



イルミナウェビナー
2013年6月19日

疾患関連遺伝子の long-range PCR Nextera 解析

国立遺伝学研究所
人類遺伝研究部門
細道一善

本Webinarについて

次世代シーケンサー MiSeq



Nextera™ DNAサンプル調製キット



比較的手軽な次世代シーケンサーの使い方

- ・数十kb程度の領域(遺伝子)をシーケンスしたい
- ・数十検体のシーケンスをしたい
- ・PCR産物をクローニングしてシーケンスするような解析

Nexteraのライブラリー調整の詳細

ILLUMINA® Webinar Series



微生物ゲノムシリーズ

Nexteraによるライブラリー調製：実験の流れ、注意点、応用例について

日程 2012年11月28日（水曜日）

時間 16:00-17:00

講師 宮崎大学医学部 感染症講座 微生物学分野
小椋 義俊先生

[Register Now](#)

遺伝子のシーケンス



タンパク質コーディング遺伝子

平均長	<u>53.6kb</u>
最小	数100bp
最大	2.4 Mb (DMD dystrophin)
平均エクソン数	9.8 (エクソン平均長 288bp)
最大	363 (TTN titin)
最小	1

ターゲットリシーケンス法

- ハイブリダイゼーションによる濃縮
 - TruSeq Custom Enrichment Kit (イルミナ)
 - SureSelect DNA Capture (アジレント)
 - Haloplex (アジレント)
 - SeqCap EZ choice library (ロシュ) など
- PCRアンプリコン
 - TruSeq カスタムアンプリコン (イルミナ)

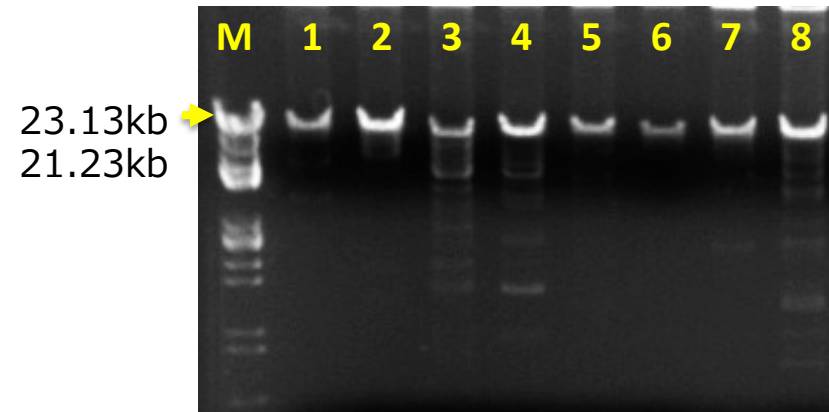
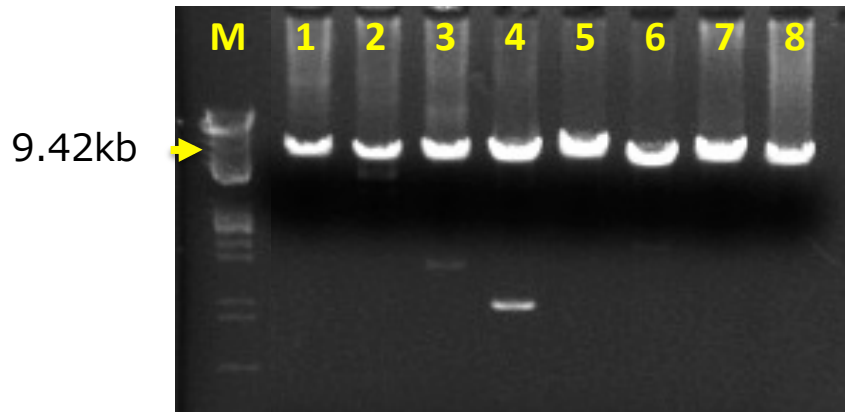
ターゲットリシーケンス法

- ハイブリダイゼーションによる濃縮
 - TruSeq Custom Enrichment Kit (イルミナ)
 - SureSelect DNA Capture (アジレント)
 - Haloplex (アジレント)
 - SeqCap EZ choice library (ロシュ) など
- PCRアンプリコン
 - TruSeq カスタムアンプリコン (イルミナ)
 - **Long-range PCR & Nextera kit**

long-range PCRの増幅長

約10kb

約20kb



	増幅長(bp)	
1	11,513	23,986
2	11,078	21,459
3	10,071	20,623
4	10,196	23,870
5	12,327	23,406
6	10,127	21,972
7	10,438	22,776
8	10,335	21,306

long-range PCR Nextera 解析 ～ 実験の流れ ～

1. ロングPCRプライマー設計
2. ロングPCR増幅
3. Nextera kitによるライブラリー調整
4. MiSeqによるラン

ロングPCRプライマー設計

- プライマー設計の条件
 - Tm値 68°C
 - 塩基数 26mer
 - プライマー配列がゲノム内でユニーク
 - プライマー設計位置にSNPが無い
- ロングPCR産物の長さ
 - 10kb-20kb (DNAの質にもよる)

2ステップ PCR (10kb)

DNA(20ng)	0.5	μl	94 °C	2 min	} 30cycles
5 x Buffer	2	μl	98 °C	10 sec	
dNTP Mixture	0.8	μl	68 °C	5 min*	
Primer F 5uM	0.4	μl	4 °C	∞	
Primer R 5uM	0.4	μl			
PrimeSTAR GXL	0.4	μl			
H ₂ O	5.5	μl			
total	10.0	μl			

※ 増幅長約20kbの場合は10min

PCR酵素

TaKaRaPrimeSTAR® GXL DNA Polymerase R050A

TOYOBO KOD FX KFX-101 など

ロングPCRプライマー設計

- プライマー設計の条件

- Tm値 68°C

- 塩基数 26mer

- プライマー配列がゲノム内でユニーク

- プライマー設計位置にSNPが無い

- ロングPCR産物の長さ

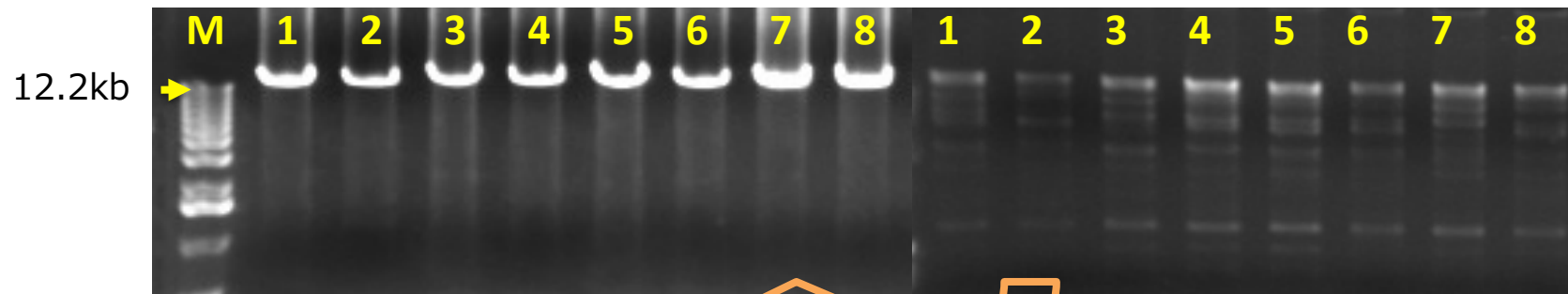
- 10kb-20kb (DNAの質にもよる)

GC%が低い領域ではTm68°C
にするためにプライマー長が
30merを超えてしまう

PCR増幅の比較 (プライマーが長い場合)

ステップダウン

2ステップ



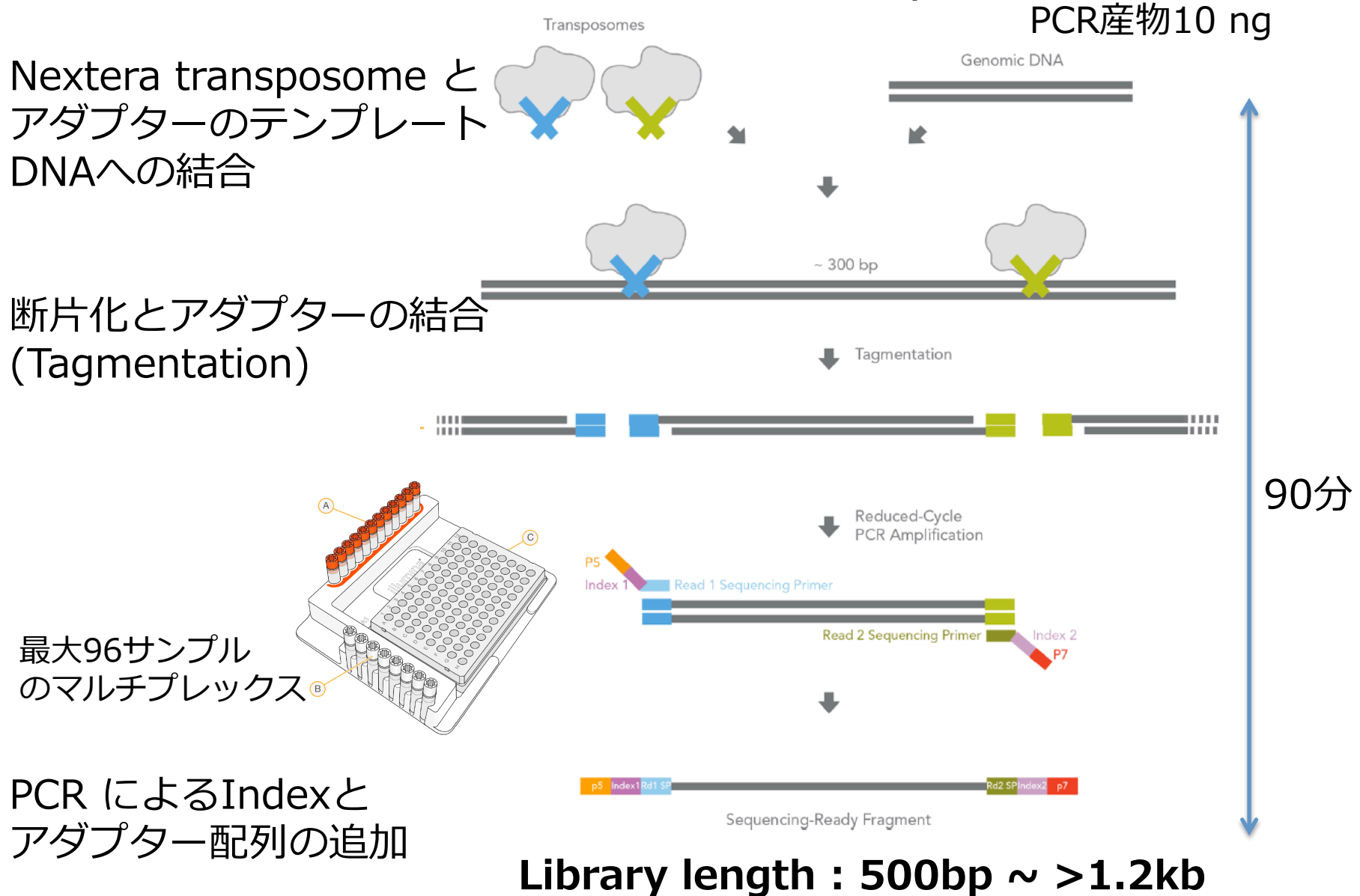
PCRサイクル条件を変更

ステップダウンPCR (プライマーが長い場合)

DNA(20ng)	0.5	μl	94 °C	2 min	
5 x Buffer	2	μl	98 °C	10 sec	} 5cycles
dNTP Mixture	0.8	μl	74 °C	5 min*	
Primer F 5uM	0.4	μl	98 °C	10 sec	} 5cycles
Primer R 5uM	0.4	μl	72 °C	5 min*	
PrimeSTAR GXL	0.4	μl	98 °C	10 sec	} 5cycles
H ₂ O	5.5	μl	70 °C	5 min*	
total	10.0	μl	98 °C	10 sec	} 20cycles
			68 °C	5 min*	
			4 °C	∞	

※ 増幅長約20kbの場合は10min

Nextera DNA Sample Preparation Kit によるライブラリ調整



Nextera kitによるライブラリー調整 (標準プロトコール)

ロングPCR増幅

↓
Tagmentation

PCR産物(50ng)	20u μ l	55°C	5min
TD	25u μ l	10°C	∞
TDE1	5u μ l		

↓
精製 Zymo-Spin1-96plate

↓
ライブラリーのPCR増幅

Index1 primer	5u μ l	72°C	3min	} 5cycles
Index2 primer	5u μ l	98°C	30sec	
NPM	15u μ l	98°C	10sec	
PPC	5u μ l	63°C	30sec	
Tagmentation DNA	20u μ l	72°C	3min	
		10°C	∞	

↓
精製 AMPure XP bead

Nextera kitによるライブラリー調整 (1/5量プロトコール)

ロングPCR増幅

↓
Tagmentation

PCR産物(10ng)	4u μ l	55°C	5min
TD	5u μ l	10°C	∞
TDE1	1u μ l		

↓
精製 Zymo-Spin1-96plate

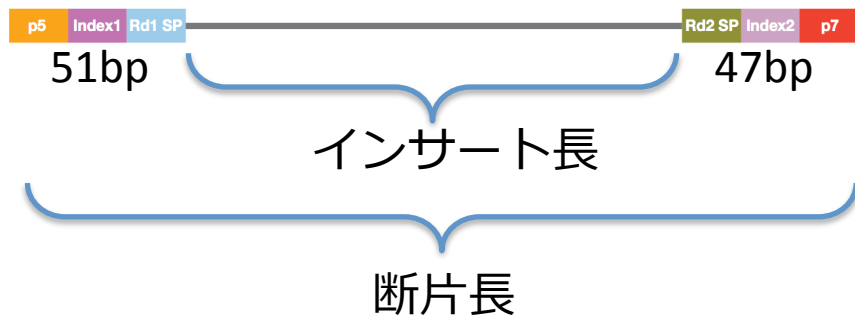
↓
ライブラリーのPCR増幅

Index1 primer	1u μ l	72°C	3min	} 7cycles
Index2 primer	1u μ l	98°C	30sec	
NPM	3u μ l	98°C	10sec	
PPC	1u μ l	63°C	30sec	
Tagmentation DNA	4u μ l	72°C	3min	
		10°C	∞	

↓
精製 AMPure XP bead

Nexteraライブラリー長の最適化

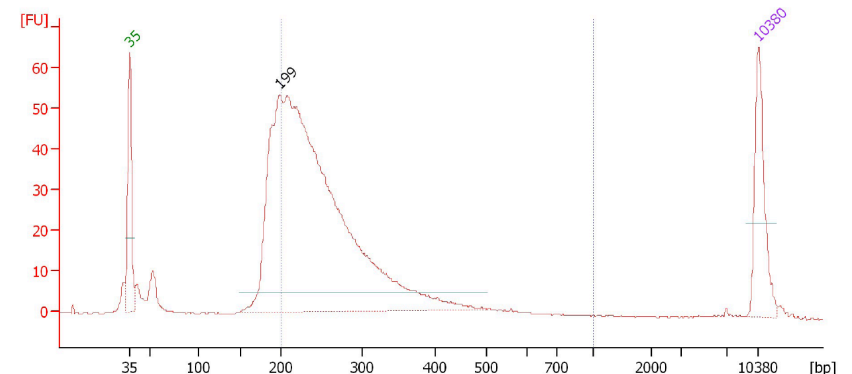
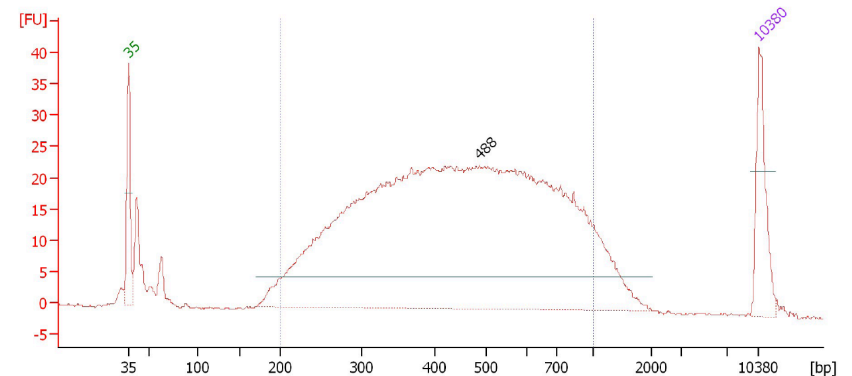
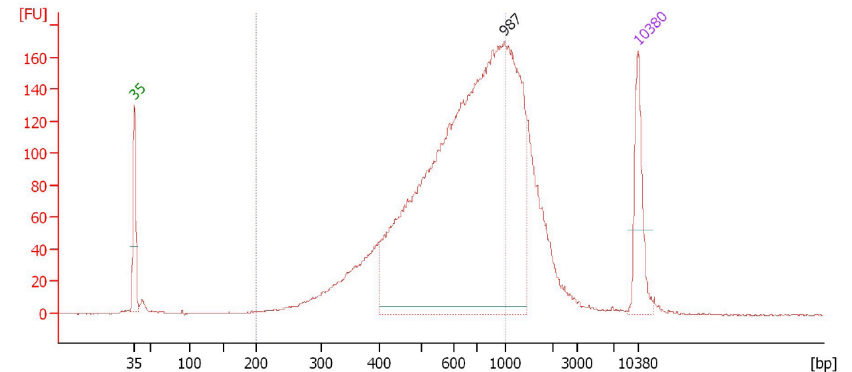
使用するPCR産物の量によりNexteraライブラリーの長さが異なるため実験デザインに応じた条件を決定



