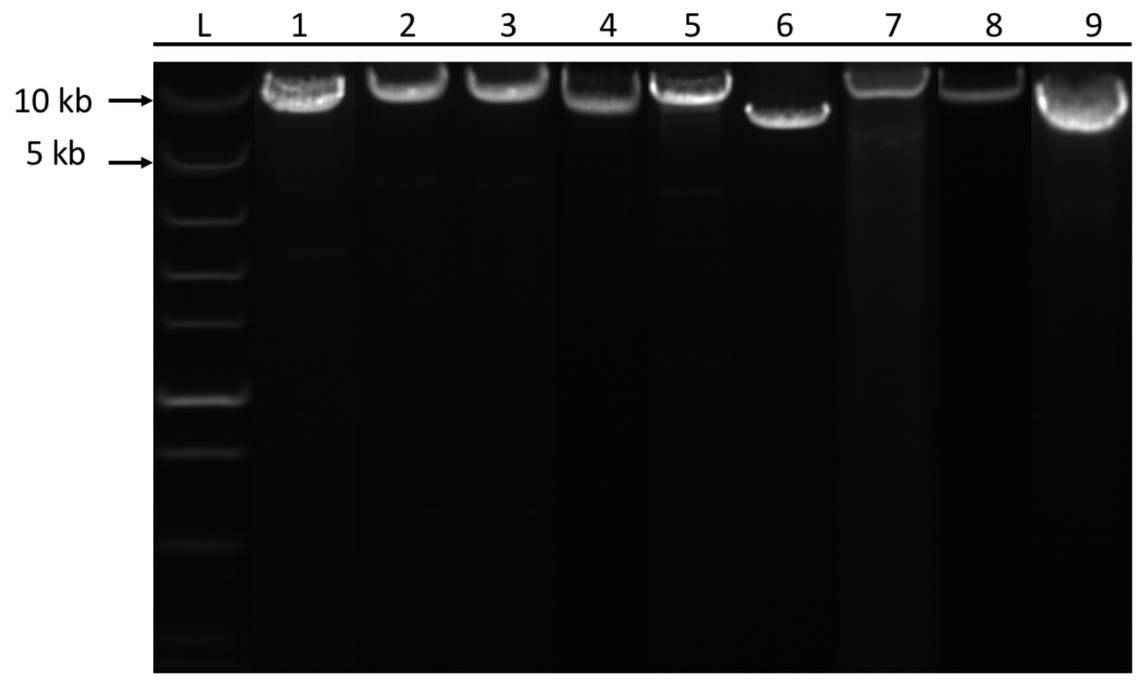
iScience, Volume *23*

Supplemental Information

Precise and Cost-Effective Nanopore Sequencing for Post-GWAS Fine-Mapping and Causal Variant Identification

Tarek Magdy, Hui–Hsuan Kuo, and Paul W. Burridge

A


L
1
2
3
4
5
6
7
8
9
10 kb
5 kb

B

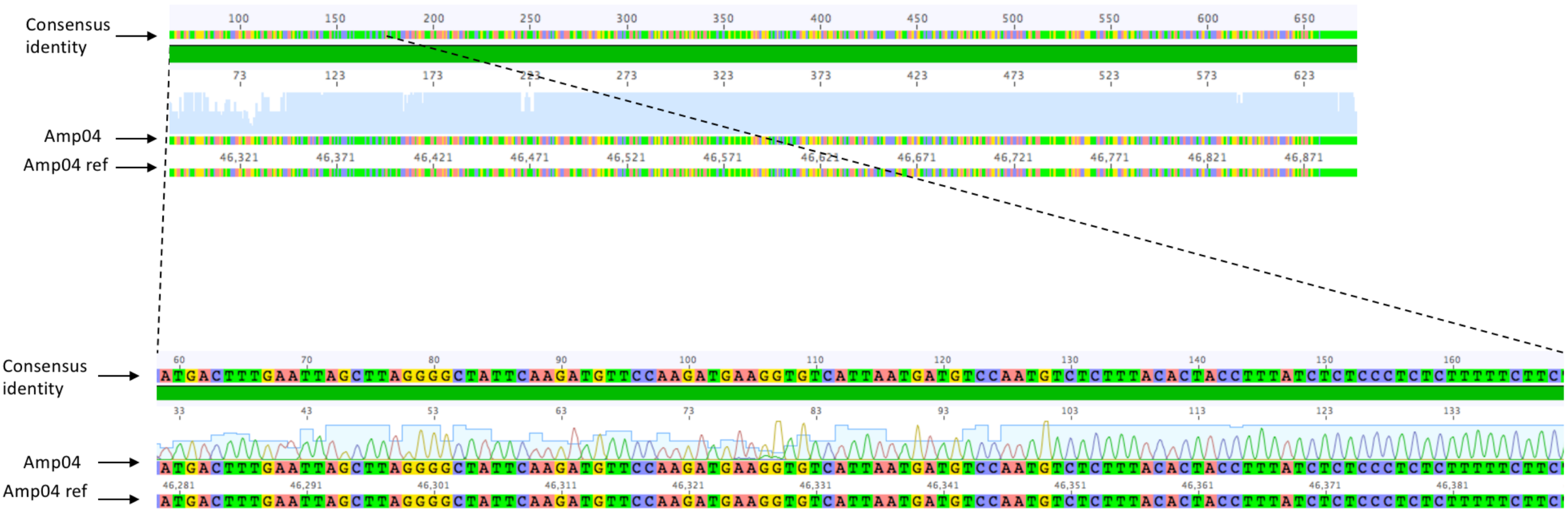
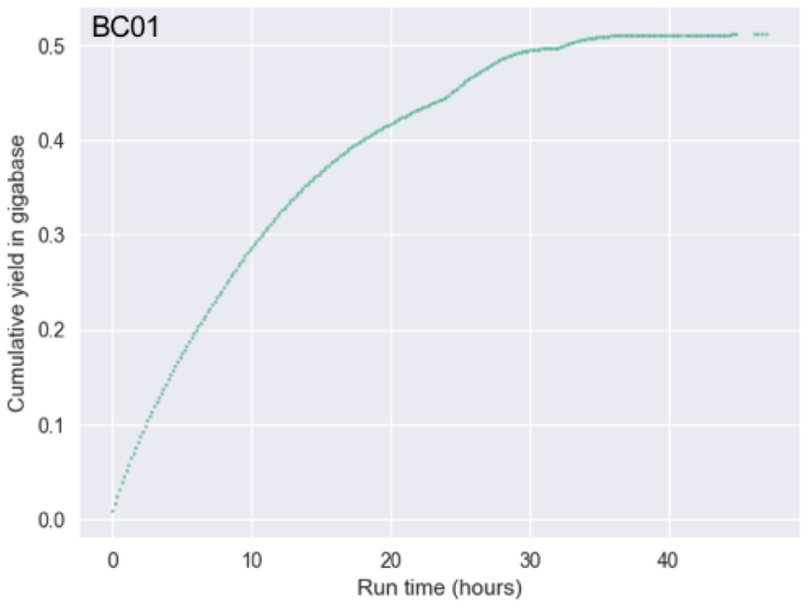
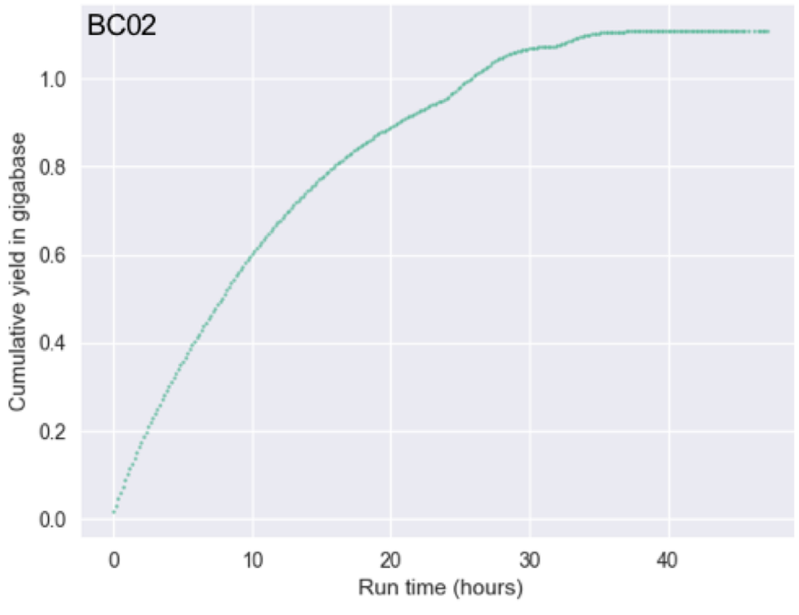
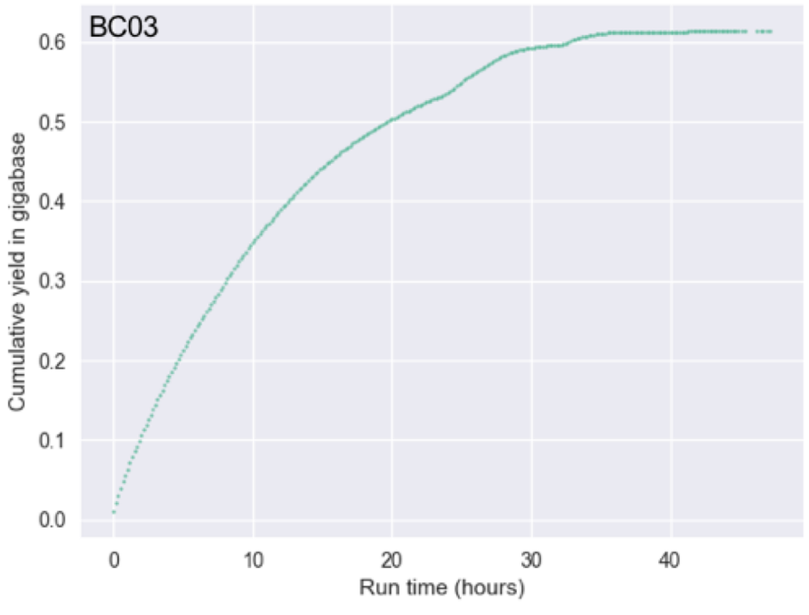

