

# 和牛のゲノム情報基盤の整備と データベースの構築

東京大学大学院新領域創成科学研究科  
生命データサイエンスセンター・センター長  
メディカル情報生命専攻・教授

鈴木 穰

## 1. 黒毛和種のゲノム情報の整備 (JRA事業)

～プラチナゲノムの決定

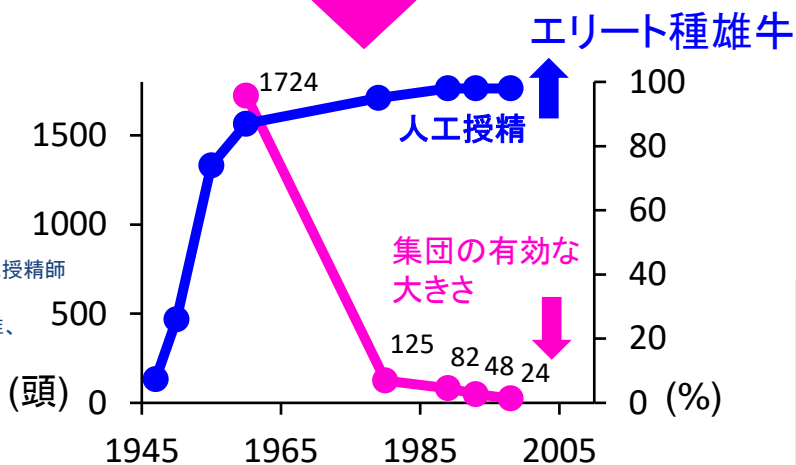
～全長RNA解析による遺伝子情報の充実

## 2. 黒毛和種のゲノムデータベースの作成

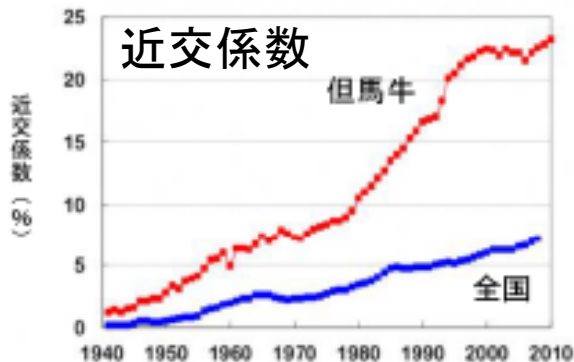
# 黒毛和種の改良と近交化、遺伝的不良形質



・ 育種価 >> ゲノム育種価 : 改良 UP

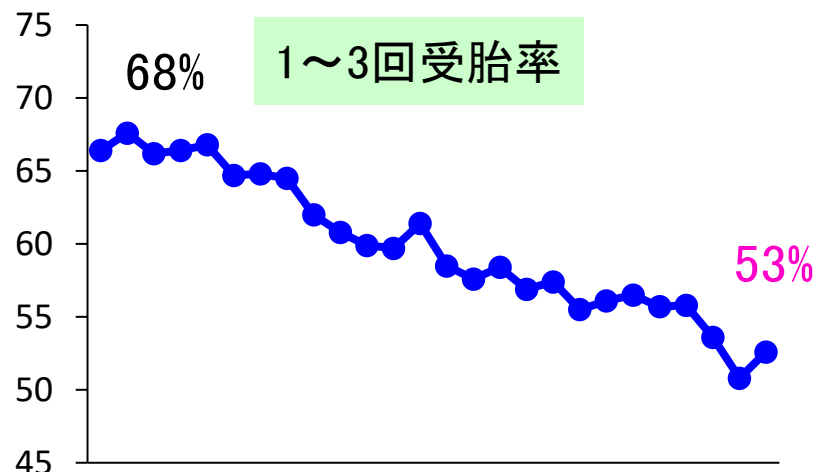


家畜人工授精師  
テキスト  
向井文雄、  
2011

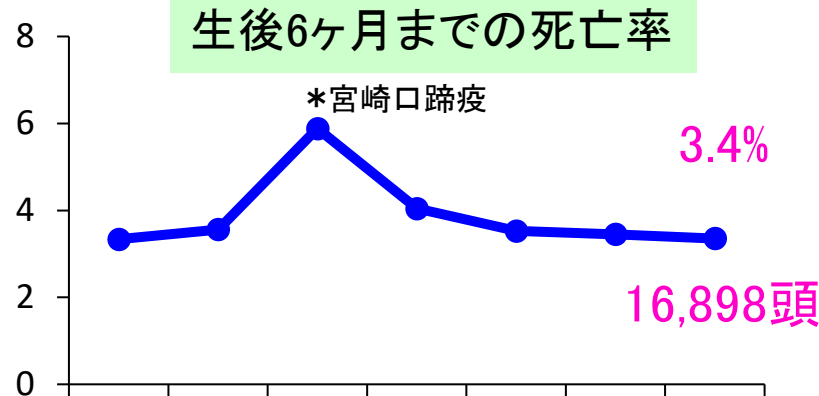


親子: 25%  
全きょうだい: 25%  
祖父母と孫: 12.5%  
いとこ: 6.25%

但馬牛の閉鎖集団での改良手法と雄性不妊牛検出法の開発と普及 (福島護之) から



H1 H4 H7 H10 H13 H16 H19 H22 H25  
家畜改良事業団



H20 H21 H22 H23 H24 H25 H26  
家畜改良センター

**生産阻害因子(有害変異)**

受胎失敗 > 170億円/年  
子牛損耗・死亡 > 200億円/年

# 育種にゲノム情報を利用：ゲノム情報を利活用する土壌

$$y = \mathbf{1}_n \mu + \sum_{i=1}^p \mathbf{X}_i g_i + \mathbf{e} \quad \text{[ゲノム育種価予測式]}$$

$$= \mathbf{1}_n \mu + \mathbf{X}_1 g_1 + \mathbf{X}_2 g_2 + \dots + \mathbf{X}_p g_p + \mathbf{e}$$

$X_i$ :  $i=1 \sim p$ までの遺伝子型 ( $X_i = \{aa, Aa, AA\} = \{0, 1, 2\}$ )

$g_i$ :  $i=1 \sim p$ までのSNP効果

全てのマーカー  
3-5万SNP/個体  
表型値

畜産技術協会(研究)

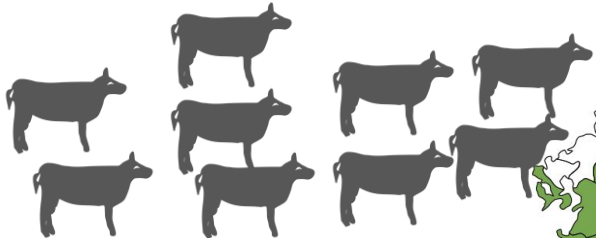
(H29~)

独立行政法人  
家畜改良センター

訓練群(学習集団)

家畜改良事業団  
(AI事業体)

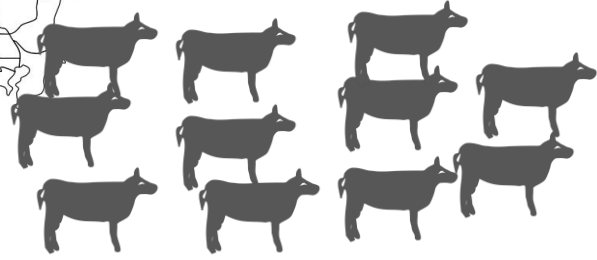
訓練群(学習集団)



表型値 + SNP型

ssGBLUP: ゲノム育種価の予測式

SNP型(+) 主に雌牛のゲノム育種価算出



表型値 + SNP型

GBLUP: ゲノム育種価の予測式

SNP型(+) 種雄牛のゲノム育種価算出

# 黒毛和種の既報の遺伝性疾患の原因変異の検出

遺伝性疾患名	略号	原因変異	BTA	UMD3.1	96 exome	RR/RA/AA	ALT Freq (%)
血液凝固第13因子欠損症	F13	c.248T>C (F13A)	23	48649432	tTc/tCc	94/2/0	1.04
チェデアックヒガシ症候群	CHS	c.6044A>G (LYST)	28	8508619	cAt/cGt	94/2/0	1.04
IARS異常症	IARS	c.235G>C (IARS)	8	85341291	Gtt/Ctt	82/14/0	7.29
前肢帯筋異常症	FMA	c.430C>T (GFRA1)	26	36951312	Caa/Taa	92/4/0	2.08
骨格粗大症	FGD3	CAC>TGC (FGD3)	8	85826989_85826990	Cac/Tac, cAc/cGc	74/21/0 74/21/0	11.05 11.05
ウシバーター症候群I型	BBS	c.1315C>T (SLC12A1)	10	62382825	cCa/cTa	95/1/0	0.52

～さらなる拡大へ

# 和牛のゲノムデータベースと地域サンプルを活用した生産阻害因子解明のための解析プラットフォームの構築

## ① 和牛ゲノムDB

畜産技術協会

- \* 事務局、管理更新

岐阜、兵庫、鳥取、島根  
鹿児島、家畜改良事業団

### 和牛ゲノムDB補強

東京大学

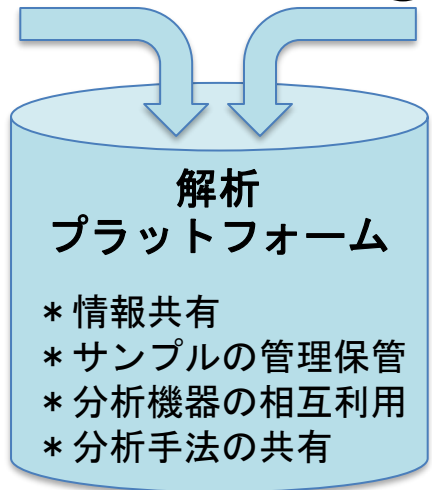
- \* プラチナゲノム (JRA)

琉球大学

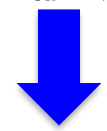
- \* データベース解析
- \* コンソ取りまとめ

- \* 和牛情報の保護  
利活用を両立  
＜管理規程＞

## ② 解析プラットフォーム



有害変異の特定  
DNA診断法



## ③ 普及・実用化支援組織

家畜改良事業団

ISO9001

- \* 全国レベルの頻度調査
- \* 実用的DNA診断法の開発
- \* 和牛生産県と全国普及、情報提供

## ウシ バイオバンク

- \* 授精記録、子牛損耗・死亡
- \* 試験交配 (ヘテロ x ヘテロ)
- \* 頻度調査、普及

和牛生産県:



### 子牛損耗・死亡解析

北里大  
東京農大

- \* サンプルまとめ、解析

生産性阻害因子のコントロール(育種、交配)

# これまでの遺伝子解析

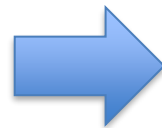
遺伝性疾患名	略号	染色体	原因変異	原因遺伝子	症状
胚死滅	CNVR_322	8	g. 378127_412061del	ANXA10	30-60日目に死亡
胚死滅	RB*	X	g.A>G	転写因子_X	AI3回不受胎、 ET不受胎
胚死滅	17_2*	17	g.G>T	CDC_X	30-60日目に死亡
骨格粗大症	FGD3	8	g. 85826989C>T g.85826990A>G	8	骨格粗大症
胎膜水腫	BBS	10	g.62382825G>A	SLC12A1	妊娠中期に 胎子死亡

\*未公開

[未解決の不良因子]

[課題]