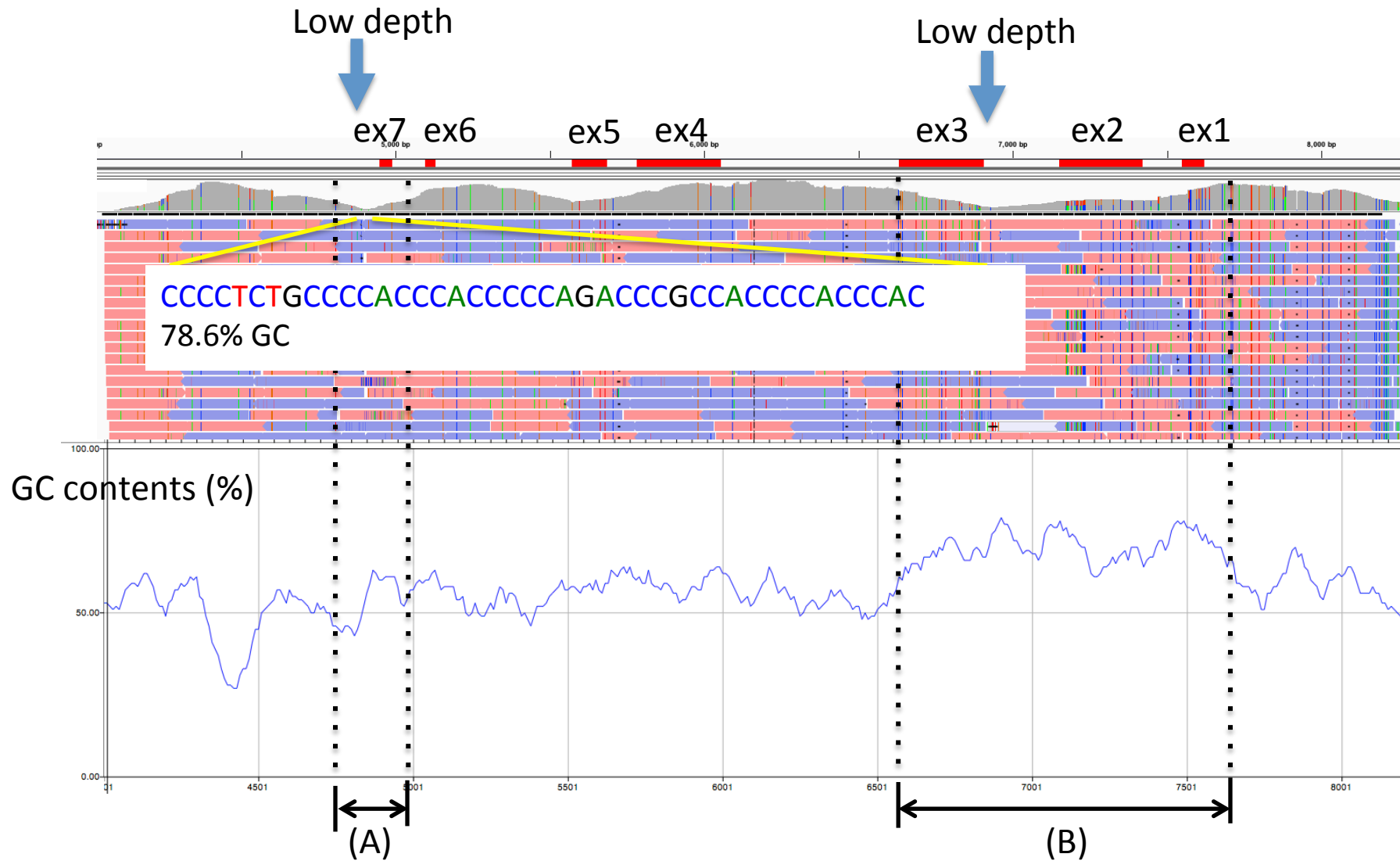
Index primer 1 (i7) 39bp + 8bp
5' CAAGCAGAAGACGGCATAACGAGAT[i7]GTCTCGTGGGCTCGG

Index primer 2 (i5) 43bp + 8bp
5' AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC

断片長 - 98bp = インサート長



GC含量の高い遺伝子



Nextera kitによるライブラリー調整 (高GC%プロトコール)

ロングPCR増幅

Tagmentation

PCR産物(10ng)	4u μ l	55°C	5min
TD	5u μ l	10°C	∞
TDE1	1u μ l		

精製 Zymo-Spin1-96plate

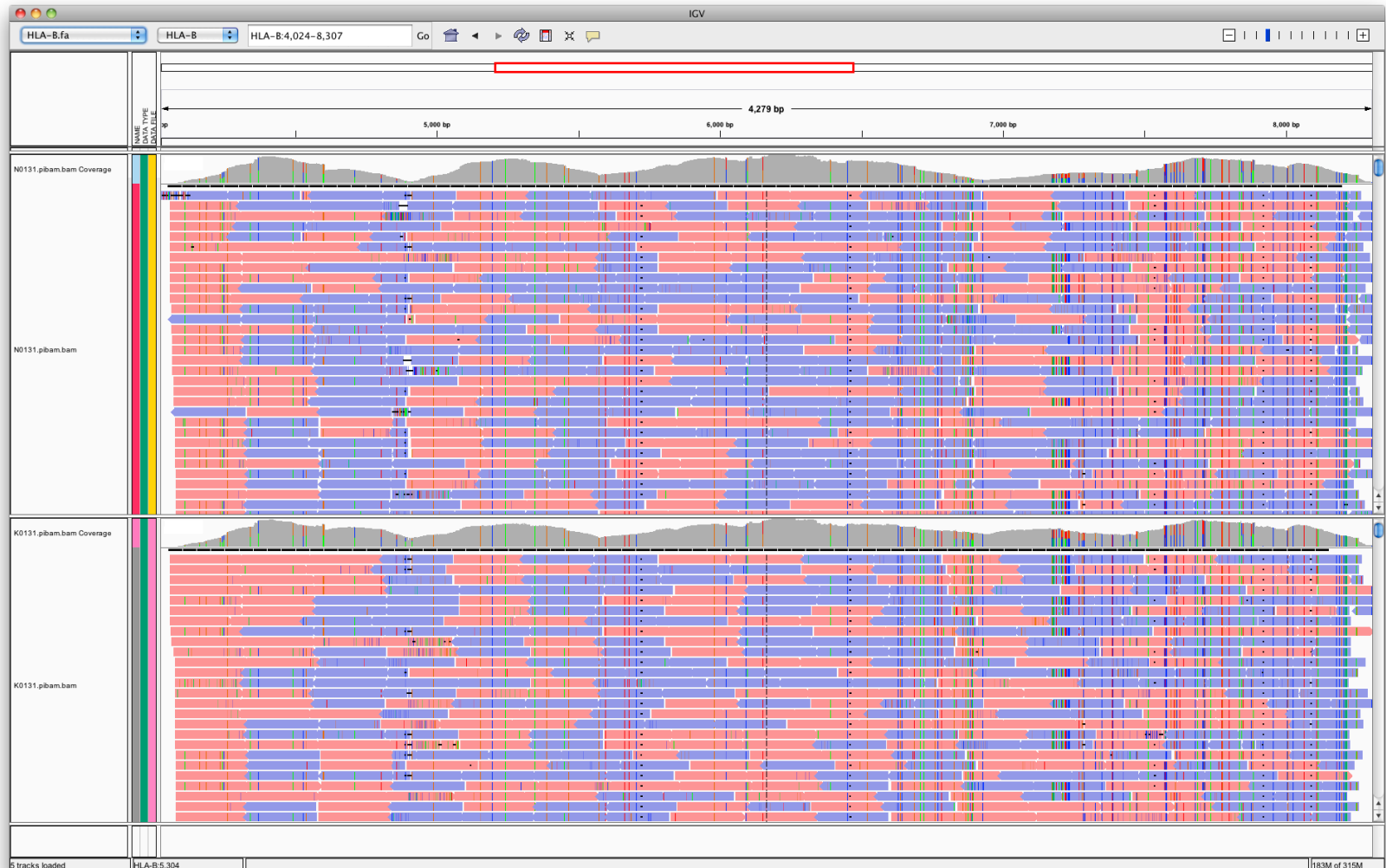
ライブラリーのPCR増幅

Index1 primer	1u μ l	72°C	3min	} 7cycles
Index2 primer	1u μ l	98°C	30sec	
2xKapa HiFi HS RM	5u μ l	98°C	10sec	
PPC	1u μ l	63°C	30sec	
Tagmentation DNA	2u μ l	72°C	3min	
		10°C	∞	

精製 AMPure XP bead

GC含量の高い遺伝子

Nextera



KAPA
LA Kit

(A)

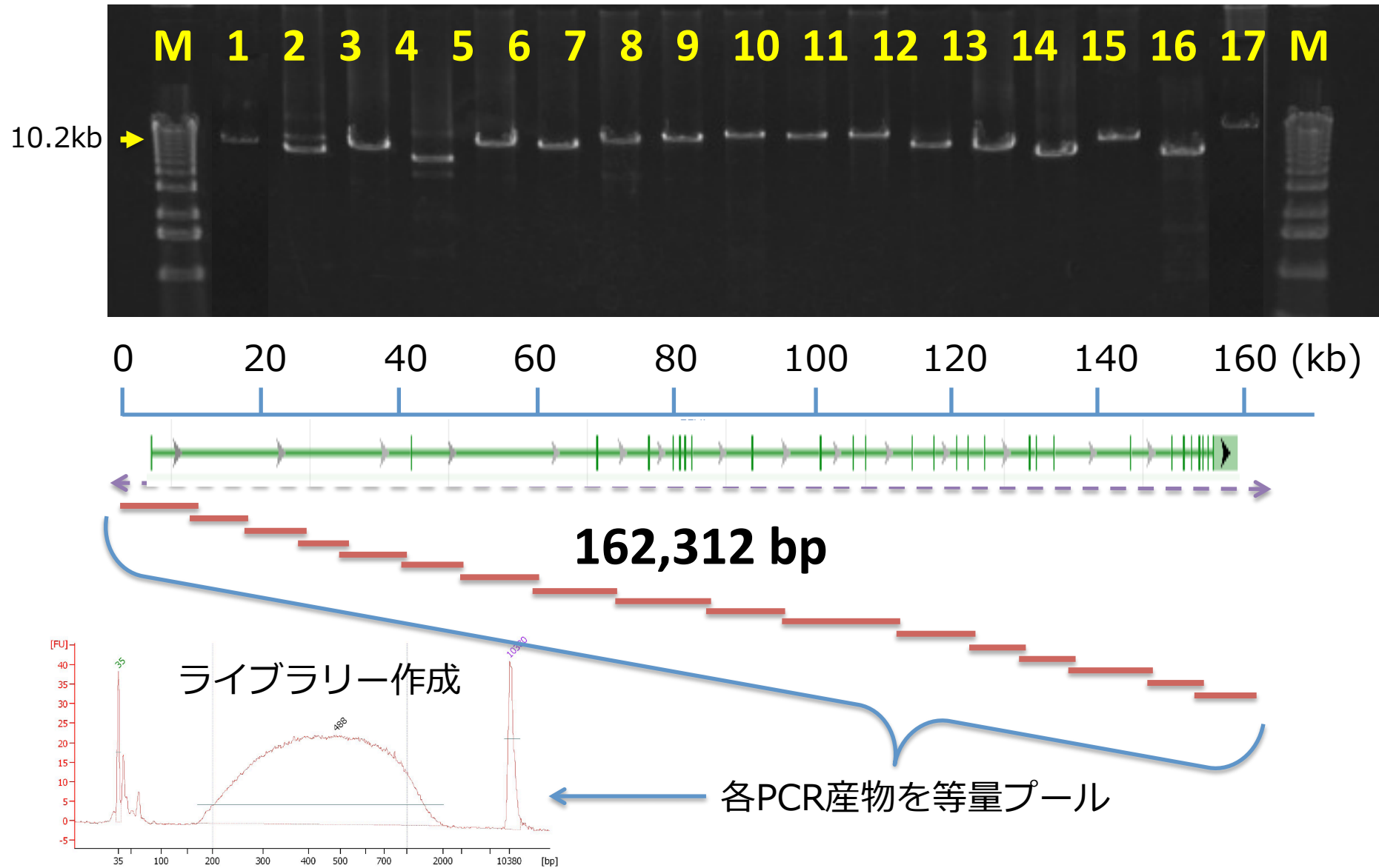
(B)

PCR酵素

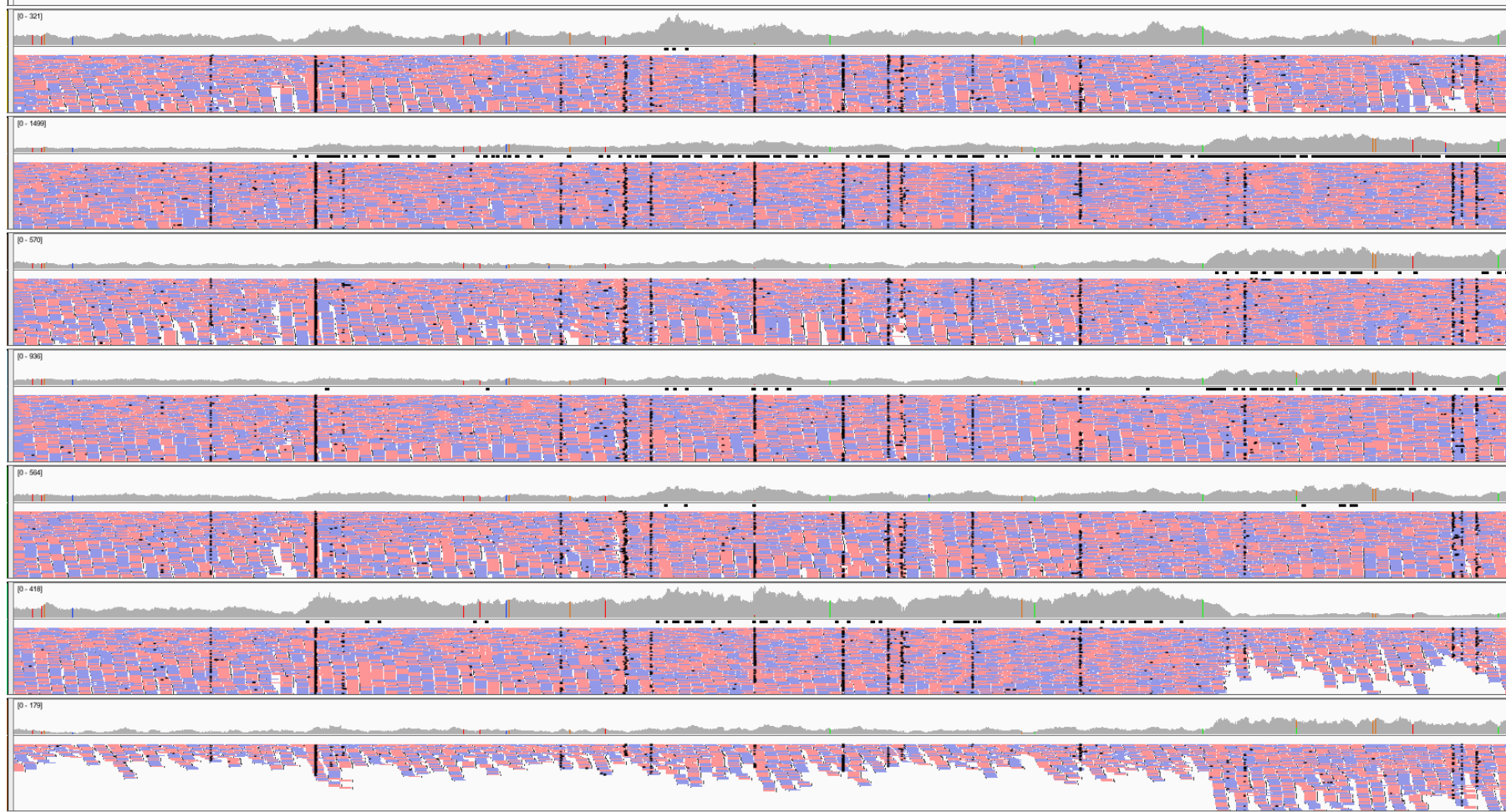
KAPA BIOSYSTEMS KAPA Library Amplification Kit KK2611

KAPA BIOSYSTEMS Kapa HiFi Hotstart mastermix KK2602 など

100kb以上の遺伝子のPCR増幅例



プールしたライブラリーのMiSeqによるラン



シーケンスの評価と原因候補変異の検出(例)