Consensus identity
Amp04
Amp04 ref
Consensus identity
Amp04
Amp04 ref

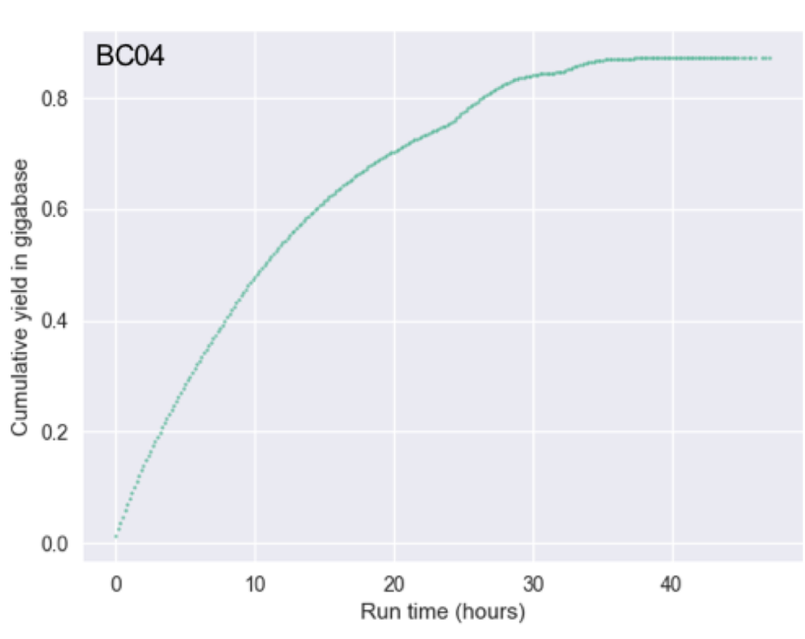
Figure S1

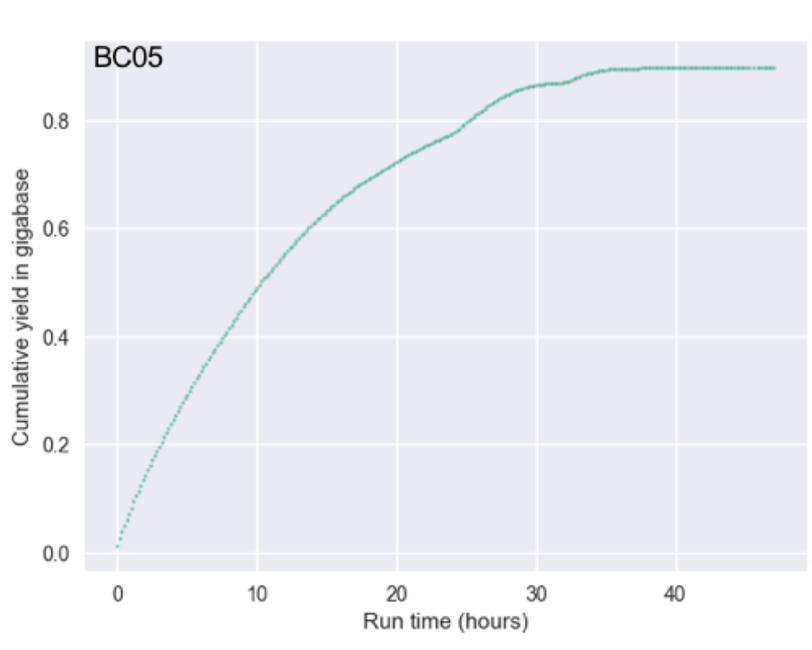
(a)

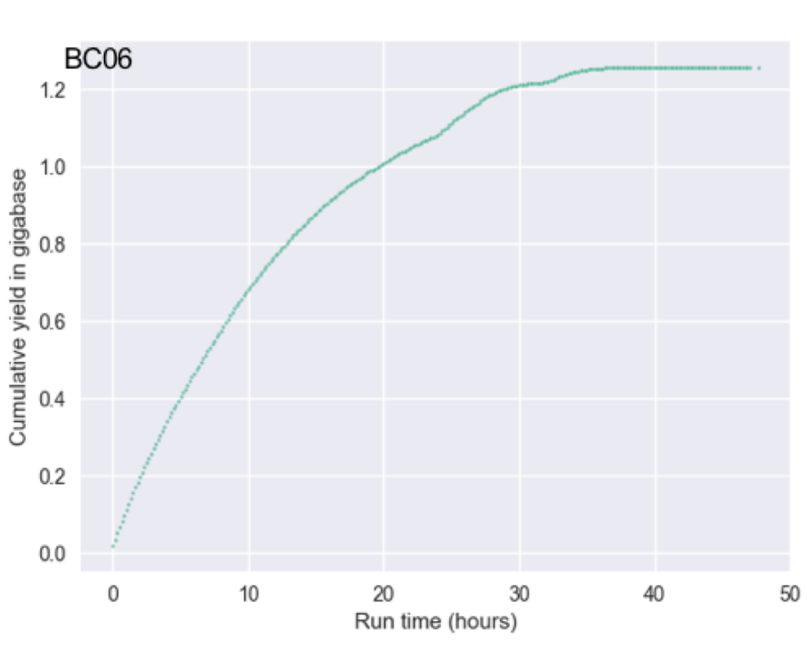

BC01
Cumulative yield in gigabase
Run time (hours)


BC02
Cumulative yield in gigabase
Run time (hours)

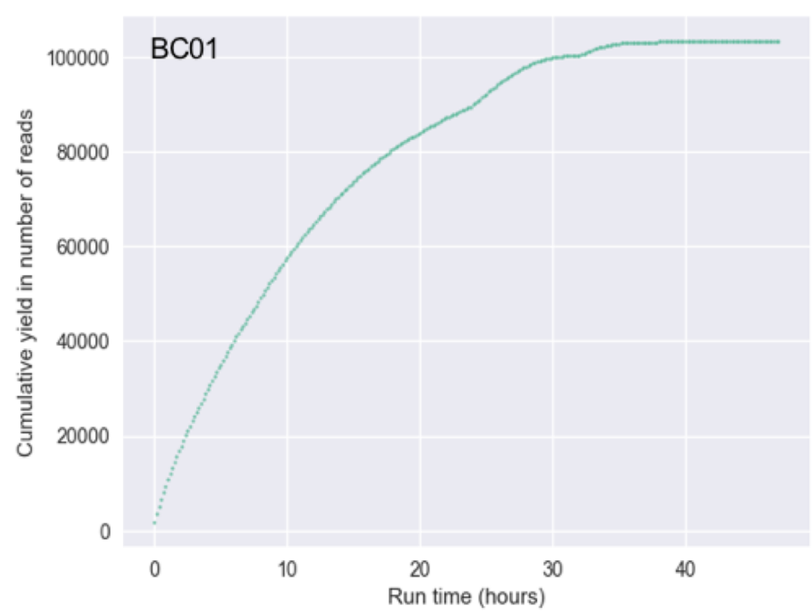

BC03
Cumulative yield in gigabase
Run time (hours)

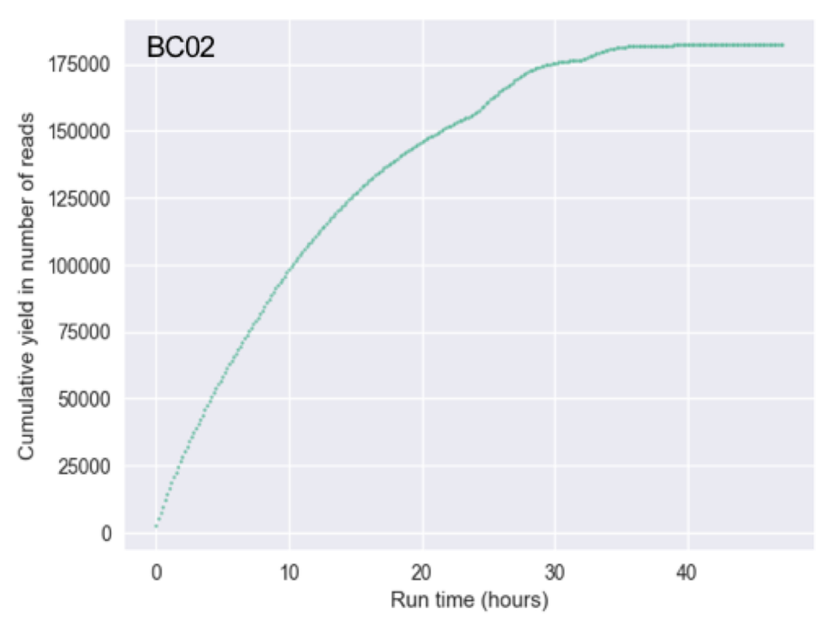

BC04
Cumulative yield in gigabase
Run time (hours)