胚死滅: 1領域  
子牛死亡: 4領域  
その他



- 1) 和牛の正確なゲノム配列がない
- 2) ウシの遺伝子情報が正確でない

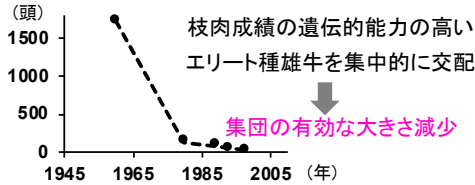
## 背景と目的

和牛の育種改良は、ゲノム育種価の利用によってスピードアップすることが期待されるが、一方で、新たな劣性遺伝病の発生が危惧される。このような背景から、育種戦略のもう一つの柱として、スピーディーに劣性遺伝病の原因変異を特定する仕組みが必要である。そこで本事業では、「1. 個々の和牛の正確なゲノム配列の構築（プラチナゲノム）」と「2. 和牛の遺伝子情報の強化」によって和牛のゲノム情報を強化し、生産阻害因子である劣性遺伝病の原因をスピーディーに特定し、安心して育種改良できる仕組みを構築する。



- ・ 育種価、(ゲノム育種価)
- ・ 人工授精

[育種改良のスピードアップ] ↑

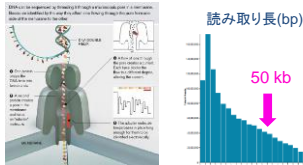


[劣性遺伝病（胚死減、子牛損耗）] ↑

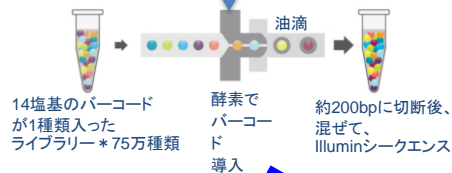
## 事業概要 [和牛のプラチナゲノム作成]

重要な和牛種雄牛 50頭

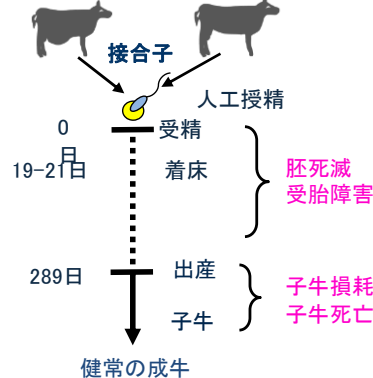
長鎖配列シーケンス: 物理的手法

Oxford  
Nanopore社  
MinION

長鎖配列シーケンス: 合成的手法

10X Genomics社  
GemCodeシステム

## [和牛の遺伝子情報の強化]



RNAseq &gt; 正確な遺伝子モデル

- (採材: 鹿児島県肉改研委託)
- ・ 成牛: 主要19臓器
  - ・ 妊娠19、30、60、90日目胎子、胎盤
  - ・ 子牛: 生後1、3週間

## 期待される成果

- ・ 西洋品種由来の参照配列に依存せず、和牛プラチナゲノムを基に変異探索
- ・ 正確な遺伝子モデルで変異のアノテーション

↓ 波及効果

スピーディーに生産阻害因子を特定し、安心して育種改良

平成29年度日本中央競馬会  
畜産振興振興事業(実施期間: H29-H31)

## 和牛のゲノム情報強化事業

- 1) 和牛のプラチナゲノム配列の構築
- 2) 遺伝子情報の強化

国立大学法人東京大学

大学院新領域創成科学研究科

メディカル情報生命専攻

教授 鈴木 穰

## 種雄牛サンプル:

岐阜県、兵庫県、鳥取県、島根県、  
鹿児島県(RNAサンプル委託契約)、  
家畜改良事業団

その他: 畜産技術協会、琉球大学

\* 畜産技術強化(共同研究契約: H29-H31)



## 1) 短鎖シーケンス (NovaSeq6000データ)

サンプル数: 64

- ・bwaでマッピング → GATKでvariation call
- ・その他、De novo assembleのpolishing等を使用

## 2) 長鎖シーケンス (PromethIONデータ)

サンプル数: 64

- ・minimap2でマッピングと集計
- ・NGMLRとSnifflesを用いてSV call

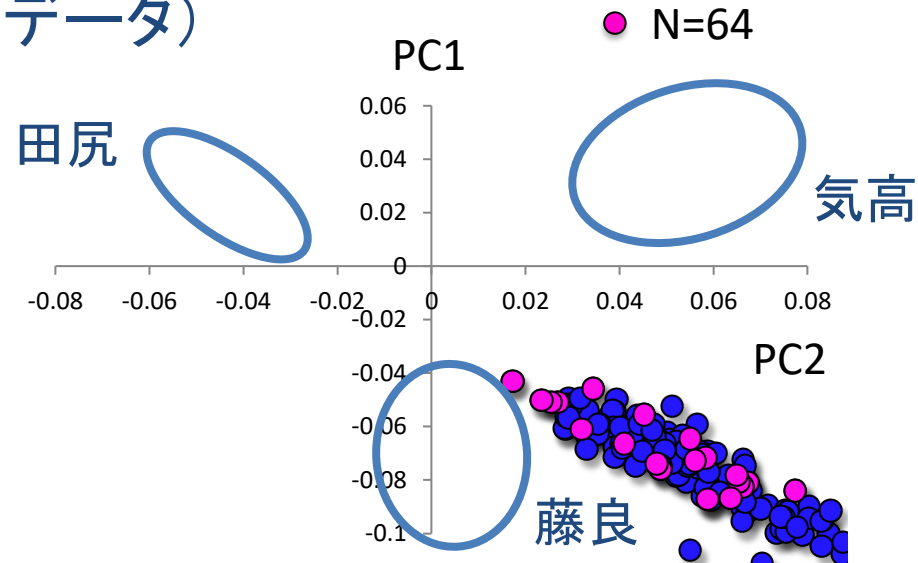
## 3) De novo assembly (和牛ゲノムを新規に作成)

1)、2)のデータを使用

## 4) RNA-seqデータ

edgeR、StringTieで処理

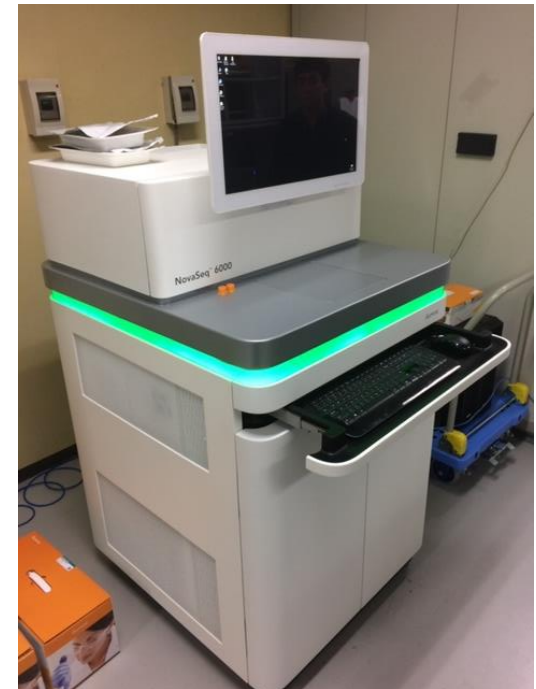
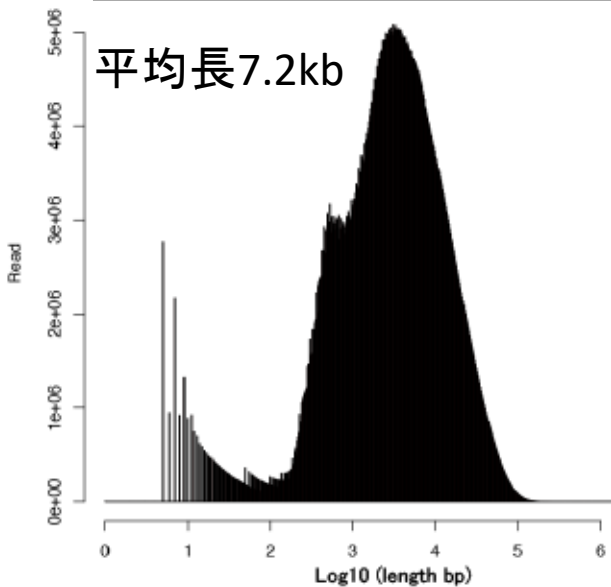
## 5) 統合ゲノムブラウザ



	PromethION (N=64)	Novaseq 6000 (N=64)
Number of reads	12,394,985	920,828,388
Max length (bp)	796,770	(150 x2)
Average length	7,212	(150 x2)
% mapped	78	97
Total read length	86,316,187,095	58,933,016,834
X Depth	32.84	52.90

## PromethIONの読み取り長

平均長7.2kb



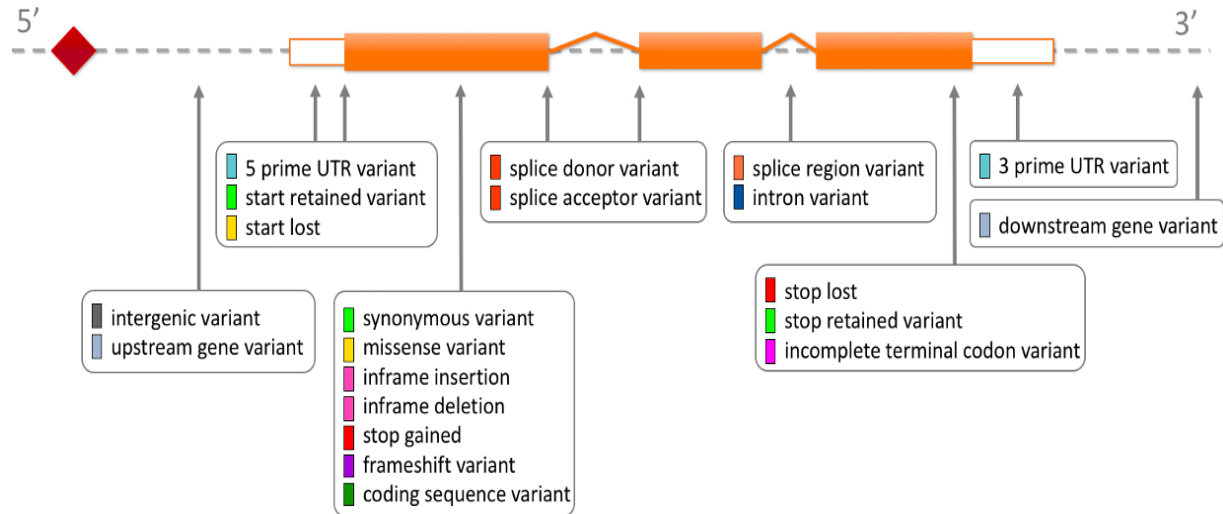
# De novo assembling

(HGC スパコン 250GBメモリ/5CPU/2～3週間)

- ①. PromethION配列を使ってwtdbg2とwtpoa-cnsで1st contigを生成
- ②. PromethION配列を使ってminimap2で1st contigにマップポリッシュ。Wtpoa-cnsで2nd contigを生成
- ③. NovaSeq配列を使ってbwa mem のゆるいgap ペナルティーで2nd contigにマップポリッシュを行いwtpoa-cnsで3rd contigを生成
- ④. NovaSeq配列を使って bwa mem (default parameter)で3rd contigにマップポリッシュを行いwtpoa-cnsで4th contigを生成し、完成contig配列とした

