values of $Y - X$ denote higher false discovery rates. Except for the HuRef blood comparison, all values are after aggregating across individuals; for individual-specific data, see Additional File 4.

Supplemental Table 14: Statistically-significant differences in the distributions of the number of CNVs detected between blood samples and non-enriched saliva or buccal samples, and between enriched saliva or buccal samples and the corresponding non-enriched samples.

Metric	Blood versus non-enriched saliva	Blood versus non-enriched buccal	Enriched saliva versus non-enriched saliva	Enriched buccal versus non-enriched buccal
Common deletions [1 kb,5 kb)	—	—	0.125	0.062
Common deletions [5 kb,10 kb)	—	—	0.854	0.062
Common deletions [10 kb,…)	0.004	0.512	0.715	0.062
Common duplications [1 kb,5 kb)	—	—	0.125	0.062
Common duplications [5 kb,10 kb)	—	—	0.062	0.062
Common duplications [10 kb,…)	—	—	0.854	0.125

For blood, non-enriched saliva, and non-enriched buccal samples, a Friedman repeated-measures test was first performed to determine whether there were significant differences among these three sample types. Metrics with Friedman P-value >0.05 are indicated with "—". For metrics with Friedman P-value $\leq$ 0.05, the P-value from the post-hoc Conover-Iman test is shown. For comparing enriched saliva and buccal samples with the corresponding non-enriched samples, the Wilcoxon signed-rank test was used, with the P-value indicated in the table. For the Wilcoxon test, a P-value of 0.062 was considered significant, as this is the smallest possible P-value with $n = 4$.

Supplemental Table 15: Consistency of CNVs detected by Canvas in the five sample types with those detected by the ERDS and CNVnator-based CNV-detection workflow in blood samples.

Sample type	**[1 kb,5 kb)**	**[5 kb,10 kb)**	**[10 kb,...)**
Deletions			
Blood	15/22	33/42	56/80
Non-enriched saliva	15/27	35/39	51/75
Enriched saliva	3/4	16/20	36/63
Non-enriched buccal	9/16	26/34	43/70
Enriched buccal	0/1	2/2	21/145
Duplications			
Blood	2/22	6/26	18/66
Non-enriched saliva	3/27	6/16	17/73
Enriched saliva	1/23	2/22	12/66
Non-enriched buccal	0/20	4/21	19/71
Enriched buccal	0/14	0/5	9/48

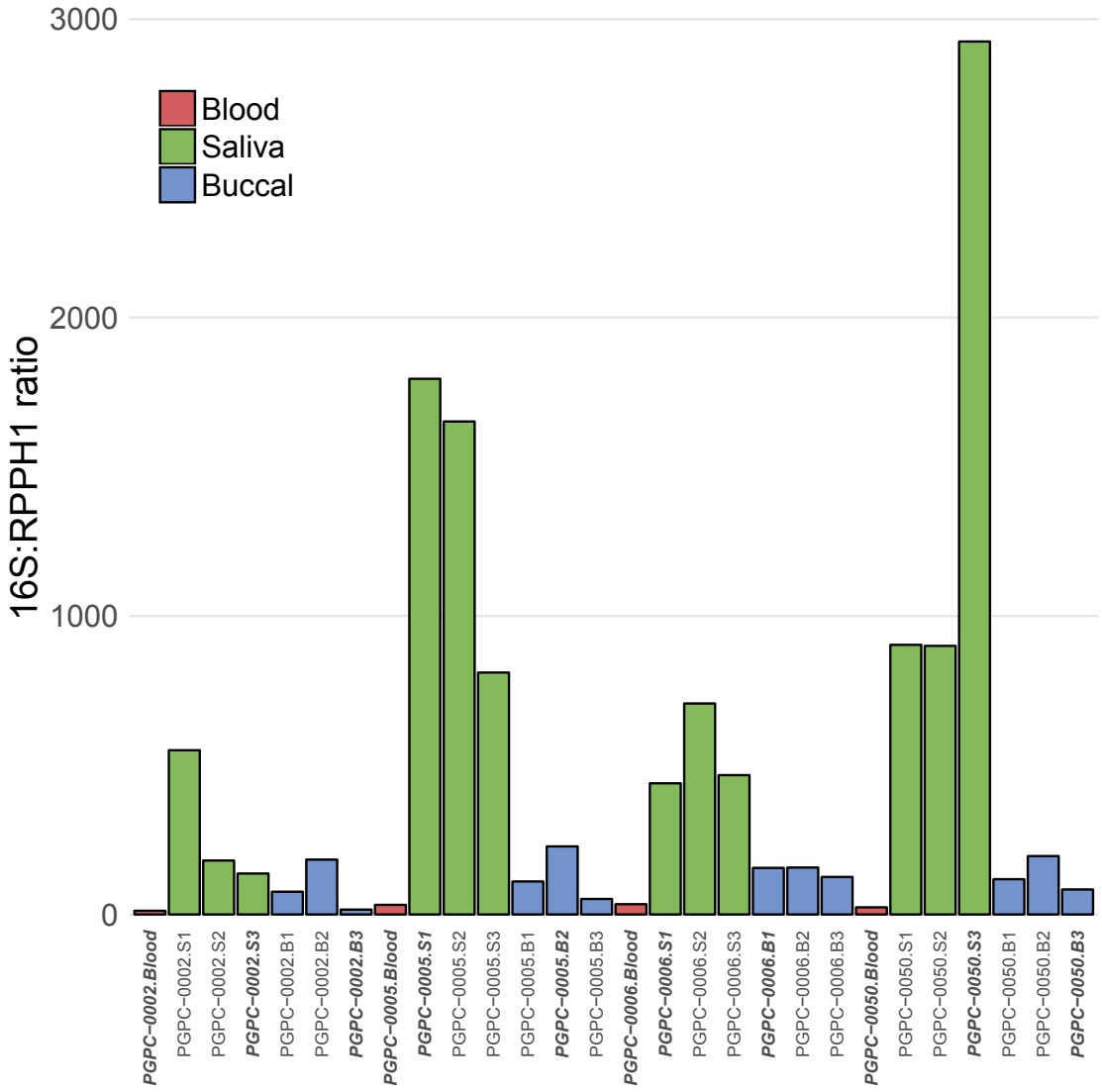
Each cell is of the form X/Y, where Y is the total number of CNVs detected and X is the number of those that were also detected by the ERDS and CNVnator-based workflow in the blood sample from the same individual. For example, across the four individuals, 27 CNVs between 1 and 5 kb were detected by Canvas in a non-enriched saliva sample, 15 of which were also detected by our ERDS/CNVnator workflow in the blood sample from the same individual. Larger values of X indicate greater sensitivity, while larger values of $Y - X$ denote higher false discovery rates. All values are after aggregating across individuals; for individual-specific data, see Additional File 4.

Supplemental Table 16: Number of structural variants of each type detected by Manta in each sample.

