

MIG-seq法： 次世代シーケンサーを用いた 手軽なゲノムワイド塩基配列分析

陶山佳久

東北大学大学院農学研究科

1. MIG-seq法の概要
2. 研究例の紹介
3. 手法の改良
4. iSeq100の利用
5. その他の応用

1. **MIG-seq法の概要**
2. 研究例の紹介
3. 手法の改良
4. iSeq100の利用
5. その他の応用

1. MIG-seq法の概要

MIG-seq

Multiplexed ISSR Genotyping by sequencing

(Suyama & Matsuki, 2015)
Scientific Reports 5: 16963



注：最新プロトコルは原著論文版から一部変更してあります。

OPEN

MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform

Yoshihisa Suyama & Yu Matsuki

Restriction-enzyme (RE)-based next-generation sequencing methods have revolutionized marker-assisted genetic studies; however, the use of REs has limited their widespread adoption, especially in field samples with low-quality DNA and/or small quantities of DNA. Here, we developed a PCR-based procedure to construct reduced representation libraries without RE digestion steps, representing *de novo* single-nucleotide polymorphism discovery, and its genotyping using next-generation sequencing. Using multiplexed inter-simple sequence repeat (ISSR) primers, thousands of genome-wide regions were amplified effectively from a wide variety of genomes, without prior genetic information. We demonstrated: 1) Mendelian gametic segregation of the discovered variants; 2) reproducibility of genotyping by checking its applicability for individual identification; and 3) applicability in a wide variety of species by checking standard population genetic analysis. This approach, called multiplexed ISSR genotyping by sequencing, should be applicable to many marker-assisted genetic studies with a wide range of DNA qualities and quantities.

The recent development of next-generation sequencing (NGS) technology has allowed the effective discovery and genotyping of large numbers of genome-wide genetic markers¹. However, many marker-assisted studies require more economical and efficient approaches, rather than the methods based on the high marker density produced by whole-genome sequencing. To optimize the cost and the amount of available information for these studies, several methods have been developed to construct reduced representation libraries (RRLs), to sample single-nucleotide polymorphisms (SNPs) from genome-wide regions, and to genotype them using NGS. Methods such as reduced representation shotgun² sequencing and restriction site-associated DNA (RAD)³ markers were later adapted for NGS-based methods by the sequencing of RRLs⁴ and RAD tags⁵. Many improved approaches have been developed in recent years, such as complexity reduction of polymorphic sequences⁶, multiplexed shotgun sequencing⁷, genotyping by sequencing (GBS)⁸, 2-enzyme GBS⁹, RAD genotyping using type IIB restriction enzymes¹⁰, double digest RAD¹¹, and restriction fragment sequencing¹². These methods have become widespread and allow marker-assisted genetic studies, such as ecological, evolutionary, phylogeographic, and genetic mapping studies, based on tens to hundreds of thousands of SNPs in hundreds of barcoded samples at the same time. However, a more simple, rapid and cost-efficient approach for smaller-scale studies is desired, especially in ecological laboratories.

Tohoku University, Kawatabi Field Science Center, Graduate School of Agricultural Science, 232-3 Yomogida, Naruko-onsen, Osaki, Miyagi 989-6711, Japan. Correspondence and requests for materials should be addressed to Y.S. (email: suyama@m.tohoku.ac.jp)