# 和牛の新規ゲノムアセンブリ

	ARS-UCD1.2 (1 sample)	WGDB_de novo (64 sample)
Assembly	GCA_002263795.2	WG_1
Constructor	USDA ARS	University of Tokyo
Genome coverage	80.0x	85.74x (平均)
Total sequence length	2,715,853,792 (chr1-29, X: 2,628,411,261)	2,548,196,231
Total assembly gap length	28,162bp (30Kbp弱)	9.5Mbp (推定) (chr1~chr29)
Gaps between scaffolds	0	テロメア領域に加え、各染色体に平均1か所程度のギャップがある
Number of contigs	2,597	5,531
Contig N50	25,896,116	10,454,269
Contig L50	32	93
SNP	-	15,934,570
indel	-	12,916 ~ 19,910
Other SV	-	344 ~ 1,477



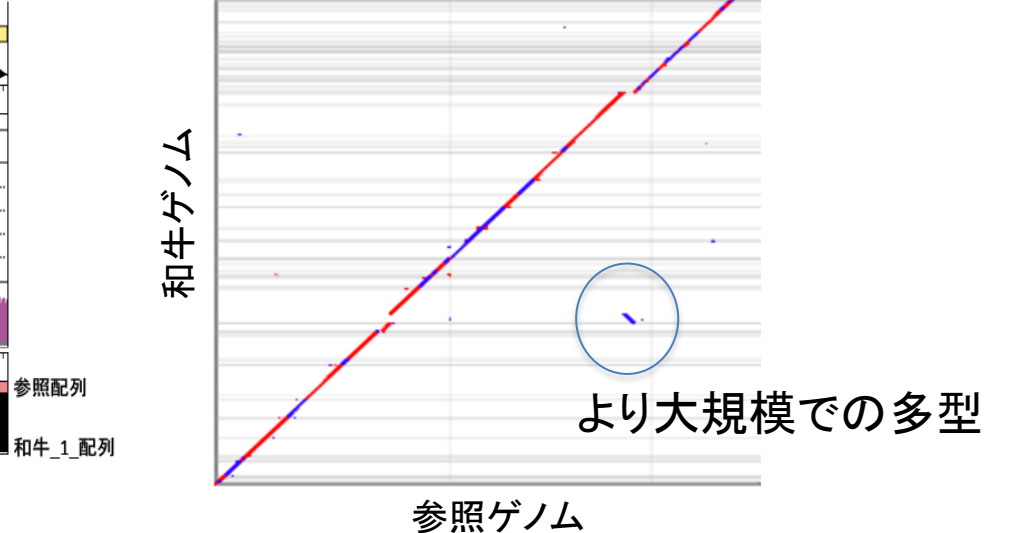
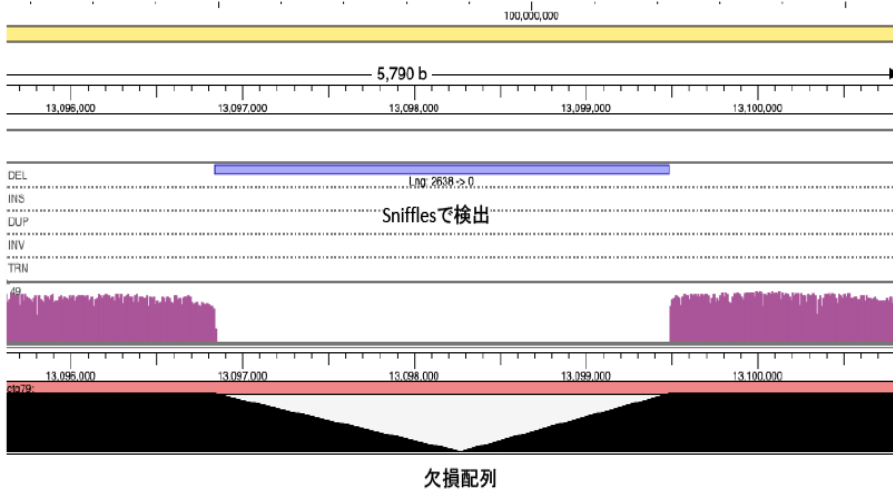
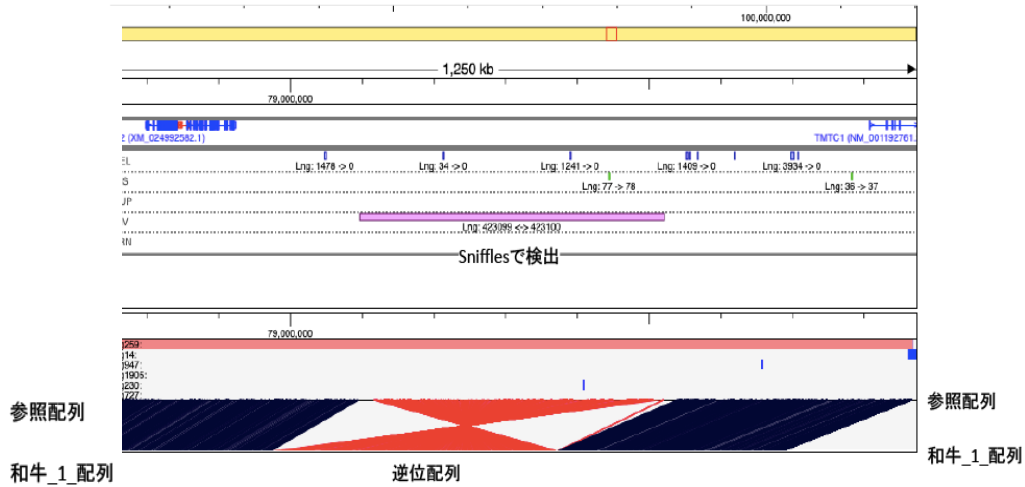
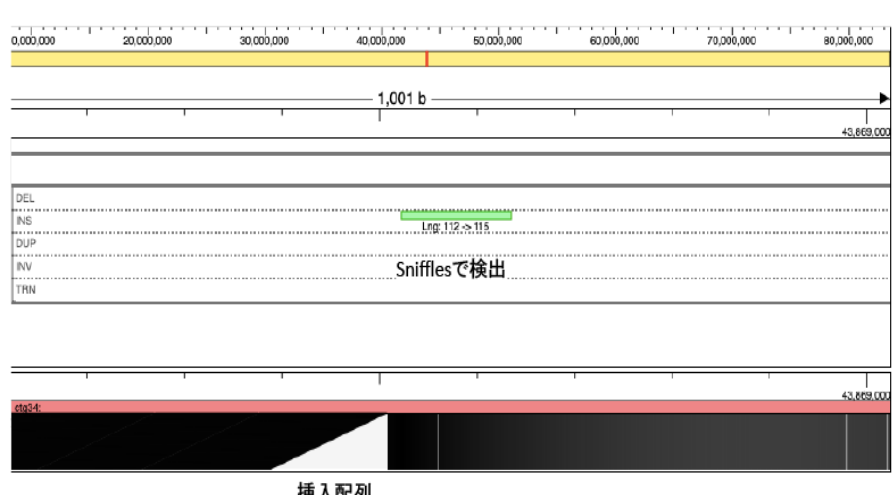
多型・変異タイプ	個数	coding
stop_gained	39,229	*
stop_lost	1,198	*
start_lost	1,898	*
missense_variant	698,961	*
stop_retained_variant	390	*
synonymous_variant	596,052	*
coding_sequence_variant	110	*
splice_acceptor_variant	8,985	
splice_donor_variant	6,331	
splice_region_variant	173,839	
mature_miRNA_variant	220	
5_prime_UTR_variant	437,651	
3_prime_UTR_variant	1,236,421	
non_coding_transcript_exon_variant	697,336	
intron_variant	98,430,471	
non_coding_transcript_variant	13,089,990	
upstream_gene_variant	9,470,836	
downstream_gene_variant	9,470,933	
intergenic_variant	17,383,561	

アミノ酸置換変異の有害度の予測 (SIFT)	個数
deleterious	124,675
tolerated	186,287

	タイプ	平均個数/個体
INS	挿入	5161.9
DEL	欠損	7753.8
DUP	重複	123.47
INV	逆位	51.078
TL	転座	169

# 和牛ゲノム解析から見出された参照ゲノムとの相違(2)

	タイプ	平均個数/個体
INS	挿入	5161.9
DEL	欠損	7753.8
DUP	重複	123.47
INV	逆位	51.078
TL	転座	169



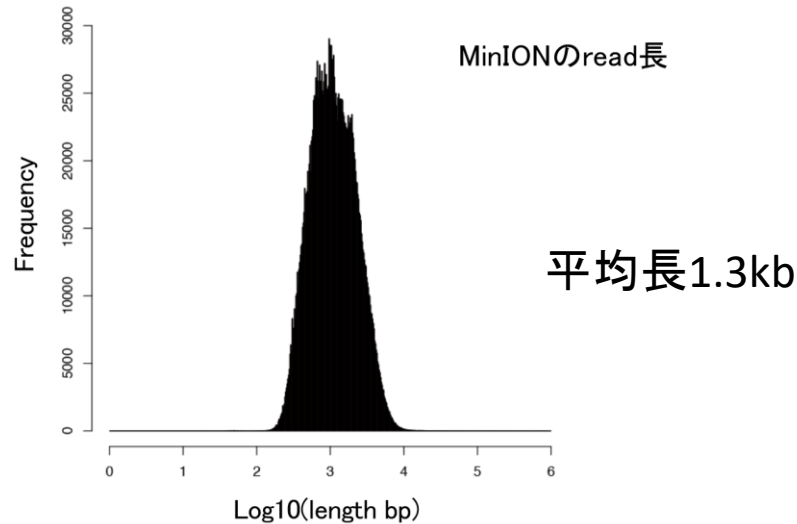


## Novaseq6000

ARS-UCD1.2		備考
総read数	50,961,865	129サンプル
参照配列へのマッピング(%)	97%	129サンプルの平均
平均検出遺伝子数 (転写方向一致)	11,845	129サンプルの平均
検出遺伝子数 (転写方向一致)	27,874	129サンプル、ARS-UCD1.2(34,256遺伝子)

