

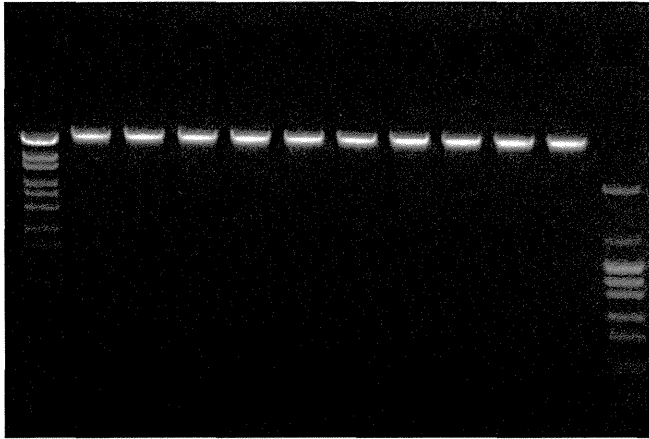
Table 9 標準ゲノムDNAにおける二本鎖DNA定量結果

サンプル	液量(ul)	濃度(ng/ul)	総量(ug)	吸光定量値/蛍光定量値
Standard DNA1	20	59.8	1.2	1
Standard DNA2	20	57.9	1.2	1
Standard DNA3	20	64.8	1.3	0.9
Standard DNA4	18	61.2	1.1	0.9
Standard DNA5	18	59.3	1.1	1
Standard DNA6	18	61.2	1.1	1
Standard DNA7	18	64.2	1.2	0.9
Standard DNA8	18	63.6	1.1	1
Standard DNA9	18	58.4	1.1	1
Standard DNA10	18	61.8	1.1	0.9
Standard DNA11	18	50.6	0.9	1.1
Standard DNA12	18	59	1.1	1
Standard DNA13	18	58.7	1.1	1
Standard DNA14	18	56.9	1	0.9
Standard DNA15	18	58	1	1
Standard DNA16	18	60.6	1.1	0.9
Standard DNA17	18	59.6	1.1	0.9
Standard DNA18	18	54.9	1	1
Standard DNA19	18	59.7	1.1	0.9
Standard DNA20	18	61.2	1.1	0.9
Total	510	40.8	20.8	-

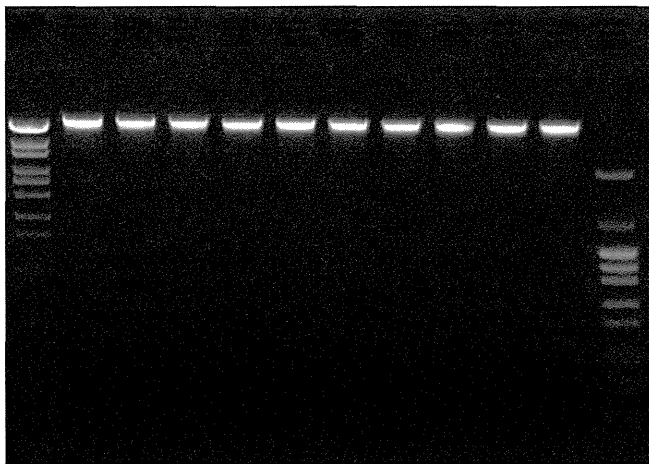
Quant-iT dsDNA BR Assay Kit (Thermo Fisher Scientific) を用いた蛍光定量が行われた。「吸光定量値/蛍光定量値」は核酸定量結果と二本鎖DNA 定量結果の乖離を示す(推奨値 ≤ 3)。評価項目として、解析に必要とされる核酸濃度、総量、及び核酸定量結果 (Table 8) と二本鎖DNA定量結果の乖離データを指標にして評価した。

Table 10 標準ゲノムDNAの定量結果 (Nanodrop を用いた吸光定量)

サンプル	液量(ul)	濃度(ng/ul)	総量(ug)	260/280	260/230
Standard DNA1	20	60.3	1.2	2.2	1.3
Standard DNA2	20	57.6	1.2	2.1	1.3
Standard DNA3	20	57.9	1.2	2.2	1.3
Standard DNA4	18	55.8	1	2.2	1.2
Standard DNA5	18	56.4	1	2.2	1.3
Standard DNA6	18	59.7	1.1	2.2	1.2
Standard DNA7	18	60.3	1.1	2	1.3
Standard DNA8	18	62.1	1.1	2.2	1.3
Standard DNA9	18	57	1	2.1	1.2
Standard DNA10	18	55.8	1	2.1	1.2
Standard DNA11	18	57.3	1	2.2	1.3
Standard DNA12	18	56.1	1	2.2	1.3
Standard DNA13	18	56.4	1	2.2	1.3
Standard DNA14	18	54	1	2.1	1.3
Standard DNA15	18	56.4	1	2.2	1.2
Standard DNA16	18	54.9	1	2.2	1.2
Standard DNA17	18	52.5	0.9	2.2	1.2
Standard DNA18	18	52.8	1	2.3	1.2
Standard DNA19	18	54.9	1	2.2	1.3
Standard DNA20	18	56.1	1	2.1	1.2

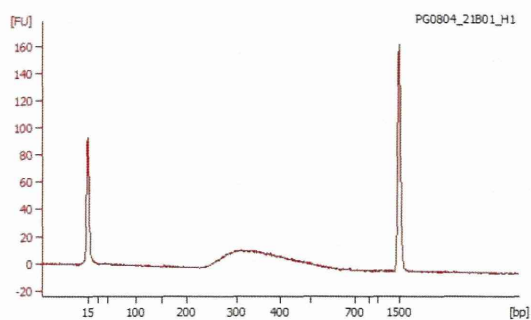


レーン (左から)	サンプル
1	λ -EcoT14 I digest
2	Standard DNA1
3	Standard DNA2
4	Standard DNA3
5	Standard DNA4
6	Standard DNA5
7	Standard DNA6
8	Standard DNA7
9	Standard DNA8
10	Standard DNA9
11	Standard DNA10
12	pHY Marker

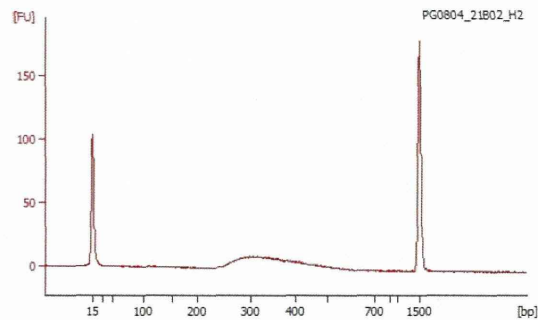


レーン (左から)	サンプル
1	λ -EcoT14 I digest
2	Standard DNA11
3	Standard DNA12
4	Standard DNA13
5	Standard DNA14
6	Standard DNA15
7	Standard DNA16
8	Standard DNA17
9	Standard DNA18
10	Standard DNA19
11	Standard DNA20
12	pHY Marker

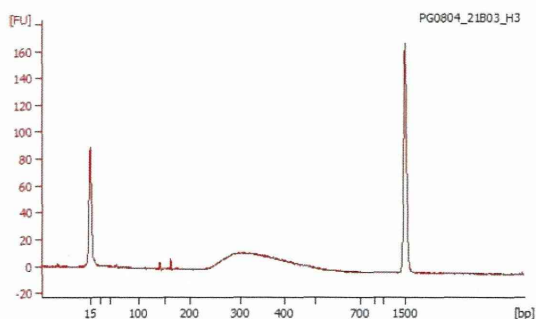
Fig. 18 アガロースゲル電気泳動によるゲノムDNAの品質確認 蛍光定量結果より50 ng 分の二本鎖DNA をアプライした.