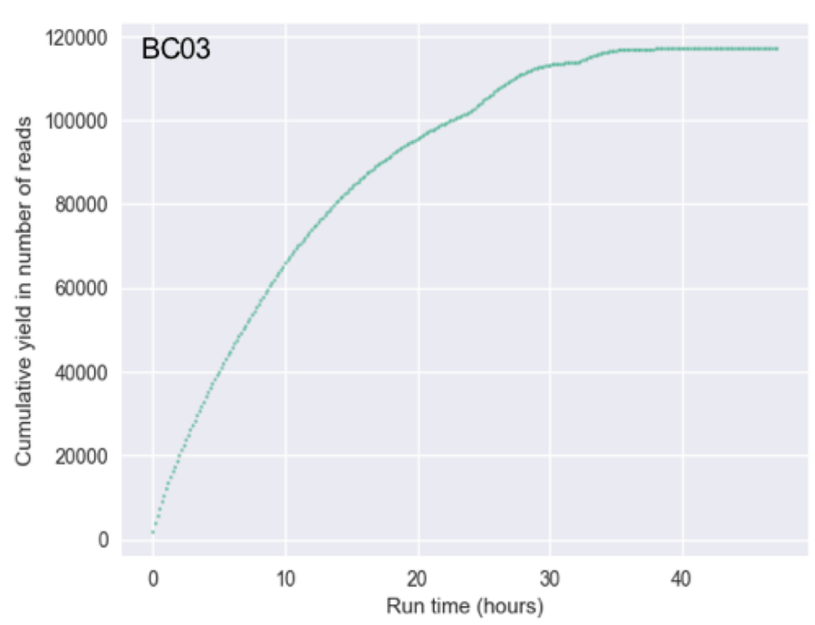

BC05
Cumulative yield in gigabase
Run time (hours)

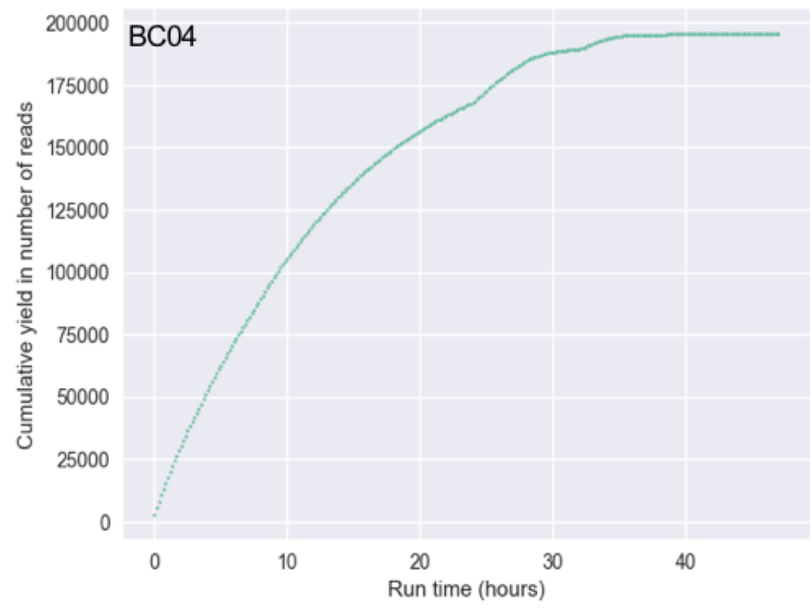

BC06
Cumulative yield in gigabase
Run time (hours)

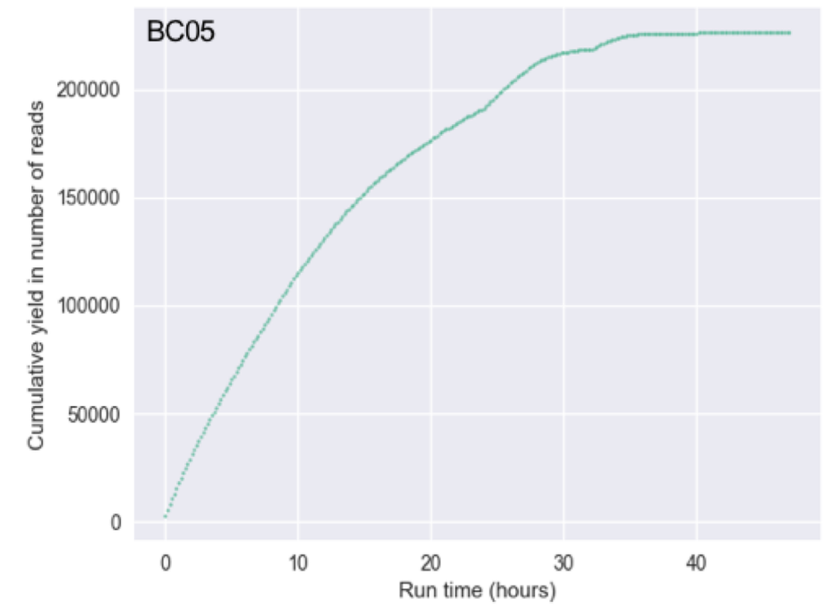
(b)


BC01
Cumulative yield in number of reads
Run time (hours)


BC02
Cumulative yield in number of reads
Run time (hours)


BC03
Cumulative yield in number of reads
Run time (hours)


BC04
Cumulative yield in number of reads
Run time (hours)


BC05
Cumulative yield in number of reads
Run time (hours)

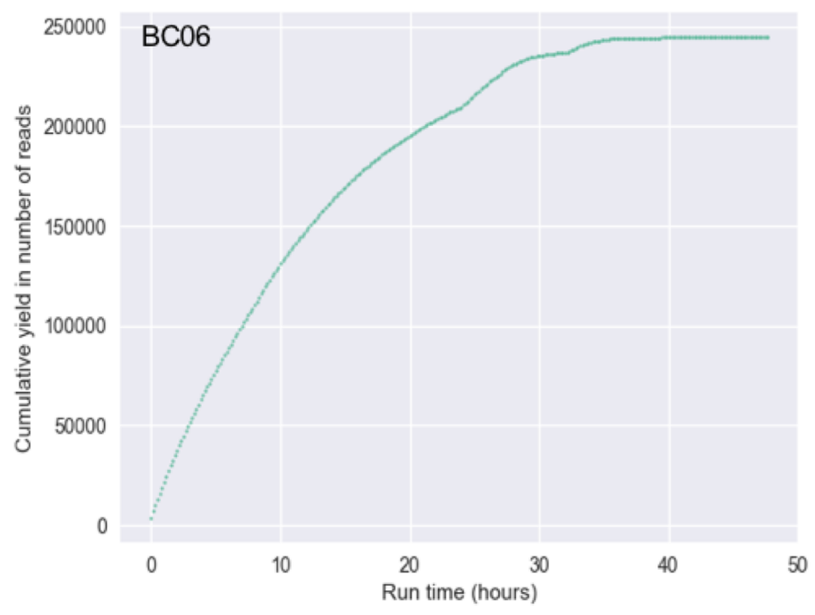

BC06
Cumulative yield in number of reads
Run time (hours)