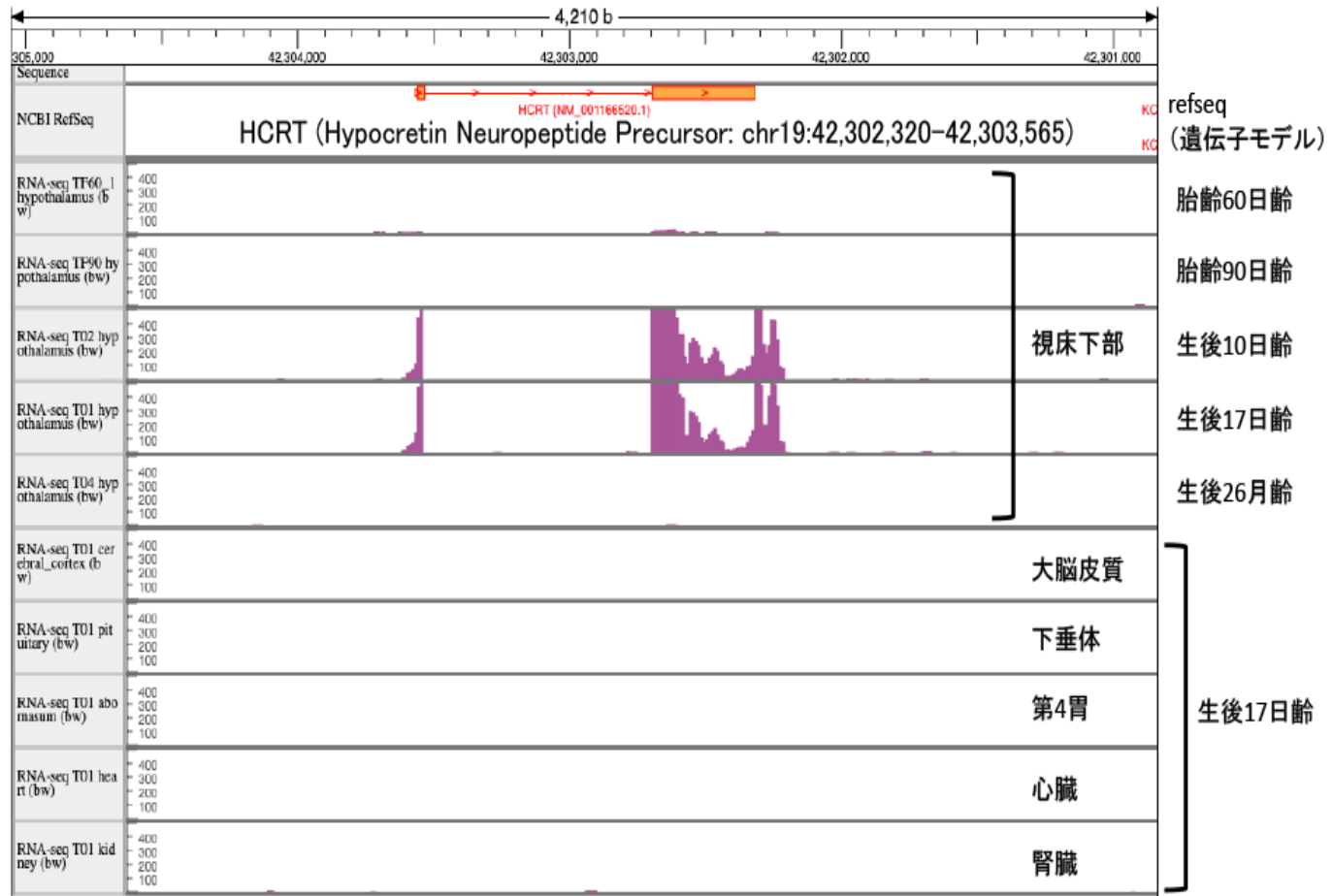
## MinION

ARS-UCD1.2		備考
総read数	2,028,224	22サンプル
平均read長	1,397bp	22サンプルの平均
最大read長	3M bp	22サンプル
参照配列へのマッピング(%)	78%	22サンプルの平均
平均検出遺伝子数 (転写方向一致)	8,631	22サンプルの平均
検出遺伝子数 (転写方向一致)	18,176	22サンプル、ARS-UCD1.2(34,256遺伝子)

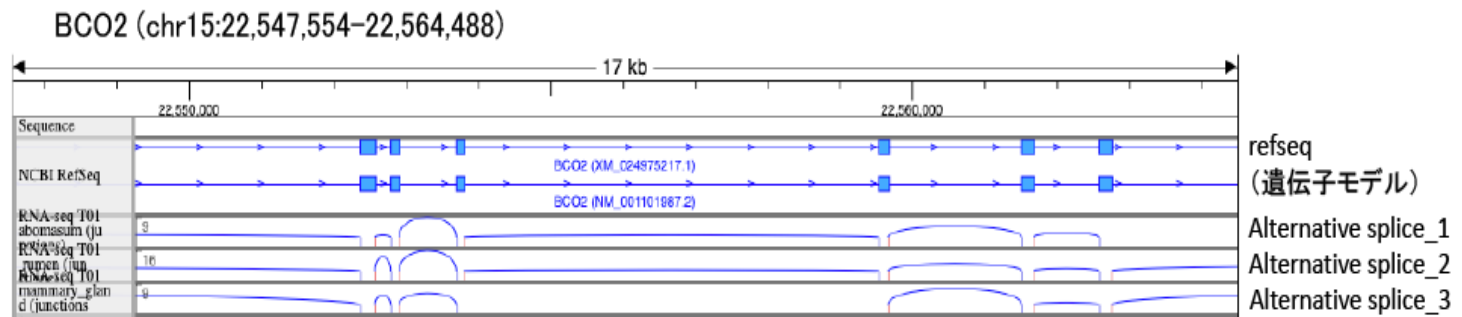
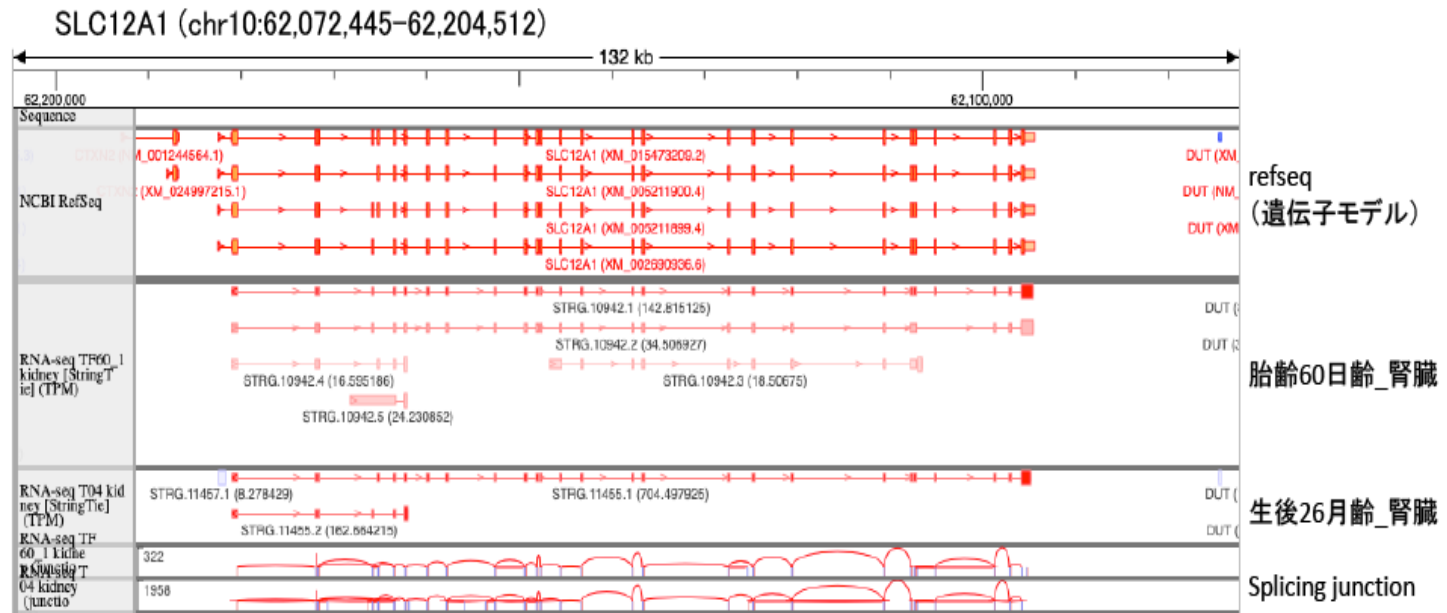




# 組織特異的発現(タグカウント)



# 選択的スプライシング(全長)



# 黒毛和種の強化ゲノム情報の改善・活用事業（JRA）

R2-R4

「1. 黒毛和種ゲノムデータベースの改善と整備」

「2. 黒毛和種ゲノムデータベースの活用による有害変異の特定」

## ① 新規の重要種雄牛

### 和牛ゲノムDB

畜産技術協会

\* 事務局、管理更新

岐阜、兵庫、鳥取、島根  
鹿児島、家畜改良事業団

### ① 和牛ゲノムDB改善

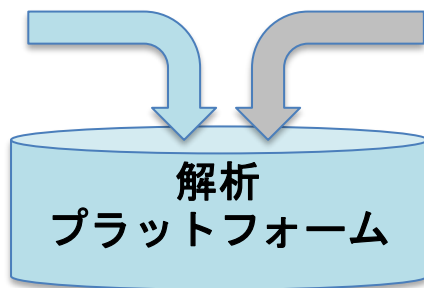
東京大学

\* プラチナゲノム(JRA)