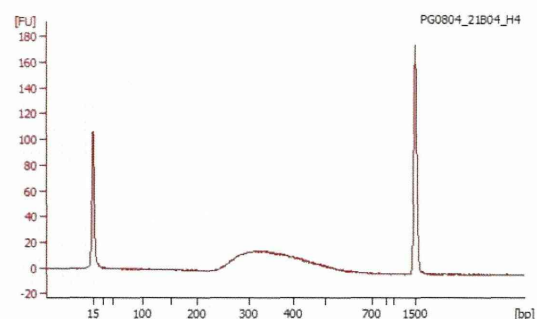
サンプル	Standard DNA Mix-1
ピークサイズ	310 bp
濃度	25.9 nmol/L



サンプル	Standard DNA Mix-2
ピークサイズ	305 bp
濃度	16.5 nmol/L



サンプル	Standard DNA Mix-3
ピークサイズ	304 bp
濃度	25.0 nmol/L



サンプル	Standard DNA Mix-4
ピークサイズ	312 bp
濃度	30.4 nmol/L

Fig. 19 Agilent 2100 Bioanalyzer を用いたリード長の検定結果 作製されたシーケンスライブラリーの品質をAgilent 2100 Bioanalyzer を用いて測定した。また、ライブラリーにはアダプター配列を付加されているため、ピークサイズからアダプターサイズ (約100base) を除いたサイズがクローニングサイズとなる。サンプルは、4種類のDepthについて解析を行うため、同サンプルで4解析分のライブラリーを独立して作製した。

Table 11 Phred クオリティスコアに基づいたシーケンス解析の精度評価

サンプル名	リード数	塩基数	Q30R1	Q30R2
Standard DNA#1	474,895,722	47,489,572,200	95.3	93.6
Standard DNA#2	529,964,716	52,996,471,600	94.8	92.0
Standard DNA#3	440,178,726	44,017,872,600	95.5	94.2
Standard DNA#4	469,051,582	46,905,158,200	95.0	93.5

Q30R1: 片鎖・両鎖解析において 1 回目に読み取られるリードの品質

Q30R2: 両鎖解析において 2 回目に読み取られるリードの品質

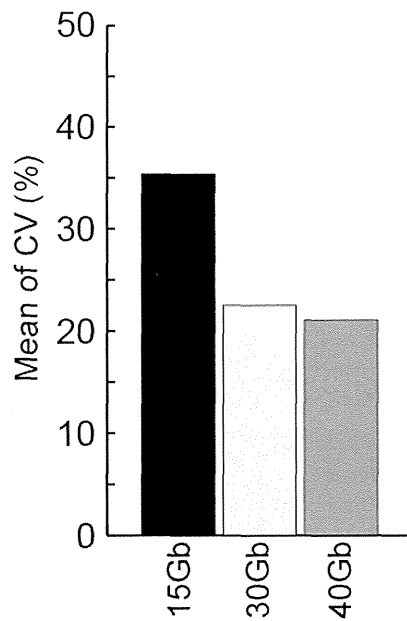


Fig. 20 シーケンス量が 15Gb, 30Gb, 40Gb の場合における変動係数 (CV) の平均

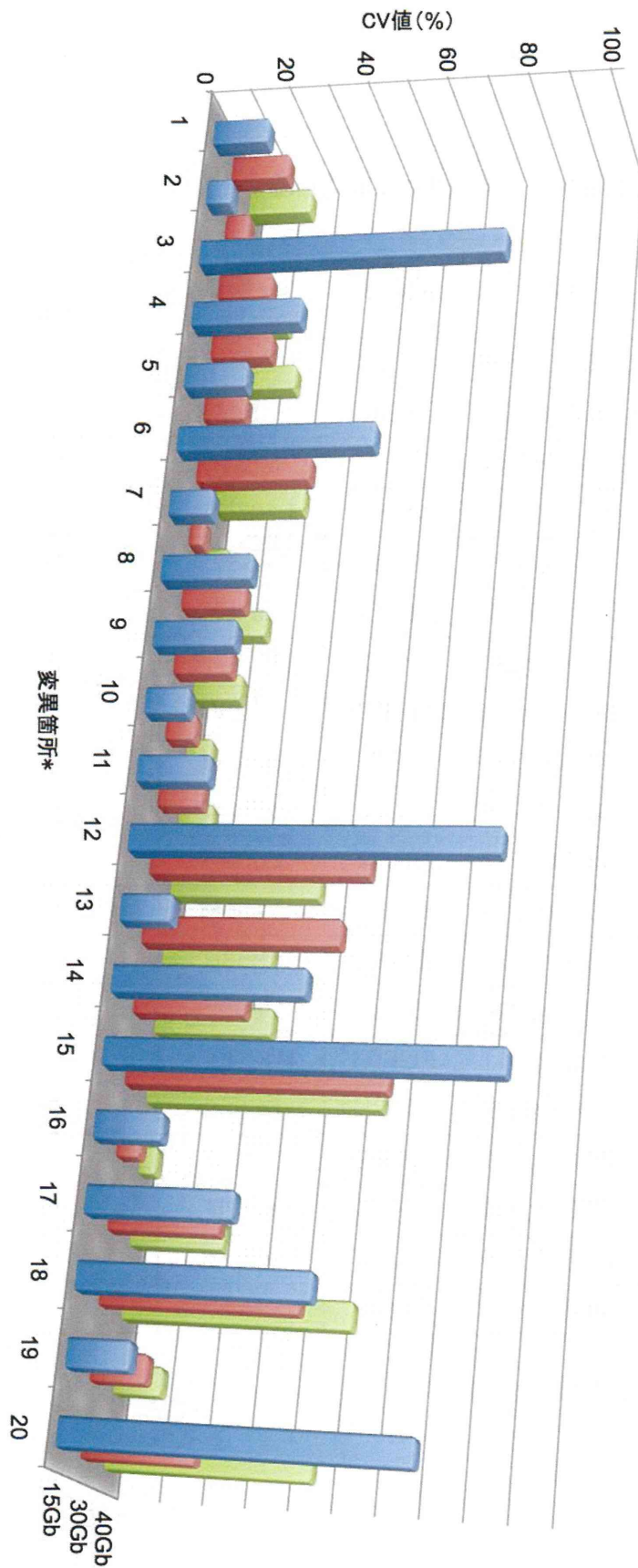


Fig. 21 15Gb、30Gb、40Gb の量でシーケンスされたときの各変異箇所における変異塩基の変動係数 (CV) CV 値は 4 回の独立したシーケンス解析から得られた変異頻度に基づいて算出された。*横軸の変異箇所については、Table 12 に示した。