アライメントされたread数 : 211,479

アライメントされたreadの割合 : 84.64%

平均 depth : 193.86

平均カバレッジ% : 94.79%

Mutations		Case (n=85)	Japanese (n=127)	1000G	PhyloP	GERP++	SIFT	PolyPhen2
c.C3784T	p.R1262W	0.206	0.004	0	0.9996	5.26	0.99	0.993
c.T3216A	p.N1072K	0.006	0.012	0.01	0.95	1.74	0	0.168
c.A3029G	p.E1010G	0.117	0.004	0.0014	0.99847	5.6	0.89	0.992
c.A2713T	p.M905L	0.017	0.035	0.02	0.88116	0.958	0.33	0

long-range PCR Nextera 解析 ～ HLA遺伝子の解析 ～

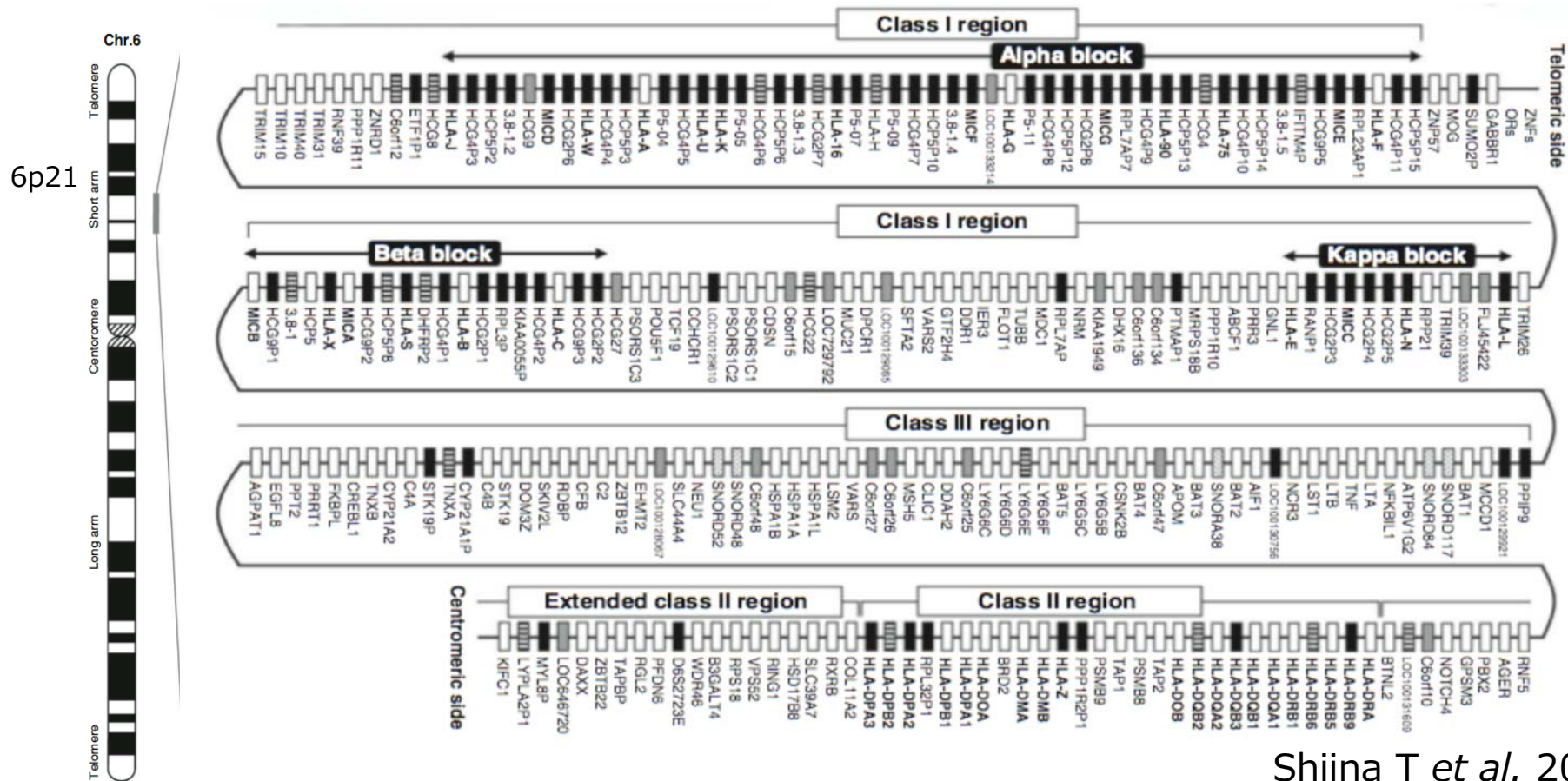
Hosomichi K *et al.* BMC Genomics. 2013
PMID: 23714642

1. HLA遺伝子シーケンスの意義
2. PCRからシーケンス
3. 物理的情報に基づく相の決定
4. HLA遺伝子解析パイプライン

HLAとは

- HLA (Human Leukocyte Antigen
= ヒト白血球抗原)
 - 1954年、白血球の血液型として発見
- 組織適合性抗原
(MHC; Major Histocompatibility Complex)
- ヒトに関しては MHC = HLA

HLA領域のゲノム配列



Shiina T et al. 2009

HLA領域の特徴

252の遺伝子、6つの古典的HLA遺伝子と少なくとも132のタンパク質をコードする遺伝子を含む

極めて高度な多型性を示す

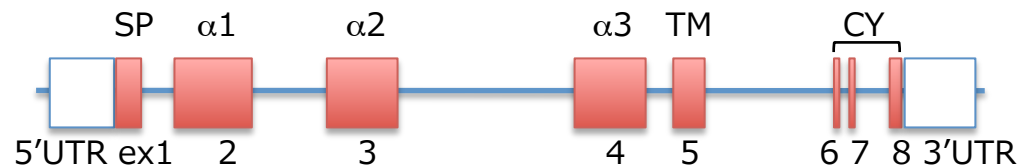
100以上の疾患および薬剤副作用と関連する

MHC分子

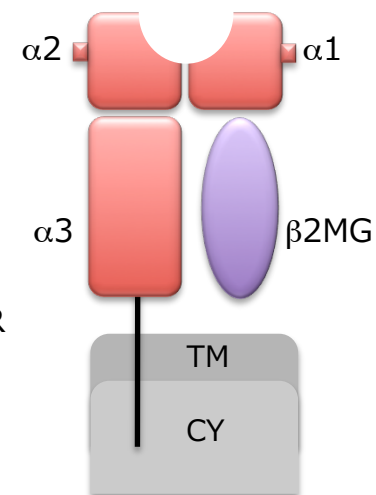
MHCクラスI

HLA-A, -C, -Bなど

内在抗原提示
細胞性免疫

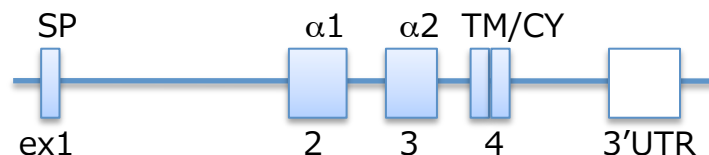


HLAクラスI分子

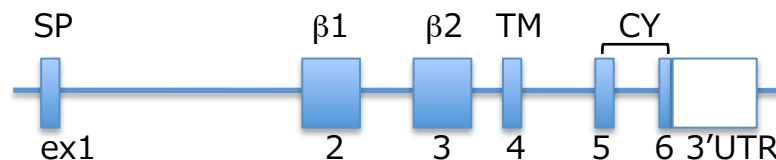


MHCクラスII

α鎖遺伝子
HLA-DRA1, -DQA1など

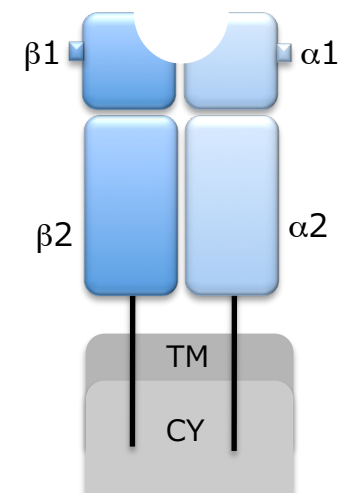


β鎖遺伝子
HLA-DRB1, -DQB1, -DPB1など

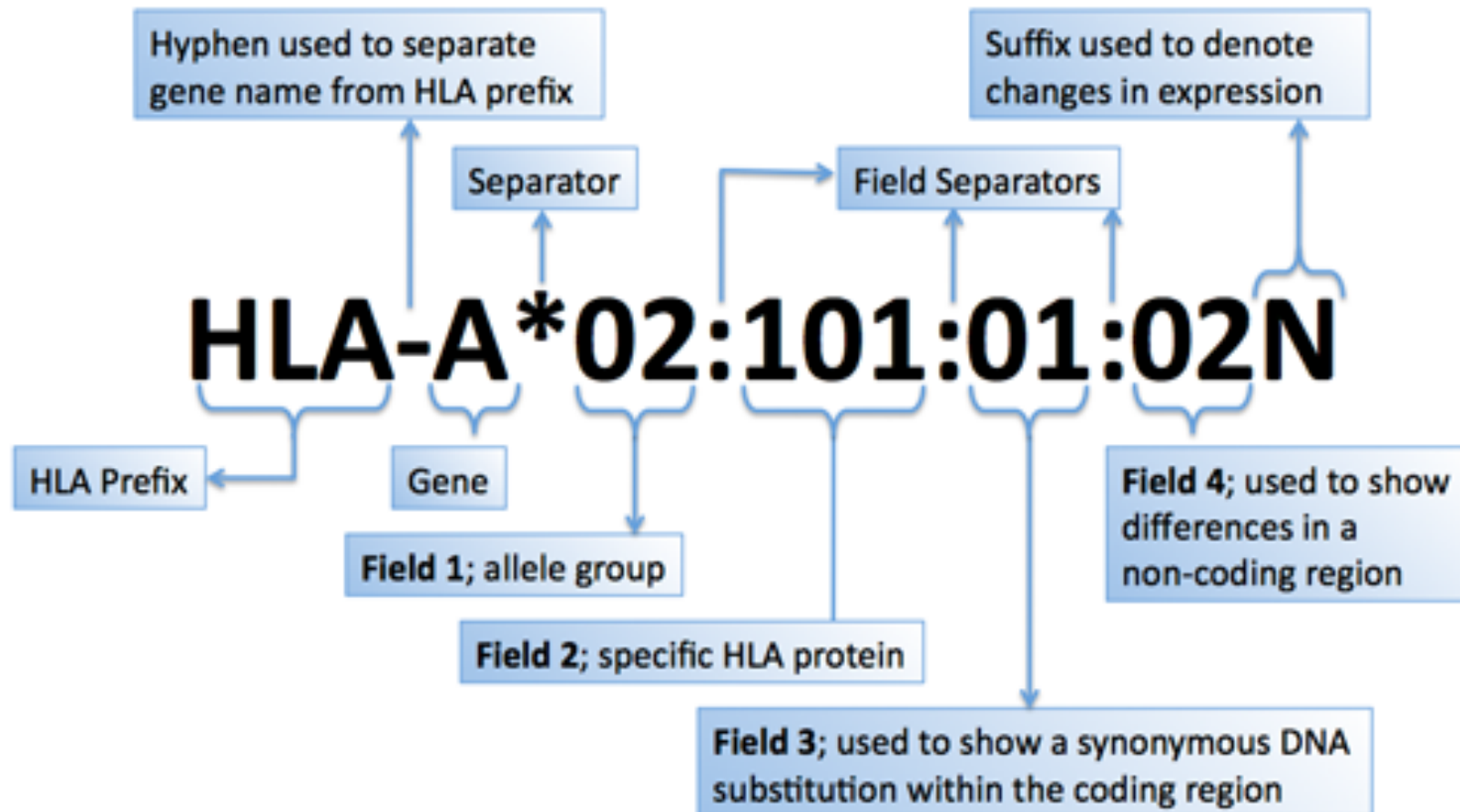


外来性抗原提示
液性免疫

HLAクラスII分子



HLAアレルの命名法



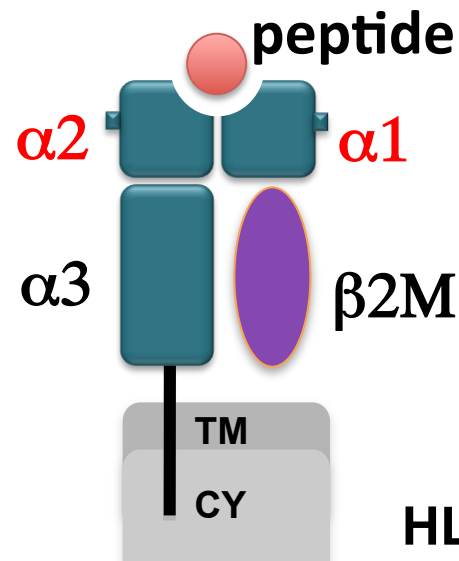
HLA クラスI α 1および α 2ドメインの多様性

5 amino acids 3

```
B52:01 AGSHSMRYFYTAMSRPGRGEPFRFI AVGVDDTQFVRFSDAASPRTEPRAPWIEQEGPEYWDRETQISKNTNTQTYRENLFIALRYYNQSEAG
B44:03 AGSHSMRYFYTAMSRPGRGEPFRFITVGYVDDTLFVRFSDA TS PRKEPRAPWIEQEGPEYWDRETQISKNTNTQTYRENLFTALRYYNQSEAG
B07:02 AGSHSMRYFYT SVSRPGRGEPFRFISVGYVDDTQFVRFSDAASPREPRAPWIEQEGPEYWDRENTQIYKQAQTDRESLFNLRGYYNQSEAG
B54:01 AGSHSMRYFYTAMSRPGRGEPFRFIAVGVDDTQFVRFSDAASPRGEPRAPWVEQEGPEYWDRENTQIYKQAQTDRESLFNLRGYYNQSEAG
B46:01 AGSHSMRYFYTAMSRPGRGEPFRFIAVGVDDTQFVRFSDAASPRMAPRAPWIEQEGPEYWDRETQKYKQAQTDREVSLFNLRGYYNQSEAG
*****.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*
```

B52:01 SHTWQTMYGCDVGP DGRLLRGHNCYAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAAREAEQLRAYLEGLCWEWLRRHLENGKETLQRAD
B44:03 SHIQRMYGCDVGP DGRLLRGYDCDAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAARVAEQLRAYLEGLCVESLRRYLENGKETLQRAD
B07:02 SHTLQSMYGCDVGP DGRLLRGHDQYAYDGKDYIALNEDLRSWTAADTAAQITQRKWEAAREAEQLRAYLEGECWEWLRRYLENGKDKLERAD
B54:01 SHTWQTMYGCDL GPDGRLLRGHNLAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAARVAEQLRAYLEGTCWEWLRRYLENGKETLQRAD
B46:01 SHTLQRMYGCDVGP DGRLLRGHDSAYDGKDYIALNEDLSSWTAADTAAQITQRKWEAAREAEQLWRAYLEGLCWEWLRRYLENGKETLQRAD
**.*

3 3 4 3 3



日本人で頻度の高いHLA-B アレル5種間

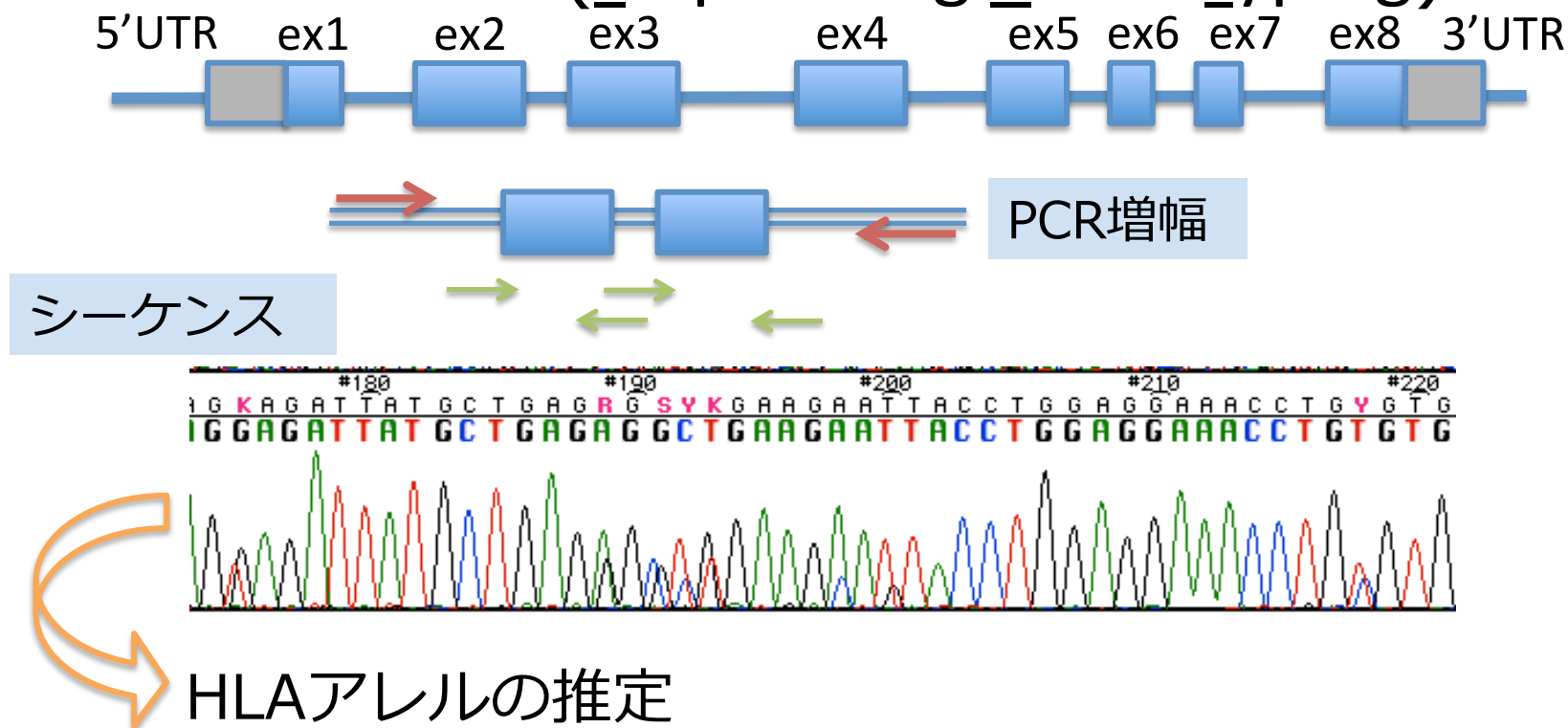
アミノ酸配列での類似性: **88.9%**
塩基配列での類似性: **93.9%**