琉球大学

\* データベース解析

\* コンソ取りまとめ



### ② ・有害変異の特定

### ②' 予備解析

・西洋品種情報の  
取り込み

・Trioサンプル



### 地域特異的サンプル

\* 授精記録、子牛損耗・死亡

\* 試験交配（ヘテロ x ヘテロ）

\* 頻度調査、普及

和牛生産県

島根

鳥取

岐阜

兵庫

鹿児島

家畜改良  
センター

家畜改良  
事業団

### 子牛損耗・死亡解析

\* サンプルまとめ、解析

北里大  
東京農大

### 普及・実用化支援組織

家畜改良事業団

ISO9001

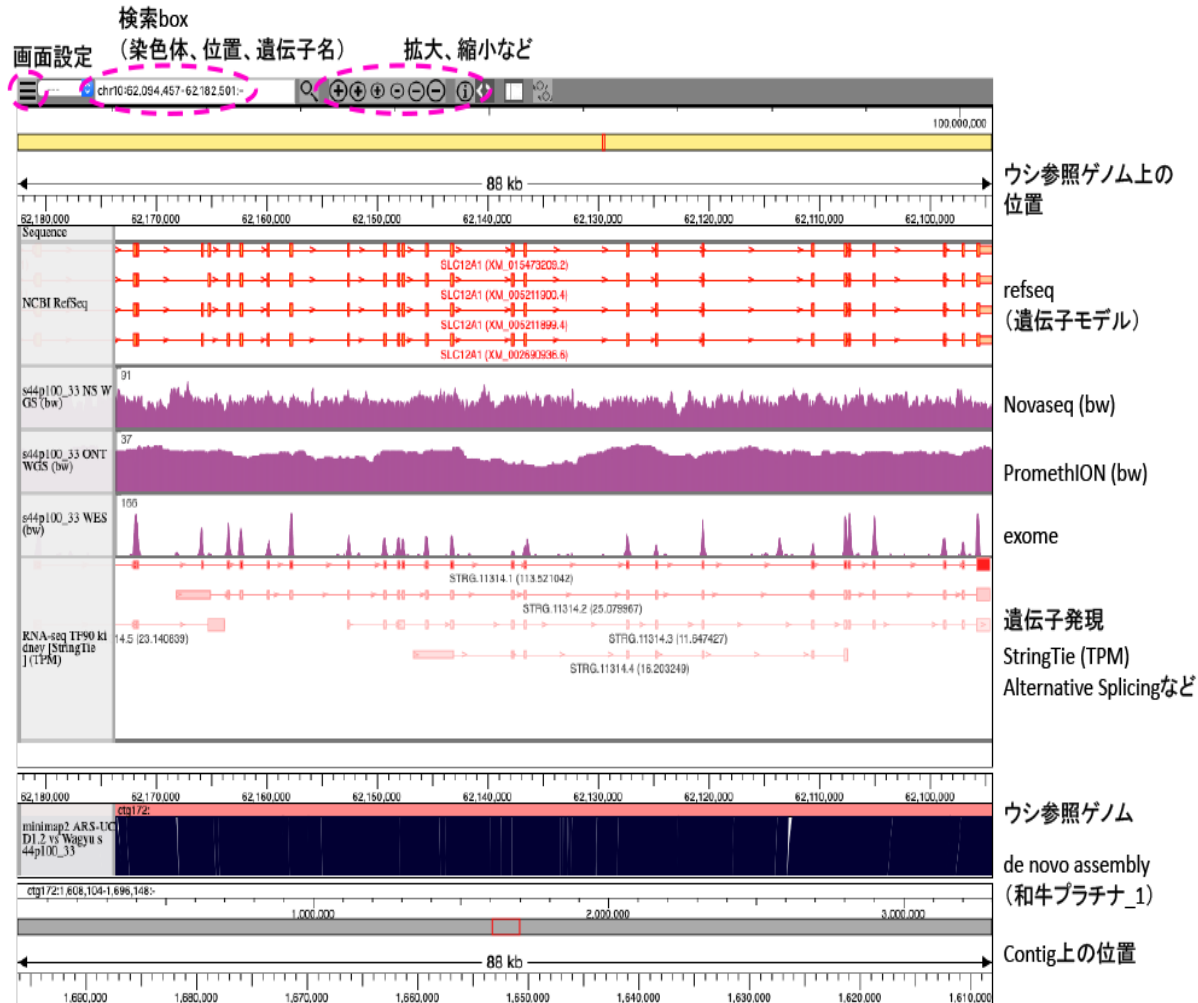
\* 全国レベルの頻度調査

\* 実用的DNA診断法の開発

\* 和牛生産県と全国普及、情報提供

情報分析がメイン  
分析の改善

データ区分	可視化データ
NovaSeq WGSデータ (short read: 150bp x2)	マッピング結果(bam, bigWig) Variationコール結果 (SNP, indel)
PromethION WGSデータ(long read)	マッピング結果(bam, bigWig) SnifflesによるSV(Structural variation) call結果
De novo assemble結果	参照ゲノムとの比較
Exomeデータ	マッピング結果(bam, bigWig) Variationコール結果 (SNP, indel)
RNAseq	StringTieのTPM値 RSEMのTPM値



# ウシゲノムブラウザ URL

---

URL:

<https://XXXXXX/>

- ログインにはパスワードが必要です！
- 各県/センターごとのページを提供  
(基本的に詳細な情報は切り分けてあります)

# 可視化サンプルについて

---

1. Exomeデータ (short read: 100bp) : 518サンプル
  - 1-1. 参照ゲノム(ARS-UCD1.2)へのマッピング結果(bam, bigWig)
  - 1-2. Variationコール結果 (SNP, (short) indel)
  - 1-3. Beagleによるimputation & phasing結果(IBD/HBD, Phased SNPs)
  
2. NovaSeq WGSデータ (short read: 150bp x2) : 64サンプル
  - 2-1. マッピング結果(bam, bigWig)
  - 2-2. Variationコール結果 (SNP, (short) indel)
  
3. PromethION WGSデータ(long read) : 64サンプル
  - 3-1. マッピング結果(bam, bigWig)
  - 3-2. SnifflesによるSV(Structural variation) call結果
  - 3-3. NovaSeqデータと共にredbeanでのde novo assemble結果
  - 3-4. NovaSeqデータと共にwhatshapでのphasing結果
  
- 
  
4. RNA-seqデータの再処理 : edgeR、StringTieで処理した結果

# SNPの可視化例

chr1:56,088,000-56,090,500 領域

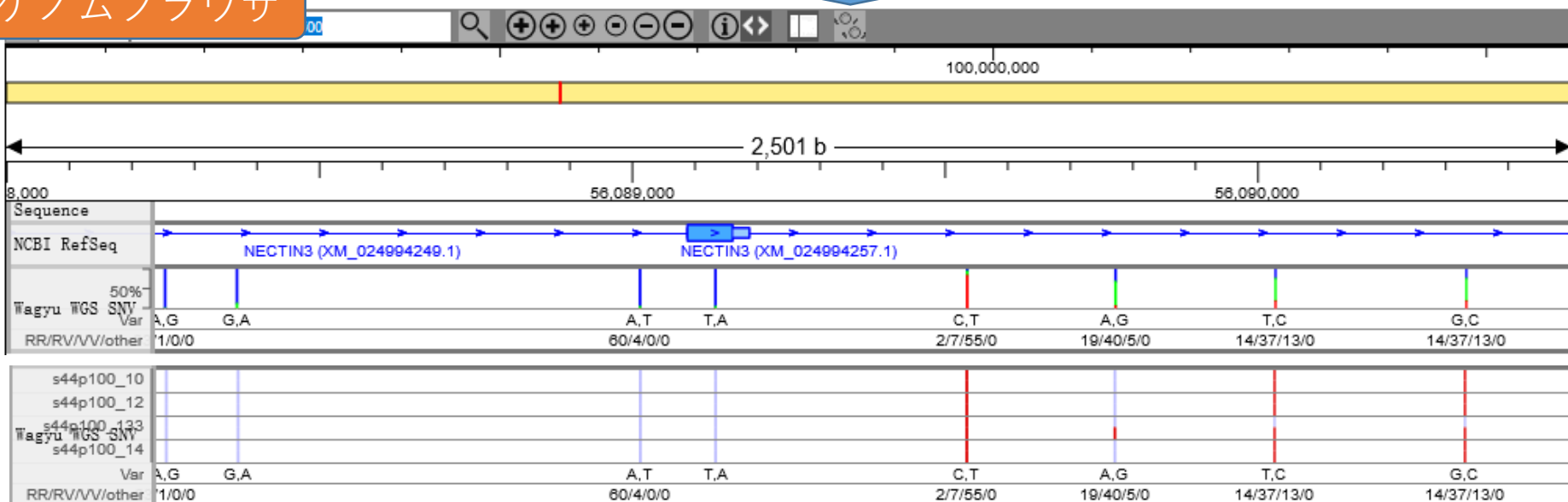
VCFファイル

chr1:56088000-56090500

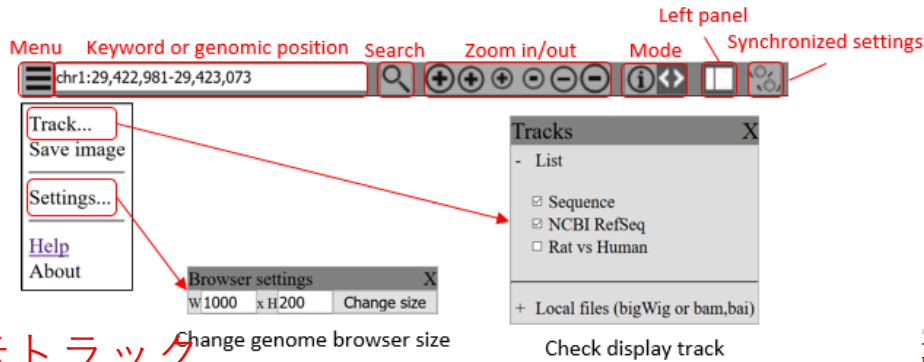
	A	B	C	D	E	F	G	H	I	J	K	L	
1	#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	s44p100_10	s44p100_12	s44p100_133	s44p100_14
2	chr1	56088253	.	A	G	314.92	PASS	AC=1;AF=7.813e-03;AN=128;A	GT:AD:DP:GQ:PL	0/0:41,0:41:99:0,105,1639	0/0:25,0:25:60:0,60,987	0/0:30,0:30:87:0,87,1305	0/0:47,0:47:
3	chr1	56088368	.	G	A	5865.5	PASS	AC=9;AF=0.070;AN=128;AN	GT:AD:DP:GQ:PL	0/0:37,0:37:99:0,99,1450	0/0:31,0:31:90:0,90,1350	0/0:29,0:29:84:0,84,1260	0/0:47,0:47:
4	chr1	56089012	.	A	T	3216.5	PASS	AC=4;AF=0.031;AN=128;AN	GT:AD:DP:GQ:PL	0/0:39,0:39:99:0,105,1477	0/0:24,0:24:72:0,72,892	0/0:31,0:31:87:0,87,1305	0/0:47,0:47:
5	chr1	56089133	.	T	A	2266.1	PASS	AC=4;AF=0.031;AN=128;AN	GT:AD:DP:GQ:PL	0/0:38,0:38:99:0,99,1470	0/0:22,0:22:60:0,60,900	0/0:31,0:31:87:0,87,1305	0/0:47,0:47:
6	chr1	56089535	.	C	T	89048	PASS	AC=117;AF=0.914;AN=128;A	GT:AD:DP:GQ:PGT:PID:PL	1/1:0,45:45:99:...:1878,135,0	1/1:0,32:32:96:...:1292,96,0	1/1:0,36:36:99:...:1408,108,0	1/1:0,61:61:61:
7	chr1	56089772	.	A	G	28889	PASS	AC=50;AF=0.391;AN=128;A	GT:AD:DP:GQ:PGT:PID:PL	0/0:39,0:39:99:...:0,99,1557	0/0:31,0:31:90:...:0,90,1224	0/1:10,13:23:99:...:474,0,348	0/0:48,0:48:
8	chr1	56090028	.	T	C	42046	PASS	AC=63;AF=0.492;AN=128;A	GT:AD:DP:GQ:PL	1/1:0,47:47:99:1867,141,0	1/1:0,40:40:99:1670,120,0	0/1:25,9:34:99:274,0,941	1/1:0,53:53:
9	chr1	56090333	.	G	C	143592	PASS	AC=63;AF=0.492;AN=128;A	GT:AD:DP:GQ:PL	1/1:0,55:55:99:2297,166,0	1/1:0,33:33:99:1344,99,0	0/1:16,13:29:99:456,0,580	1/1:0,71:71:

