

Supplemental Tables

Table S1: Mapping Statistics for WGS data from coffee cup samples, related to STAR Methods and Figure 3.

Sample Name	# of Total Reads	Percentage of Mapped Reads
IndA-Cup1	127,850,309	82.36
IndA-Cup2	137,723,180	82.14
IndA-Cup3	160,138,265	61.50
IndA-Cup4	215,379,662	75.98
IndA-Cup5	157,769,539	83.56
IndA-Cup6	121,329,654	83.54
IndB-Cup1	207,414,742	83.58
IndB-Cup2	151,139,298	62.65
IndB-Cup3	122,164,133	83.76
IndB-Cup4	220,846,681	77.93
IndB-Cup5	140,335,759	85.26
IndB-Cup6	172,391,527	88.97

Table S2: Genotyping statistics of data from coffee cups for different technologies, related to STAR Methods.

samples	WGS			Genotyping Array				ONT		
	Called variants	Correct variants	Overlap with RNA-Seq	Called variants	Correct variants	Overlap with RNA-Seq	Call rate	Called variants	Correct variants	Overlap with RNA-Seq
indA-cup1	219,887	127,389	2,843	781,807	544,932	4	84%	14	0	0
indA-cup2	526,798	366,539	5,764	805,034	572,318	11	86%	287	9	0
indA-cup3	252,452	104,710	2,314	761,566	532,966	1	81%	79	14	0
indA-cup4	217,840	84,418	2,323	672,857	452,859	2	72%	12	4	4
indA-cup5	25,822	6,984	272	808,505	600,175	5	86%	0	0	0
indA-cup6	56,577	30,580	601	743,598	574,474	3	79%	0	0	0
indA-cup1	267,844	127,606	9,872	751,349	534,890	12	80%	122	15	3
indA-cup2	243,783	136,480	9,831	788,263	562,419	15	84%	0	0	0
indA-cup3	35,133	6,848	949	823,790	608,231	5	88%	0	0	0
indA-cup4	85,613	56,716	5,014	821,752	500,897	11	88%	0	0	0
indA-cup5	228,913	93,103	7,450	696,447	515,111	8	74%	67	12	2
indA-cup6	229,167	102,538	7,734	702,263	474,838	6	75%	83	6	2

Table S3: The functional genomics experiments used in this study with their total coverage, related to STAR Methods and Figure 4.