IMGT/HLAデータベースに登録されている HLAアレル数

Numbers of HLA Alleles			
HLA Class I Alleles		7,089	
HLA Class II Alleles		2,065	
HLA Alleles		9,154	
HLA Class I			
Gene	A	B	C
Alleles	2,244	2,934	1,788
Proteins	1,612	2,211	1,280
Nulls	109	97	47
HLA Class II			
Gene	DRB	DQB1	DPB1
Alleles	1,418	323	185
Proteins	1,051	216	153
Nulls	32	7	6

現行のHLAタイピング法

PCR-SBT (sequencing based typing)



複数の組み合わせ候補

Combination 1 $B^*07:021 + B^*35:011$

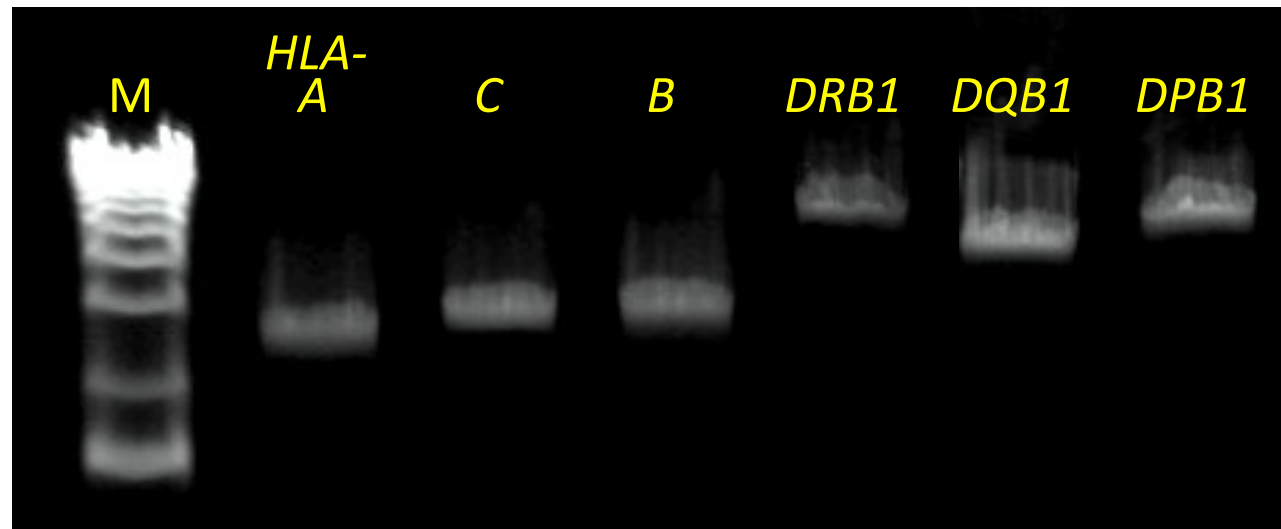
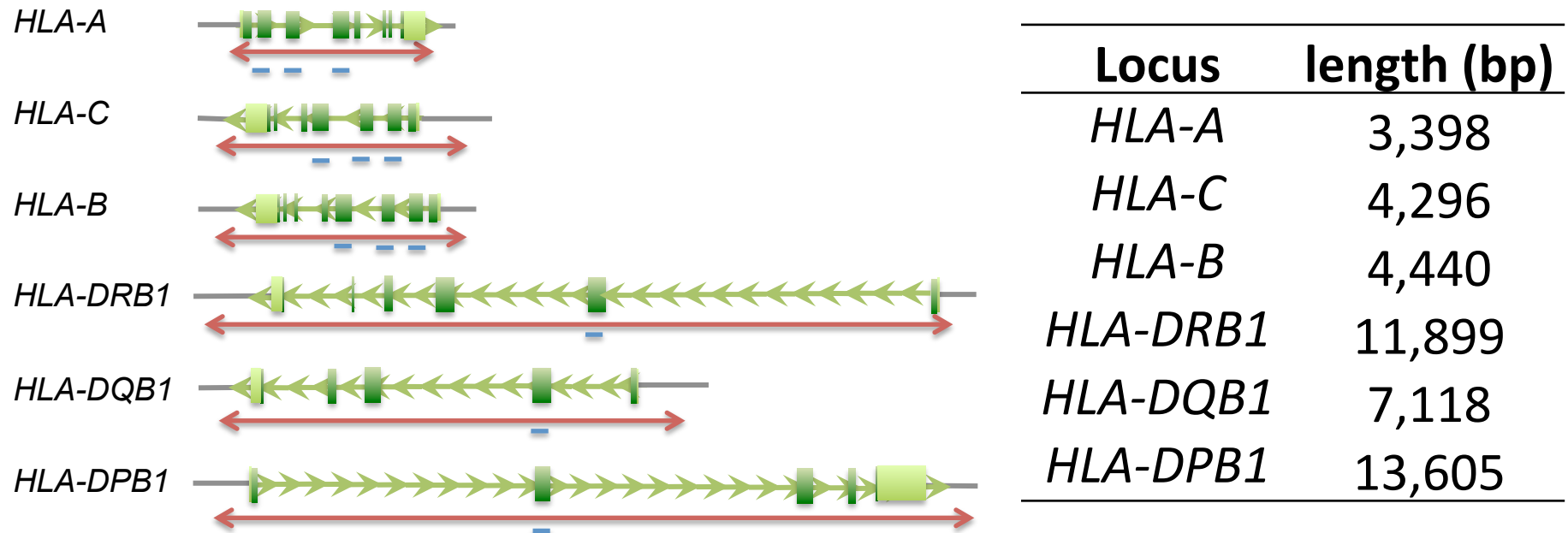
Combination 2 $B^*07:18 + B^*35:05$

Combination 3 $B^*07:09 + B^*35:34$

Combination 4 $B^*07:24 + B^*35:15$

Ambiguity

HLA遺伝子のロングPCR



NexteraとMiSeqによる HLA遺伝子のシーケンシング

PCR増幅

locus	Length (bp)
HLA-A	3,398
HLA-B	4,296
HLA-C	4,440
HLA-DRB1	11,899
HLA-DQB1	7,118
HLA-DPB1	13,605

ライブラリ調整

Nextera DNA Sample Preparation Kit
(illumina)

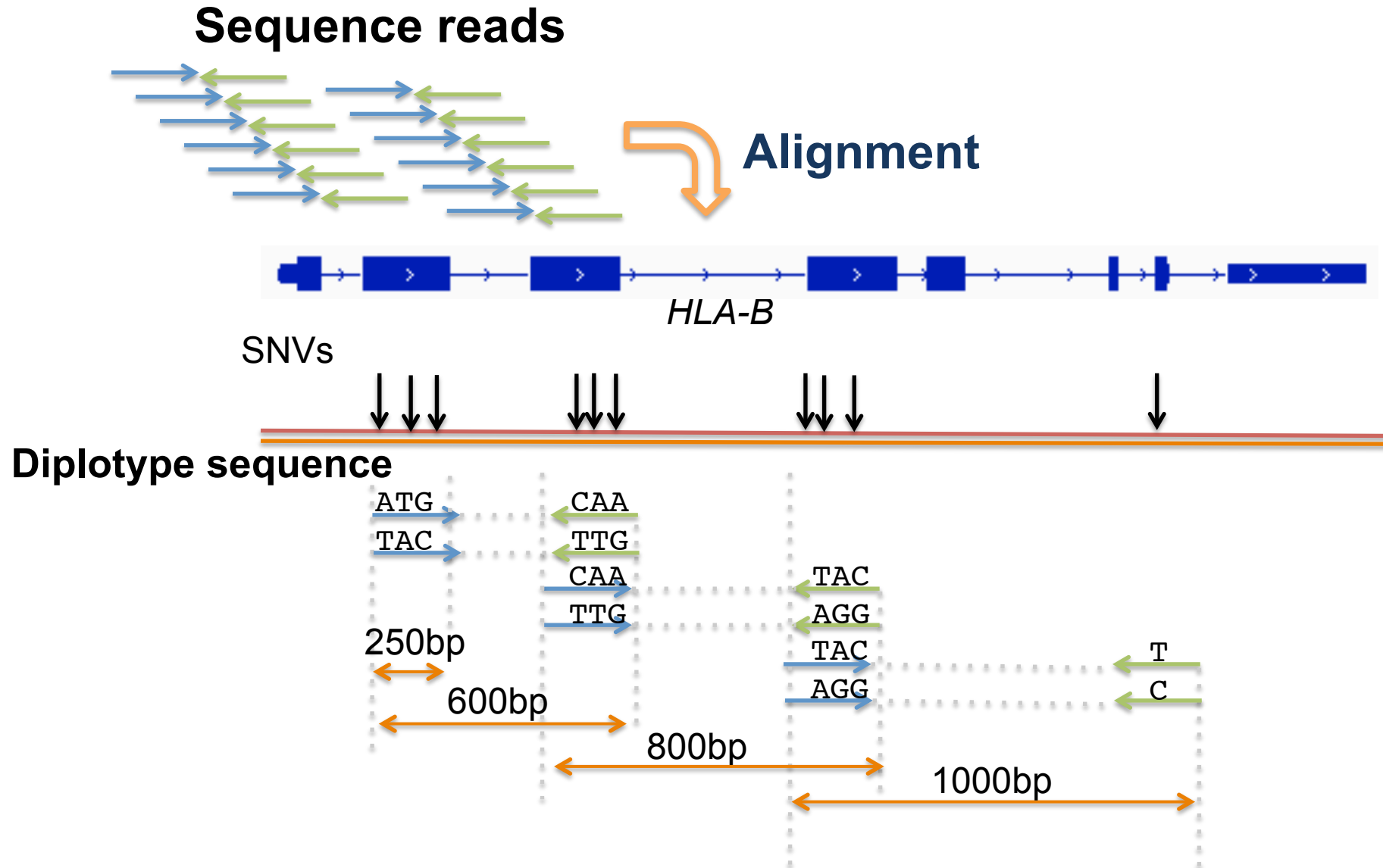


シーケンシング MiSeq v2 (illumina)

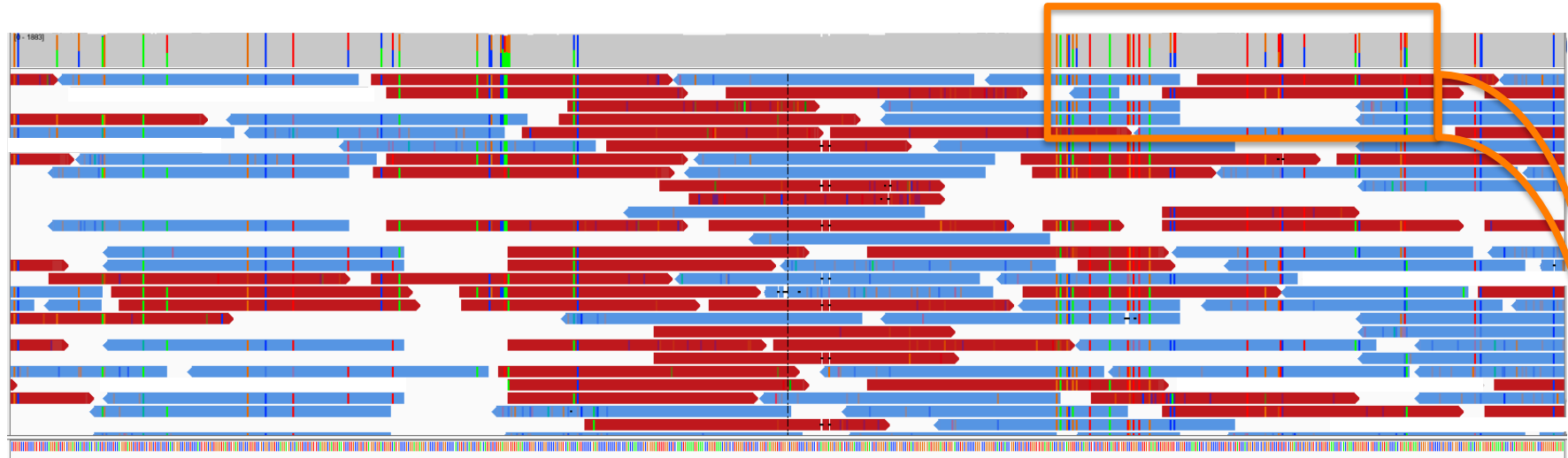
Read length 2 x 250 bp
Output 7.5-8.5 Gb
Total time 39 hr