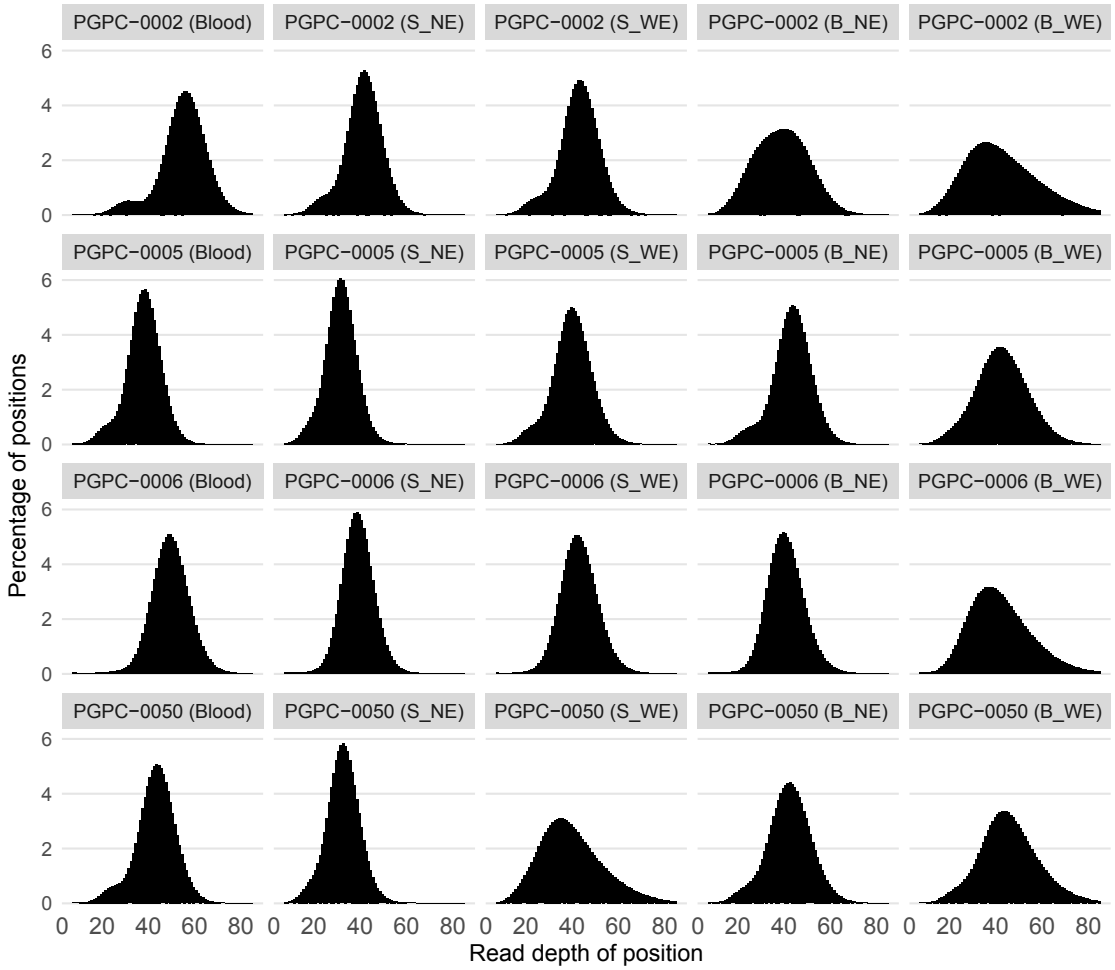
Sample	Deletions [1 kb,5 kb)	Deletions [5 kb,10 kb)	Deletions [10 kb,...)	Duplications [1 kb,5 kb)	Duplications [5 kb,10 kb)	Duplications [10 kb,...)	Inversions	Insertions	Breakends
PGPC-0002 (blood)	461	126	196	57	16	164	462	1417	3108
PGPC-0002 (non-enriched saliva)	444	123	188	67	15	142	438	1356	3040
PGPC-0002 (enriched saliva)	424	122	158	71	12	141	402	1352	2988
PGPC-0002 (non-enriched buccal)	397	116	137	48	13	108	365	1373	2826
PGPC-0002 (enriched buccal)	421	119	166	57	12	136	361	1713	3142
PGPC-0005 (blood)	446	125	185	70	14	167	467	1391	3204
PGPC-0005 (non-enriched saliva)	415	122	169	71	19	154	404	1244	2874
PGPC-0005 (enriched saliva)	417	117	183	66	18	162	416	1509	3106
PGPC-0005 (non-enriched buccal)	415	129	172	69	17	163	437	1362	3044
PGPC-0005 (enriched buccal)	418	127	157	62	17	140	396	1611	3032
PGPC-0006 (blood)	440	123	194	74	14	157	417	1289	3028
PGPC-0006 (non-enriched saliva)	437	130	173	74	15	146	434	1406	3014
PGPC-0006 (enriched saliva)	426	122	175	64	19	149	419	1471	2972
PGPC-0006 (non-enriched buccal)	416	120	163	65	13	117	396	1428	2896
PGPC-0006 (enriched buccal)	443	123	169	60	19	127	407	1843	3160
PGPC-0050 (blood)	425	132	199	73	16	160	403	1408	3134
PGPC-0050 (non-enriched saliva)	395	121	169	56	12	129	344	1233	2782
PGPC-0050 (enriched saliva)	412	120	146	61	18	130	380	1613	3112
PGPC-0050 (non-enriched buccal)	416	133	181	68	13	131	369	1394	2956
PGPC-0050 (enriched buccal)	412	128	147	56	15	127	379	1639	3028

Counts for deletions and duplications are stratified by size (between 1 and 5 kb, between 5 and 10 kb, and >10 kb), whereas counts for inversions, insertions, and breakends are not.