Figure S2

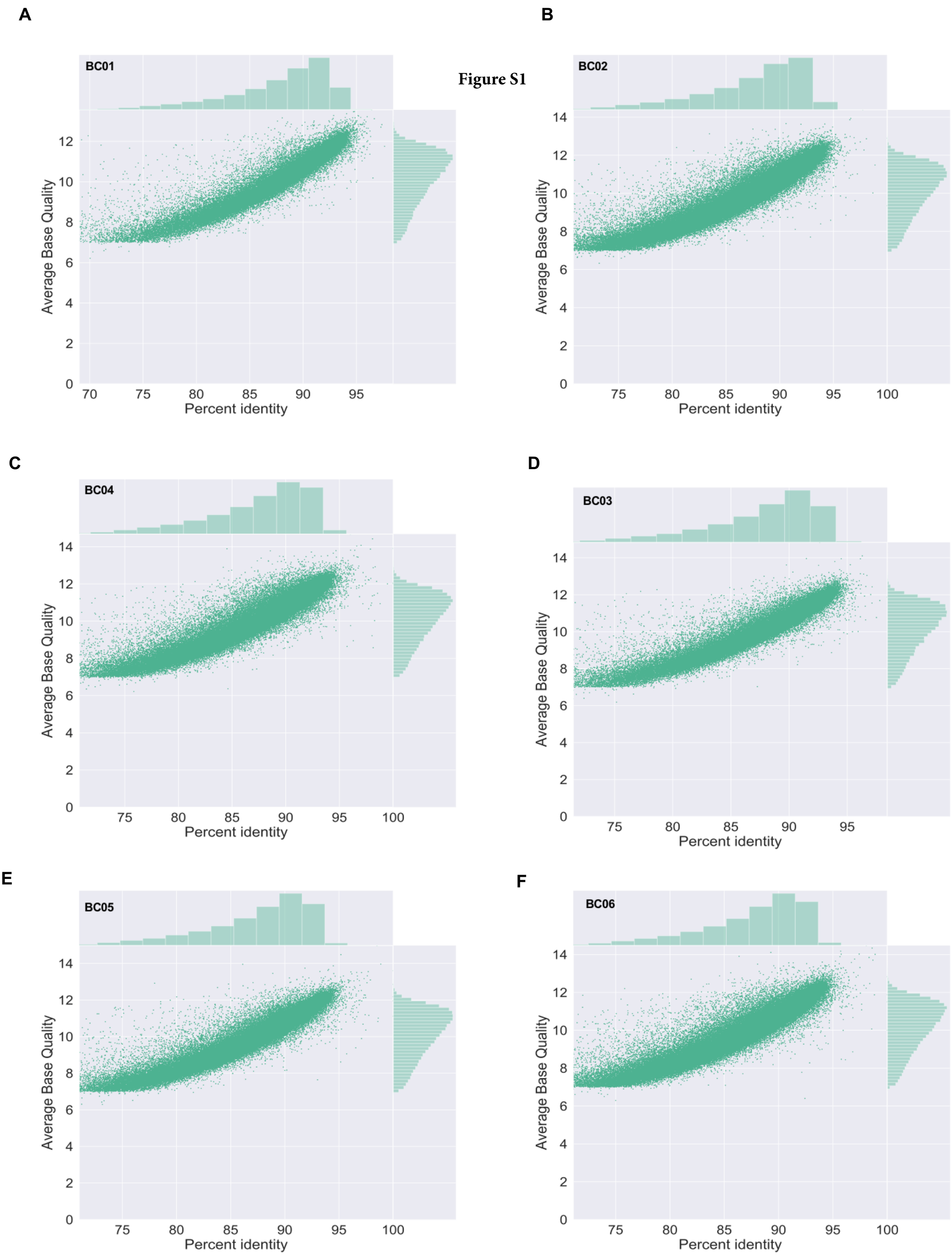


Figure S1

Figure S3

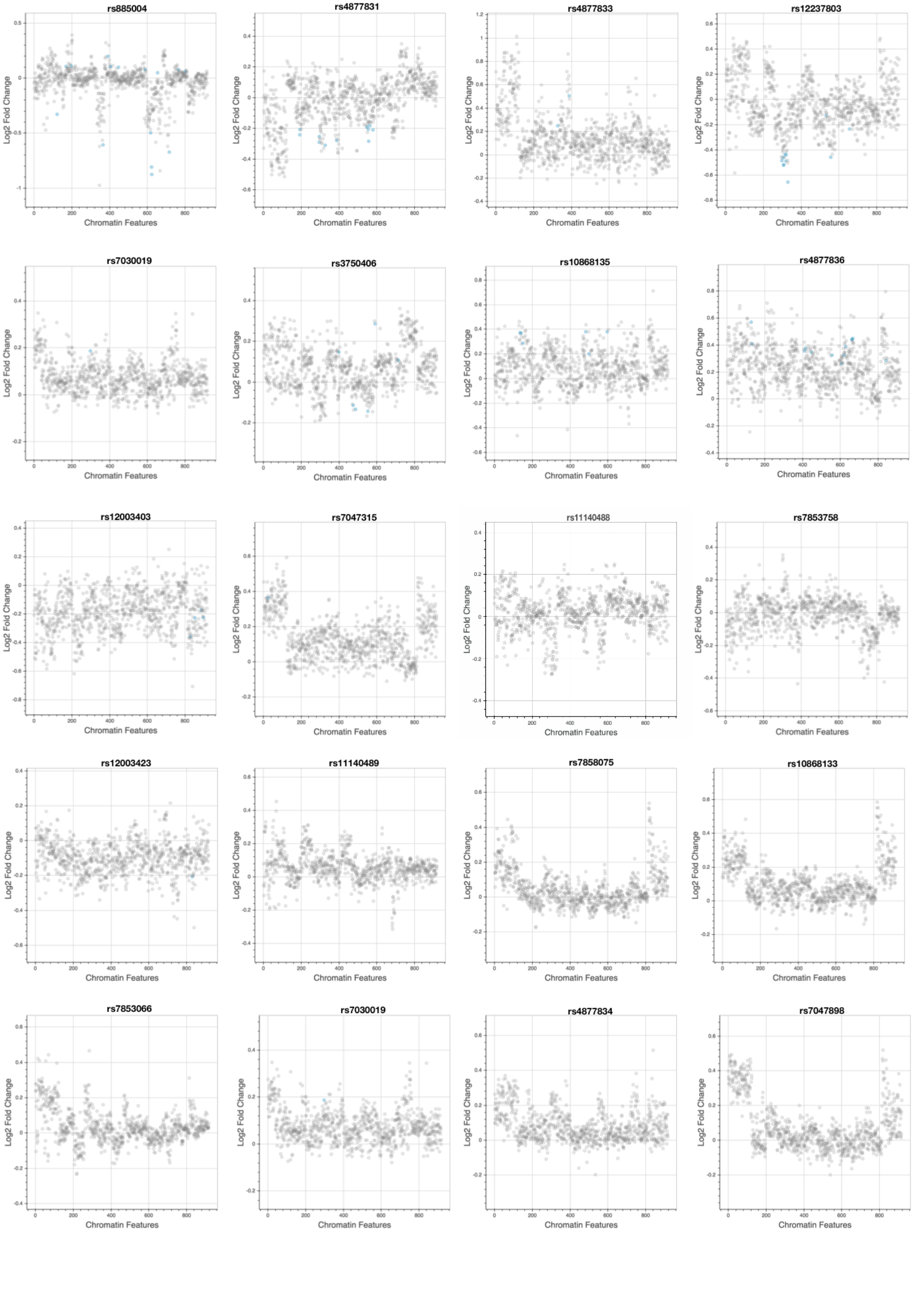

rs885004
rs4877831
rs4877833
rs12237803
rs7030019
rs3750406
rs10868135
rs4877836
rs12003403
rs7047315
rs11140488
rs7853758
rs12003423
rs11140489
rs7858075
rs10868133
rs7853066
rs7030019
rs4877834
rs7047898
Log2 Fold Change
Chromatin Features

Figure S4

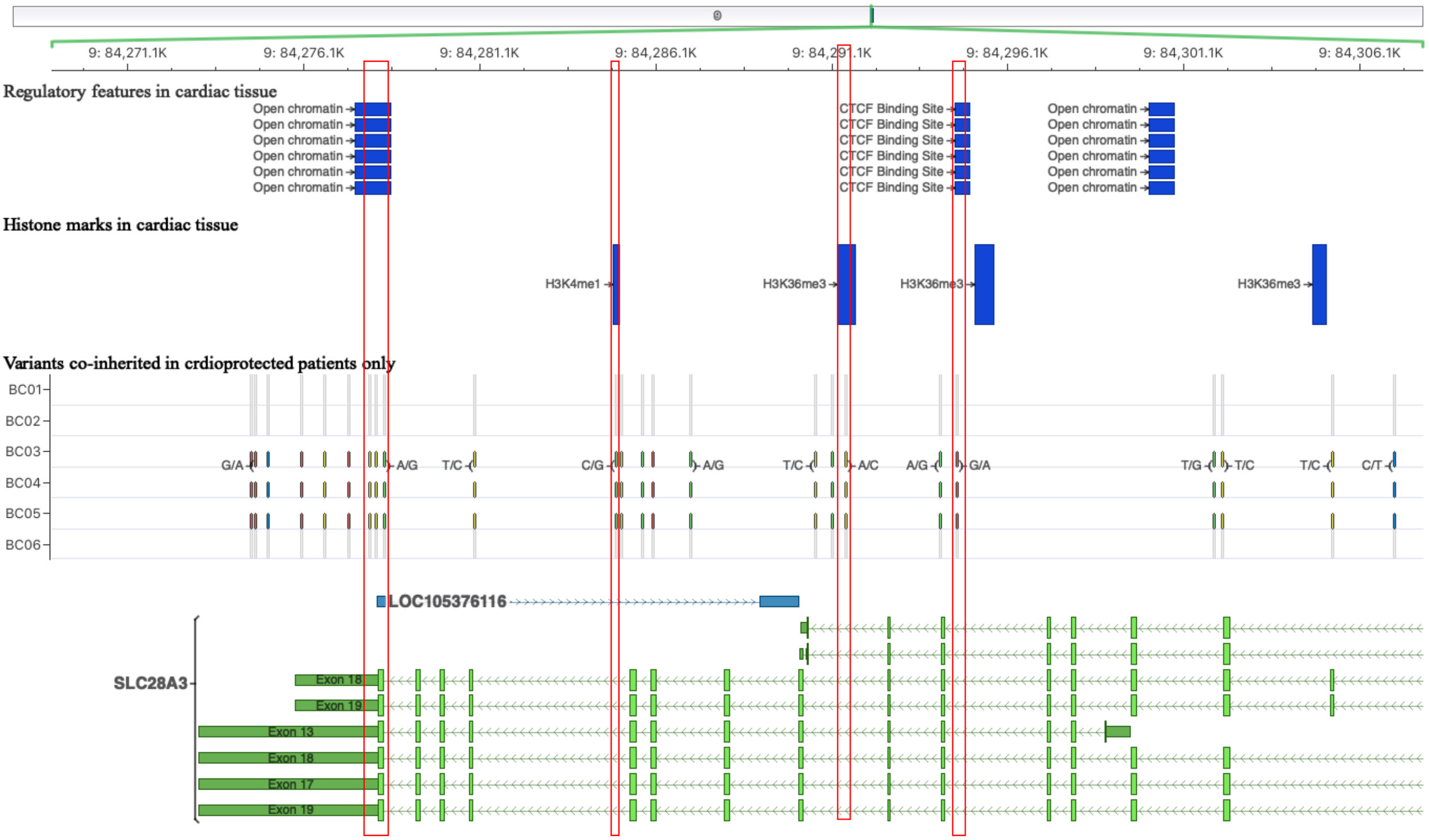

9
9: 84,271.1K
9: 84,276.1K
9: 84,281.1K
9: 84,286.1K
9: 84,291.1K
9: 84,296.1K
9: 84,301.1K
9: 84,306.1K
Regulatory features in cardiac tissue
Open chromatin
CTCF Binding Site
Histone marks in cardiac tissue
H3K4me1
H3K36me3
Variants co-inherited in crdioprotected patients only
BC01
BC02
BC03
BC04
BC05
BC06
G/A
A/G
T/C
C/G
A/C
T/G
C/T
LOC105376116
SLC28A3
Exon 18
Exon 19
Exon 13
Exon 17

Figure S5

Supplementary Figures legends

Figure S1. Generation of *SLC28A3* amplicons, related to Figure 1. (A) Exemplary agarose gel picture for all nine overlapping *SLC28A3* amplicons generated from one sample. L, ladder; 1-9, amplicons one to nine. **(B)** Exemplary pre-nanopore sequencing amplicon validation by sanger sequencing for amplicon number four (Amp04). Top panel shows the first ~ 600 bp of generated amplicon four (AMP04) aligned to its reference sequence (AMP04 ref). Bottom panel shows a zoom-in view for the first ~ 100 bp of generated amplicon four perfectly matching its reference sequence.