Table 12 標準ゲノムDNA上で確認された変異箇所

	リファレンス ID	位置	遺伝子	アミノ酸変化	リファレンス配列 (塩基)	変異塩基
1	chr1	115,256,530	NRAS	Q61K	G	T
2	chr2	29,416,326	ALK	P1543S	G	A
3	chr2	209,106,786	IDH1	S261L	G	A
4	chr3	37,061,883	MLH1	L187M**	C	A
5	chr3	41,266,101	CTNNB1	S33Y	C	A
6	chr3	178,936,091	PIK3CA	E545K	G	A
7	chr3	178,952,085	PIK3CA	H1047R	A	G
8	chr4	55,138,600	PDGFRA	G426D	G	A
9	chr4	55,599,321	KIT	D816V	A	T
10	chr5	112,179,431	APC	R2714C	C	T
11	chr7	55,241,707	EGFR	G719S	G	A
12	chr7	55,249,071	EGFR	T790M	C	T
13	chr7	55,259,515	EGFR	L858R	T	G
14	chr7	140,453,136	BRAF	V600E	A	T
15	chr8	38,285,611	FGFR1	P124L***	G	A
16	chr9	139,409,754	NOTCH1	P668S	G	A
17	chr12	25,398,281	KRAS	G13D	C	T
18	chr12	25,398,284	KRAS	G12D	C	T
19	chr13	28,626,706	FLT3	V197A	A	G
20	chr16	68,867,462	CDH1	N/A	T	C

変異頻度については、Table 8を参照.

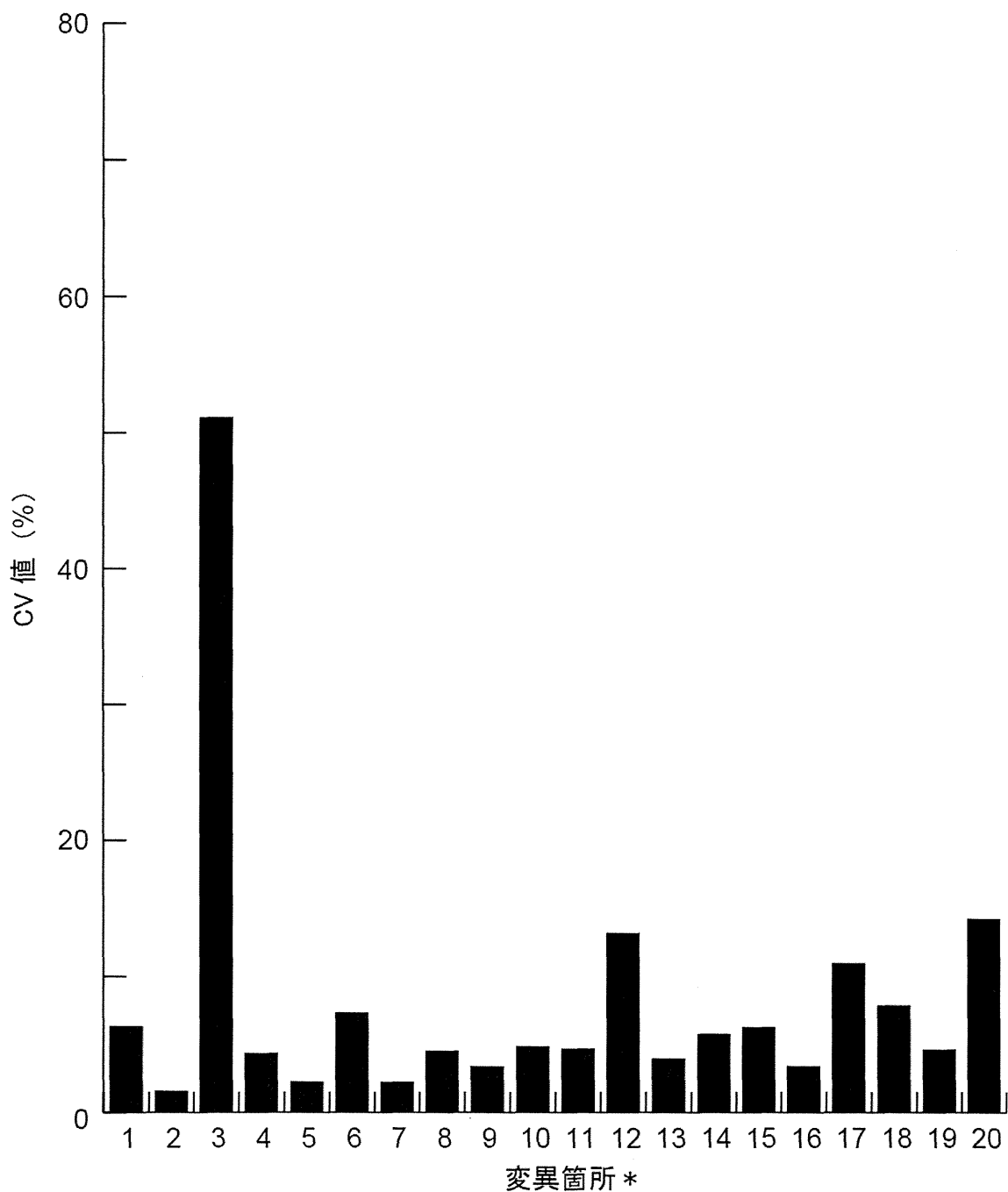


Fig. 22 全マッピング結果を合わせた時の各変異箇所における変異頻度の変動係数 (CV) 全マッピング結果をまとめた後、40Gb、80Gb、120Gb、160Gb 相当のデータとなるようにダウンサンプリングし、各変異箇所におけるそれぞれのシーケンス量 (40Gb、80Gb、120Gb、160Gb) に相当する塩基頻度を測定後、CV 値を算出した。*横軸の変異箇所については、Table 12 に示した。