データ解析

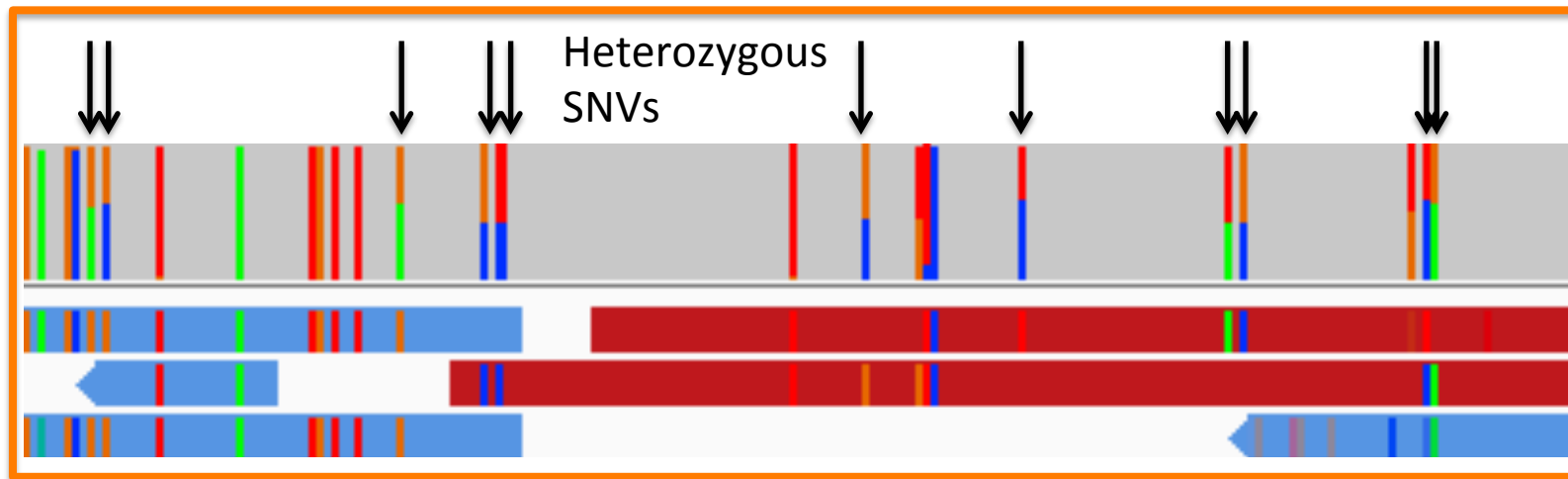
HLA遺伝子配列決定の概要



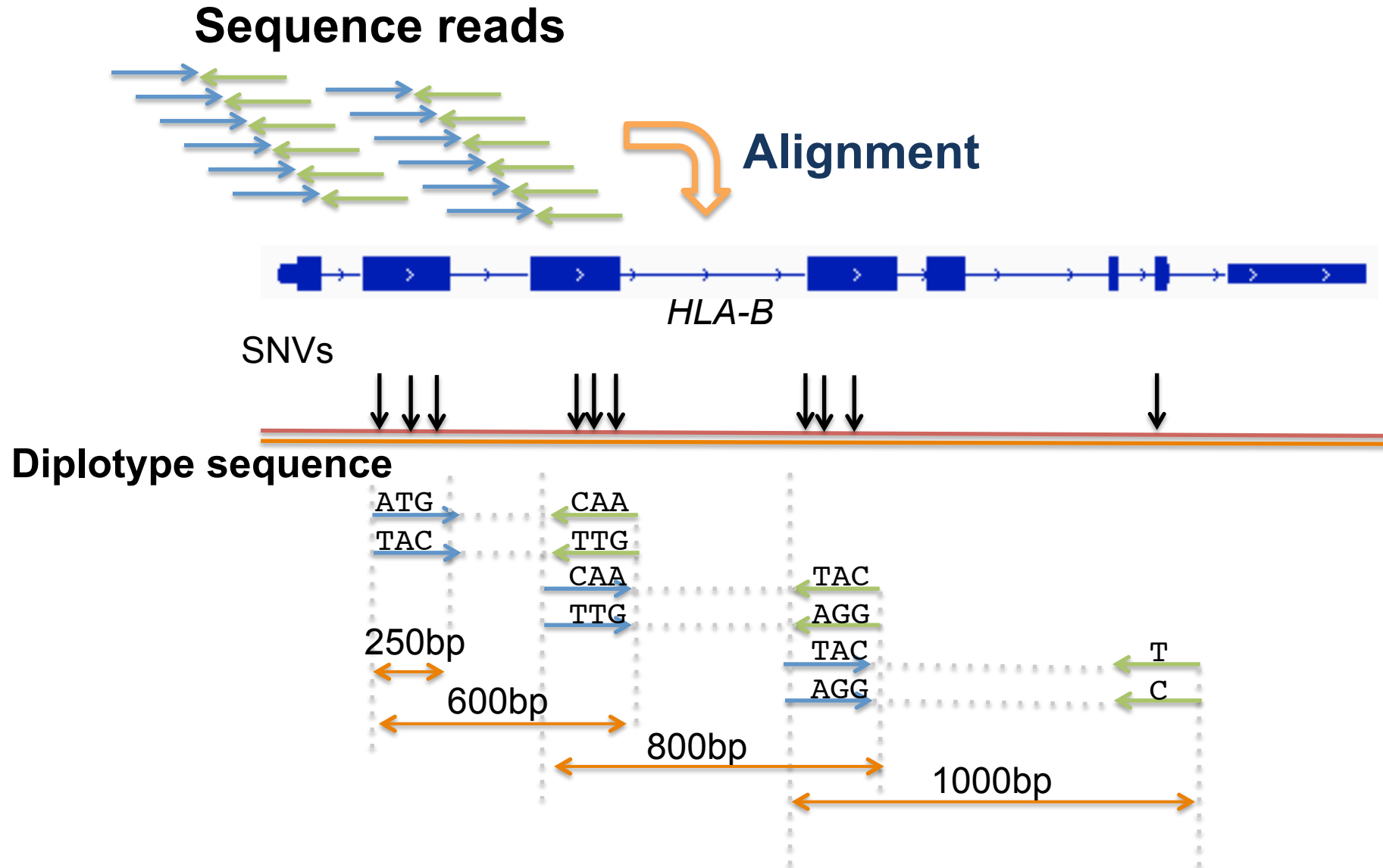
HLA-B エクソン2および3



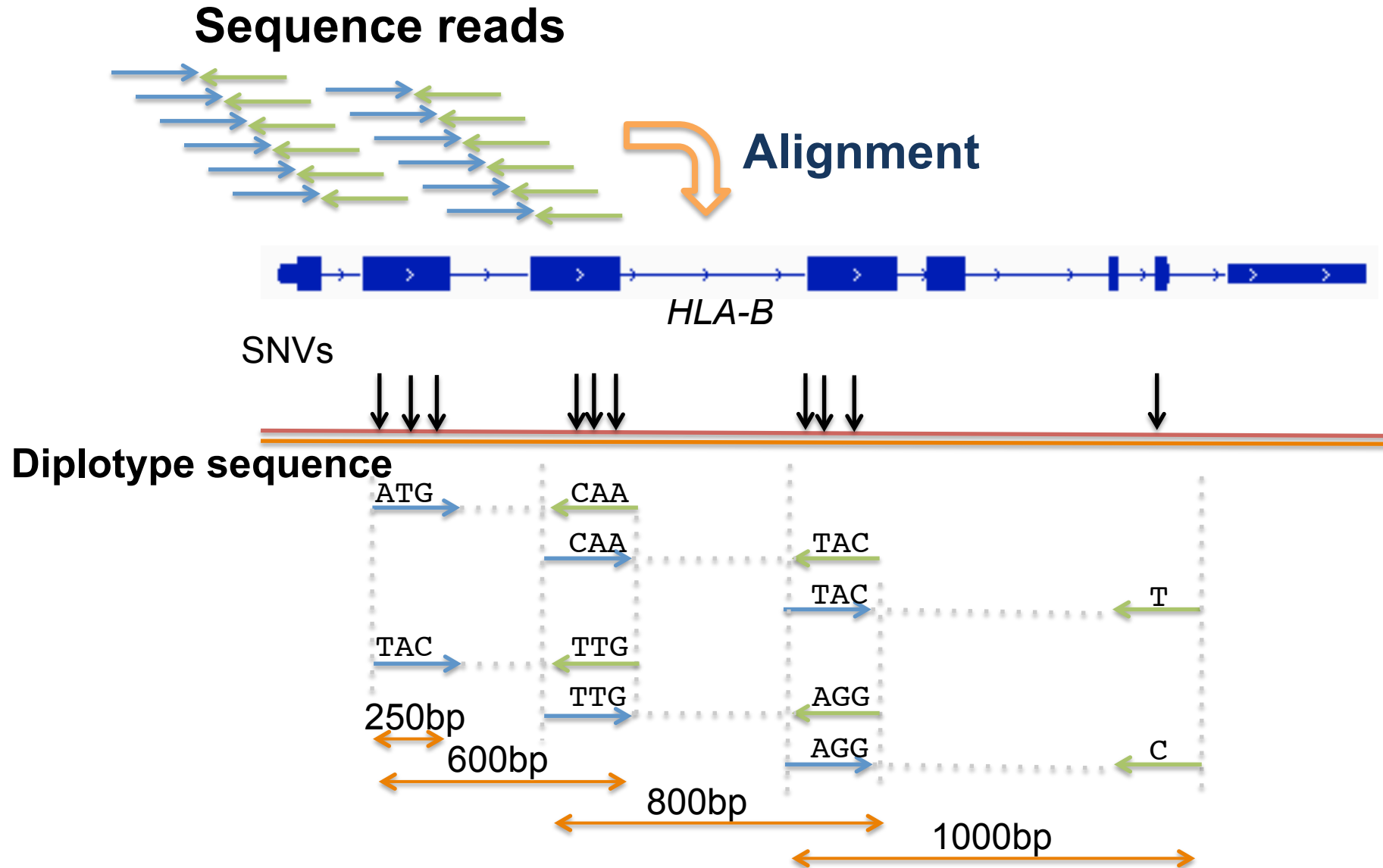
maximum differences = 80 in the read(250bp)