可視化

ゲノムブラウザ



# ゲノムブラウザの使い方



## トラック選択画面

## 表示トラック



Category buttons  
Show public data

Public data  
Sequence NCBI RefSeq

Wagyu WGS

Strain	Exome(bam, bw)	Exome IBD/HBD	NovaSeq WGS(bam, bw)	PromethION WGS(bam, bw)	PromethION WhatsHap phasing
e452			e452 NS WGS (bam)	e452 ONT WGS (bam)	e452 ONT phased SIV
e44p100_3	e44p100_3 WES (bw) e44p100_3 WES (bam)	e44p100_3 Beagle BD e44p100_3 Beagle HBD	e44p100_3 NS WGS (bw) e44p100_3 NS WGS (bam)	e44p100_3 ONT WGS (bw) e44p100_3 ONT WGS (bam)	e44p100_3 ONT phased SIV e44p100_3 ONT phased region
e44p100_6	e44p100_6 WES (bw) e44p100_6 WES (bam)	e44p100_6 Beagle BD e44p100_6 Beagle HBD	e44p100_6 NS WGS (bw) e44p100_6 NS WGS (bam)	e44p100_6 ONT WGS (bw) e44p100_6 ONT WGS (bam)	e44p100_6 ONT phased SIV e44p100_6 ONT phased region
e44p100_8	e44p100_8 WES (bw) e44p100_8 WES (bam)	e44p100_8 Beagle BD e44p100_8 Beagle HBD	e44p100_8 NS WGS (bw) e44p100_8 NS WGS (bam)	e44p100_8 ONT WGS (bw) e44p100_8 ONT WGS (bam)	e44p100_8 ONT phased SIV e44p100_8 ONT phased region
e44p100_10	e44p100_10 WES (bw) e44p100_10 WES (bam)	e44p100_10 Beagle BD e44p100_10 Beagle HBD	e44p100_10 NS WGS (bw) e44p100_10 NS WGS (bam)	e44p100_10 ONT WGS (bw) e44p100_10 ONT WGS (bam)	e44p100_10 ONT phased SIV e44p100_10 ONT phased region
e44p100_12	e44p100_12 WES (bw) e44p100_12 WES (bam)	e44p100_12 Beagle BD e44p100_12 Beagle HBD	e44p100_12 NS WGS (bw) e44p100_12 NS WGS (bam)	e44p100_12 ONT WGS (bw) e44p100_12 ONT WGS (bam)	e44p100_12 ONT phased SIV e44p100_12 ONT phased region
e44p100_14	e44p100_14 WES (bw) e44p100_14 WES (bam)	e44p100_14 Beagle BD e44p100_14 Beagle HBD	e44p100_14 NS WGS (bw) e44p100_14 NS WGS (bam)	e44p100_14 ONT WGS (bw) e44p100_14 ONT WGS (bam)	e44p100_14 ONT phased SIV e44p100_14 ONT phased region
e44p100_15	e44p100_15 WES (bw) e44p100_15 WES (bam)	e44p100_15 Beagle BD e44p100_15 Beagle HBD	e44p100_15 NS WGS (bw) e44p100_15 NS WGS (bam)	e44p100_15 ONT WGS (bw) e44p100_15 ONT WGS (bam)	e44p100_15 ONT phased SIV e44p100_15 ONT phased region
e44p100_16	e44p100_16 WES (bw) e44p100_16 WES (bam)	e44p100_16 Beagle BD e44p100_16 Beagle HBD	e44p100_16 NS WGS (bw) e44p100_16 NS WGS (bam)	e44p100_16 ONT WGS (bw) e44p100_16 ONT WGS (bam)	e44p100_16 ONT phased SIV e44p100_16 ONT phased region
e44p100_17	e44p100_17 WES (bw) e44p100_17 WES (bam)	e44p100_17 Beagle BD e44p100_17 Beagle HBD	e44p100_17 NS WGS (bw) e44p100_17 NS WGS (bam)	e44p100_17 ONT WGS (bw) e44p100_17 ONT WGS (bam)	e44p100_17 ONT phased SIV e44p100_17 ONT phased region
e44p100_18	e44p100_18 WES (bw) e44p100_18 WES (bam)	e44p100_18 Beagle BD e44p100_18 Beagle HBD	e44p100_18 NS WGS (bw) e44p100_18 NS WGS (bam)	e44p100_18 ONT WGS (bw) e44p100_18 ONT WGS (bam)	e44p100_18 ONT phased SIV e44p100_18 ONT phased region
e44p100_19	e44p100_19 WES (bw) e44p100_19 WES (bam)	e44p100_19 Beagle BD e44p100_19 Beagle HBD	e44p100_19 NS WGS (bw) e44p100_19 NS WGS (bam)	e44p100_19 ONT WGS (bw) e44p100_19 ONT WGS (bam)	e44p100_19 ONT phased SIV e44p100_19 ONT phased region
e44p100_22	e44p100_22 WES (bw) e44p100_22 WES (bam)	e44p100_22 Beagle BD e44p100_22 Beagle HBD	e44p100_22 NS WGS (bw) e44p100_22 NS WGS (bam)	e44p100_22 ONT WGS (bw) e44p100_22 ONT WGS (bam)	e44p100_22 ONT phased SIV e44p100_22 ONT phased region
e44p100_23	e44p100_23 WES (bw) e44p100_23 WES (bam)	e44p100_23 Beagle BD e44p100_23 Beagle HBD	e44p100_23 NS WGS (bw) e44p100_23 NS WGS (bam)	e44p100_23 ONT WGS (bw) e44p100_23 ONT WGS (bam)	e44p100_23 ONT phased SIV e44p100_23 ONT phased region
e44p100_24	e44p100_24 WES (bw) e44p100_24 WES (bam)	e44p100_24 Beagle BD e44p100_24 Beagle HBD	e44p100_24 NS WGS (bw) e44p100_24 NS WGS (bam)	e44p100_24 ONT WGS (bw) e44p100_24 ONT WGS (bam)	e44p100_24 ONT phased SIV e44p100_24 ONT phased region

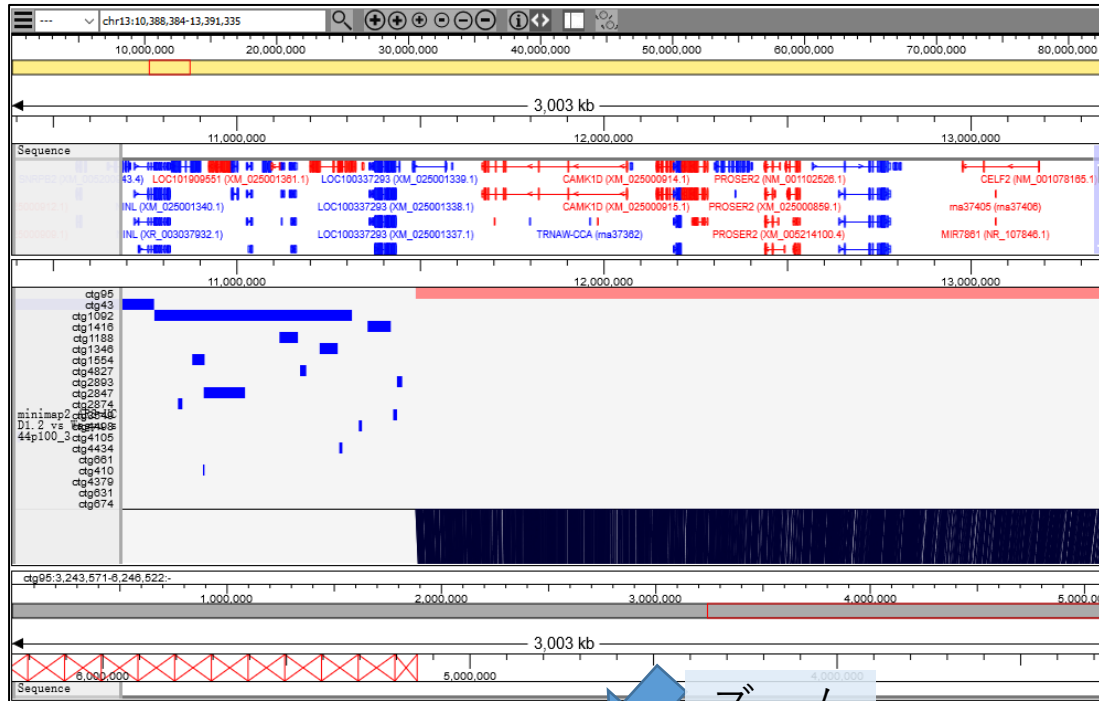
Exome解析の統合	Short read	Long read	構造変異	De novo assemble
------------	------------	-----------	------	------------------

Strain	Exome(bam, bw)	Exome IBD/HBD	NovaSeq WGS(bam, bw)	PromethION WGS(bam, bw)	PromethION WhatsHap phasing	PromethION Sniffles SV	de novo contigs
e452			e452 NS WGS (bam)	e452 ONT WGS (bam)	e452 ONT phased SIV	e452 Sniffles SV	ARS-UCD1.2 vs Wagyu e452
e44p100_3	e44p100_3 WES (bw) e44p100_3 WES (bam)	e44p100_3 Beagle BD e44p100_3 Beagle HBD	e44p100_3 NS WGS (bw) e44p100_3 NS WGS (bam)	e44p100_3 ONT WGS (bw) e44p100_3 ONT WGS (bam)	e44p100_3 ONT phased SIV e44p100_3 ONT phased region	e44p100_3 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_3
e44p100_6	e44p100_6 WES (bw) e44p100_6 WES (bam)	e44p100_6 Beagle BD e44p100_6 Beagle HBD	e44p100_6 NS WGS (bw) e44p100_6 NS WGS (bam)	e44p100_6 ONT WGS (bw) e44p100_6 ONT WGS (bam)	e44p100_6 ONT phased SIV e44p100_6 ONT phased region	e44p100_6 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_6
e44p100_8	e44p100_8 WES (bw) e44p100_8 WES (bam)	e44p100_8 Beagle BD e44p100_8 Beagle HBD	e44p100_8 NS WGS (bw) e44p100_8 NS WGS (bam)	e44p100_8 ONT WGS (bw) e44p100_8 ONT WGS (bam)	e44p100_8 ONT phased SIV e44p100_8 ONT phased region	e44p100_8 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_8
e44p100_10	e44p100_10 WES (bw) e44p100_10 WES (bam)	e44p100_10 Beagle BD e44p100_10 Beagle HBD	e44p100_10 NS WGS (bw) e44p100_10 NS WGS (bam)	e44p100_10 ONT WGS (bw) e44p100_10 ONT WGS (bam)	e44p100_10 ONT phased SIV e44p100_10 ONT phased region	e44p100_10 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_10
e44p100_12	e44p100_12 WES (bw) e44p100_12 WES (bam)	e44p100_12 Beagle BD e44p100_12 Beagle HBD	e44p100_12 NS WGS (bw) e44p100_12 NS WGS (bam)	e44p100_12 ONT WGS (bw) e44p100_12 ONT WGS (bam)	e44p100_12 ONT phased SIV e44p100_12 ONT phased region	e44p100_12 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_12
e44p100_14	e44p100_14 WES (bw) e44p100_14 WES (bam)	e44p100_14 Beagle BD e44p100_14 Beagle HBD	e44p100_14 NS WGS (bw) e44p100_14 NS WGS (bam)	e44p100_14 ONT WGS (bw) e44p100_14 ONT WGS (bam)	e44p100_14 ONT phased SIV e44p100_14 ONT phased region	e44p100_14 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_14
e44p100_15	e44p100_15 WES (bw) e44p100_15 WES (bam)	e44p100_15 Beagle BD e44p100_15 Beagle HBD	e44p100_15 NS WGS (bw) e44p100_15 NS WGS (bam)	e44p100_15 ONT WGS (bw) e44p100_15 ONT WGS (bam)	e44p100_15 ONT phased SIV e44p100_15 ONT phased region	e44p100_15 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_15
e44p100_16	e44p100_16 WES (bw) e44p100_16 WES (bam)	e44p100_16 Beagle BD e44p100_16 Beagle HBD	e44p100_16 NS WGS (bw) e44p100_16 NS WGS (bam)	e44p100_16 ONT WGS (bw) e44p100_16 ONT WGS (bam)	e44p100_16 ONT phased SIV e44p100_16 ONT phased region	e44p100_16 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_16
e44p100_17	e44p100_17 WES (bw) e44p100_17 WES (bam)	e44p100_17 Beagle BD e44p100_17 Beagle HBD	e44p100_17 NS WGS (bw) e44p100_17 NS WGS (bam)	e44p100_17 ONT WGS (bw) e44p100_17 ONT WGS (bam)	e44p100_17 ONT phased SIV e44p100_17 ONT phased region	e44p100_17 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_17
e44p100_18	e44p100_18 WES (bw) e44p100_18 WES (bam)	e44p100_18 Beagle BD e44p100_18 Beagle HBD	e44p100_18 NS WGS (bw) e44p100_18 NS WGS (bam)	e44p100_18 ONT WGS (bw) e44p100_18 ONT WGS (bam)	e44p100_18 ONT phased SIV e44p100_18 ONT phased region	e44p100_18 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_18
e44p100_19	e44p100_19 WES (bw) e44p100_19 WES (bam)	e44p100_19 Beagle BD e44p100_19 Beagle HBD	e44p100_19 NS WGS (bw) e44p100_19 NS WGS (bam)	e44p100_19 ONT WGS (bw) e44p100_19 ONT WGS (bam)	e44p100_19 ONT phased SIV e44p100_19 ONT phased region	e44p100_19 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_19
e44p100_22	e44p100_22 WES (bw) e44p100_22 WES (bam)	e44p100_22 Beagle BD e44p100_22 Beagle HBD	e44p100_22 NS WGS (bw) e44p100_22 NS WGS (bam)	e44p100_22 ONT WGS (bw) e44p100_22 ONT WGS (bam)	e44p100_22 ONT phased SIV e44p100_22 ONT phased region	e44p100_22 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_22
e44p100_23	e44p100_23 WES (bw) e44p100_23 WES (bam)	e44p100_23 Beagle BD e44p100_23 Beagle HBD	e44p100_23 NS WGS (bw) e44p100_23 NS WGS (bam)	e44p100_23 ONT WGS (bw) e44p100_23 ONT WGS (bam)	e44p100_23 ONT phased SIV e44p100_23 ONT phased region	e44p100_23 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_23
e44p100_24	e44p100_24 WES (bw) e44p100_24 WES (bam)	e44p100_24 Beagle BD e44p100_24 Beagle HBD	e44p100_24 NS WGS (bw) e44p100_24 NS WGS (bam)	e44p100_24 ONT WGS (bw) e44p100_24 ONT WGS (bam)	e44p100_24 ONT phased SIV e44p100_24 ONT phased region	e44p100_24 Sniffles SV	ARS-UCD1.2 vs Wagyu e44p100_24