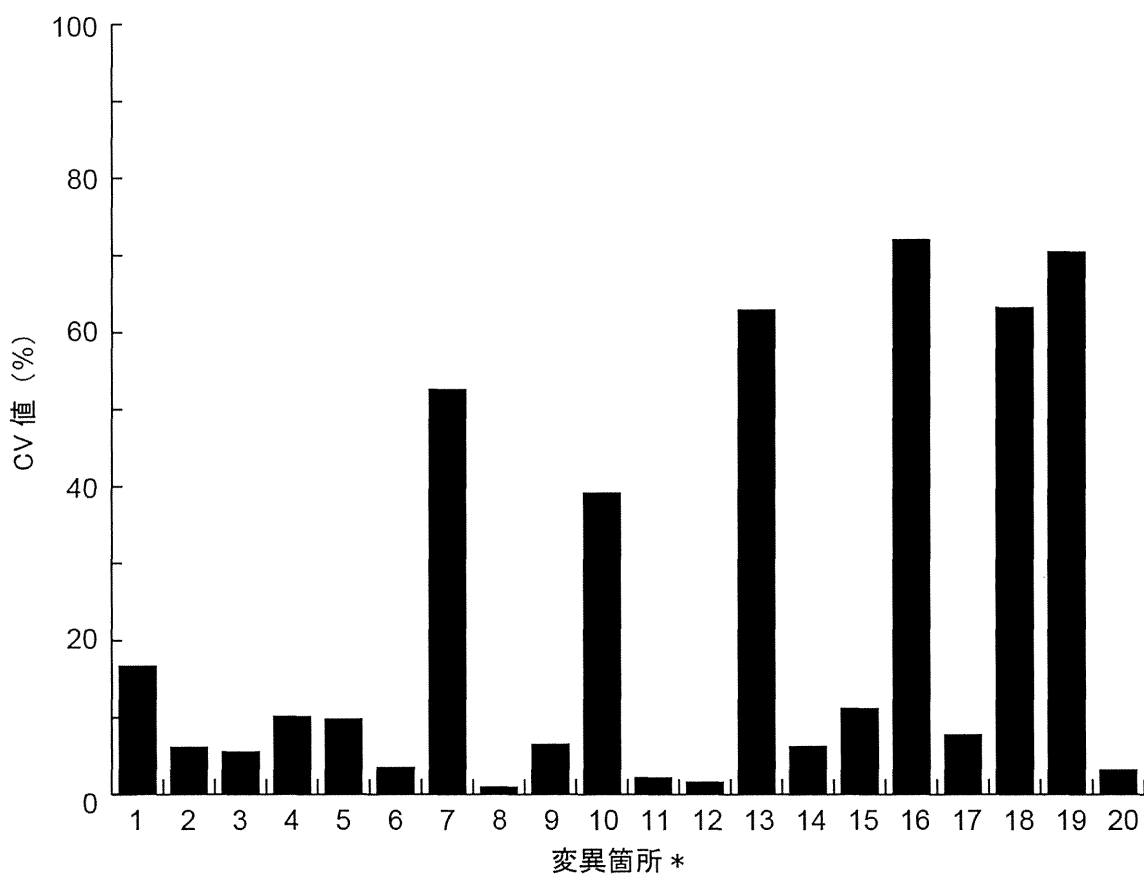


Fig. 23 全マッピング結果を合わせた時の各変異箇所における変異頻度の公表値に対する変動係数 (CV) (公表値に対する実測値のバラツキ) 全マッピング結果をまとめた後、40Gb, 80Gb, 120Gb, 160Gb 相当のデータとなるようにダウンサンプリングし、各変異箇所におけるそれぞれのシーケンス量 (40Gb, 80Gb, 120Gb, 160Gb) に相当する塩基頻度を測定後、CV 値を算出した。さらに、公表値を含めた CV 値を算出し、実測値および公表値の 2 群間の CV 値を求めた。*横軸の変異箇所については、Table 12 に示した。