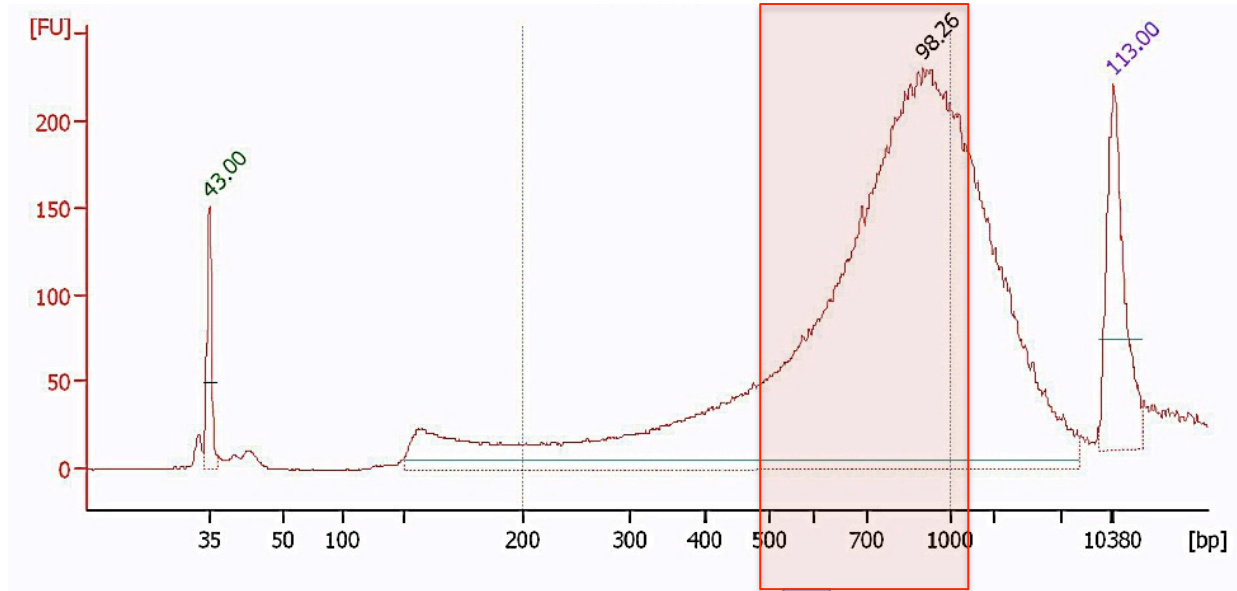
HLA遺伝子配列決定の概要



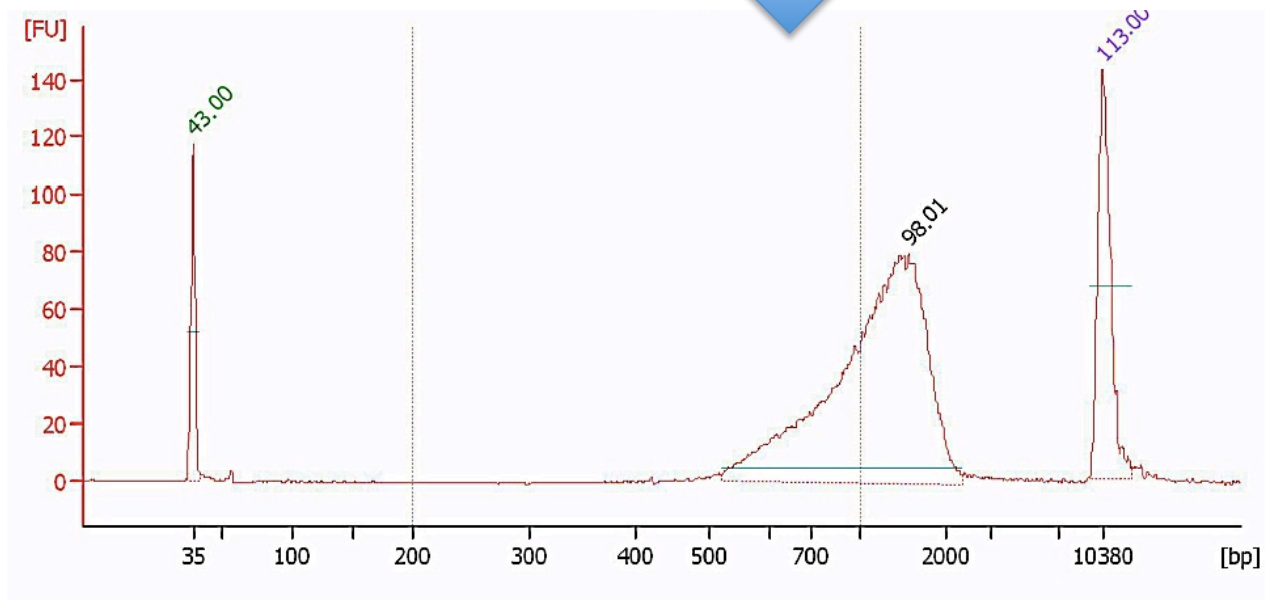
HLA遺伝子配列決定の概要



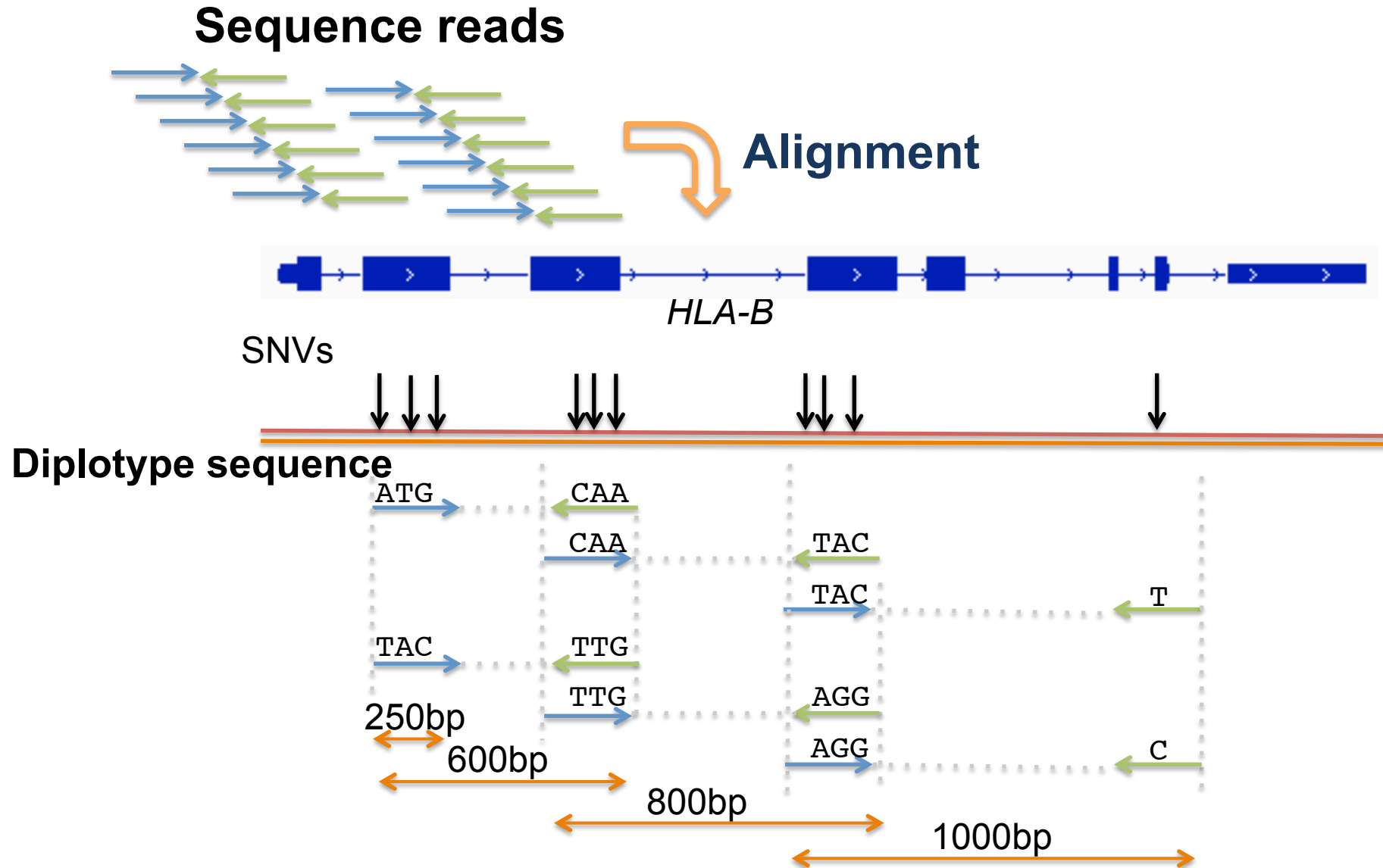
Nextera DNA ライブラリのサイズ選択



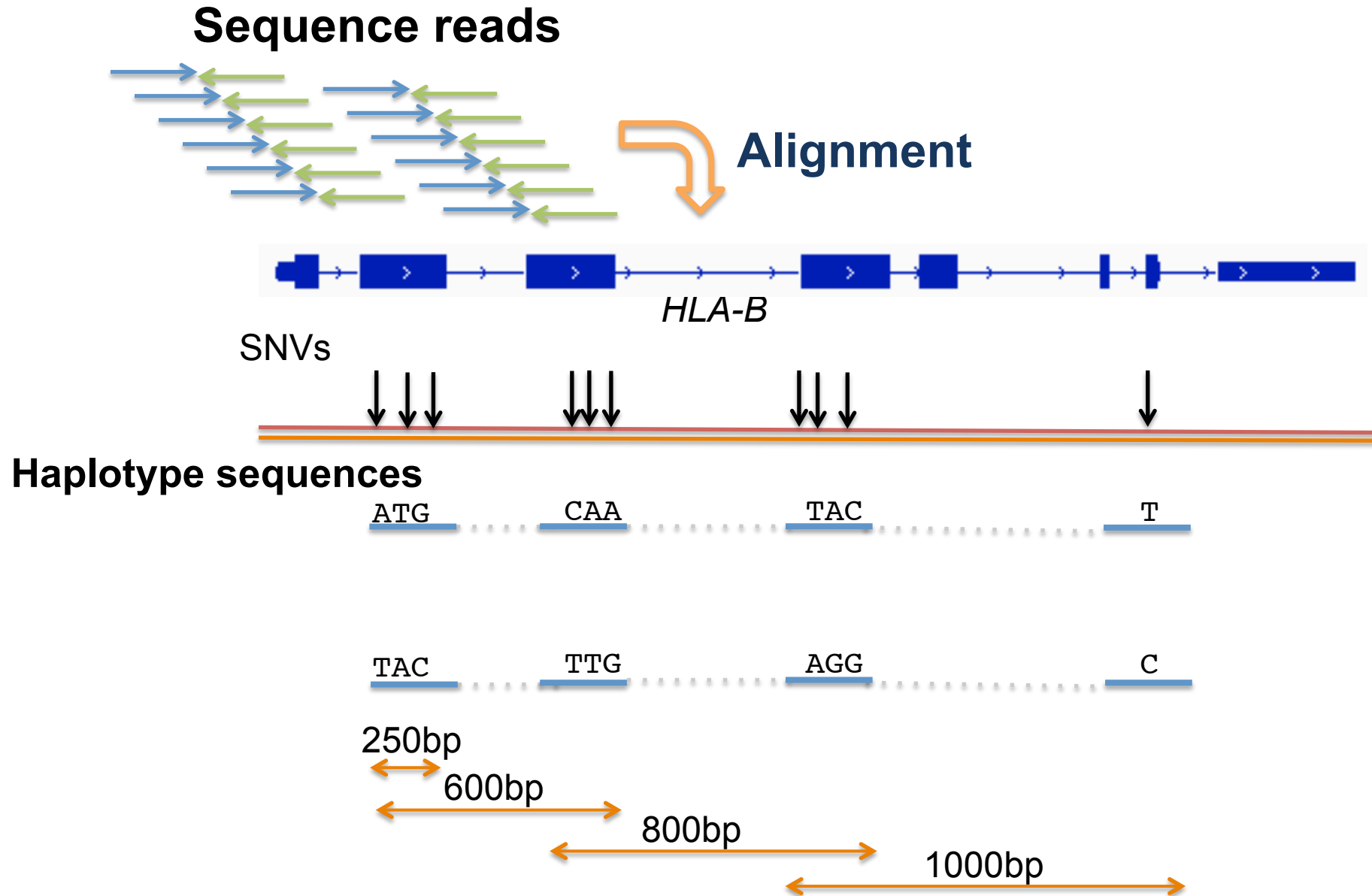
Agarose gel size selection



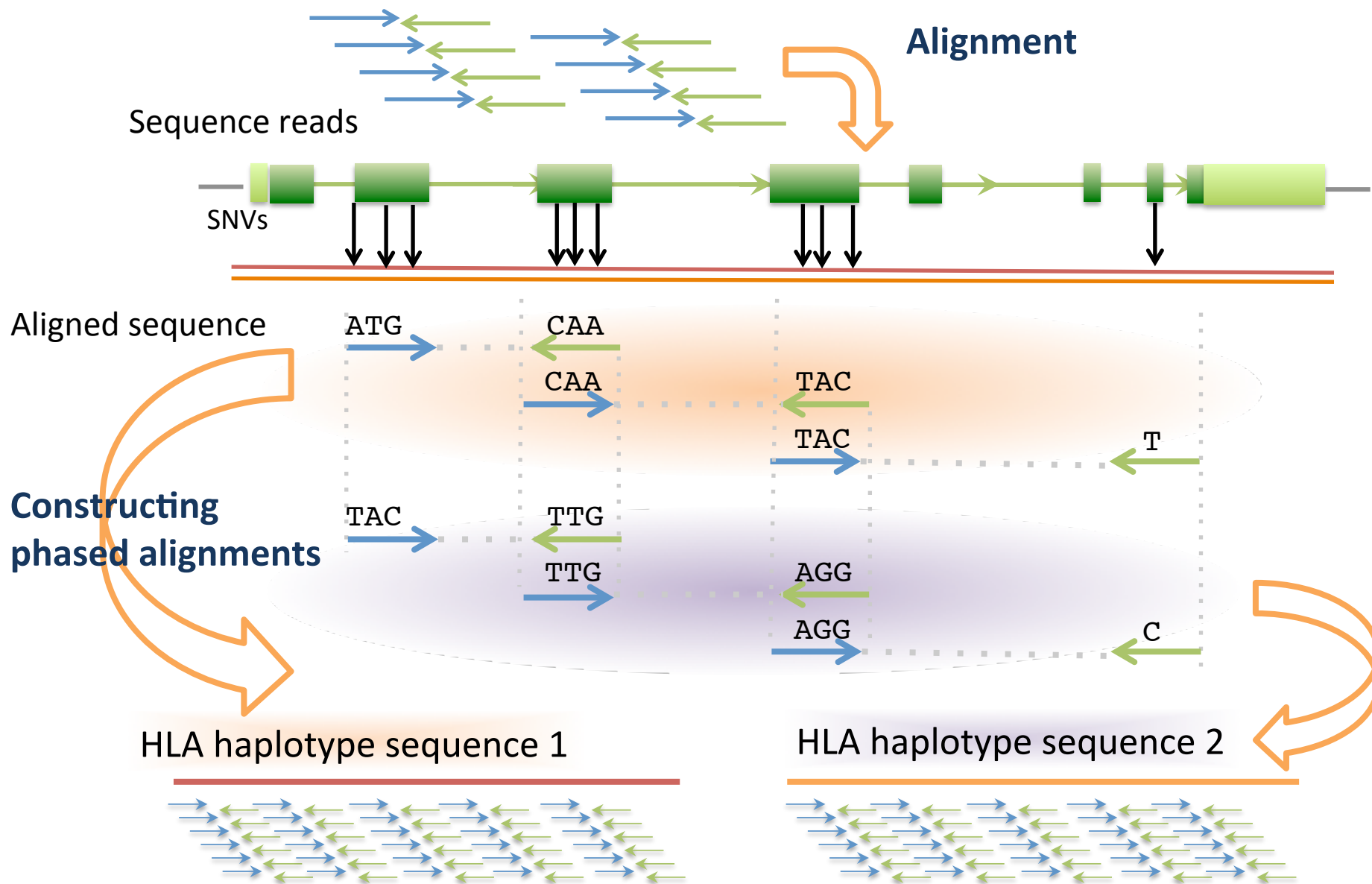
HLA遺伝子配列決定の概要



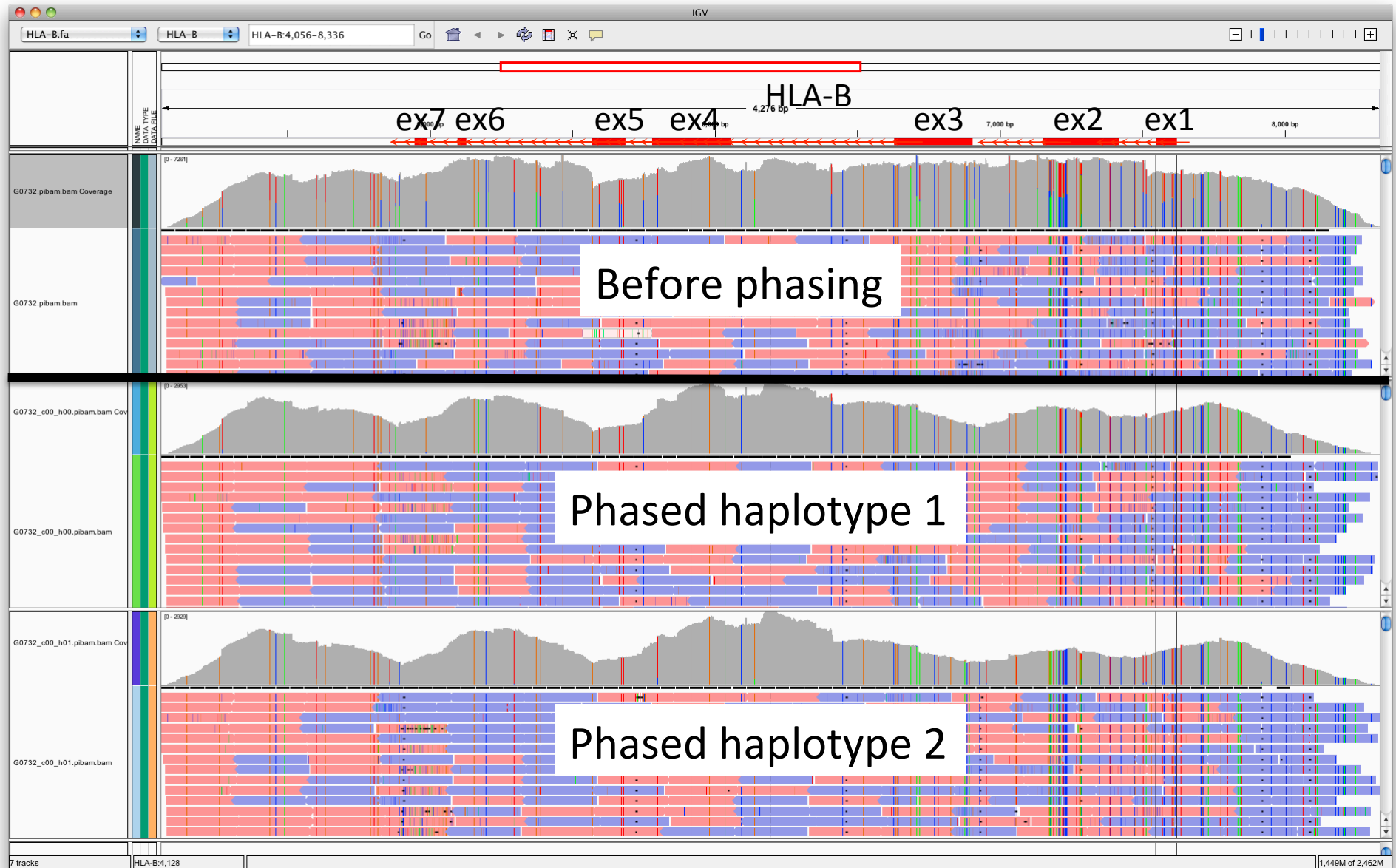
HLA遺伝子配列決定の概要



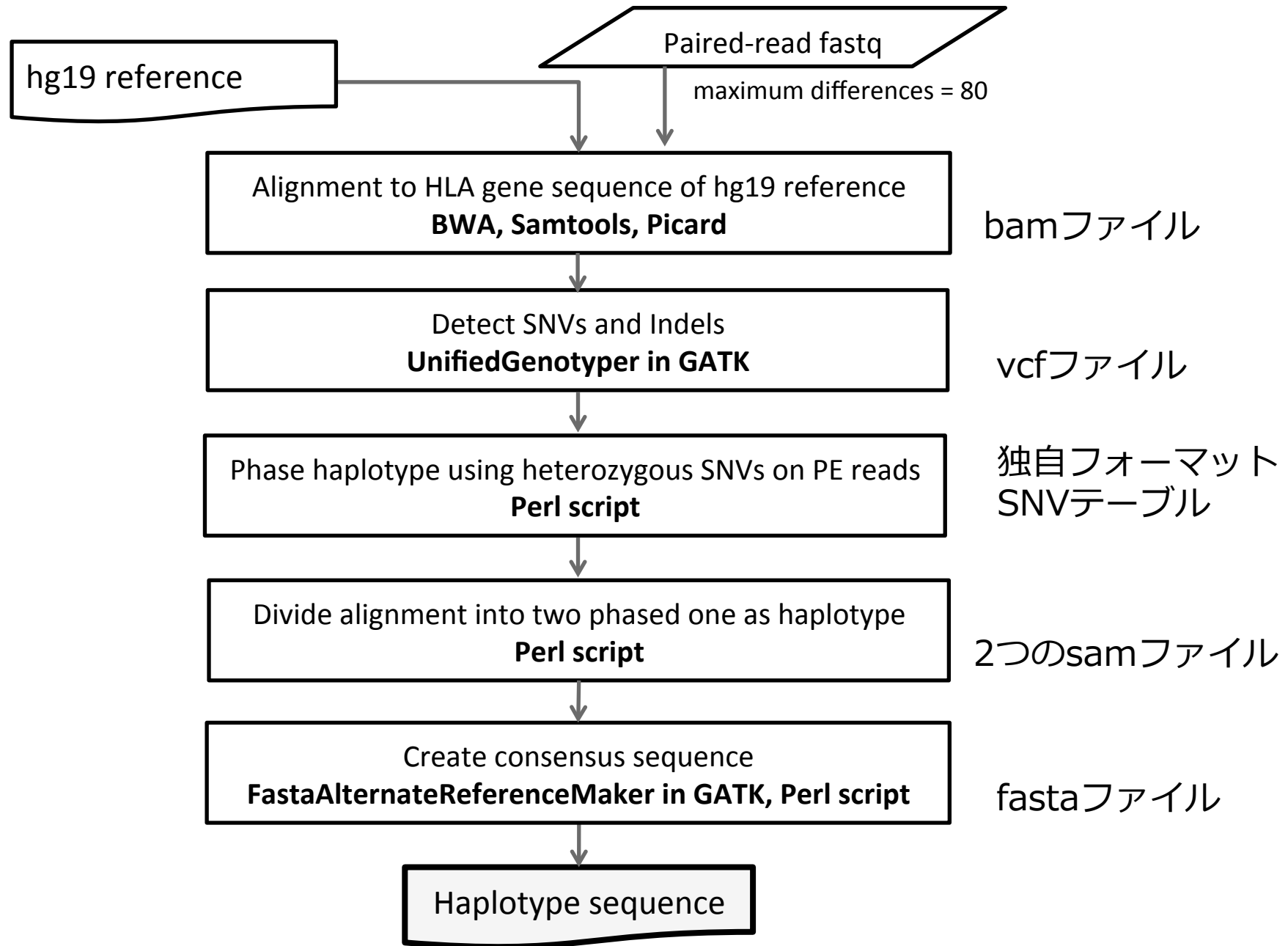
相を特定した2つのHLA遺伝子配列の アライメントの構築



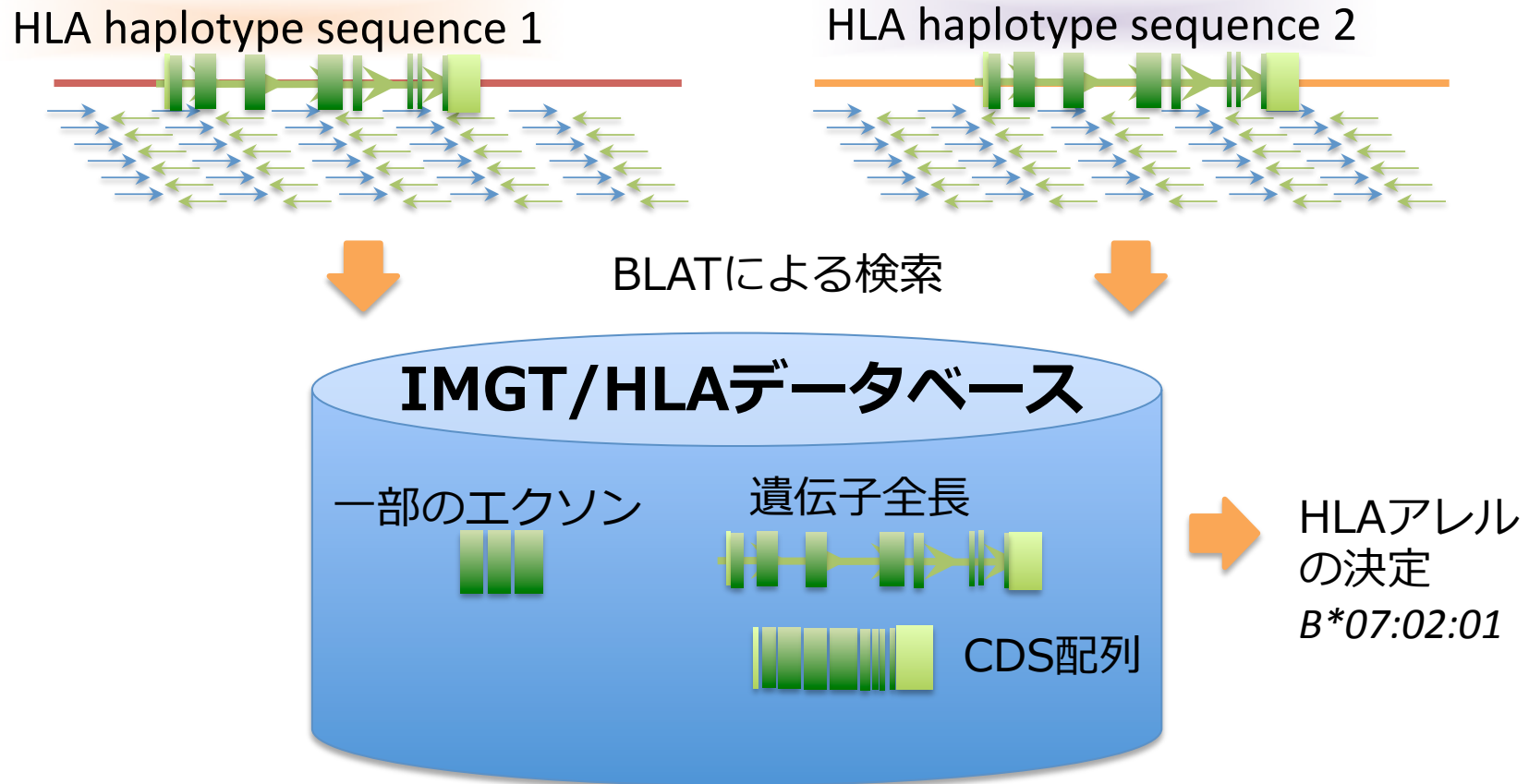
ハプロタイプに分類したアライメント



HLAハプロタイプシーケンス決定のワークフロー



決定したHLA遺伝子完全配列の HLAアレル決定



HLAアレルを決定することよりもHLA遺伝子の塩基配列を完全に決定することが本質

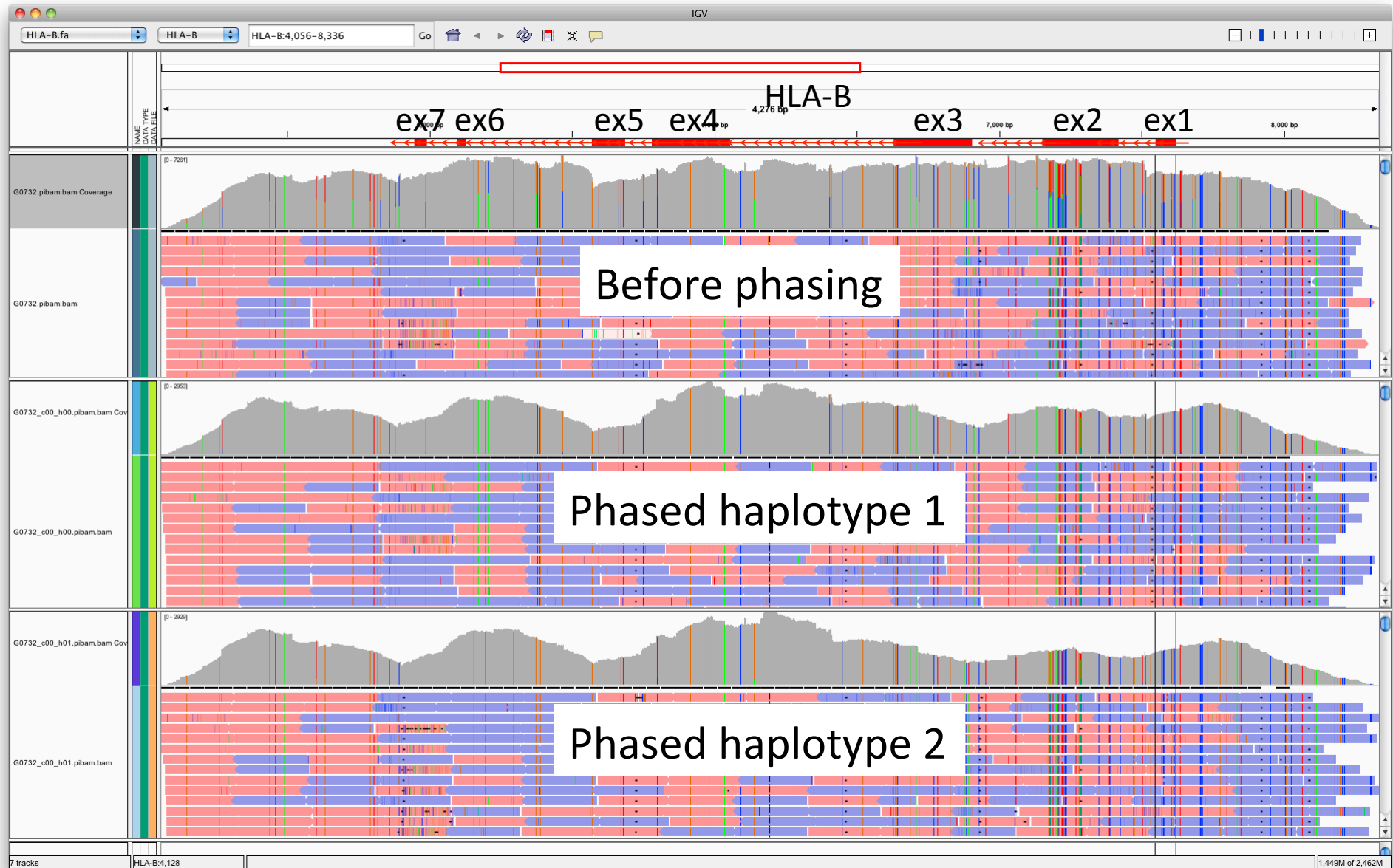
決定したHLA遺伝子配列

Sample ID	PCR-SSO		Next-gen SBT		Average read depth (allele1 / allele2)	Determined sequence (%) (allele1 / allele2)
1	<i>B*51:01:01</i>	<i>B*15:02</i>	<i>B*51:01:01</i>	<i>B*15:02:01</i>	2593.5 / 2431.3	100 / 100
2	<i>B*07:02:01</i>	<i>B*15:02</i>	<i>B*07:02:01</i>	<i>B*15:02:01</i>	3391.2 / 4206.2	100 / 100
3	<i>B*35:01:01</i>	<i>B*15:02</i>	<i>B*35:01:01:01</i>	<i>B*15:02:01</i>	2096.2 / 2431.8	100 / 100
4	<i>B*15:21</i>	<i>B*15:02:01</i>	<i>B*15:21</i>	<i>B*15:02:01</i>	4308.2 / 4693.6	100 / 100
5	<i>B*35:05</i>	<i>B*15:02:01</i>	<i>B*35:05:01</i>	<i>B*15:02:01</i>	4264.3 / 3620.2	100 / 100
6	<i>B*15:18</i>	<i>B*15:02</i>	<i>B*15:18:01</i>	<i>B*15:02:01</i>	1239.9 / 1277.3	100 / 100
7	<i>B*40:06:01:01</i>	<i>B*15:02</i>	<i>B*40:06:01:01</i>	<i>B*15:02:01</i>	5552.1 / 5139.5	100 / 100
8	<i>B*52:01:01</i>	<i>B*15:02</i>	<i>B*52:01:01:01</i>	<i>B*15:02:01</i>	4651.7 / 5085.9	100 / 100
9	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	2326.4 / 2195.1	100 / 100
10	<i>B*54:01</i>	<i>B*57:01:01</i>	<i>B*54:01:01</i>	<i>B*57:01:01</i>	1999.6 / 2034.6	100 / 100
11	<i>B*40:06:01:01</i>	<i>B*57:01:01</i>	<i>B*40:06:01:01</i>	<i>B*57:01:01</i>	3087.4 / 2833.9	100 / 100
12	<i>B*15:11:01</i>	<i>B*57:01:01</i>	<i>B*15:11:01</i>	<i>B*57:01:01</i>	1948.4 / 1721.9	100 / 100
13	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	1660.0 / 1482.0	100 / 100
14	<i>B*44:03:01</i>	<i>B*57:01:01</i>	<i>B*44:03:01</i>	<i>B*57:01:01</i>	806.0 / 668.2	100 / 100
15	<i>B*55:07</i>	<i>B*58:01</i>	<i>B*55:07</i>	<i>B*58:01:01</i>	771.6 / 1000.2	100 / 100
16	<i>B*38:01:01</i>	<i>B*58:01</i>	<i>B*38:01:01</i>	<i>B*58:01:01</i>	1565.1 / 2139.7	100 / 100
17	<i>B*48:01:01</i>	<i>B*58:01</i>	<i>B*48:01:01</i>	<i>B*58:01:01</i>	5555.3 / 4610.8	100 / 100
18	<i>B*35:01:01</i>	<i>B*58:01:01</i>	<i>B*35:01:01:01</i>	<i>B*58:01:01</i>	819.5 / 778.1	100 / 100
19	<i>B*15:25:01</i>	<i>B*58:01:01</i>	<i>B*15:25:01</i>	<i>B*58:01:01</i>	3441.2 / 3852.0	100 / 100
20	<i>B*54:01</i>	<i>B*58:01:01</i>	<i>B*54:01:01</i>	<i>B*58:01:01</i>	1308.3 / 1454.1	100 / 100
21	<i>B*51:02:01</i>	<i>B*58:01:01</i>	<i>B*51:02:01</i>	<i>B*58:01:01</i>	1464.0 / 1132.7	100 / 100
22	<i>B*51:01:01</i>	<i>B*58:01:01</i>	<i>B*51:01:01</i>	<i>B*58:01:01</i>	462.9 / 636.3	100 / 100
23	<i>B*51:02:01</i>	<i>B*58:01:01</i>	<i>B*51:02:02</i>	<i>B*58:01:01</i>	637.0 / 663.7	100 / 100
24	<i>B*39:23</i>	<i>B*58:01:01</i>	<i>B*39:23</i>	<i>B*58:01:01</i>	2112.6 / 2553.3	100 / 100

決定したHLA遺伝子配列

Sample ID	PCR-SSO		Next-gen SBT		Average read depth (allele1 / allele2)	Determined sequence (%) (allele1 / allele2)
1	<i>B*51:01:01</i>	<i>B*15:02</i>	<i>B*51:01:01</i>	<i>B*15:02:01</i>	2593.5 / 2431.3	100 / 100
2	<i>B*07:02:01</i>	<i>B*15:02</i>	<i>B*07:02:01</i>	<i>B*15:02:01</i>	3391.2 / 4206.2	100 / 100
3	<i>B*35:01:01</i>	<i>B*15:02</i>	<i>B*35:01:01:01</i>	<i>B*15:02:01</i>	2096.2 / 2431.8	100 / 100
4	<i>B*15:21</i>	<i>B*15:02:01</i>	<i>B*15:21</i>	<i>B*15:02:01</i>	4308.2 / 4693.6	100 / 100
5	<i>B*35:05</i>	<i>B*15:02:01</i>	<i>B*35:05:01</i>	<i>B*15:02:01</i>	4264.3 / 3620.2	100 / 100
6	<i>B*15:18</i>	<i>B*15:02</i>	<i>B*15:18:01</i>	<i>B*15:02:01</i>	1239.9 / 1277.3	100 / 100