SUPPLEMENTAL FIGURES

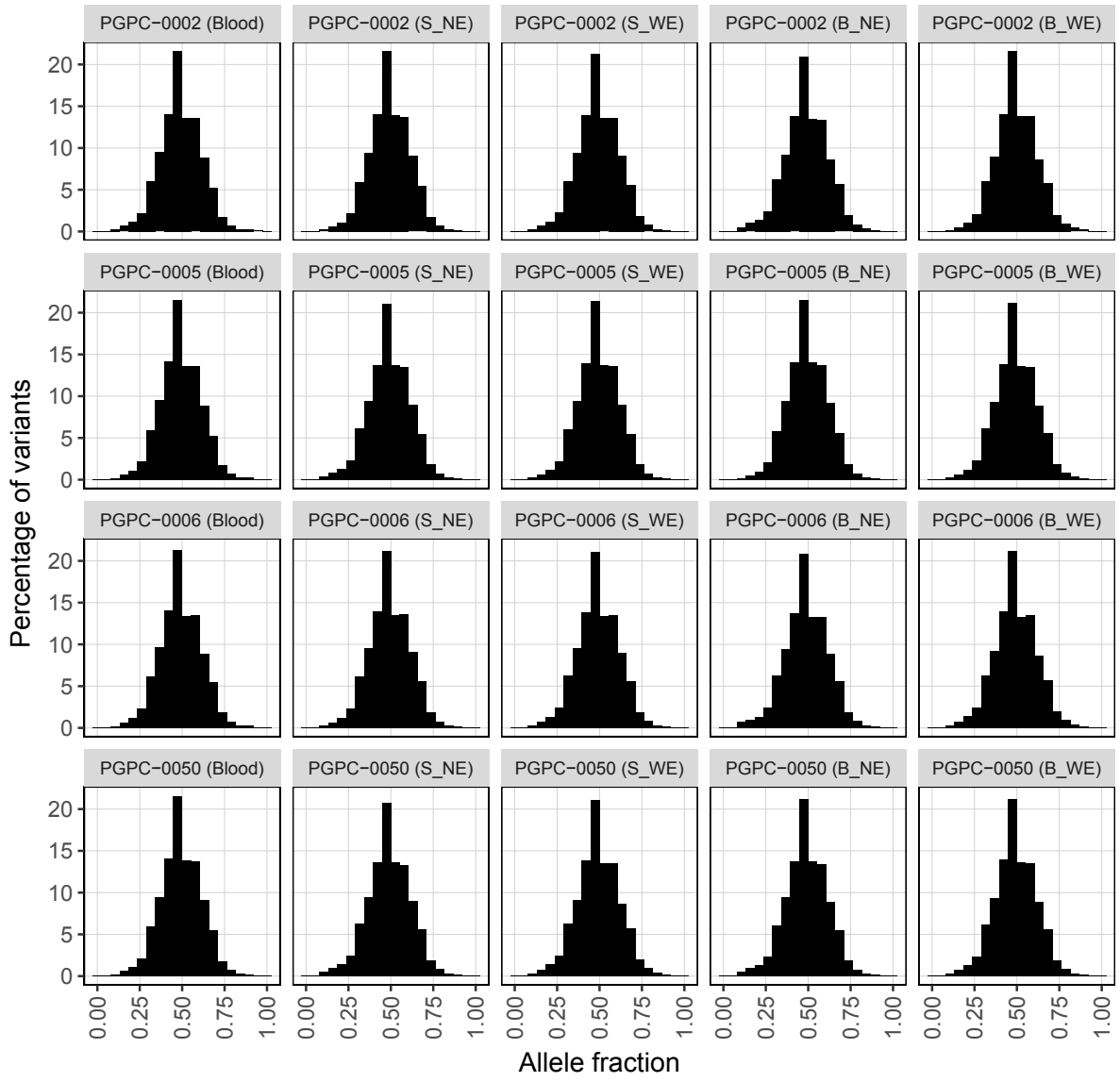


Supplemental Figure 1: Bacterial DNA quantification. One blood, three saliva (S1, S2, and S3), and three buccal (B1, B2, and B3) samples were collected from each of four individuals (PGPC-0002, PGPC-0005, PGPC-0006, and PGPC-0050), and their concentrations of bacterial DNA quantified. Higher 16S:RPPH1 ratios indicate higher bacterial DNA concentrations. Samples in bold were selected for further analysis.

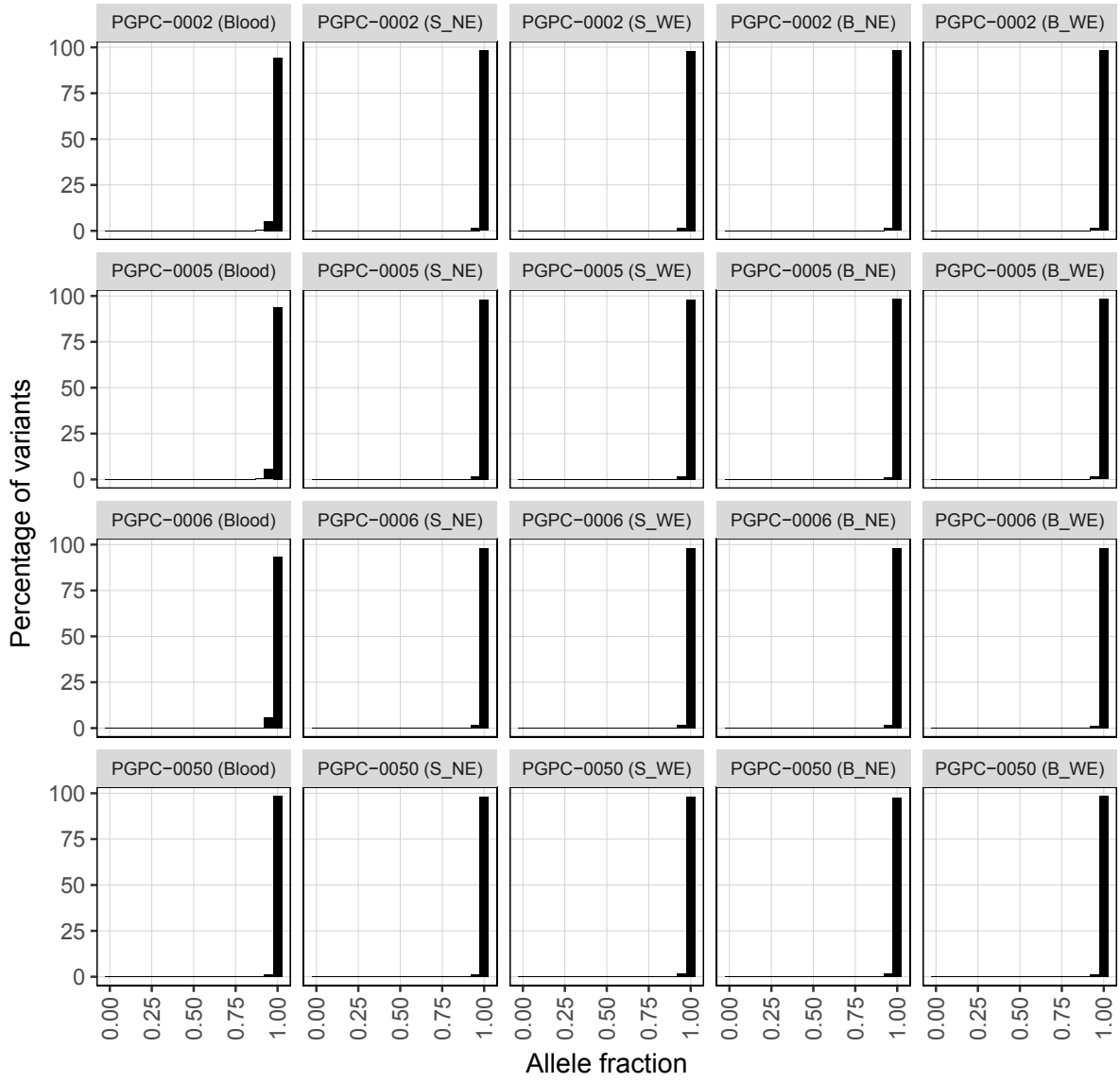


Supplemental Figure 2: Read-depth uniformity of each sample. Each histogram indicates the percentage of genomic positions sequenced to a given depth, with wider peaks indicating lower read-depth uniformity. S_NE, non-enriched saliva; S_WE, enriched saliva; B_NE, non-enriched buccal; B_WE, enriched buccal.