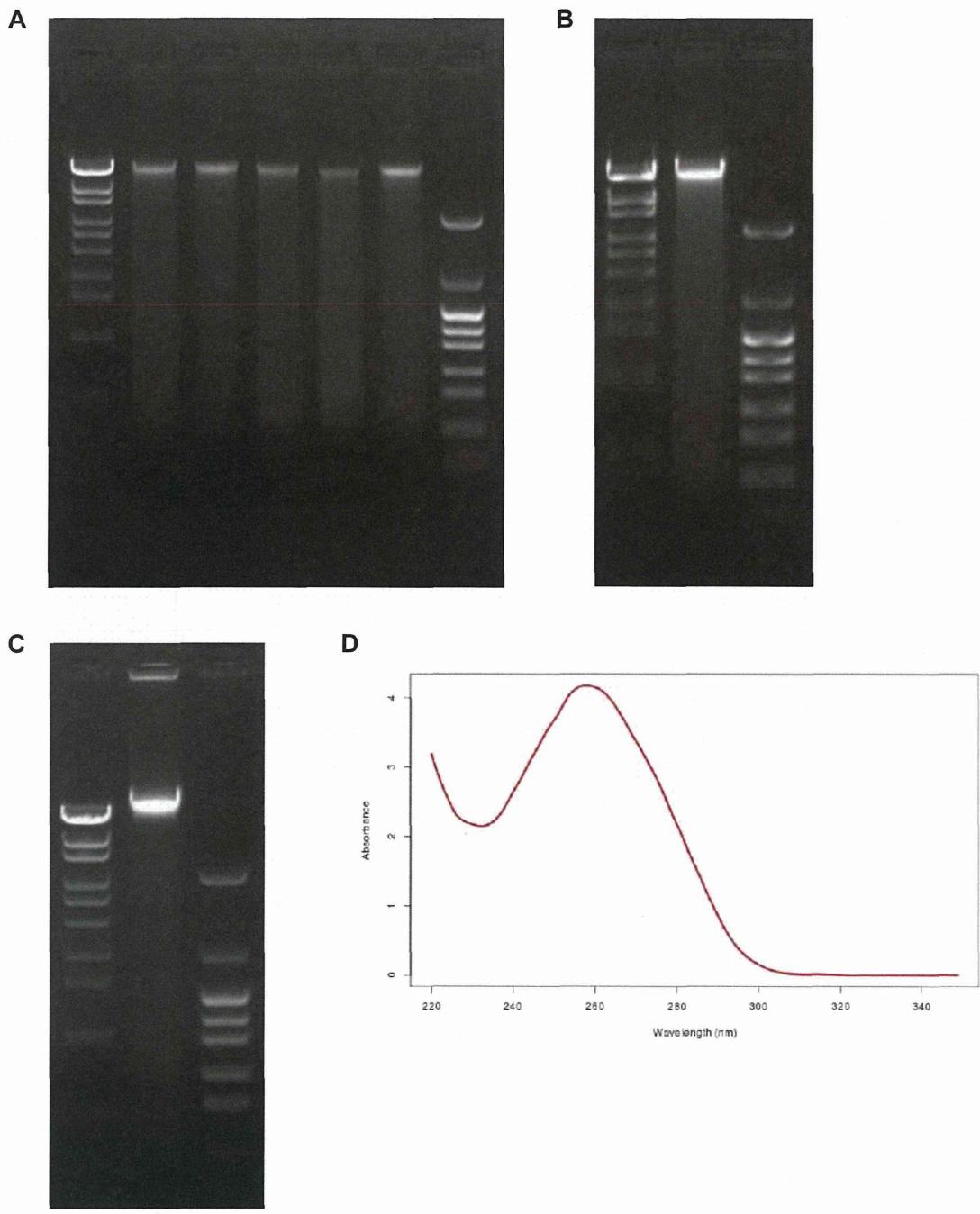


Fig. 24 ゲノム DNA 抽出結果

(A, B) 6 人の健常日本人男性(JOM001, JOM002, JOM003, JOM004, JOM005, JOM006)の精液 250 μ L から抽出したゲノム DNA の一部を電気泳動した。収量は左側の検体からそれぞれ, 3.2 μ g, 1.1 μ g, 2.5 μ g, 7.7 μ g, 3.1 μ g, 4.0 μ g であった。(C) 左から 5 番目の個人(JOM005) について, さらに 1750 μ L の精液から 73.0 μ g のゲノム DNA を得た。サイズマーカーは λ -EcoT14I および pHY を用いた。(D) C に示したサンプルに対して吸光定量を行い, 品質の確認を行った。

Table 13 SNP タイピング 2,294,794 サイトの結果

ID	Homo (Ref/Ref)	Homo (Alt/Alt)	Hetero
Sample 1	1,599,066	305,672	385,367
Sample 2	1,599,828	307,376	384,500
Sample 3	1,600,578	307,857	383,589
Sample 4	1,597,215	307,531	387,059
Sample 5	1,600,864	308,225	381,883

Table 14

ID	SNPs1	SNPs2	SNPs3	SNPs4
Sample 1	17	41	4	6
Sample 2	19	37	12	1
Sample 3	24	39	9	2
Sample 4	13	23	7	1
Sample 5	25	30	3	2

SNPs1: 全体と比較して日本人に稀な SNP 群

SNPs2: 欧州人と比較して日本人に稀な SNP 群

SNPs3: 東アジア人と比較して日本人に稀な SNP 群

SNPs4: 東アジア人には稀だが、日本人に多い SNP 群

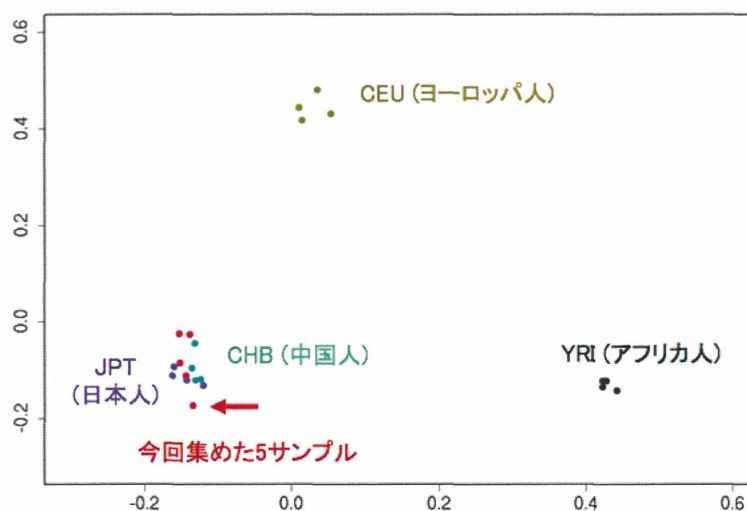


Fig. 25 SNP タイピングのデータをもとに行った主成分分析

1000 人ゲノムから選んだ 16 サンプルと、今回 SNP タイピングを行った 5 サンプル(赤色)との比較. アセンブリ対象の Sample 5 を赤矢印で示した.

Table 15 Illumina HiSeq から得られた塩基数 (Mb: mega base)

Library type	Insert size	Mb	Mb (Q > 15)	Read length
Paired end	260 bp	269,140	261,298	150 nt × 2
Paired end	360 bp	137,934	132,559	125 nt × 2
Paired end	660 bp	161,552	154,689	150 nt × 2
Mate pair	2 kb	102,230	97,729	150 nt × 2
Mate pair	5 kb	45,245	43,738	150 nt × 2
Mate pair	9 kb	46,134	44,512	150 nt × 2
Total	N/a	762,235	601,966	N/a

Table 16 トリミング後の塩基数

Library type	Insert size	Trimmed Mb	Mb (Q > 15)	Actual size
Paired end	260 bp	259,596	257,686	240 bp
Paired end	360 bp	135,509	131,901	330 bp
Paired end	660 bp	154,125	152,814	640 bp
Mate pair	2 kb	59,506	58,999	1.6 kb
Mate pair	5 kb	27,084	26,920	4.6 kb
Mate pair	9 kb	27,584	27,410	8.6 kb
Total	N/a	663,404	655,730	N/a

Table 17 ALLPATHS-LG によるゲノムサイズ等の推定

一倍体ゲノムサイズ *	2,741,890,408 bp
G+C 含量	40.2%
反復配列の割合 (K = 25)	27.0%

* GRCh37/hg19 の参考値: 3,137,161,264 bp

Table 18 ALLPATHS-LG による de novo アセンブリの結果

contig の最短長	1000 bp
全 contig 数	485,318
全 scaffold 数	333,852
総 contig 長	2,254,772,438 bp
総 scaffold 長 (gap を含む)	2,535,926,213 bp
N50 contig	9.3 kb
N50 scaffold	39 kb
N50 scaffold (gap を含む)	45 kb

Table 19 PacBio RS II から得られたデータ

PacBio RS II	ライブラリ S1	ライブラリ S2	S1 + S2
塩基数	9,146 Mb	35,694 Mb	44,840 Mb
リード数	757,979	3,673,817	4,431,796
平均リード長	12,066 bp	9716 bp	N/a
最長リード長	49,865 bp	53,506 bp	53,506 bp