Figure S2. Cumulative yield of demultiplexed sequence reads, related to Figure 1. (A) Cumulative sequence yield in gigabase (GB) generated over 48 h. **(B)** Cumulative number of sequence reads generated over 48 h.

Figure S3. Alignment of Nanopore sequence reads, related to Figure 1. Dot plot showing reads percentage identity versus reads average base quality for each study sample. Histogram on the x-axis shows number of reads with relevant percentage identity. Histogram on the y-axis shows number of reads with relevant average base quality.

Figure S4. Effect of candidate SNPs on chromatin features binding sites, related to Figure 3. Effect of each SNP on altering chromatin features (transcription factors, DNase hypersensitive site, and histone marks) binding sites. Log2 fold change measure the fold change in the probability of observing a binding site for relevant chromatin feature between reference and alternative allele for a particular SNP**.**

Figure S5. Candidate SNPs located at regulatory regions in human cardiac tissue, related to Figure 3. Regulatory regions assessed included transcription factors, histone mark, and DNase hypersensitive regions in human cardiac tissue integrated in ensemble regulatory build. SNPs marked by red rectangles are SNPs that are located in regulatory regions binding sites.

Table S1. Inclusion and exclusion criteria, related to Figure 1.

Inclusion Criteria	Exclusion Criteria
Diagnosis of cancer	Unwilling to consent/assent to ≤10 ml blood draw
Treatment with doxorubicin (Adriamycin)	
Age ≤18 years at time treatment	
Documentation of pre–chemotherapy shortening fraction of ≥30%	No documentation of pre-chemotherapy echocardiography shortening fraction
For affected patients only: SF of ≤24% or signs and symptoms of cardiac compromise requiring intervention based on CTCAEv3. Only echos ≥ 21 d after doxorubicin dose are to be considered.	
For control patients: SF of ≥24% and no symptoms of cardiac compromise for at least 5 years after treatment	

Table S2. Doxorubicin–treated patients recruited in this study, related to Figure 1.

ID	Anthracycline	Cardiotoxicity	Gender	Age at Treatment	Cancer diagnosis	Heart radiation	rs7853758
BC01	Yes	Yes	Male	5.1	Wilm's Tumor	Yes	GG
BC02	Yes	Yes	Male	1.6	ALL	No	GG
BC06	Yes	Yes	Female	4.3	ALL	No	GG
BC03	Yes	No	Female	2.7	ALL	No	AG
BC04	Yes	No	Female	2.2	ALL	No	AG
BC05	Yes	No	Male	1.6	ALL	No	AG

Radiation therapy includes significant radiation exposure to the heart or surrounding tissue. This includes mantle and mediastinal radiation, whole–lung radiation, whole–abdomen or upper abdominal radiation, left–side flank radiation and total–body irradiation. SF, shortening fraction; NA, not applicable

Table S3: primers for *SLC28A3* amplicons amplification, related to Figure 1.

Primer ID	Sequence 5'>3'	Direction	Amplicon length (bp)
Amp01_fw	AGTTGCATGTTGCCATTCTG	Forward	9218
Amp01_rw	GTTGCTGTAGCCCTCAGCTC	Reverse	
Amp02_fw	CTCCCCAGGAGTGCAAATAG	Forward	9908
Amp02_rv	TCAAGGGGAATCACTTCAGG	Reverse	
Amp03_fw	TCAAGTTTGCATGATCACACC	Forward	8979
Amp03_rv	CAGGAAATATGGCTTCAGCTC	Reverse	
Amp04_fw	AAGGAAGATCCCACGTTGTG	Forward	9286
Amp04_rv	AAGTGATGCTTCCCATCAGG	Reverse	
Amp05_fw	GCTGTTTGTTGAATCGGATG	Forward	9306
Amp05_rv	TCCAACTGTCTGAGCACCAG	Reverse	
Amp06_fw	TGTTGCAGGTGTTTGGAAAG	Forward	5732
Amp06_rv	ACATTATGAGCCCACCGAAG	Reverse	
Amp07_fw	CGGCCGCTGGTGAGGTCCCCCAA	Forward	8668
Amp07_rv	TGGGCAGTGGTGCTGGCAAGCGT	Reverse	
Amp08_fw	TTGGCAATGTCCGGATTC	Forward	9420
Amp08_rv	TTCCCCTTTCCAGGGATAAC	Reverse	
Amp09_fw	GGACCTCTTCTCCCTGGAAC	Forward	9509
Amp09_rv	AGACCCTAAGGCCTCTCCAG	Reverse	

Table S4: PCR reaction mixture and conditions, related to Figure 1.

Amplicon ID	Composition of reaction mixture	PCR condition
Amp01, Amp02, Amp04, Amp05, and Amp09	10 µl 5X PrimeSTAR GXL Buffer, 4 µl dNTP Mixture (2.5 mM each), 1 µl of 100 µM primer, 300 ng DNA template, and 1 µl PrimeSTAR GXL DNA Polymerase 1.25 U/50, and Sterile distilled water to 50 µl	30 cycles 98℃ 10 sec 60℃ 15 sec 68℃ 10 min Hold at 4 ℃
Amp03, and Amp08	10 µl 5X PrimeSTAR GXL Buffer, 4 µl dNTP Mixture (2.5 mM each), 1 µl of 100 µM primer, 300 ng DNA template, and 1 µl PrimeSTAR GXL DNA Polymerase 1.25 U/50, and Sterile distilled water to 50 µl	30 cycles 98℃ 10 sec 58℃ 15 sec 68℃ 10 min Hold at 4 ℃
Amp06	10 µl 5X PrimeSTAR GXL Buffer, 4 µl dNTP Mixture (2.5 mM each), 1 µl of 100 µM primer, 300 ng DNA template, and 1 µl PrimeSTAR GXL DNA Polymerase 1.25 U/50, and Sterile distilled water to 50 µl	30 cycles 98℃ 10 sec 60℃ 15 sec 68℃ 6 min Hold at 4 ℃
Amp07	10 µl 5X PrimeSTAR GXL Buffer, 4 µl dNTP Mixture (2.5 mM each), 1 µl of 100 µM primer, 300 ng DNA template, and 1 µl PrimeSTAR GXL DNA Polymerase 1.25 U/50, and Sterile distilled water to 50 µl	30 cycles 98℃ 10 sec 66℃ 15 sec 68℃ 6 min Hold at 4 ℃

Table S5: Quality assessment of *SLC28A3* amplicons, related to Figure 1.

Sample ID	Amplicon	Conc (ng/µl)	A260	A280	260/280	260/230
BC01	Amp01	117.2	2.344	1.232	1.9	1.72
	Amp02	7.525	0.15	0.095	1.58	1.69
	Amp03	74.43	1.489	0.771	1.93	1.73
	Amp04	48.37	0.967	0.514	1.88	2.08
	Amp05	128.3	2.566	1.342	1.91	1.94
	Amp06	59.79	1.196	0.649	1.84	1.62
	Amp07	90.3	1.806	0.971	1.86	1.69
	Amp08	7.217	0.144	0.074	1.95	1.52
	Amp09	146.8	2.936	1.546	1.9	2.13
BC02	Amp01	148.3	2.967	1.57	1.89	2.14
	Amp02	80.62	1.612	0.88	1.83	1.72
	Amp03	51.97	1.039	0.549	1.89	1.75
	Amp04	107.7	2.154	1.142	1.89	1.96
	Amp05	98.96	1.979	1.06	1.87	1.95
	Amp06	122.6	2.453	1.288	1.9	1.94
	Amp07	103.4	2.069	1.123	1.84	1.88
	Amp08	19.44	0.389	0.217	1.79	2.04
	Amp09	81.35	1.627	0.855	1.9	2.02
BC06	Amp01	96.78	1.936	1.016	1.91	2.09
	Amp02	60.44	1.209	0.672	1.8	1.57
	Amp03	73.76	1.475	0.796	1.85	1.72
	Amp04	134.1	2.681	1.424	1.88	2.07
	Amp05	84.22	1.684	0.887	1.9	1.92
	Amp06	41.24	0.825	0.439	1.88	1.86
	Amp07	104.5	2.089	1.095	1.91	2.11
	Amp08	45.57	0.911	0.496	1.84	1.79
	Amp09	73.19	1.464	0.81	1.81	1.67
BC03	Amp01	120.1	2.403	1.284	1.87	2.07
	Amp02	40.18	0.804	0.449	1.79	1.61
	Amp03	58.98	1.18	0.631	1.87	1.8
	Amp04	91.2	1.824	0.973	1.87	1.92
	Amp05	114.6	2.293	1.241	1.85	1.72
	Amp06	93.09	1.862	0.979	1.9	2.01
	Amp07	106.4	2.128	1.11	1.92	1.77
	Amp08	64.44	1.289	0.711	1.81	1.91
	Amp09	61.63	1.233	0.651	1.89	2.08
BC04	Amp01	147	2.94	1.551	1.9	1.79
	Amp02	38.41	0.768	0.405	1.9	1.78
	Amp03	35.4	0.708	0.387	1.83	1.56
	Amp04	155.2	3.105	1.664	1.87	1.91
	Amp05	101.4	2.028	1.045	1.94	1.98
	Amp06	46.66	0.933	0.494	1.89	1.88
	Amp07	82.67	1.653	0.911	1.81	1.9
	Amp08	14.07	0.281	0.15	1.88	1.81
	Amp09	61.4	1.228	0.663	1.85	2.11
BC05	Amp01	72.96	1.459	0.771	1.89	1.72
	Amp02	116.3	2.327	1.246	1.87	2.13

	Amp03	125.4	2.509	1.312	1.91	2.2
	Amp04	59.99	1.2	0.644	1.86	1.55
	Amp05	33.31	0.666	0.353	1.89	1.79
	Amp06	141.3	2.826	1.515	1.87	1.73
	Amp07	27.08	0.542	0.295	1.84	1.59
	Amp08	41.91	0.838	0.44	1.91	1.72
	Amp09	90.51	1.81	0.978	1.85	1.83

Table S7: Genotype concordance between Nanopore sequencing and genotyping on InfiniumOmniExpress-24v1-2, related to Figure 2.