# ゲノムブラウザの使い方(西洋種ゲノムとの比較)

Reference chr13:10,388,384-13,391,335 vs s44p100\_3 contigs:

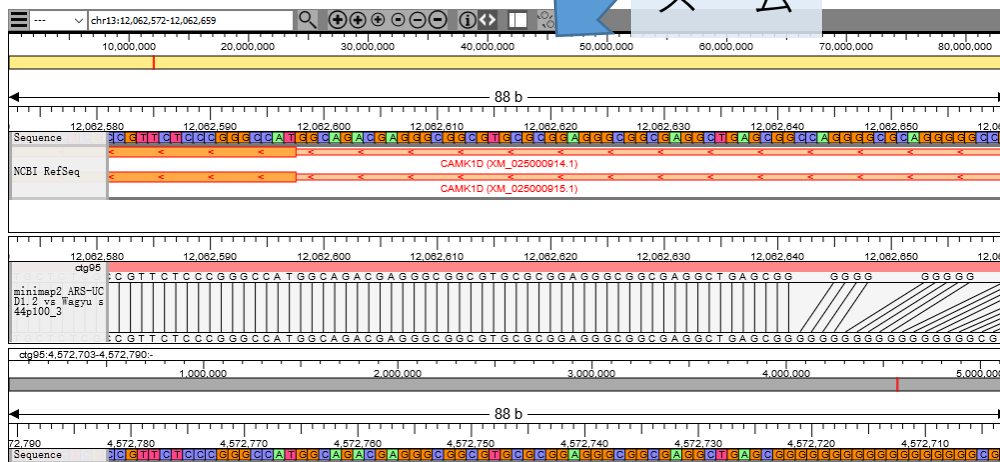


Referenceの表示領域にヒットする contig の位置。赤が alignment を表示している contig

Reference と contig の Alignment

align された contig の領域情報

ズームすると reference genome と de novo contig の塩基レベルのアラインメントまで確認することができます。



# 和牛は我が国の遺伝資源



和牛遺伝資源国内活用協議会

和牛の種雄牛造成機関、生産者、関連団体

和牛を海外へ出されては困る

和牛の遺伝資源(デジタル配列情報は?)

>和牛を海外へ出されては困る  
>何が起るかわからない



公共データベース  
(フリーアクセス)

閉鎖型  
研究支援型  
データベース

海外の研究者

「海外で特許>国内生産者が特許料を払う」  
「和牛の優位性が侵害される報告」など

締約国会議 (Conference of the Parties; COP) COP15(北京)

<生物多様性条約及び名古屋議定書 ABS>

- ・ 遺伝資源へのアクセス (A: Access)
- ・ 遺伝資源から得られる利益配分 (B: Benefit, S: Share)

>そもそもデジタル配列情報が遺伝資源か？  
>デジタル配列情報の取扱いが議論中！

和牛研究者グループ(管理規程)  
「国内で保護しつつ、研究者が研究できるように整備する」

- ・ 研究成果物を「人類の知」として公開
- ・ 知財確保して国内産業を保護

# 和牛データベース協議会・管理規程の制定へ

我が国の重要な遺伝資源であり、かつ地域ごとの遺伝資源である和牛のゲノム情報を保護しつつ、迅速な生産阻害因子の特定に効果的に利活用するため、代表機関である畜産技術協会が事務局となって日本国内で、「閉鎖・研究支援型」の和牛ゲノムデータベースを管理、利活用する体制を構築する。

次期JRA事業：  
和牛ゲノム解析は途中段階で、  
継続が必要

東京大学学内

## 和牛ゲノムデータベース (WGDB) 協議会

[WGDB管理規程]

データベースの管理、構築、更新  
利用申請規程

遺伝的不良形質の解析のみ(現時点)  
利用申請規程に則り申請

国内公的  
サーバー

\* 代表機関のみアクセス  
\* 契約し、データベース  
解析者にアクセス権を与える

代表機関・事務局：  
畜産技術協会

国内利用者

利用申請規程に則り申請

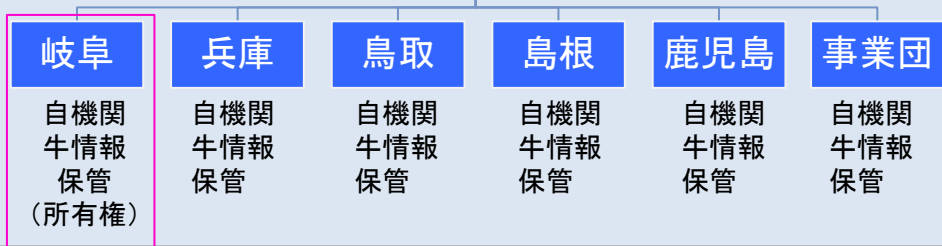
現時点では、解析者は東京大学  
琉球大学が想定される  
(委託)。

サンプル提供  
依頼書

提供  
通知  
審議、承諾

\* 各機関は、  
自機関の情報  
にアクセス

現時点で  
は、解析途  
中のため、  
非対応



\* 今後、参画県を増やし、協議会を安定化(15-20機関)ですが、費用負担があると参画機関が増えないだろう、との意見あり。

遺伝的  
不良形  
質以外  
の解析に  
個別に利用可能

左と同じ 左と同じ 左と同じ 左と同じ 左と同じ

所有権のある自機関の情報は、個別に利用可能(協議会、WGDB管理規程に縛られない)

共同研究者

\* 代表機関が借りる  
48TB 使用費：  
234万円/年

- \* 研究段階では研究費で賄っている。今後は協議。(イノベ、JRA事業)
- \* 設備費、電気代、安全性により国内の公的機関のサーバーが安く、よい。
- \* 現在は東大内スーパーコンピューター、サーバーを利用している。
- \* 計算機のリソース利用可

将来の学術的・産業的課題の解決へ

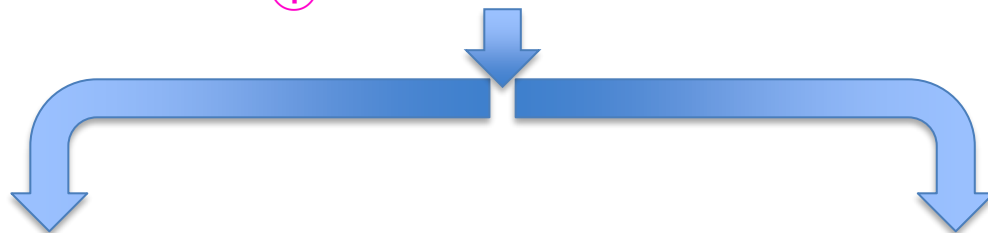
# 黒毛和種の既報の遺伝性疾患の原因変異の検出

遺伝性疾患名	略号	原因変異	BTA	UMD3.1	96 exome	RR/RA/AA	ALT Freq (%)
血液凝固第13因子欠損症	F13	c.248T>C (F13A)	23	48649432	tTc/tCc	94/2/0	1.04
チェデアックヒガシ症候群	CHS	c.6044A>G (LYST)	28	8508619	cAt/cGt	94/2/0	1.04
IARS異常症	IARS	c.235G>C (IARS)	8	85341291	Gtt/Ctt	82/14/0	7.29
前肢帯筋異常症	FMA	c.430C>T (GFRA1)	26	36951312	Caa/Taa	92/4/0	2.08
骨格粗大症	FGD3	CAC>TGC (FGD3)	8	85826989_85826990	Cac/Tac, cAc/cGc	74/21/0 74/21/0	11.05 11.05
ウシバーター症候群I型	BBS	c.1315C>T (SLC12A1)	10	62382825	cCa/cTa	95/1/0	0.52

～さらなる拡大へ

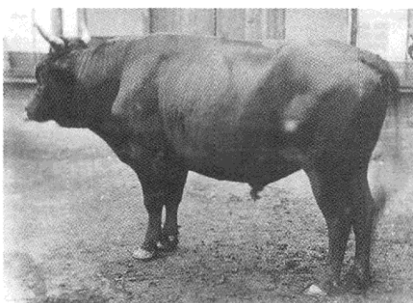
# 遺伝形質、西洋品種との交雑、変異率は？

① Common ancestor



1000bull genome project  
3,800 cattle

Native Japanese cattle



Ayrshire



Devon



Simmental (Fleckvieh)



Brown swiss

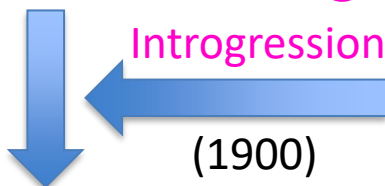


Holstein



②

Introgression



(1900)