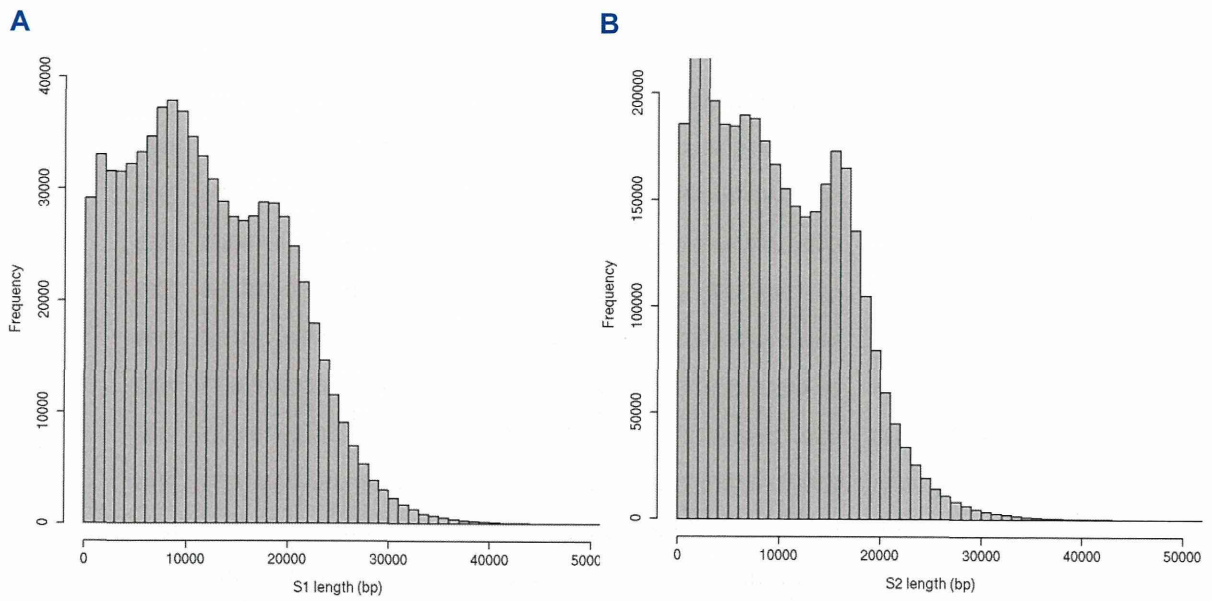


Fig. 26 PacBio RS II から得られたロングリードの長さの分布

(A) ライブラリ S1 から得られたリード長の分布. (B) ライブラリ S2 から得られたリード長の分布. S2 から得られた配列データの方が多いため、縦軸のスケールが一致していないことに注意.

HiSeq	1748	AGGCGATAGAAATTGAAACCTGGCGCAATAGATATAGTACCG--CAAGGGAAAGA-TG--	1802
PacBio	1824	AGGCGATAGAAATTGAAACCTGGCGCAATAGATATAGTACCGCCAAGGG-AAGATTGAA	1882
HiSeq	1803	--AAAAATTATAACCAAGCA-TAATATA-GCAAGGACTAACCCCTATACCTTCTGCATAA	1858
PacBio	1883	AAAAAAATTATAACCA-GCATTAAATATAGGCAAGGACTAACCCCTATACCTTCTGCATAA	1941
HiSeq	1859	TGAATT-AACTAGAAATAACTTTGCAAGGAGAGCCAAAGCTAAGACCCCGAAACCAGAC	1917
PacBio	1942	T-AATTAACCTAGAAATAACTTTGCAAGGAGAGCC-AAGCTAAGACCCCGAAACCAGAC	1999
HiSeq	1918	GA-GCTACCTAAGAACAGCT--AAAAGAGCACACCCGTCTATGTAGCAAATAGT--GGG	1972
PacBio	2000	GAGGCTACCTAAGAACAGCTTAAAAGAGCACACCCGTCTATGTAGCAAATAGTGGGGG	2059
HiSeq	1973	AAGATTTATAGGT--A-GAGGGACAAACCT-ACCGA-GCC-TGGTGATAGC-TGGTTGT	2025
PacBio	2060	AAGATTTATAGGTTTATGAGGGACAAACCTTACCGAGGCCTTGGTGATAGCTTGGTTGT	2119
HiSeq	2026	CC-AAGATAGAATCTTAGTTCAAC-TTAAATTTGCCACAGAACCCTCTAAATCCCCCT	2083
PacBio	2120	CCAAAGATAGAATCTTAGTTCAACTTTAAATTTGCCACAGAACCCTCTAA-TGCCCTT	2178
HiSeq	2084	GTAATTTAACTGTTAGTCCAAAGAGGAACAGCTCTTTGGACACTAGG--AAAAACCTT	2141
PacBio	2179	GT-AATTTAACTGTTAGTCC-AAGAGGAACAGCTCTTTGGACACTAGGGCAAAAAACCTT	2236
HiSeq	2142	-GTAGAGAGAGTAAAAAATTTAACCCCATAGTAGGCCTAAAAGCAGCCACCAATTAAGA	2200
PacBio	2237	CGTAGAGAGAGTAAAAAATTTAACCCCATAGTAGGCCT-AAAGCAGCCACCAATTAAGA	2295
HiSeq	2201	AAGCGTT-CAAGCTCAACACCCACTACCTAAAAATCCCAACATATAA-CTG-AACTCC	2257
PacBio	2296	AAGCGTTCCAAGCTCAACACCCACTACCTAAAAATCCCAACATATAATCTGAAACTCC	2355
HiSeq	2258	TCACACCCAATTGGACCAA--TCTATCACC-C---TATAGAAGAACTAATGTTAGTATA	2310
PacBio	2356	TCACACCCAATTG-ACCAATTTCTATCACCTCTATGTATAGAAGAACTAATTTAGTATA	2414
HiSeq	2311	AGTAACAT-GAAAAATTCTCCTCGCATAAGCCTGGGT---CAG-ATTAATAA-CACTGA	2364
PacBio	2415	AGTAACATTGAAAAC-TTCTCCTCGCATAAGCCTGGGTTTCCAGTATTAATAACCACTGA	2473
HiSeq	2365	ACTGACAATTAACAGCCCAA-TATCT-ACAATCAACCAACAAGTCATTATTACCCTCACT	2422
PacBio	2474	ACTGACATTAACAGCCCAAATATCTAACAATCAACCAACAAGTCATTATTACCCTCACT	2533
HiSeq	2423	GTCAACCCAACACAGGC-ATGCTCATAAGGAAAGG--TTaaaaaaGTAAAAGGAACT-C	2478
PacBio	2534	GTCAACCCAACACAGGCTATGCTCATAAGGAAAGGCTTAAAAAAGTAAAAGGAACTTC	2593
HiSeq	2479	GGCAAATCT--TACCCCGCCTGTTTACC-AAAAACAT-CA-CCTCTAGCATC-ACCAGTA	2532
PacBio	2594	GGCAAATCTGATA-CCCGCCTGTTTACCCAAAAACATCCAGCCTCTAGCATCGACCAGTA	2652
HiSeq	2533	TTAGAGGCACCGCC-TGCCAGT-GACACATGTTTAAACGGCCGGGTACCTAACCGTGC	2590
PacBio	2653	TTAGAGGCACCGCCCTGCCAGTGGAC-CATGTTTAA-GGCCGGGTACCTAACCGTGC	2710
HiSeq	2591	AAAGGTA-GCATAATCACTTGTTCCCTTAAATAGGGACCTGT-ATGAAT-GGCTCCACGAG	2647
PacBio	2711	AAAGGTAGGCATAATCACTTGTTCCCTTAAATAGGGAC-TGTAATGAATGGGCTCCACGAG	2769
HiSeq	2648	GG-TTCAGCTGTCTCTTACTTTTAAACAGTGAAATTGACCT-GCCCGTGAAGAGCGGGG-	2704
PacBio	2770	GGCTTCAGCTGTCTTAC--TTAACAGTGAAATTGACCTGGCCCGTGAAGAGCGGGG	2827

Fig. 27 HiSeq で得られたショートリードをアセンブルした結果と、PacBio から得られた 1 つのローングリッドを mtDNA の一部で比較. BLASTN によって計算された塩基の identity は mtDNA 全体で 92%であった.