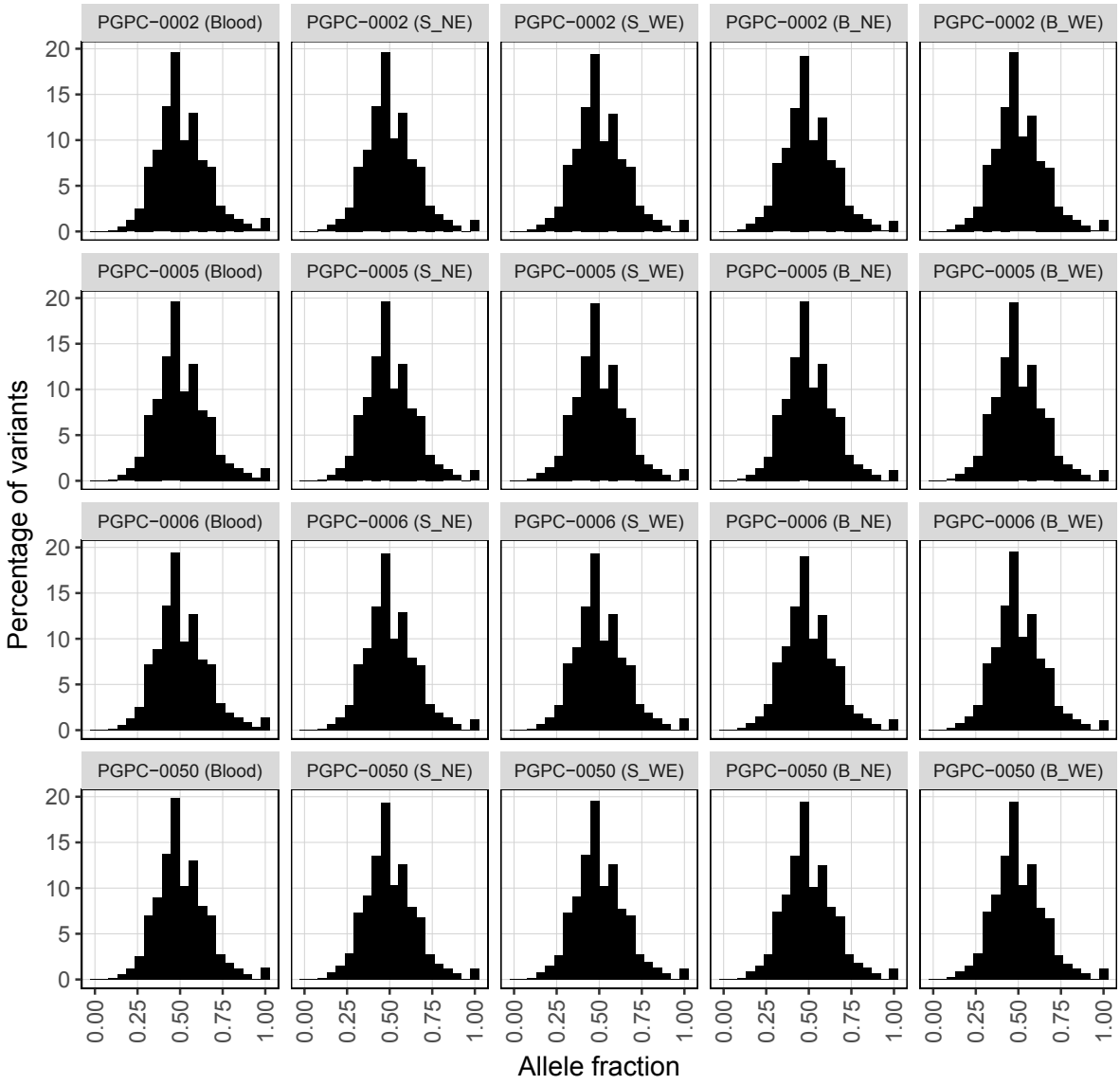
(A)


PGPC−0002 (Blood)
PGPC−0002 (S_NE)
PGPC−0002 (S_WE)
PGPC−0002 (B_NE)
PGPC−0002 (B_WE)
PGPC−0005 (Blood)
PGPC−0005 (S_NE)
PGPC−0005 (S_WE)
PGPC−0005 (B_NE)
PGPC−0005 (B_WE)
PGPC−0006 (Blood)
PGPC−0006 (S_NE)
PGPC−0006 (S_WE)
PGPC−0006 (B_NE)
PGPC−0006 (B_WE)
PGPC−0050 (Blood)
PGPC−0050 (S_NE)
PGPC−0050 (S_WE)
PGPC−0050 (B_NE)
PGPC−0050 (B_WE)
Percentage of variants
0
5
10
15
20
0.00
0.25
0.50
0.75
1.00
Allele fraction

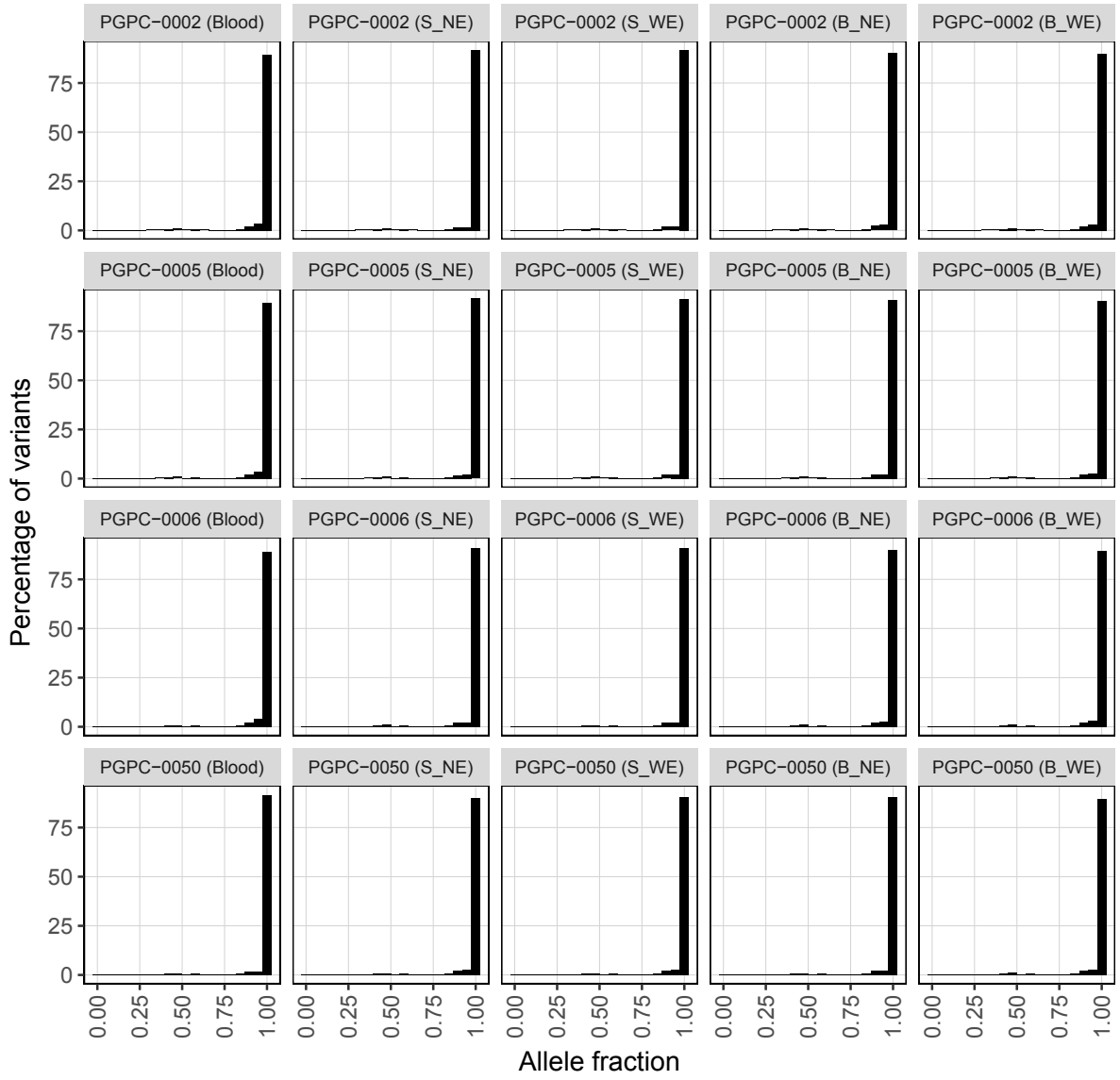
(B)