Type	Count	Percent of total
REF/REF	421	74.78
REF/ALT1	0	0.00
REF/ALT2	0	0.00
ALT1/REF	46	8.17
ALT1/ALT1	74	13.14
ALT1/ALT2	0	0.00
ALT2/REF	3	0.53
ALT2/ALT1	0	0.00
ALT2/ALT2	19	3.37
Total	563	100.00

The genotypes are: homozygous reference (REF), heterozygous (ALT1), and homozygous non-reference (ALT2). The pairs of genotypes, e.g., REF/REF, on each row represent genotype calls by nanopore sequence and by the Omni chip respectively. REF/REF, ALT1/ALT1, and ALT2/ALT2 indicate concordant genotypes. Line graph represents genotype percentage of total

Table S8: Regulatory properties of *SLC28A3* SNPs coinherited only in cardio protected patients, related to Figure 3.

rs ID	No. of altered chromatin feature binding sites
rs11140490	206
rs4877835	204
rs4877836	141
rs7867504	134
rs4877272	107
rs885004	105
rs12237803	52
rs3750406	41
rs12003403	40
rs10868135	33
rs4877831	32
rs4877833	31
rs10868137	30
rs7853758	11
rs7858075	6
rs7047315	4
rs7853066	4
rs7030019	3
rs12003423	2
rs7047898	2
rs11140488	1
rs4877834	1

Table S9: *SLC28A3* SNPs coinherited only in cardio protected patient affecting chromatin feature binding sites (showing only SNPs with Log2 fold change value >=1), related to Figure 3.

rs ID	Cell type| chromatin| treatment	E-value	Log2 fold change
rs4877272	ECC-1|ERalpha|BPA_100nM	0.01	-1.01
	H1-hESC|TEAD4|None	0.01	-1.60
	NT2-D1|DNase|None	0.03	-1.22
	NHEK|DNase|None	0.03	-1.01
	H7-hESC|DNase|None	0.03	-1.22
	H1-hESC|DNase|None	0.04	-1.22
	RWPE1|DNase|None	0.05	-1.09
rs7867504	GM12878|JunD|None	0.00	-1.11
	PrEC|DNase|None	0.01	-1.55
	GM12878|BATF|None	0.01	-1.34
	GM12865|DNase|None	0.01	-1.04
	GM12864|DNase|None	0.01	-1.00
	SAEC|DNase|None	0.01	-1.62
	HMEC|DNase|None	0.01	-1.12
	HEEpiC|DNase|None	0.01	-1.52
	pHTE|DNase|None	0.01	-1.06
	NHEK|DNase|None	0.01	-1.17
	HRCEpiC|DNase|None	0.02	-1.18
	HRE|DNase|None	0.02	-1.21
	HPDE6-E6E7|DNase|None	0.02	-1.22
	MCF10A-Er-Src|STAT3|4OHTAM_1uM_12hr	0.02	-1.30
	MCF10A-Er-Src|STAT3|EtOH_0.01pct_12hr	0.02	-1.25
	MCF10A-Er-Src|c-Fos|4OHTAM_1uM_12hr	0.02	-1.81
	MCF10A-Er-Src|c-Myc|4OHTAM_1uM_4hr	0.02	-1.06
	MCF10A-Er-Src|STAT3|EtOH_0.01pct_4hr	0.02	-1.18
	MCF10A-Er-Src|STAT3|4OHTAM_1uM_36hr	0.02	-1.21
	MCF10A-Er-Src|STAT3|EtOH_0.01pct	0.02	-1.00

	MCF10A-Er-Src|c-Fos|4OHTAM_1uM_4hr	0.02	-1.67
	RWPE1|DNase|None	0.02	-1.17
	HUVEC|c-Fos|None	0.03	-1.09
	MCF10A-Er-Src|c-Fos|EtOH_0.01pct	0.03	-1.66
	MCF10A-Er-Src|c-Fos|4OHTAM_1uM_36hr	0.03	-1.82
	HMVEC-dBl-Ad|DNase|None	0.03	-1.13
	RPTEC|DNase|None	0.03	-1.01
	HMVEC-dLy-Neo|DNase|None	0.03	-1.02
	WI-38|DNase|4OHTAM_20nM_72hr	0.04	-1.28
	HMVEC-LBl|DNase|None	0.04	-1.18
	HUVEC|c-Jun|None	0.04	-1.01
	HFF-Myc|DNase|None	0.05	-1.03
	NHLF|DNase|None	0.05	-1.16
rs11140490	Melano|DNase|None	0.00	1.05
	HSMM_emb|DNase|None	0.00	1.22
	HSMMtube|DNase|None	0.00	1.42
	NHDF-neo|DNase|None	0.00	1.76
	NHDF-Ad|DNase|None	0.00	1.73
	AG10803|DNase|None	0.00	1.56
	ProgFib|DNase|None	0.00	1.34
	FibroP|DNase|None	0.00	1.25
	HGF|DNase|None	0.00	1.56
	HPdLF|DNase|None	0.00	1.58
	Stellate|DNase|None	0.00	1.36
	HCF|DNase|None	0.00	1.42
	AG09319|DNase|None	0.00	1.46
	HSMM|DNase|None	0.00	1.36
	SK-N-SH|TAF1|None	0.00	1.07
	HFF|DNase|None	0.00	1.35
	BJ|DNase|None	0.00	1.42
	HCM|DNase|None	0.00	1.42

	AG09309|DNase|None	0.00	1.45
	Myometr|DNase|None	0.00	1.16
	AG04449|DNase|None	0.00	1.37
	HPF|DNase|None	0.00	1.51
	AoAF|DNase|None	0.00	1.36
	AoSMC|DNase|None	0.00	1.40
	SKMC|DNase|None	0.00	1.29
	PanIsletD|DNase|None	0.00	1.18
	HMF|DNase|None	0.00	1.42
	HPAF|DNase|None	0.00	1.31
	HConF|DNase|None	0.00	1.37
	HAc|DNase|None	0.00	1.07
	HFF-Myc|DNase|None	0.00	1.08
	HBMEC|DNase|None	0.00	1.29
	WI-38|DNase|4OHTAM_20nM_72hr	0.00	1.17
	NH-A|DNase|None	0.01	1.16
	WI-38|DNase|None	0.01	1.25
	NHLF|DNase|None	0.01	1.15
	AG04450|DNase|None	0.01	1.21
	HCFaa|DNase|None	0.01	1.13
	HNPCEpiC|DNase|None	0.01	1.16
	HVMF|DNase|None	0.01	1.26
	HCPEpiC|DNase|None	0.01	1.05
	HIPEpiC|DNase|None	0.01	1.06
	HAEpiC|DNase|None	0.01	1.10
rs4877835	NHDF-Ad|DNase|None	0.01	1.07
	NHDF-neo|DNase|None	0.01	1.06
	BE2_C|DNase|None	0.01	1.08
	SK-N-SH_RA|DNase|None	0.01	1.08
rs10868137	H1-hESC|TCF12|None	0.00	1.07
	GM12878|ZEB1|None	0.00	1.06

E-value, expect value stands for the significance of each individual chromatin feature predicted score; Log2 fold change, measure the fold change in the probability of observing a binding site for relevant chromatin feature between reference and alternative allele for a particular SNP (Zhou and Troyanskaya, 2015).

Table S10: *SLC28A3* SNPs coinherited only in cardio protected patient located at regulatory regions and histone marks in cardiac tissues, and at transcription factor binding sites using ensemble regulatory build, related to Figure 3.