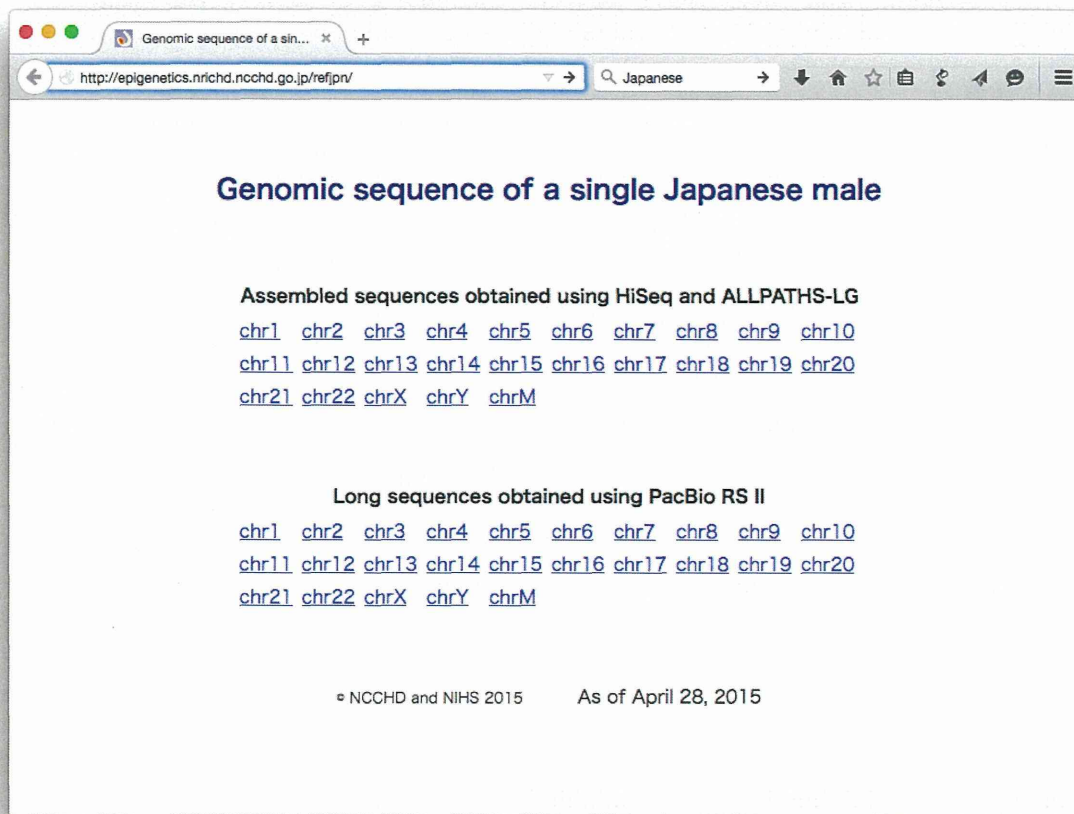


Fig. 28 JOM005 配列データ公開サイト <http://epigenetics.nrichd.ncchd.go.jp/refjpn/>

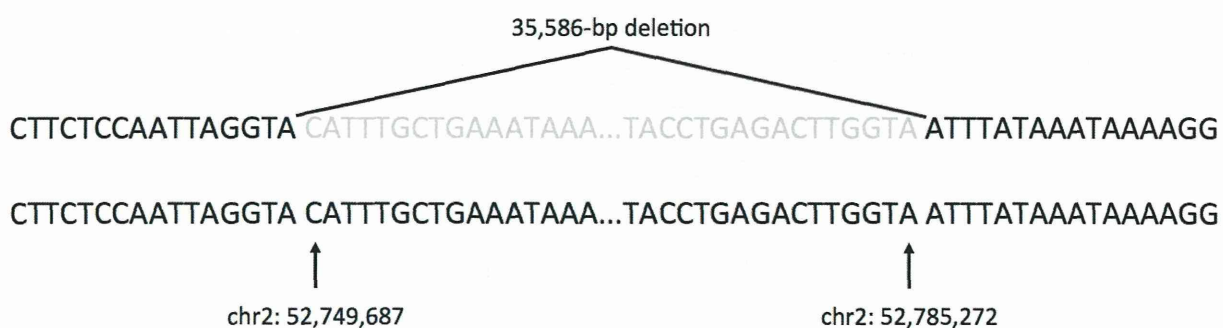


Fig. 29 多くの健常日本人が保有する第2番染色体短腕 52.77 Mb (2p16.2) 26 kb 欠失のブレイクポイントを JOM005 の配列データから決めることができた。

Table 20 ヒト iPS 細胞の神経分化誘導プロトコール

分化誘導	培地	添加試薬	
Day 0	KSR 培地 ¹	10 μM SB431542	500 ng/ml Noggin
Day 1	KSR 培地	10 μM SB431542	500 ng/ml Noggin
Day 2	KSR 培地	10 μM SB431542	500 ng/ml Noggin
Day 4	KSR 培地 75% + N2 培地 ² 25%	10 μM SB431542	500 ng/ml Noggin
Day 6	KSR 培地 50% + N2 培地 50%	—	500 ng/ml Noggin
Day 8	KSR 培地 25% + N2 培地 75%	—	500 ng/ml Noggin
Day 10	細胞回収, RNA 抽出		

¹KSR 培地 : DMEM/F12 (Gibco), 20% KSR (Gibco), 0.1 mM β-Mercaptoethanol

²N2 培地 : Neurobasal Medium (Gibco), 100 × N2 supplement (Gibco)

Table 21 ヒト iPS 細胞の心筋細胞分化誘導プロトコール

分化誘導	培地	添加試薬
Day 0	RPMI/B27 (w/o insulin) 培地 ¹	12 μM CHIR99021
Day 1	RPMI/B27 (w/o insulin) 培地	—
Day 3	RPMI/B27 (w/o insulin) 培地	5 μM IWP-4
Day 5	RPMI/B27 (w/o insulin) 培地	—
Day 7	RPMI/B27 培地 ²	—
Day 10	RPMI/B27 培地	—
Day 13	RPMI/B27 培地	—
Day 15	細胞回収, RNA 抽出	