PGPC−0002 (Blood)
PGPC−0002 (S_NE)
PGPC−0002 (S_WE)
PGPC−0002 (B_NE)
PGPC−0002 (B_WE)
PGPC−0005 (Blood)
PGPC−0005 (S_NE)
PGPC−0005 (S_WE)
PGPC−0005 (B_NE)
PGPC−0005 (B_WE)
PGPC−0006 (Blood)
PGPC−0006 (S_NE)
PGPC−0006 (S_WE)
PGPC−0006 (B_NE)
PGPC−0006 (B_WE)
PGPC−0050 (Blood)
PGPC−0050 (S_NE)
PGPC−0050 (S_WE)
PGPC−0050 (B_NE)
PGPC−0050 (B_WE)
Percentage of variants
100
75
50
25
0
0.00
0.25
0.50
0.75
1.00
Allele fraction

(C)

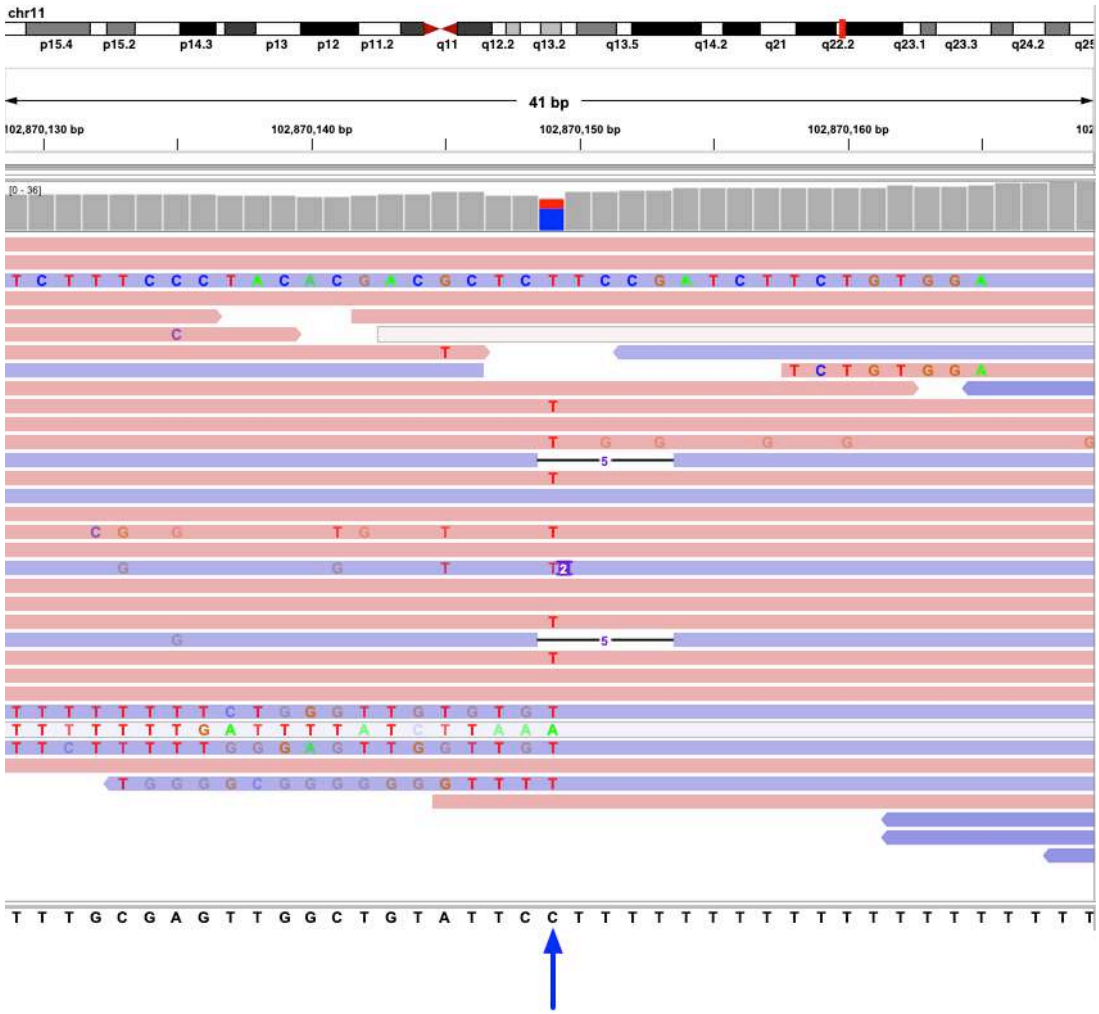

PGPC−0002 (Blood)
PGPC−0002 (S_NE)
PGPC−0002 (S_WE)
PGPC−0002 (B_NE)
PGPC−0002 (B_WE)
PGPC−0005 (Blood)
PGPC−0005 (S_NE)
PGPC−0005 (S_WE)
PGPC−0005 (B_NE)
PGPC−0005 (B_WE)
PGPC−0006 (Blood)
PGPC−0006 (S_NE)
PGPC−0006 (S_WE)
PGPC−0006 (B_NE)
PGPC−0006 (B_WE)
PGPC−0050 (Blood)
PGPC−0050 (S_NE)
PGPC−0050 (S_WE)
PGPC−0050 (B_NE)
PGPC−0050 (B_WE)
Percentage of variants
0
5
10
15
20
0.00
0.25
0.50
0.75
1.00
Allele fraction

(D)