SNPs	Position	Histone marks in cardiac tissue	Regulatory region in cardiac tissue	Motifs present at SNP locus
rs3750406	84277979	—	Open chromatin	TEAD4::RFX5, FOXJ3::TBX21, SOX6::TBX21, ELK1::FOXI1, ETV2::FOXI1, MGA, TBX2, TBX4, TBX5, ONECUT1, ONECUT2, ONECUT3, HOXB2::EOMES,HOXB2::TBX21, HOXB2::TBX3, MGA::DLX2, MGA::DLX3, MGA::EVX1, PITX1::HES7, E2F3::ONECUT2,TFAP2C::ONECUT2, ETV2::SREBF2, CUX1::SOX15, HOXB13::EOMES, HOXB13::TBX21,HOXD12::TBX21, TBX20, KLF13, KLF14,SREBF2, GLIS1, EOMES, SNAI2, TCF3, TCF4, THRB (n=36)
rs7858075	84278156	—	Open chromatin	TEAD4::FOXI1, IRF3, ETV2::SOX15, POU2F1::FOXO6, POU2F1::DLX2, TEAD4::FOXI1 (n=6)
rs11140490	84278398	—	Open chromatin	CLOCK::FIGLA, TEAD4::EOMES, TEAD4::TBX21, ETV2::DRGX, ZIC1, ZIC3, ZIC4, HOXB2::NHLH1,TEAD4::TCF3, GCM2::SOX15, and TEAD4::FIGLA (n=11)
rs4877831	84284969	H3K4me1	—	—
rs7047898	84291502	H3K36me3	—	—
rs10868137	84294167	—		TFAP2C::DLX3, FOXO1::HOXB13, MGA::DLX3, HOXB2::TCF3 (n=4)
rs885004	84294635	—	CTCF binding site	THRB, TEAD4::CEBPD, ERF::PITX1, ETV2::GSC2, ERF::ONECUT2, ETV2::ONECUT2, FLI1::ONECUT2, POU2F1::DLX2, R, X3::SRF, TEAD4::PAX5, PITX1::HES7, HESX1, LHX9, HOXD12::HOXA3, ZBED1, BARHL2, E2F1, E2F2, E2F3, BARX1, MSX1, MSX2, TBX1, TBX20, HOXB13::EOMES, HOXB13::TBX21, TEAD4::HOXB13, PBX4::HOXA1, PBX4::HOXA10, ONECUT1, ONECUT2, HMX1, HMX2, HMX3, CUX1::SOX15, TFAP2C::ONECUT2 (n=36)
rs4877835	84301936	—		POU2F1::FOXO6, POU2F1::EOMES, CLOCK::BHLHA15, MAX, TFAP4::MAX, HOXD12::EOMES, FOXO1, FOXO3, FOXO4, FOXO6, CTCF, ZNF238, ASCL2, BHLHA15, BHLHE22, BHLHE23,MESP2, MSC, MYF6, NEUROD2, NEUROG2, NHLH1, OLIG1, OLIG2, OLIG3, TCF15, TFAP4, ESRRA, ESRRG, FOXJ2::HOXB13 (n=30)
rs4877836	84302173	—		MYBL1, MYBL2, IRF4, IRF5, IRF8, IRF9, ELK1::FOXI1, ERF::FOXI1, ETV2::FOXI1, ETV5::FOXI1, FLI1::FOXI1, FOXO1::ELF1, FOXO1::ELK1, ELK1::HOXA3 (n=14)

Table S11: eQTL (expression quantitative trait loci) functional annotation of *SLC28A3* SNPs coinherited only in cardio protected patients, related to Figure 3.

SNP Id	*P*-value	NES	Tissue
rs10868133	2.10E-07	-0.22	Cells - Cultured fibroblasts
	4.50E-07	0.21	Thyroid
rs10868135	4.10E-07	-0.22	Cells - Cultured fibroblasts
	0.0000034	0.2	Thyroid
rs10868137	3.80E-07	0.23	Thyroid
	6.70E-07	-0.22	Cells - Cultured fibroblasts
rs11140488	1.60E-07	-0.22	Cells - Cultured fibroblasts
	2.60E-07	0.22	Thyroid
rs11140489	1.50E-07	-0.22	Cells - Cultured fibroblasts
	4.20E-07	0.21	Thyroid
rs11140490	1.40E-07	-0.22	Cells - Cultured fibroblasts
	6.30E-07	0.21	Thyroid
rs12003403	1.60E-07	-0.22	Cells - Cultured fibroblasts
	2.60E-07	0.22	Thyroid
rs12003423	1.60E-07	-0.22	Cells - Cultured fibroblasts
	7.20E-07	0.21	Thyroid
rs12237803	4.70E-08	-0.24	Cells - Cultured fibroblasts
	7.50E-08	0.24	Thyroid
rs3750406	1.40E-07	-0.22	Cells - Cultured fibroblasts
	6.30E-07	0.21	Thyroid
rs4877272	6.40E-08	-0.23	Cells - Cultured fibroblasts
	5.60E-07	0.21	Thyroid
rs4877831	6.00E-09	-0.21	Cells - Cultured fibroblasts
	0.000021	0.16	Thyroid
rs4877833	5.30E-07	-0.21	Cells - Cultured fibroblasts
	8.90E-07	0.21	Thyroid
rs4877834	4.70E-08	0.24	Thyroid
	7.40E-07	-0.21	Cells - Cultured fibroblasts
rs4877835	4.20E-07	0.23	Thyroid
	5.20E-07	-0.22	Cells - Cultured fibroblasts
rs4877836	3.10E-07	-0.23	Cells - Cultured fibroblasts
	3.30E-07	0.23	Thyroid
rs7030019	1.70E-08	0.25	Thyroid
	8.70E-08	-0.23	Cells - Cultured fibroblasts
	0.000014	0.64	Brain - Amygdala

rs7047315	3.80E-07	0.23	Thyroid
	6.70E-07	-0.22	Cells - Cultured fibroblasts
rs7047898	3.80E-07	0.23	Thyroid
	6.70E-07	-0.22	Cells - Cultured fibroblasts
rs7853066	1.50E-07	0.23	Thyroid
	7.00E-07	-0.21	Cells - Cultured fibroblasts
rs7853758	3.10E-08	0.23	Thyroid
	0.0000019	-0.2	Cells - Cultured fibroblasts
	0.000014	0.61	Brain - Amygdala
rs7867504	0.000003	-0.16	Cells - Cultured fibroblasts
rs885004	1.30E-07	-0.23	Cells - Cultured fibroblasts
	1.90E-07	0.23	Thyroid

NES, normalized effect size; This analysis was done using GTEX eQTL database

Table S12: Quality Control of Genotype Imputation, related to Figure 4.

Posterior probability intervals	Number of Genotypes	% Concordance	Posterior probability intervals	% Called	% Concordance
[0.0-0.1]	0	0	[>= 0.0]	100	95.7
[0.1-0.2]	0	0	[>= 0.1]	100	95.7
[0.2-0.3]	0	0	[>= 0.2]	100	95.7
[0.3-0.4]	0	0	[>= 0.3]	100	95.7
[0.4-0.5]	1	0	[>= 0.4]	100	95.7
[0.5-0.6]	0	0	[>= 0.5]	99.8	96.4
[0.6-0.7]	0	0	[>= 0.6]	99.8	96.4
[0.7-0.8]	3	0	[>= 0.7]	99.8	96.4
[0.8-0.9]	4	100	[>= 0.8]	99.8	97.8
[0.9-1.0]	130	100	[>= 0.9]	99.5	98.5

Table S14: Cost estimates of Nanopore candidate loci sequencing compared to Illumina targeted sequencing, , related to Figure 4.

Steps	**Cost per ~100kb / sample in USD**							
	Nanopore MinION Sequencing				**Illumina Targeted Sequencing**			
Number of multiplexed samples	*n* = 6	*n* = 12	*n* = 24	*n* = 96	*n* = 6	*n* = 12	*n* = 24	*n* = 96
Library preparation	16.6[a]	8.33[a]	4.16[a]	1.04[a]	35.83[b]	35.83[b]	35.83[b]	35.83[b]
Samples barcoding	4[c]	4[c]	4[c]	1.77[c]	7[d]	7[d]	7[d]	7[d]
Sequencing	5.20[e]	5.20[e]	5.20[e]	5.20[e]	160[f]	80[f]	40[f]	10[f]
Total cost	**25.8**	**17.54**	**13.37**	**8.02**	**313.67**	**178.25**	**110.54**	**57.99**

[a]Library preparation using SQK-LSK109 kit. [b]Library preparation using Nextera XT DNA Library Preparation Kit. [c] Sample barcoding using EXP-NBD104 or EXP-NBD 114 or EXP-PBC096 kits. [d] Sample barcoding using Illumina Nextera DNA Unique Indexes. [e]Sequencing using Nanopore MinION (1 flow cell). Unlike Miseq, real-time data analysis provides the privilege of controlling the utilization of the Minion flow cell. Once sufficient coverage is reached, the sequencing is stopped, flow cell is washed and stored until the next experiment. [f]Sequencing using Miseq 600-cycle (one flow cell). The cost of the devices and kits are adopted from https://nanoporetech.com/products/comparison and https://www.illumina.com/products.html

Transparent Methods

Patient recruitment and hiPSC-CM generation.

SNP rs7853758 is associated with cardioprotective effect after doxorubicin treatment in a Canadian pediatric patient population, and has been validated in both Dutch and US patient populations (Aminkeng et al., 2015). With written consent, six well–phenotyped, doxorubicin–exposed patients from the Canadian cohort were specifically re–recruited according to the original inclusion criteria (**Table S1 and S2**). This study was approved by the individual ethics committees or institutional review boards of the universities and institutions where patients were enrolled. Written informed consent or assent was obtained from patients or their parents or legal guardians in accordance with the Declaration of Helsinki as revised in 2008. Peripheral blood was drawn from three pediatric patients who carried a heterozygous SNP and were protected from doxorubicin-induced cardiotoxicity (BC03, BC04, and BC05), and three control patients who did not carry this protective SNP and developed cardiotoxicity upon same doxorubicin therapy (BC01, BC02, and BC06). Detailed patient demographics and phenotype, including age, sex, ethnicity, type of cancer, treatment regime and cardiovascular function, are all well documented (Visscher et al., 2012, Visscher et al., 2013). Genomically stable hiPSC lines from each individual have been established after non-integrating (Sendai virus-based) reprogramming (Burridge et al., 2016, Burridge et al., 2015).

DNA extraction and purification.