¹RPMI/B27 (w/o insulin) 培地 : RPMI1640 (Sigma), 50 × B27 supplement without insulin (Gibco), 1% Penicillin/Streptomycin (Gibco)

²RPMI/B27 培地 : RPMI1640 (Sigma), 50 × B27 supplement (Gibco), 1% Penicillin/Streptomycin (Gibco)

Table 22 外胚葉分化関連候補 miRNA と標的候補遺伝子

miRNA	遺伝子名
miR-373	<i>BCLAF1</i>
	<i>BRWD1</i>
	<i>CHD9</i>
	<i>CORO2A</i>
	<i>GALNT3</i>
	<i>LAMP5</i>
	<i>LHX8</i>
	<i>MBNL1</i>
	<i>MBNL3</i>
	<i>MRPL43</i>
	<i>NTN4</i>
	<i>POLQ</i>
	<i>SLC6A15</i>
	<i>ST8SIA3</i>
	<i>STXBP5L</i>
<i>TRIM2</i>	
<i>ZFX</i>	
miR-371-5p	<i>BCLAF1</i>
	<i>CASZ1</i>
	<i>KIAA0101</i>
	<i>NME5</i>
	<i>NRP2</i>
	<i>SPAG9</i>
	<i>SRSF3</i>
	<i>STXBP5L</i>
	<i>TFPI</i>
	<i>YAF2</i>
miR-371-3p	<i>DLG2</i>

Table 23 中胚葉分化関連候補 miRNA と標的候補遺伝子

miRNA	遺伝子名
miR-524-5p	<i>CD47</i>

Table 24 内胚葉分化関連候補 miRNA と標的候補遺伝子

miRNA	遺伝子名
miR-4739	<i>APCDD1L</i>
	<i>XPNPEP3</i>
miR-4505	<i>APCDD1L</i>
	<i>ATP8B3</i>
	<i>EFCAB2</i>
miR-4521	<i>ST8SIA4</i>
miR-520g	<i>BAGE2</i>
	<i>C17orf51</i>
	<i>EFCAB2</i>
	<i>FBN1</i>
miR-3714	<i>RAB3B</i>
	<i>SLC25A4</i>
miR-367	<i>CDK6</i>
	<i>CPNE8</i>
	<i>DCC</i>
	<i>FBN1</i>
	<i>KLF12</i>
	<i>PCDH9</i>
	<i>RAB3B</i>
	<i>XPNPEP3</i>

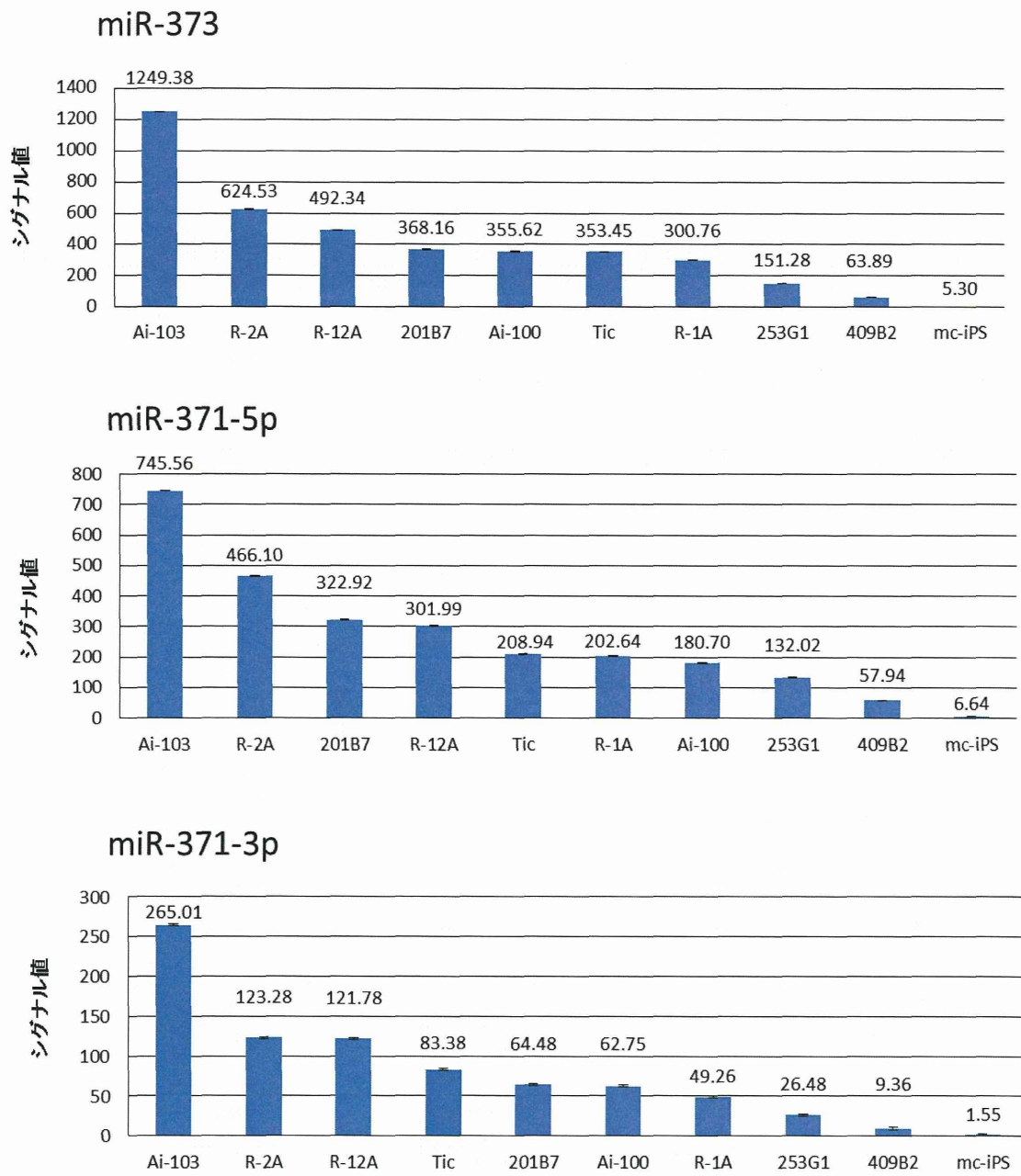


Fig. 30 外胚葉分化プロペンシティ予測マーカーmiRNAの発現量

縦軸にマイクロアレイ解析から得た遺伝子発現シグナル値, 横軸にヒト iPS 細胞株を示した。データは平均値±SEM (n = 6) で示した。