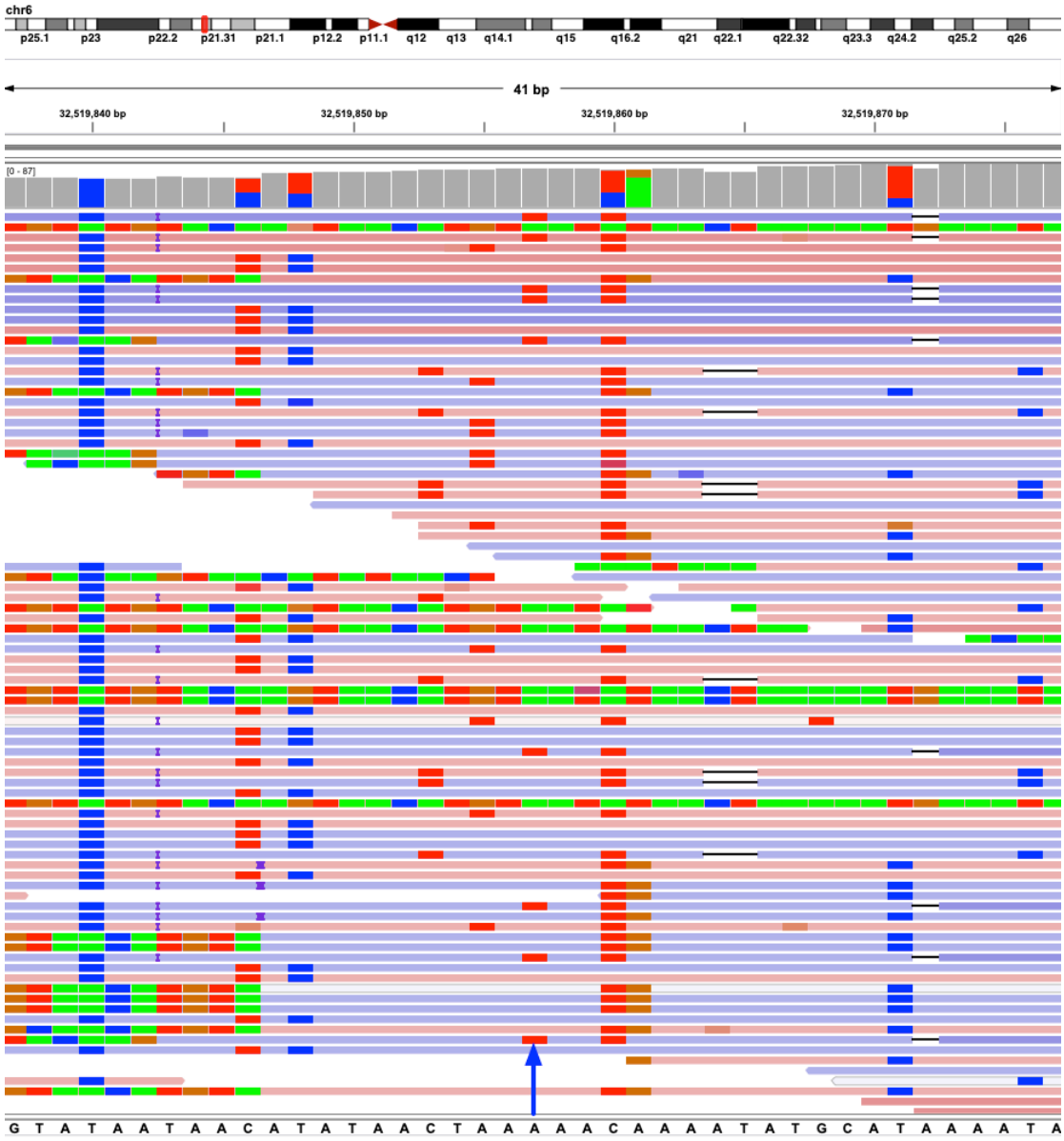


Supplemental Figure 3: Allele fraction distributions for each sample for (A) heterozygous SNVs, (B) homozygous SNVs, (C) heterozygous indels, and (D) homozygous indels. S_NE, non-enriched saliva; S_WE, enriched saliva; B_NE, non-enriched buccal; B_WE, enriched buccal.

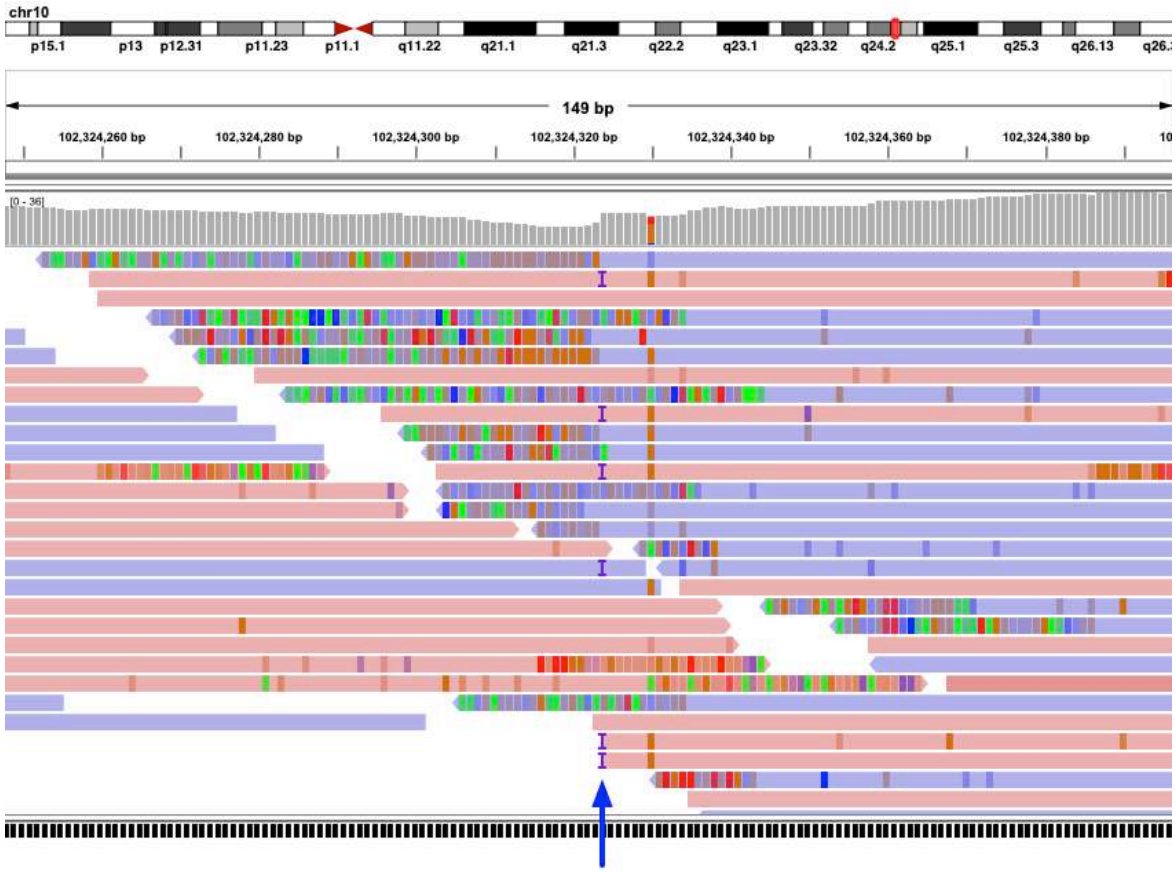
(A)


chr11
p15.4 p15.2 p14.3 p13 p12 p11.2 q11 q12.2 q13.2 q13.5 q14.2 q21 q22.2 q23.1 q23.3 q24.2
41 bp
102,870,130 bp
102,870,140 bp
102,870,150 bp
102,870,160 bp
[0 - 36]

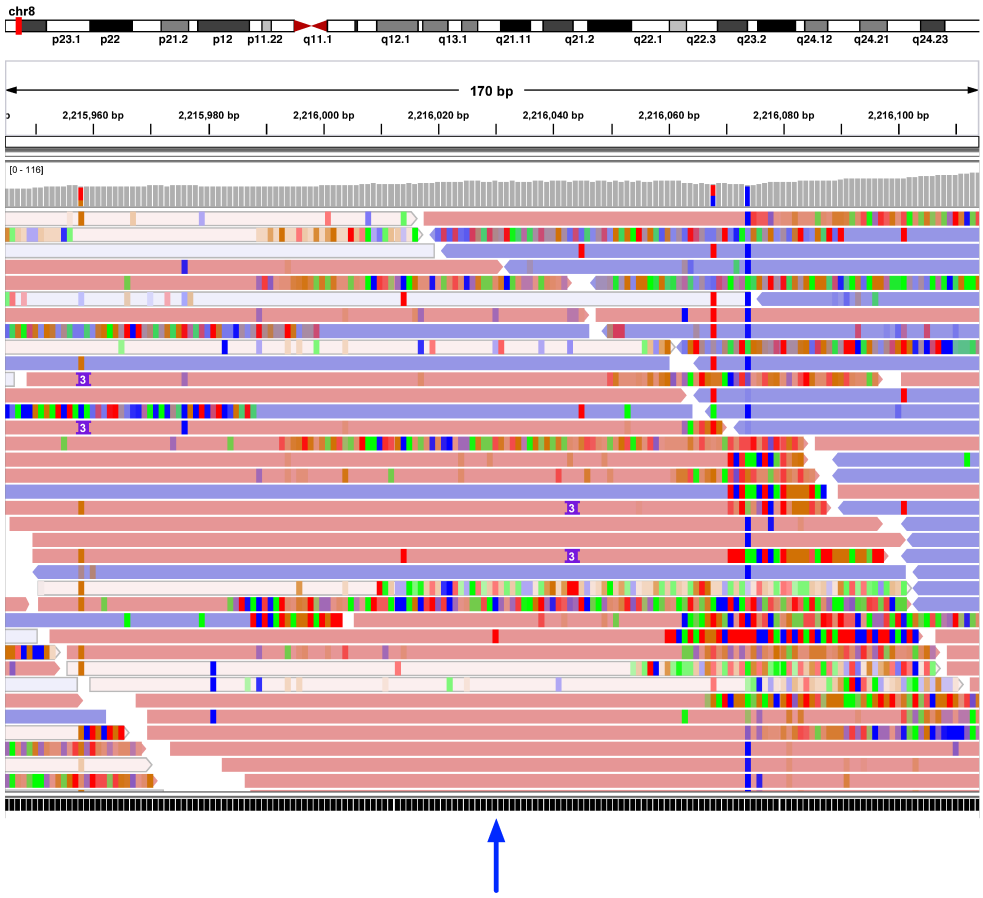
(B)