DNA was isolated from six patients derived human induced pluripotent stem cells, using QuickExtract DNA Extraction Solution (Epicenter) according to manufacturer protocol. Isolated DNA was then purified using Genomic DNA Clean & Concentrator-10 (Zymo research) according to manufacturer protocol.

SLC28A3 locus amplification and amplicons validation.

About 75 kb located on chr9: 84,274,029-84,349,802 (NC_000009.12, GRCh38.p7) encompassing the coding region of *SLC28A3* gene (chr9: 84,340,634-84,278,218) plus 9 kb and 5 kb at the 5'UTR and 3`UTR, respectively were amplified using long range PCR. A set of primer pairs were designed to amplify nine overlapping amplicons covering the target region whereas, the

length of amplicons ranged between 5732 and 9908 bp (**Table S3**). Generation of overlapping amplicons help compensate for the low depth of coverage at the start and the end of each sequence read. Using ~200 ng of DNA per reaction, amplicons were amplified using PrimeSTAR GXL DNA Polymerase (Takara) via three steps-PCR. PCR conditions were optimized for each amplicon to avoid any unspecific amplification. PCR reaction mixture components and cycling conditions are mentioned in **Table S4.** Amplified amplicons were then purified using PureLink PCR Purification Combo Kit (Thermo Scientific) to get rid of contaminants that could damage the pores of the Nanopore flow cell, and eventually decrease the number of the sequencing reads.

Amplicon validation prior sequencing

PCR products (amplicons) were run on 1% agarose gel and visualized by staining with GelGreen Nucleic Acid Stain (Biotium) (**Figure S1A**). Gel bands corresponding to target amplicons size were confirmed for all amplified amplicons. For further confirmation, ~ 1 kb at the start and the end of each purified amplicon were then Sanger sequenced, and *in silico* aligned to its relevant reference sequence (**Figure S1B**). The quality and concentration of the generated amplicons were assessed using NanoDrop 8000 and Qubit 3.0 fluorometer, respectively **(Table S5)**. It is important to generate amplicons with reasonable purity to avoid ruining the pores of the flow cell which decreases the number of generated sequence reads. Thus, amplicons with 260/280 and 260/230 absorbance ratios of less than 1.8 and 1.5, respectively were excluded and regenerated.

MinION library preparation and flow cell loading

Library preparation was done using ligation sequencing (Oxford, Nanopore, SQK-LSK108) and 1D Native barcoding (Oxford, Nanopore, EXP-NBD103) kits. All amplicons were pooled together in an equimolar amount and repaired using NEBNext End repair / dA-tailing Module (New England Biolabs, E7546). Reaction mix was prepared by adding 45 µl eluted DNA to 7 µl Ultra II End-prep reaction buffer, 10 µl Ultra II End-prep enzyme mix, and 5 µl nuclease-free water. Reaction mix was then incubated for 5 min at 20 °C followed by 5 min at 65 °C. DNA was then purified using AMPure XP beads (see above). Finally, 25 µl clear elute was transferred into DNA LoBind tube.

Each sample was then barcoded using 1D Native barcoding (Oxford, Nanopore, EXP-NBD103), by adding 2.5 µl native Barcode to 22.5 µl end-prepped DNA, and 25 µl Blunt/TA

Ligase Master Mix (New England Biolabs, M0367). Reaction mix was then incubated for 10 min at room temperature, DNA was then purified using AMPure XP beads (see above), and 26 µl of clear elute was transferred into Eppendorf DNA LoBind tube.

Barcoded samples were pooled in an equimolar amount to a final concentration of 700 ng, then diluted by adding 24 µl nuclease free water. Adapters were then ligated using NEBNext Quick Ligation Module (New England Biolabs, E6056). Pooled DNA (700 ng) was then mixed with 20 µl Barcode Adapter Mix, 20 µl NEBNext Quick Ligation Reaction Buffer, and 10 µl Quick T4 DNA Ligase. Reaction mix was then incubated for 10 min at room temperature, and DNA was then purified by adding 62 µl AMPure XP beads Beckman Coulter, A63880), incubated on a hula mixer at room temperature for 5 min, spun down, and pelleted on a magnet, and supernatant was discarded. Beads were then resuspended in 140 µl Adapter Bead Buffer (ABB) by flicking the tube, pelleted on magnet, and supernatant was discarded (resuspension step was repeated). Pellet was resuspended in 15 µl Elution Buffer, incubated for 10 min at room temperature, pelleted on magnet until the elute is clear, and finally 15 µl clear elute was transferred into Eppendorf DNA LoBind tube.

Priming mix was prepared by adding 576 µl RBF to 624 µl nuclease-free water, then 800 µl priming mix was loaded on the flow cell using priming port dropwise to avoid the introduction of air bubbles. Five minutes later, SpotON sample cover on MinION was opened and 200 µl priming mix was loaded. DNA library was prepared for loading by adding 12 µl DNA library to 35 µl RBF, 25.5 µl LLB, and 2.5 µl nuclease-free water. DNA library was gently mixed, loaded on the flow cell (FLO-MIN 106 R9 version, FAF19356) through SpotON port. Library was then sequenced for 48 hours with live base-calling.

Raw sequencing data and SNPs functional analysis

Raw barcoded sequence reads were demultiplexed into six fastaq files using Porechop (Wylie et al., 1996). Quality of demultiplexed sequence reads were assessed using Nanopack (De Coster et al., 2018). Sequence reads were then aligned to reference human genome (GRCh38.p92) using minimap2 (Li, 2018) "-ax map-ont", sam files were then sorted and converted into bam files using SAMtools (Li et al., 2009). Bam files were down-sampled using SAMtools "-s 0.1 to -s 0.9", and the quality of aligned reads were assessed using Nanopack (De Coster et al., 2018). Depth of

coverage analysis was done using deepTools2 (Ramirez et al., 2016). Sequence reads were indexed and variants were called using Nanopolish (Loman et al., 2015). Variant call format files containing called SNPs were processed and analyzed using several tools including VCFtools (Danecek et al., 2011), SnpSift (Cingolani et al., 2012), and BCFtools (Narasimhan et al., 2016). SNPs functional annotation analysis was done using DeepSEA (Zhou and Troyanskaya, 2015), R (RCoreTeam) and BiomaRt (Durinck et al., 2009) Bioconductor package that includes multiple ensemble gene regulation database. Conservation analysis was done using SnpSift (Cingolani et al., 2012) and PhastCons dataset that includes genome-wide multiple alignments with other 99 vertebrate species. (http://hgdownload.cse.ucsc.edu/goldenpath/hg38/phastCons100way)

Genotype imputation analysis

First the vcf file containing the original GWAS dataset (23 genotypes in six samples) (Visscher et al., 2012) was converted to plink bed format that is compatible with the downstream analysis in SHAPEIT, using Plink " --make-bed " (Purcell et al., 2007). We then checked the alignment of SNPs between the GWAS dataset and the1000 genomes phase I reference panel using SHAPEIT "shapeit -check" (Delaneau et al., 2011). GWAS genotypes were then phased using SHAPEIT using 1000 genomes phase I haplotype reference panel. Additional SNPs that are not present on the GWAS genotyping platform were then imputed using the pre-phased GWAS genotypes with 1000 genomes phase I haplotype reference panel using " impute2 use_prephased_g -known_haps_g " in IMPUTE2 (Howie et al., 2009).

Supplementary References

Aminkeng, F., et al. 2015. A coding variant in RARG confers susceptibility to anthracycline-induced cardiotoxicity in childhood cancer. *Nat Genet,* 47**,** 1079-84.

Burridge, P. W., Diecke, S., Matsa, E., Sharma, A., Wu, H. & Wu, J. C. 2016. Modeling Cardiovascular Diseases with Patient-Specific Human Pluripotent Stem Cell-Derived Cardiomyocytes. *Methods Mol Biol,* 1353**,** 119-30.

Burridge, P. W., Holmstrom, A. & Wu, J. C. 2015. Chemically Defined Culture and Cardiomyocyte Differentiation of Human Pluripotent Stem Cells. *Curr Protoc Hum Genet,* 87**,** 21 3 1-15.

Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M. & Lu, X. 2012. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front Genet,* 3**,** 35.

Danecek, P., et al. 2011. The variant call format and VCFtools. *Bioinformatics,* 27**,** 2156-8.

De Coster, W., D'hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics,* 34**,** 2666-2669.

Delaneau, O., Marchini, J. & Zagury, J. F. 2011. A linear complexity phasing method for thousands of genomes. *Nat Methods,* 9**,** 179-81.

Durinck, S., Spellman, P. T., Birney, E. & Huber, W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols,* 4**,** 1184-1191.

Howie, B. N., Donnelly, P. & Marchini, J. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet,* 5**,** e1000529.

Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics,* 34**,** 3094-3100.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-9.

Loman, N. J., Quick, J. & Simpson, J. T. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods,* 12**,** 733-5.

Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C. & Durbin, R. 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics,* 32**,** 1749-51.

Purcell, S., et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**,** 559-75.

Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dundar, F. & Manke, T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res,* 44**,** W160-5.

Rcoreteam R: A Language and Environment for Statistical Computing.

Visscher, H., et al. 2012. Pharmacogenomic prediction of anthracycline-induced cardiotoxicity in children. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology,* 30**,** 1422-1428.

Visscher, H., et al. 2013. Validation of variants in SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. *Pediatric Blood & Cancer,* 60**,** 1375-1381.

Wylie, C., Kofron, M., Payne, C., Anderson, R., Hosobuchi, M., Joseph, E. & Heasman, J. 1996. Maternal beta-catenin establishes a 'dorsal signal' in early Xenopus embryos. *Development,* 122**,** 2987-96.

Zhou, J. & Troyanskaya, O. G. 2015. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods,* 12**,** 931-4.