7	<i>B*40:06:01:01</i>	<i>B*15:02</i>	<i>B*40:06:01:01</i>	<i>B*15:02:01</i>	5552.1 / 5139.5	100 / 100
8	<i>B*52:01:01</i>	<i>B*15:02</i>	<i>B*52:01:01:01</i>	<i>B*15:02:01</i>	4651.7 / 5085.9	100 / 100
9	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	2326.4 / 2195.1	100 / 100
10	<i>B*54:01</i>	<i>B*57:01:01</i>	<i>B*54:01:01</i>	<i>B*57:01:01</i>	1999.6 / 2034.6	100 / 100
11	<i>B*40:06:01:01</i>	<i>B*57:01:01</i>	<i>B*40:06:01:01</i>	<i>B*57:01:01</i>	3087.4 / 2833.9	100 / 100
12	<i>B*15:11:01</i>	<i>B*57:01:01</i>	<i>B*15:11:01</i>	<i>B*57:01:01</i>	1948.4 / 1721.9	100 / 100
13	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	<i>B*15:01:01:01</i>	<i>B*57:01:01</i>	1660.0 / 1482.0	100 / 100
14	<i>B*44:03:01</i>	<i>B*57:01:01</i>	<i>B*44:03:01</i>	<i>B*57:01:01</i>	806.0 / 668.2	100 / 100
15	<i>B*55:07</i>	<i>B*58:01</i>	<i>B*55:07</i>	<i>B*58:01:01</i>	771.6 / 1000.2	100 / 100
16	<i>B*38:01:01</i>	<i>B*58:01</i>	<i>B*38:01:01</i>	<i>B*58:01:01</i>	1565.1 / 2139.7	100 / 100
17	<i>B*48:01:01</i>	<i>B*58:01</i>	<i>B*48:01:01</i>	<i>B*58:01:01</i>	5555.3 / 4610.8	100 / 100
18	<i>B*35:01:01</i>	<i>B*58:01:01</i>	<i>B*35:01:01:01</i>	<i>B*58:01:01</i>	819.5 / 778.1	100 / 100
19	<i>B*15:25:01</i>	<i>B*58:01:01</i>	<i>B*15:25:01</i>	<i>B*58:01:01</i>	3441.2 / 3852.0	100 / 100
20	<i>B*54:01</i>	<i>B*58:01:01</i>	<i>B*54:01:01</i>	<i>B*58:01:01</i>	1308.3 / 1454.1	100 / 100
21	<i>B*51:02:01</i>	<i>B*58:01:01</i>	<i>B*51:02:01</i>	<i>B*58:01:01</i>	1464.0 / 1132.7	100 / 100
22	<i>B*51:01:01</i>	<i>B*58:01:01</i>	<i>B*51:01:01</i>	<i>B*58:01:01</i>	462.9 / 636.3	100 / 100
23	<i>B*51:02:01</i>	<i>B*58:01:01</i>	<i>B*51:02:01</i>	<i>B*58:01:01</i>	637.0 / 663.7	100 / 100
24	<i>B*39:23</i>	<i>B*58:01:01</i>	<i>B*39:23</i>	<i>B*58:01:01</i>	2112.6 / 2553.3	100 / 100

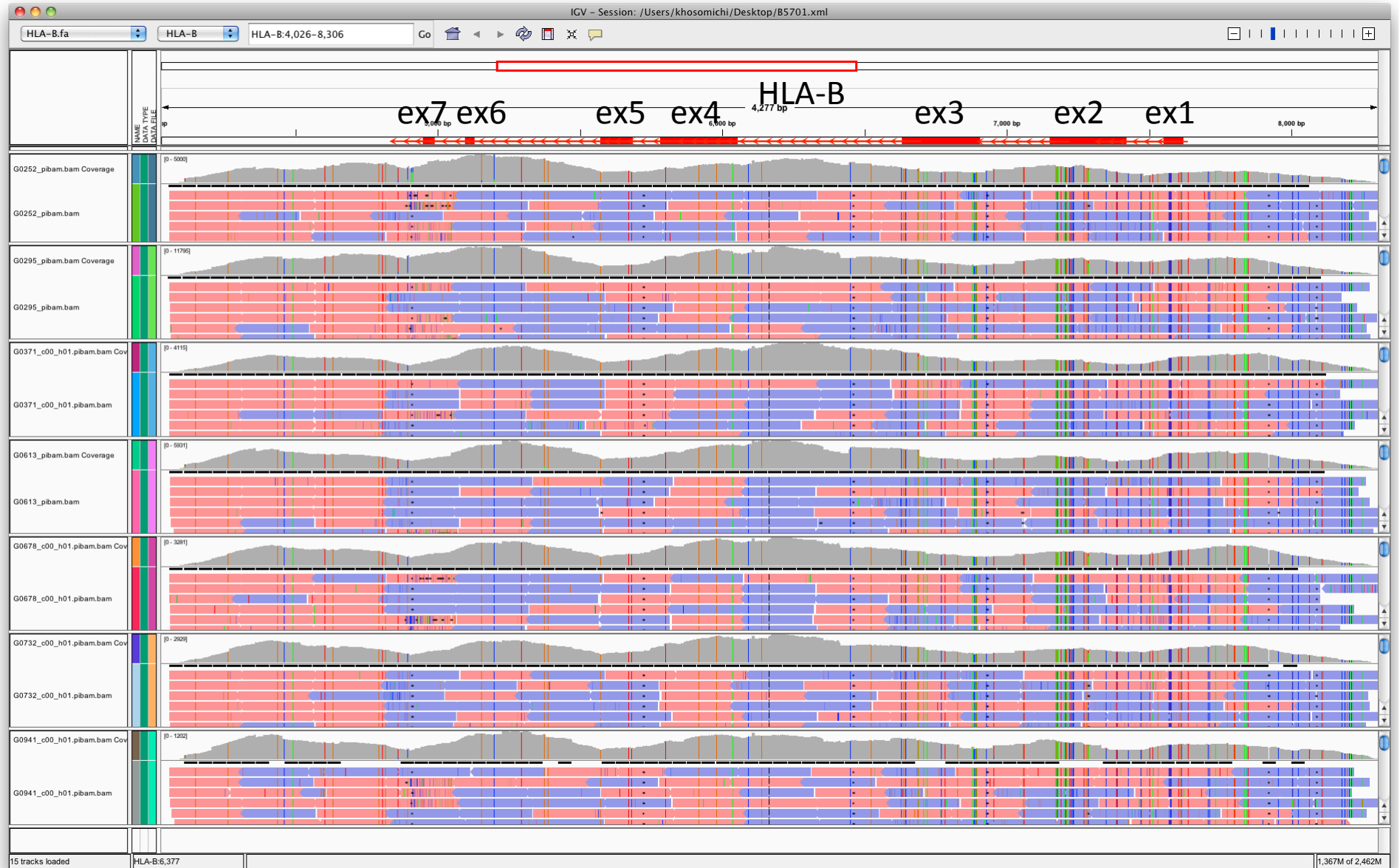
ハプロタイプに分類したアライメント



決定したHLA遺伝子配列

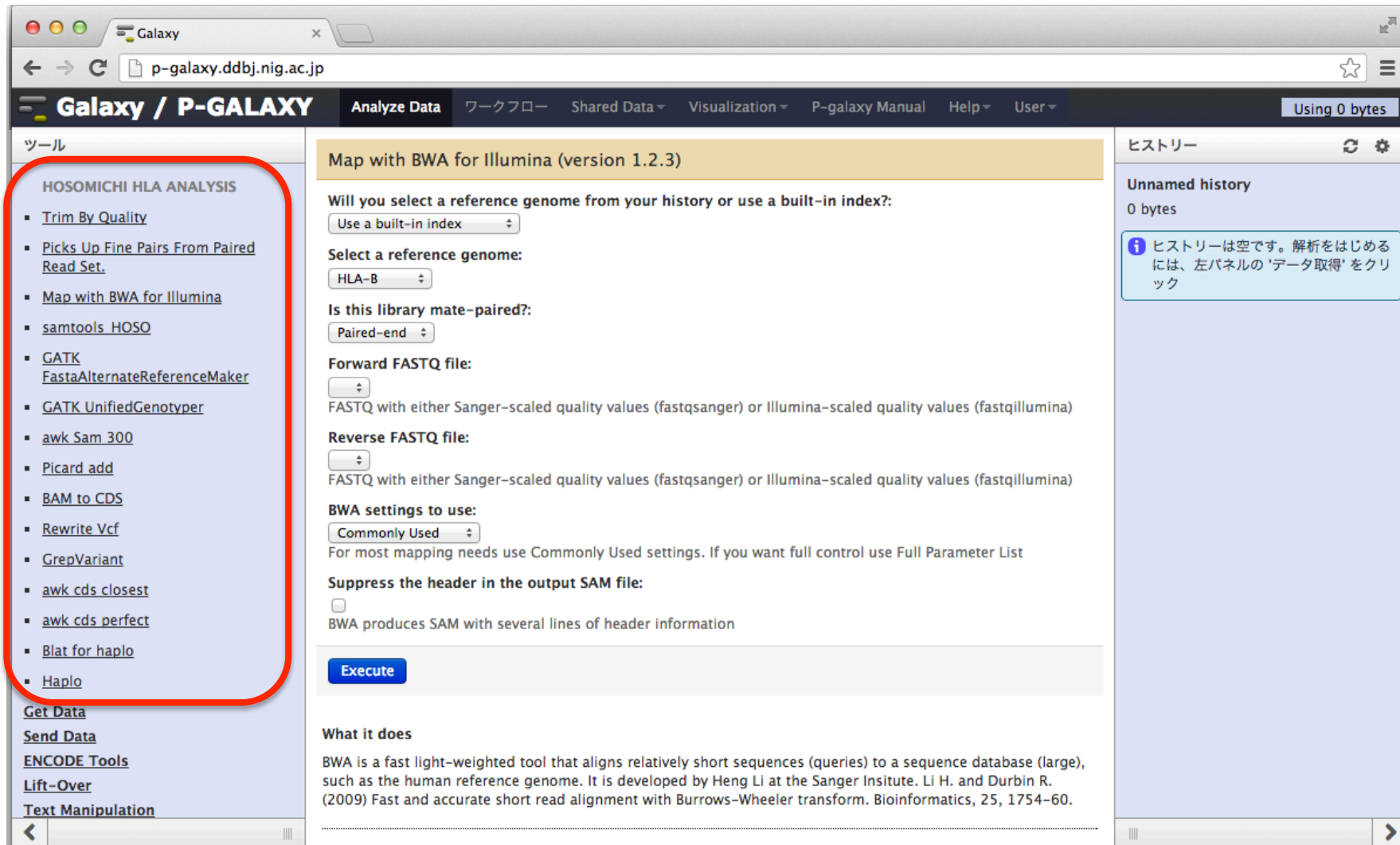
Sample ID	PCR-SSO		Next-gen SBT		Average read depth (allele1 / allele2)	Determined sequence (%) (allele1 / allele2)
1	B*51:01:01	B*15:02	B*51:01:01	B*15:02:01	2593.5 / 2431.3	100 / 100
2	B*07:02:01	B*15:02	B*07:02:01	B*15:02:01	3391.2 / 4206.2	100 / 100
3	B*35:01:01	B*15:02	B*35:01:01:01	B*15:02:01	2096.2 / 2431.8	100 / 100
4	B*15:21	B*15:02:01	B*15:21	B*15:02:01	4308.2 / 4693.6	100 / 100
5	B*35:05	B*15:02:01	B*35:05:01	B*15:02:01	4264.3 / 3620.2	100 / 100
6	B*15:18	B*15:02	B*15:18:01	B*15:02:01	1239.9 / 1277.3	100 / 100
7	B*40:06:01:01	B*15:02	B*40:06:01:01	B*15:02:01	5552.1 / 5139.5	100 / 100
8	B*52:01:01	B*15:02	B*52:01:01:01	B*15:02:01	4651.7 / 5085.9	100 / 100
9	B*15:01:01:01	B*57:01:01	B*15:01:01:01	B*57:01:01	2326.4 / 2195.1	100 / 100
10	B*54:01	B*57:01:01	B*54:01:01	B*57:01:01	1999.6 / 2034.6	100 / 100
11	B*40:06:01:01	B*57:01:01	B*40:06:01:01	B*57:01:01	3087.4 / 2833.9	100 / 100
12	B*15:11:01	B*57:01:01	B*15:11:01	B*57:01:01	1948.4 / 1721.9	100 / 100
13	B*15:01:01:01	B*57:01:01	B*15:01:01:01	B*57:01:01	1660.0 / 1482.0	100 / 100
14	B*44:03:01	B*57:01:01	B*44:03:01	B*57:01:01	806.0 / 668.2	100 / 100
15	B*55:07	B*58:01	B*55:07	B*58:01:01	771.6 / 1000.2	100 / 100
16	B*38:01:01	B*58:01	B*38:01:01	B*58:01:01	1565.1 / 2139.7	100 / 100
17	B*48:01:01	B*58:01	B*48:01:01	B*58:01:01	5555.3 / 4610.8	100 / 100
18	B*35:01:01	B*58:01:01	B*35:01:01:01	B*58:01:01	819.5 / 778.1	100 / 100
19	B*15:25:01	B*58:01:01	B*15:25:01	B*58:01:01	3441.2 / 3852.0	100 / 100
20	B*54:01	B*58:01:01	B*54:01:01	B*58:01:01	1308.3 / 1454.1	100 / 100
21	B*51:02:01	B*58:01:01	B*51:02:01	B*58:01:01	1464.0 / 1132.7	100 / 100
22	B*51:01:01	B*58:01:01	B*51:01:01	B*58:01:01	462.9 / 636.3	100 / 100
23	B*51:02:01	B*58:01:01	B*51:02:01	B*58:01:01	637.0 / 663.7	100 / 100
24	B*39:23	B*58:01:01	B*39:23	B*58:01:01	2112.6 / 2553.3	100 / 100