Individual	ENCODE ID/Source	Experiment	# of Reads	Read Length
NA12878	IkG	WGS	757,704,193	255
NA12878	IkG	WES	212,461,381	76
NA12878	Rao et al. 2014 (?)	Hi-C exp 1 PE1	219,616,072	101
NA12878	Rao et al. 2014	Hi-C exp 1 PE2	220,087,882	101
NA12878	Rao et al. 2014	Hi-C exp 2 PE1	448,843,710	101
NA12878	Rao et al. 2014	Hi-C exp 2 PE2	451,088,484	101
NA12878	Rao et al. 2014	Hi-C exp 3 PE1	536,684,803	101
NA12878	Rao et al. 2014	Hi-C exp 3 PE2	536,101,709	101
NA12878	ENCSR000CVT	Total RNA-Seq	227,501,266	202
NA12878	ENCSR000COQ	PolyA RNA-Seq	267,602,146	76
NA12878	ENCSR000AJA	Single-cell RNA-Seq1	38,377,124	100
NA12878	ENCSR000AJH	Single-cell RNA-Seq2	47,896,396	100
NA12878	ENCSR000AKF	H3K4me1	42,763,056	36
NA12878	ENCSR145XQO	HDGF	41,626,373	101
NA12878	ENCSR387QUV	RELB	25,652,682	101
NA12878	ENCSR000DZN	CTCF-Snyder	25,463,397	36
NA12878	ENCSR000AKA	H3K4me3	20,221,959	36
NA12878	ENCSR000DYS	JUND	18,701,295	36
NA12878	ENCSR000AOW	H3K79me2	16,073,184	36
NA12878	ENCSR000AKE	H3K36me3	15,239,685	51
NA12878	ENCSR000AOV	H2AFZ	14,724,790	36
NA12878	ENCSR000AOX	H3K9me3	14,049,420	36
NA12878	ENCSR000AKB	CTCF-Broad	11,026,086	51
NA12878	ENCSR000BIF	rnap2	10,428,778	36
NA12878	ENCSR000AKC	H3K27ac	10,410,928	51
NA12878	ENCSR000AKG	H3K4me2	9,815,194	51
NA12878	ENCSR000AKI	H4K20me1	9,757,368	51
NA12878	ENCSR000AKD	H3K27me3	8,454,639	51
NA12878	ENCSR000AKH	H3K9ac	7,981,456	51
NA12878	ENCSR000DKV	CTCF-Iyer	7,614,943	35
NA12878	ENCSR000BGD	rnap2	7,516,461	36
NA12878	ENCSR000BGR	PBX3	6,119,046	36
NA19239	ENCSR018VOS	ChIA-PET (H3K4me1)	335,232,702	PE 101
NA19239	ENCSR332ZHA	ChIA-PET (H3K4me2)	289,328,492	PE 101
NA19239	ENCSR952NXC	ChIA-PET (H3K4me3)	322,739,907	PE 101
NA19239	ENCSR761FUE	ChIA-PET (H3K27ac)	271,351,477	PE 101
NA19239	ENCSR000DLE	CTCF-Iyer	9,999,915	36
NA19238	ENCSR823TEV	ChIA-PET (H3K4me1)	286,387,111	PE 101
NA19238	ENCSR380UPB	ChIA-PET (H3K4me2)	294,881,881	PE 101
NA19238	ENCSR029IXY	ChIA-PET (H3K27ac)	289,564,091	PE 101
NA19238	ENCSR527RXH	ChIA-PET (RAD21)	339,707,301	PE 101
NA19238	ENCSR000DLD	CTCF-Iyer	16,368,229	36
NA12812	ENCSR281KLF	Repli-Seq	8,082,874	36
NA12813	ENCSR834FTN	Repli-Seq	9,999,915	36
NA10847	ENCSR000DYO	POLR2A	6,476,857	28
NA10847	ENCSR000DYM	RELA	19,376,644	28
NA18505	ENCSR000EAU	POLR2A	28,951,453	28
NA18505	ENCSR000EAW	RELA	22,274,656	28
NA18526	ENCSR000EAY	POLR2A	5,058,348	28
NA18526	ENCSR000EBA	RELA	6,353,939	28
NA18951	ENCSR000EBC	POLR2A	8,729,371	28
NA18951	ENCSR000EBD	RELA	5,514,493	28
NA19099	ENCSR000EBG	POLR2A	7,759,177	28
NA19099	ENCSR000EBI	RELA	11,961,302	28

Table S4: p-tools performance and associated file sizes, related to Figure 7 and STAR Methods

Experiment	Total RNA-Seq	PolyA RNA-Seq	ChIP-Seq (CTCF)	ChIP-Seq (H3K4me1)
BAM size (bytes)	35,219,346,385	16,301,017,652	954,993,667	2,230,202,265
pBAM size (bytes)	31,986,293,946	14,057,962,755	876,237,603	1,838,044,304
CRAM size (bytes)	21,317,709,703	10,084,587,074	425,193,485	1,120,941,477
pCRAM size (bytes)	20,436,579,905	966,7758,517	378,736,942	883,158,959
.diff size (bytes)	623,943,745	260,978,063	6,991,519	19,438,692
BAM to .diff compression	99.98%	99.98%	99.99%	99.99%
CRAM to .diff compression	96.94%	97.30%	99.98%	99.98%
BAM+hg to pBAM+.diff CPU time	23:21:17	17:05:26	00:37:11	01:08:00
pBAM+.diff+hg to BAM CPU time	12:34:23	04:21:05	00:27:47	00:46:50