chr6
p25.1 p23 p22.2 p21.31 p21.1 p12.2 p11.1 q12 q13 q14.1 q15 q16.2 q21 q22.1 q22.32 q23.3 q24.2 q25.2 q26
41 bp
32,519,840 bp
32,519,850 bp
32,519,860 bp
32,519,870 bp
[0 - 87]
G T A T A A T A A C A T A T A A C T A A A A A C A A A A T A T G C A T A A A A T A

(C)

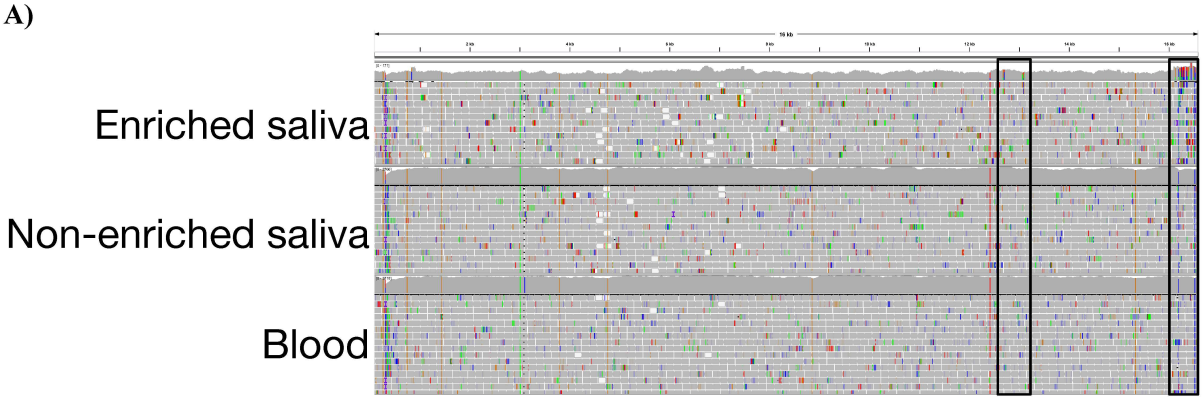

chr10
p15.1
p13
p12.31
p11.23
p11.1
q11.22
q21.1
q21.3
q22.2
q23.1
q23.32
q24.2
q25.1
q25.3
q26.13
149 bp
102,324,260 bp
102,324,280 bp
102,324,300 bp
102,324,320 bp
102,324,340 bp
102,324,360 bp
102,324,380 bp
[0 - 36]

(D)