Japanese Black cattle



1) 西洋品種の情報を有効活用

- ・ 有害変異、haplotype情報など

2) 変異率 (trioサンプル): 予備的検討



# 謝辞～和牛のゲノム研究ネットワーク

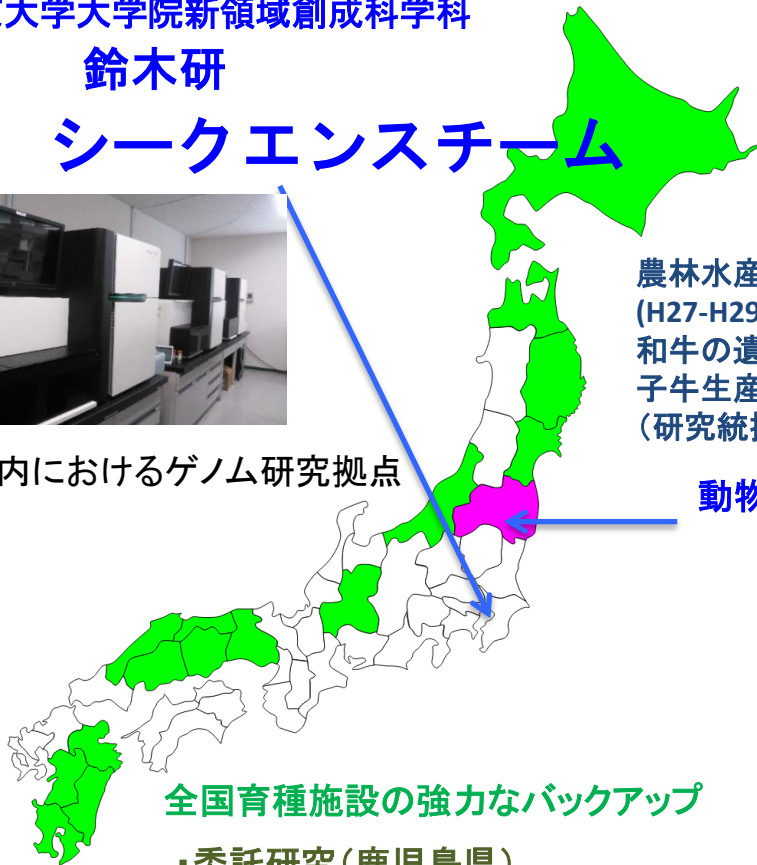
東京大学大学院新領域創成科学科

鈴木研

シーケンスチーム



東大:国内におけるゲノム研究拠点



全国育種施設の強力なバックアップ

- ・委託研究(鹿児島県)
- ・共同研究契約(動物遺伝研)

北海道 青森県 岩手県 宮城県  
岐阜県 兵庫県 島根県 鳥取県  
広島県 岡山県 大分県 熊本県  
宮崎県 鹿児島県 沖縄県 事業団

岡山大 東農大 北里大  
NOSAI岩手 NOSAI山形 民間牧場

農林水産業・食品産業科学技術研究推進事業  
(H27-H29)  
和牛の遺伝子多様体データベースの構築による  
子牛生産阻害因子の迅速な解明  
(研究統括:杉本喜憲)



動物遺伝研究所

- ・研究統括、デザイン、管理
- ・シーケンス分析

平成29年5月26日  
於:鹿児島県肉用牛改良研究所

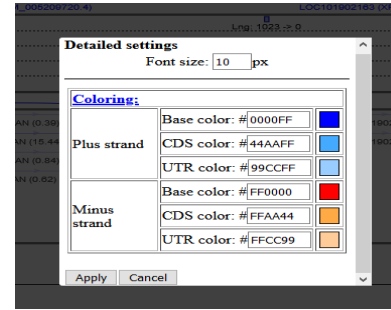
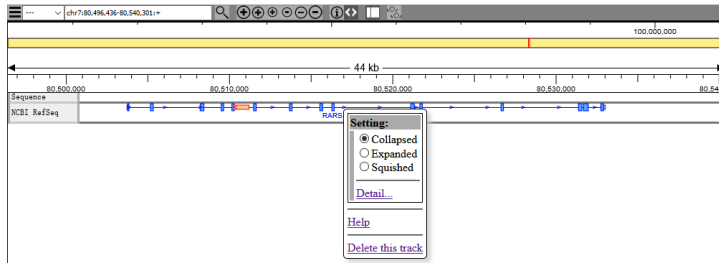


農研機構生研支援センター  
「イノベーション創出強化研究推進事業」

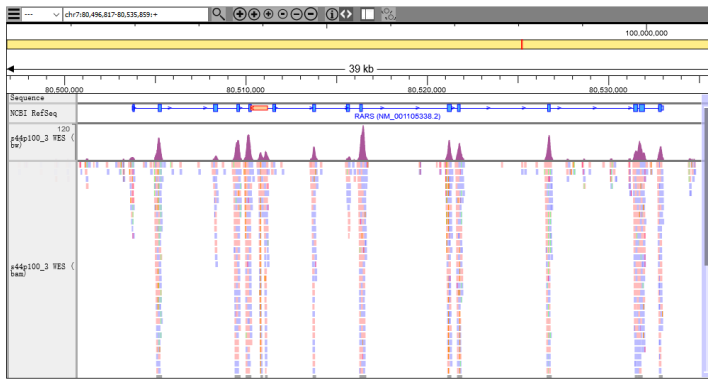
JRA  
「黒毛和種の強化ゲノム情報の改善・活用事業」

# ゲノムブラウザの使い方 (各トラック)

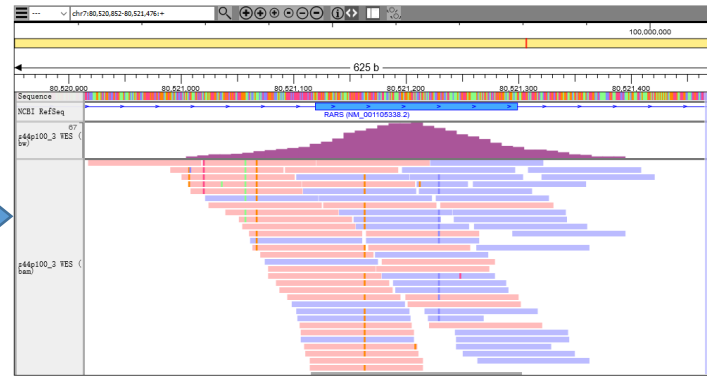
右クリックでトラックの設定ができます



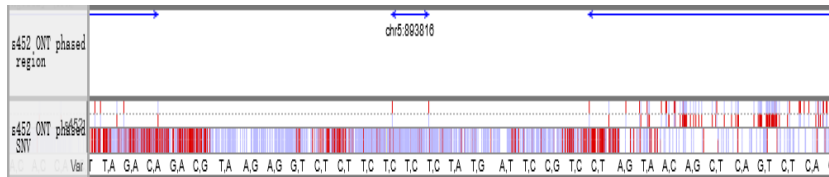
Read depthとマッピング結果表示・拡大するとindel/塩基置換が確認できます



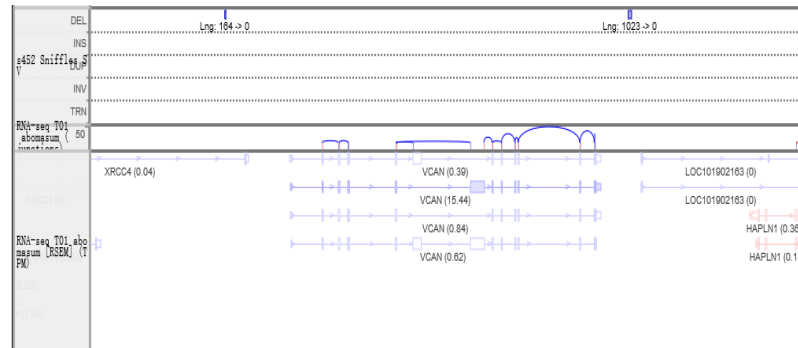
拡大



Phasing 領域と Phaseパターン

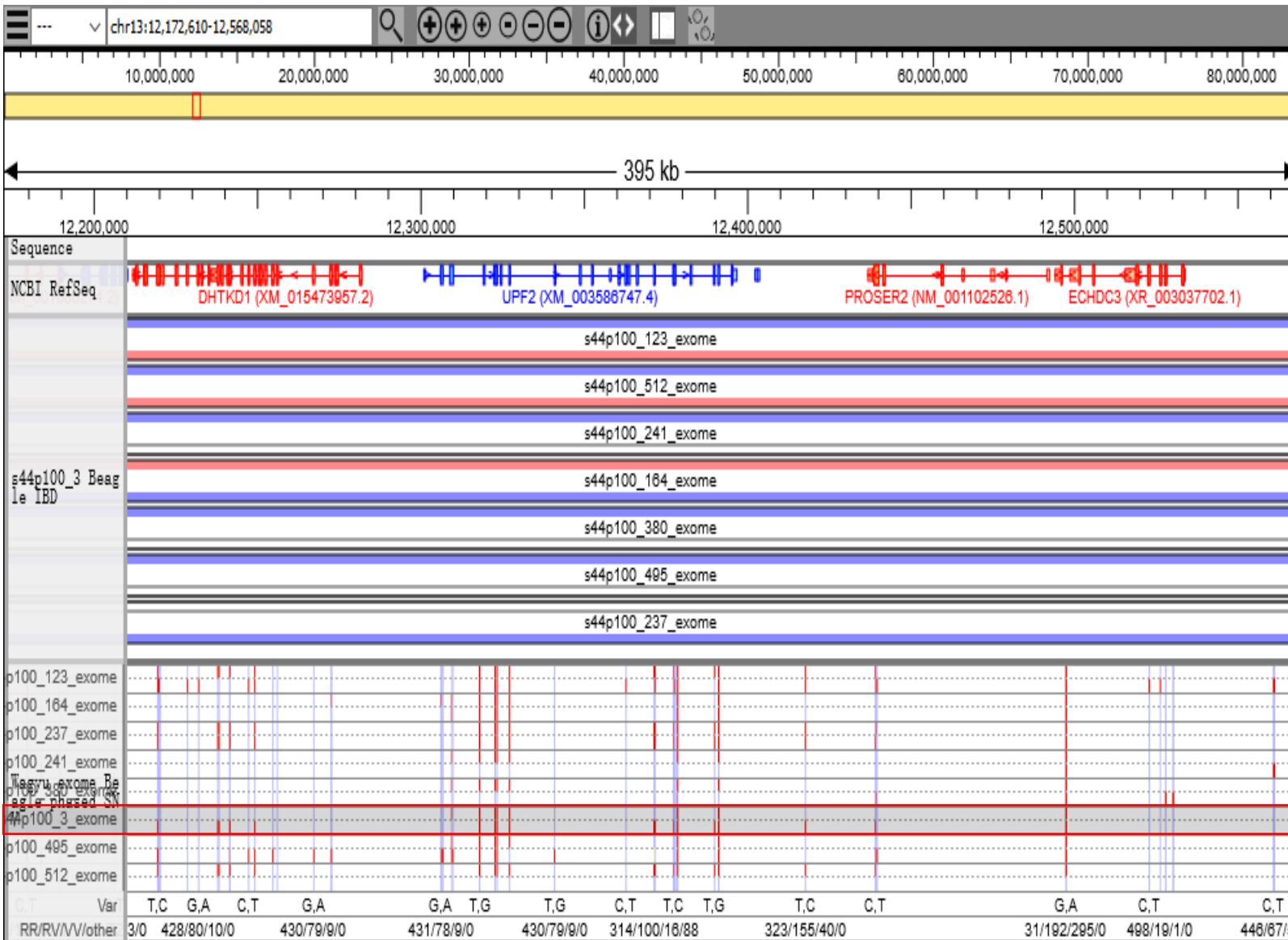


構造変異(SV)と発現量、exon/intron junction  
※SVについては別途ブラウザを開発中





# ゲノムブラウザの使い方(フェーズパターンとHBD)



青が相手。  
赤が自分。  
この場合、  
s44p100\_3の下側の  
フェーズパターン  
と、s44p100\_123の  
上側のフェーズパ  
ターンが一致

一番下の  
s44p100\_237は自分  
相手ともに下側の  
フェーズパターン  
が一致。(つまり赤  
青が重なって青色  
になっています)