HLA-B*58:01 の塩基配列の比較



HLA解析ツール (1)

- p-galaxy <http://p-galaxy.ddbj.nig.ac.jp>



The screenshot shows the p-galaxy web interface. The browser address bar displays `p-galaxy.ddbj.nig.ac.jp`. The main navigation bar includes "Galaxy / P-GALAXY", "Analyze Data", "ワークフロー", "Shared Data", "Visualization", "P-galaxy Manual", "Help", and "User". A status bar on the right indicates "Using 0 bytes".

The left sidebar contains a "ツール" (Tools) section with a red box highlighting the "HOSOMICHI HLA ANALYSIS" category. The tools listed are:

- Trim By Quality
- Picks Up Fine Pairs From Paired Read Set.
- Map with BWA for Illumina
- samtools_HOSO
- GATK
- FastaAlternateReferenceMaker
- GATK UnifiedGenotyper
- awk Sam 300
- Picard add
- BAM to CDS
- Rewrite Vcf
- GrepVariant
- awk cds closest
- awk cds perfect
- Blat for haplo
- Haplo

The main content area displays the "Map with BWA for Illumina (version 1.2.3)" tool configuration page. The configuration options are:

- Will you select a reference genome from your history or use a built-in index?:
- Select a reference genome:
- Is this library mate-paired?:
- Forward FASTQ file:
- Reverse FASTQ file:
- BWA settings to use:
- Suppress the header in the output SAM file:

An "Execute" button is located at the bottom of the configuration area. Below the configuration, the "What it does" section provides a description of BWA: "BWA is a fast light-weighted tool that aligns relatively short sequences (queries) to a sequence database (large), such as the human reference genome. It is developed by Heng Li at the Sanger Institute. Li H. and Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25, 1754-60."

The right sidebar shows the "ヒストリー" (History) section, which is currently empty. A message box states: "ヒストリーは空です。解析をはじめするには、左パネルの 'データ取得' をクリック"

DDBJパイプラインの詳細

ILLUMINA® Webinar Series

de novoシリーズ 第3回

「DDBJパイプラインによるRNA-seq配列のde novoアセンブル」

イルミナウェビナー de novoシリーズ 第3回は、国立遺伝学研究所から3名の先生をお招きして、DDBJの解析パイプラインに関する内容をデモを交えてご紹介いただきます。

de novoシリーズの第1回「非モデル生物のRNA-seq解析 -- 実験デザインから解析パイプラインまで」、第2回「Miseqでのde novo実験」の録画および資料は、[イルミナウェビナーページ](#)でご覧いただけます。

皆様のご参加、お待ちしております。

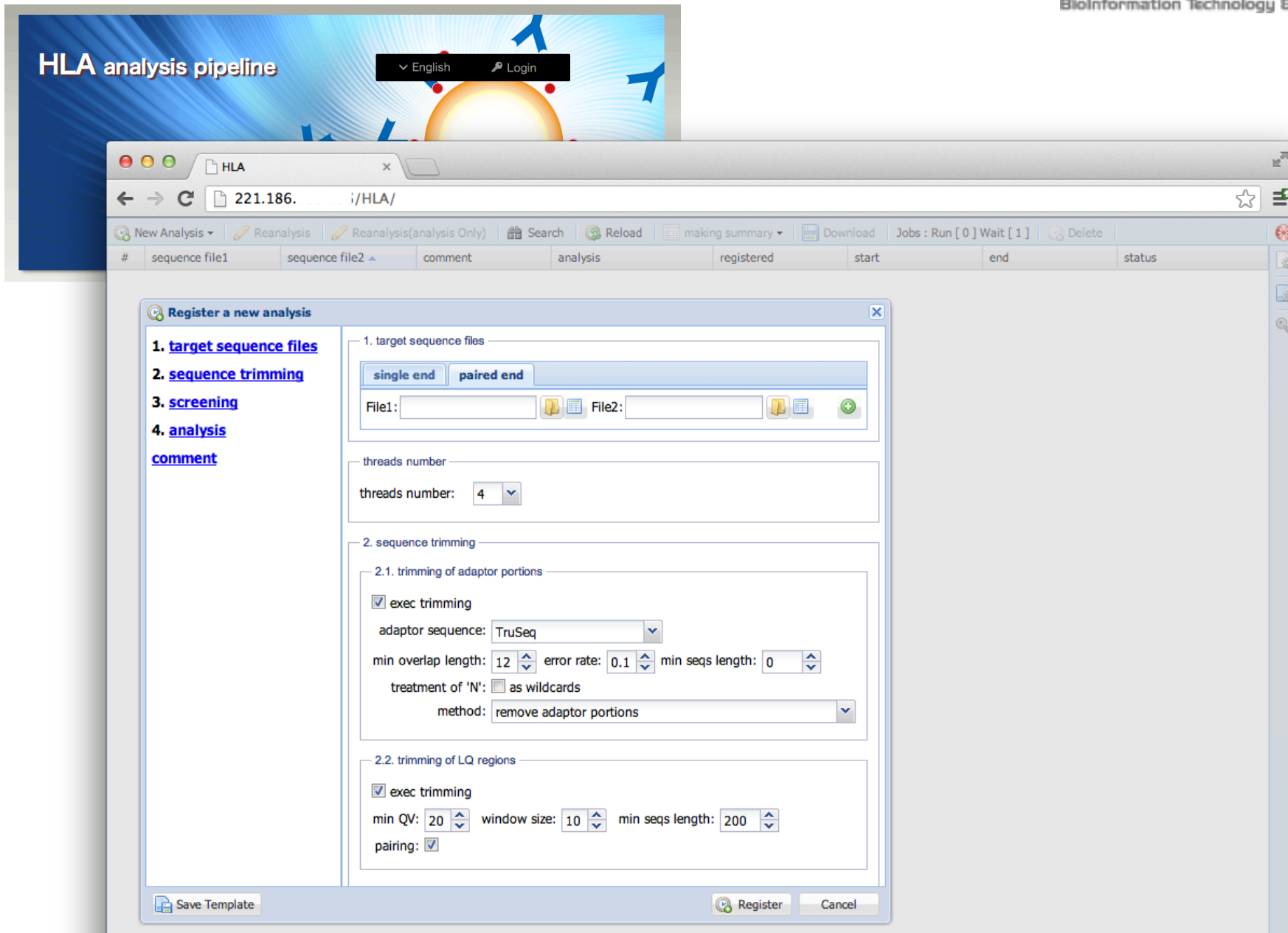
講演タイトル：「DDBJパイプラインによるRNA-seq配列のde novoアセンブル」

講演者： 国立遺伝学研究所 大量遺伝情報研究室
中村 保一 先生、長崎 英樹 先生、谷沢 靖洋 先生

日程： 2013年5月8日（水曜日）

時間： 16:00 - 17:00

HLA解析ツール (2)



The screenshot displays the HLA analysis pipeline web interface. At the top left, a banner reads "HLA analysis pipeline" with a navigation menu showing "English" and "Login". Below this is a browser window with the URL "221.186. ... /HLA/". The browser's toolbar includes options like "New Analysis", "Reanalysis", "Reanalysis(analysis Only)", "Search", "Reload", "making summary", "Download", "Jobs : Run [0] Wait [1]", and "Delete". A table below the toolbar lists analysis jobs with columns for "#", "sequence file1", "sequence file2", "comment", "analysis", "registered", "start", "end", and "status".

The main content area features a "Register a new analysis" dialog box with the following sections:






- 1. target sequence files**
 - Buttons for "single end" and "paired end".
 - Input fields for "File1:" and "File2:" with file selection icons and a "+" button.
- threads number**
 - Input field: "threads number: 4" with a dropdown arrow.
- 2. sequence trimming**
 - 2.1. trimming of adaptor portions**
 - exec trimming
 - adaptor sequence: TruSeq (dropdown)
 - min overlap length: 12 (spinners) error rate: 0.1 (spinners) min seqs length: 0 (spinners)
 - treatment of 'N': as wildcards
 - method: remove adaptor portions (dropdown)
 - 2.2. trimming of LQ regions**
 - exec trimming
 - min QV: 20 (spinners) window size: 10 (spinners) min seqs length: 200 (spinners)
 - pairing:

At the bottom of the dialog are buttons for "Save Template", "Register", and "Cancel".


HLA解析パイプライン

1. target sequence files

single end paired end

File1:   File2:   

threads number

threads number: 

HLA解析パイプライン

1. target sequence files

single end paired end

2. sequence trimming

2.1. trimming of adaptor portions

exec trimming

adaptor sequence: TruSeq

min overlap length: 12 error rate: 0.1 min seqs length: 0

treatment of 'N': as wildcards

method: remove adaptor portions

2.2. trimming of LQ regions

exec trimming

min QV: 20 window size: 10 min seqs length: 200

pairing:

HLA解析パイプライン

1. target sequence files

single end | **paired end**

2. sequence trimming

2.1. trimming of adaptor portions

exec trimming

adaptor sequence: TruSeq

min overlap length: 12 error rate: 0.1 min seqs length: 0

treatment of 'N': as wildcards

method: remove adaptor portions

2.2. trimming of LQ regions

exec trimming

min QV: 20 window size: 10 min seqs length: 200

pairing:

4. analysis

Program: bwa (aln)

DB: HLA-A Release: Apr. 2013

4.1. mapping, SNP call, phasing

bwa options

max edit distance: 80

other bwa aln/bwasw options:

other bwa samse/sampe options:

GATK options

contamination rate: 0.1 stand call confidence: 80

stand emit confidence: 50 depth: 1000

other GATK options:

re-calling options

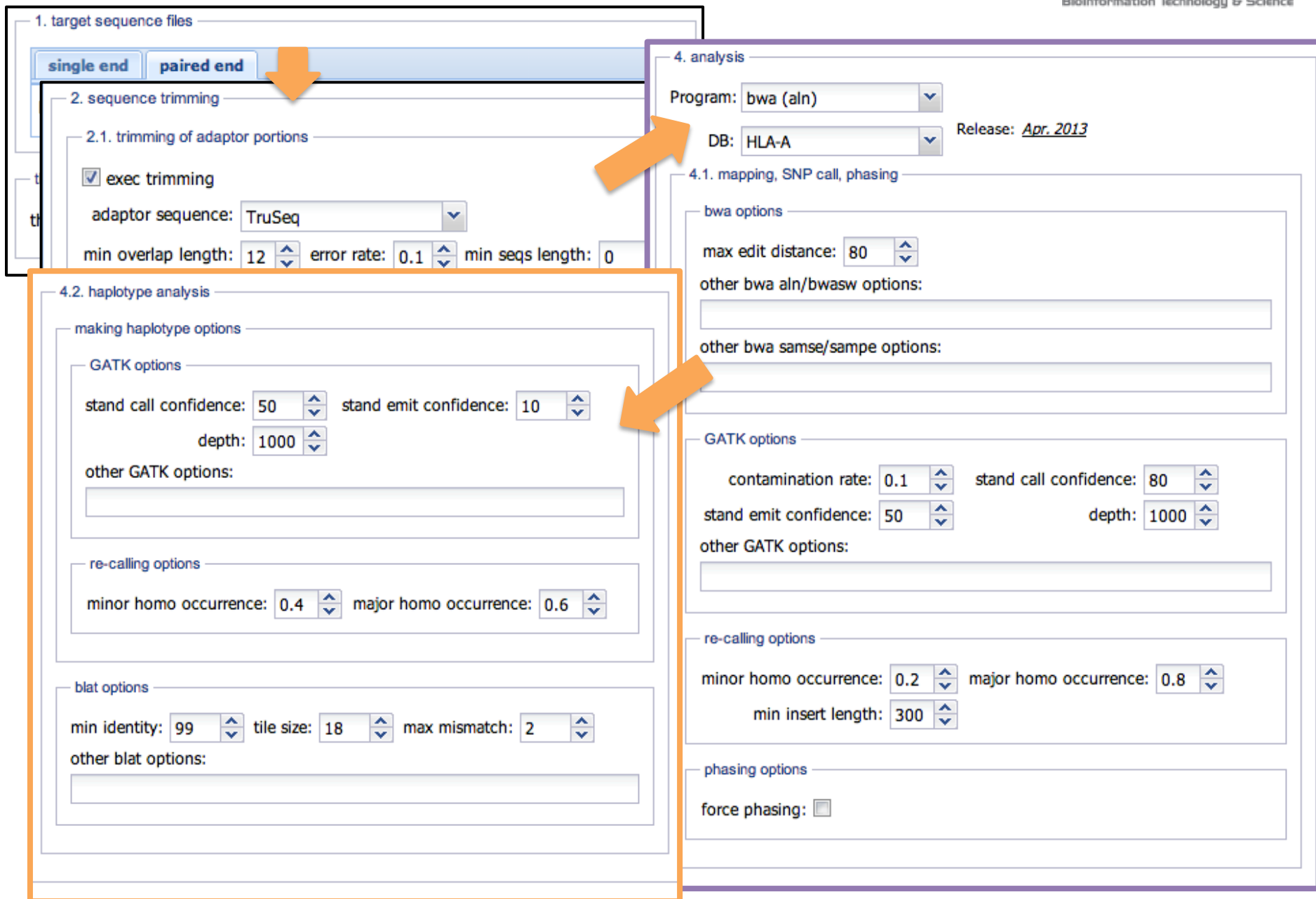
minor homo occurrence: 0.2 major homo occurrence: 0.8

min insert length: 300

phasing options

force phasing:

HLA解析パイプライン



1. target sequence files

single end | paired end

2. sequence trimming

2.1. trimming of adaptor portions

exec trimming

adaptor sequence: TruSeq

min overlap length: 12 error rate: 0.1 min seqs length: 0

4. analysis

Program: bwa (aln) DB: HLA-A Release: Apr. 2013

4.1. mapping, SNP call, phasing

bwa options

max edit distance: 80

other bwa aln/bwasw options:

other bwa samse/sampe options:

GATK options

contamination rate: 0.1 stand call confidence: 80

stand emit confidence: 50 depth: 1000

other GATK options:

re-calling options

minor homo occurrence: 0.2 major homo occurrence: 0.8

min insert length: 300

phasing options

force phasing:

4.2. haplotype analysis

making haplotype options

GATK options

stand call confidence: 50 stand emit confidence: 10

depth: 1000

other GATK options:

re-calling options

minor homo occurrence: 0.4 major homo occurrence: 0.6

blat options

min identity: 99 tile size: 18 max mismatch: 2

other blat options:

ご質問等は

khosomic@nig.ac.jp

までご連絡ください