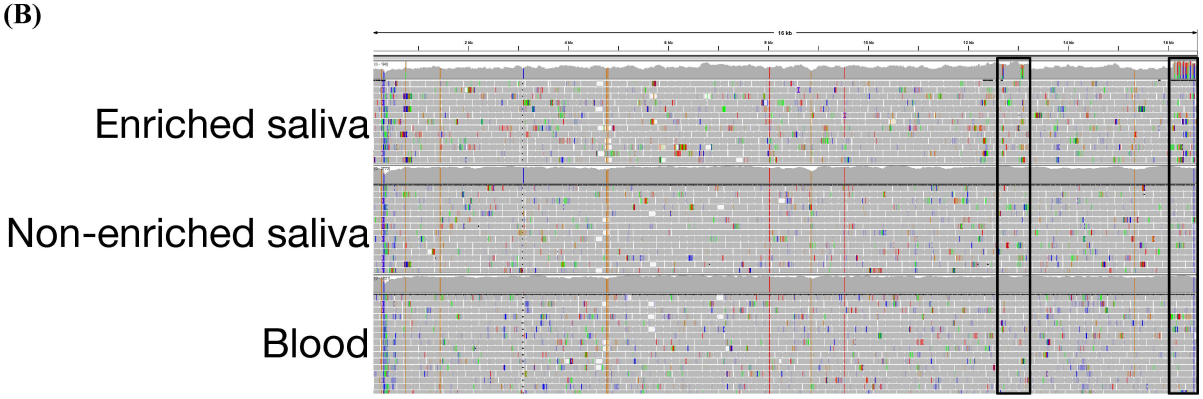


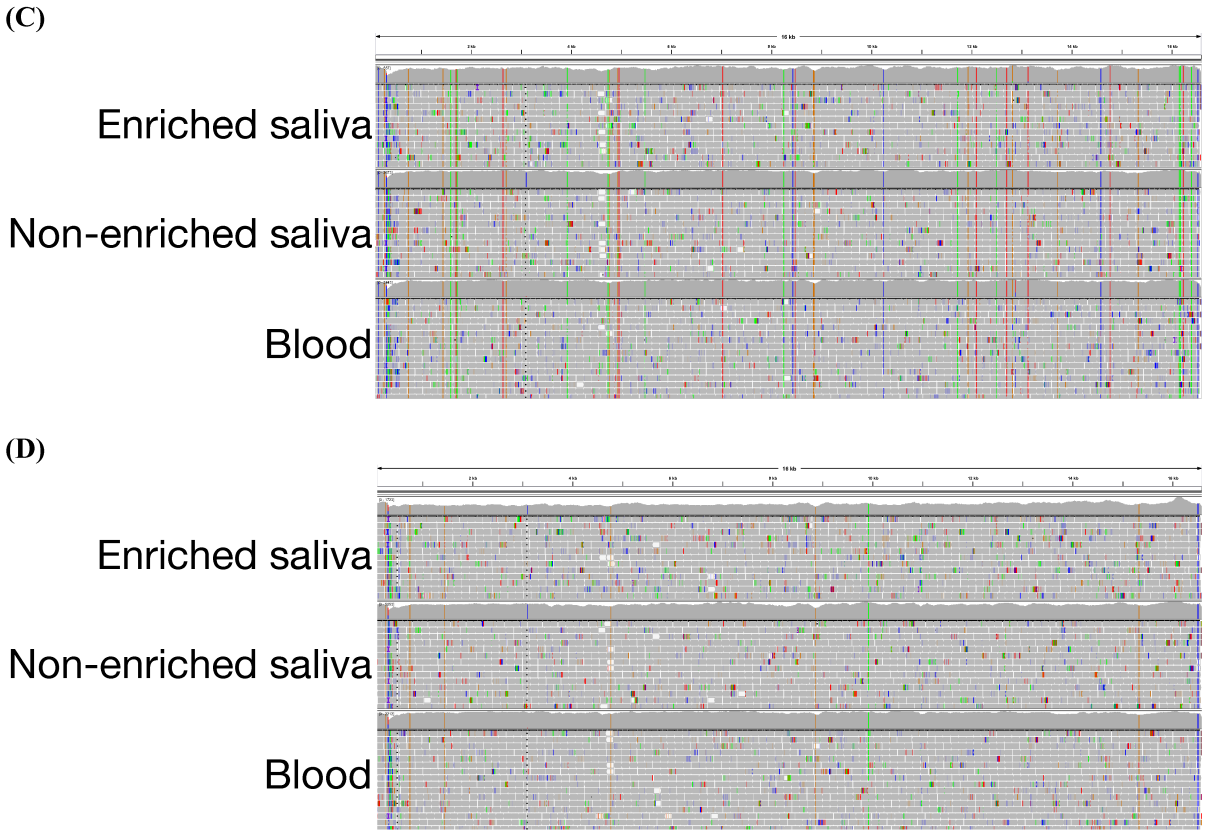
Supplemental Figure 4: Examples of evidence suggesting that a given discordant SNV or indel is false. (A) The SNV chr11.102870149C>T was detected in the blood sample from individual PGPC-0006 but not the non-enriched buccal sample from the same individual. Indications that this SNV is false include the low allele fraction and the inconsistency of other variant calls surrounding the SNV. (B) The SNV chr6:32519857A>T was detected in the enriched buccal sample from individual PGPC-0050 but not the non-enriched buccal sample from the same individual. Indications that this SNV is false include strand bias (almost all reads supporting the variant are from the same strand, indicated by the blue reads) and an abundance of surrounding soft-clipped reads and inconsistent SNVs and indels. (C) The insertion chr10:102324323insG was detected in the blood sample from individual PGPC-0005 but not the non-enriched buccal sample from the same individual. Indications that this insertion is false include an abundance of soft-clipped reads that are inconsistently aligned. (D) The deletion chr8:2216030delGCAGCAGAGGGCTTGCTGTAGGTCTGGGGGAGTAATGCGACCTTGTGGGGTGTTGGGAGGATCTACCAGGCAAAGCGCCTCCAGAAT was detected in the enriched saliva sample from individual PGPC-0006 but not the non-enriched saliva sample from the same individual. Indications that this deletion is false include low supporting evidence, abundant soft-clipping, and inconsistent variant calls in the region. In each panel, the variant being examined is indicated by a blue arrow.

A)

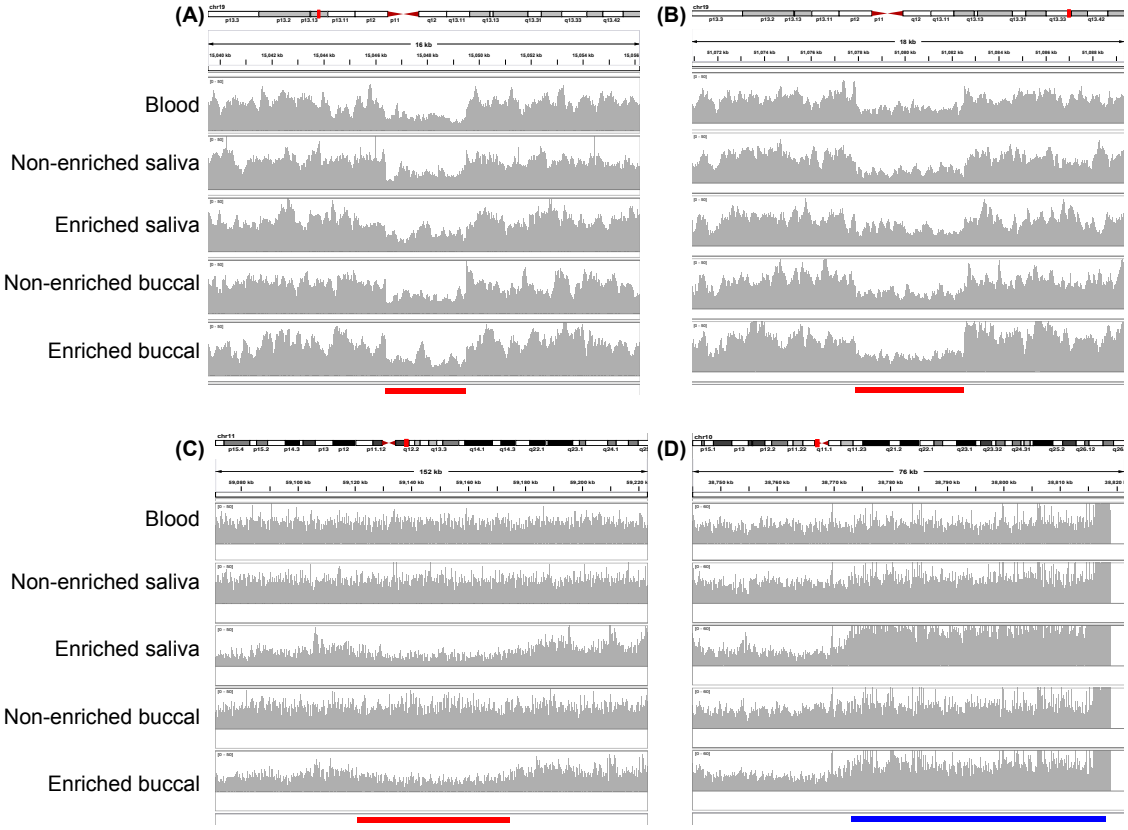

Enriched saliva
Non-enriched saliva
Blood

(B)


Enriched saliva
Non-enriched saliva
Blood



Supplemental Figure 5: Integrative Genomics Viewer visualization of the mitochondrial genome for three sample types (enriched saliva, non-enriched saliva, and blood) for A) PGPC-0002, B) PGPC-0005, C) PGPC-0006, and D) PGPC-0050. The black boxes denote the regions containing clusters of apparent SNVs in the enriched saliva samples from PGPC-0002 and PGPC-0005 but not the non-enriched saliva or blood samples from those individuals or the enriched saliva samples from PGPC-0006 and PGPC-0050.



Supplemental Figure 6: Integrative Genomics Viewer coverage plots of selected false-negative and false-positive CNVs detected in saliva and buccal samples that were enriched for eukaryotic DNA. The Y-axis of each plot indicates read depth. (A) True deletion (coordinates chr19:15,046,393-15,049,477; red bar) detected in all sample types from individual PGPC-0005 except the enriched saliva and buccal samples. (B) True deletion (coordinates chr19:51,077,892-51,082,511; red bar) detected in all sample types from individual PGPC-0006 except the enriched saliva and buccal samples. (C) False deletion (coordinates chr11:59,120,801-59,174,600; red bar) detected only in the enriched buccal sample from individual PGPC-0050. (D) False duplication (coordinates chr10:38,773,001-38,818,000; blue bar) detected only in the enriched saliva and buccal samples from individual PGPC-0050. We considered the deletions in panels A and B to be true because they exhibited clear decreases in read depth and had support from anomalously-mapped paired-end reads (not shown). We considered the CNVs in panels C and D to be false because the non-enriched samples and the blood sample did not exhibit changes in read depth in these regions and because they lacked support from split reads or anomalously-mapped paired-end reads.