1. MIG-seq法の概要

MIG-seq: Multiplexed ISSR Genotyping by sequencing

PCRのみによってゲノム縮約ライブラリーを構築し、NGSを用いた塩基配列分析によりSNPジェノタイピング等を行うための新しい手法

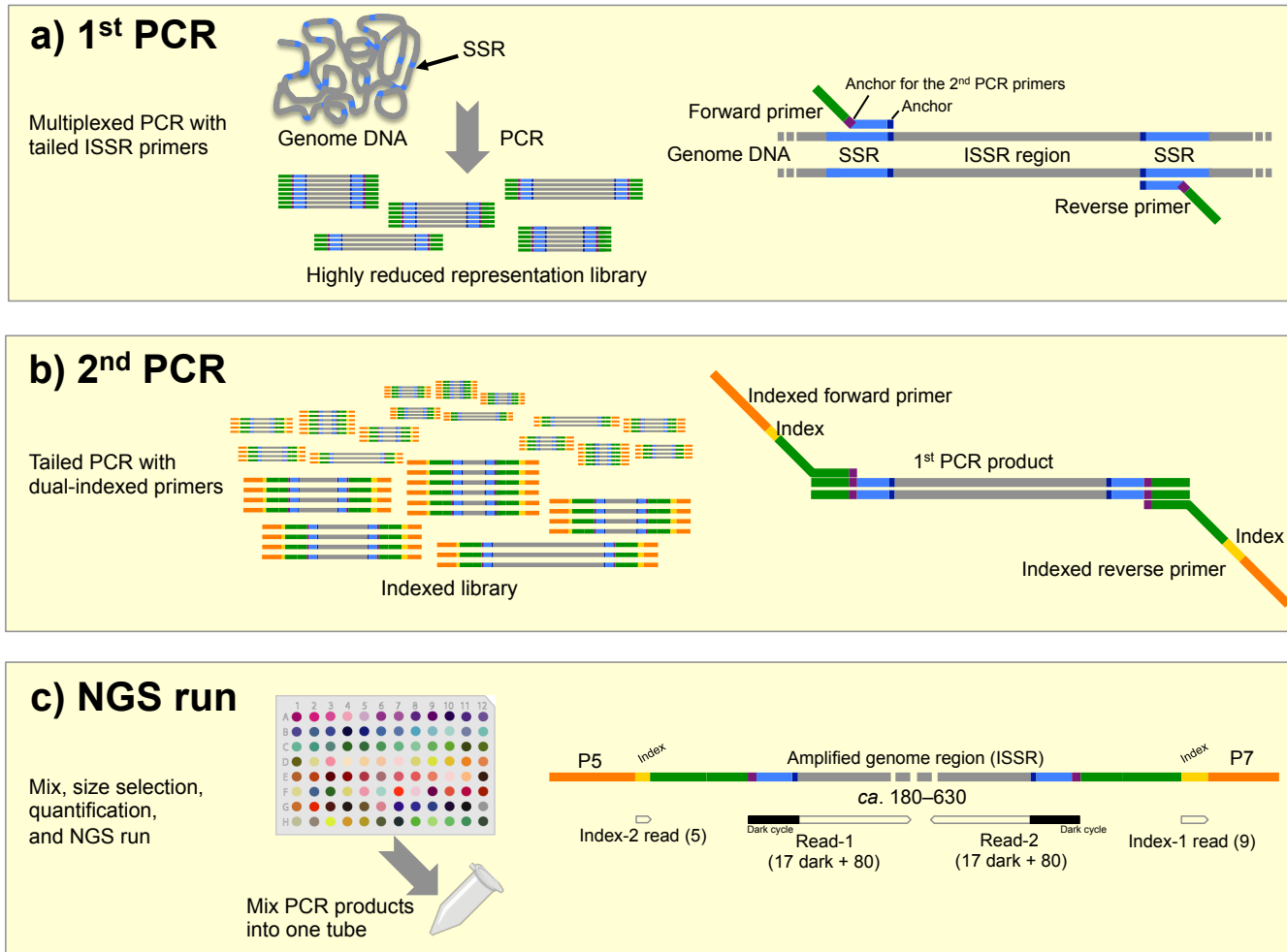


図 MIG-seq法の3ステップ (Suyama & Matsuki 2015を改変)

RADseq法との違い

ゲノム内の縮約方法

(「**ほどほど**」の領域を取り出して分析対象とする方法)
が違う。

RAD : **制限酵素**

v.s.

MIG : **PCR**

制限酵素処理を行う



十分な量の高品質DNAが必要



そもそもRADほどのデータは必要ない場合も。

非モデルの野外サンプルでは「無理！」というケースがしばしばある。



例) 熱帯で採取した劣化サンプル



例) 希少種で、ごく少量の枯死寸前の葉などしか採取できない場合

<その他の例>

微生物

標本

古代試料

何か解決方法はないのか？



制限酵素を使わずにゲノムの一部を選び出せばいい
(縮約ゲノムライブラリーの作成法)



PCRだけでうまく選び出す方法はないか？

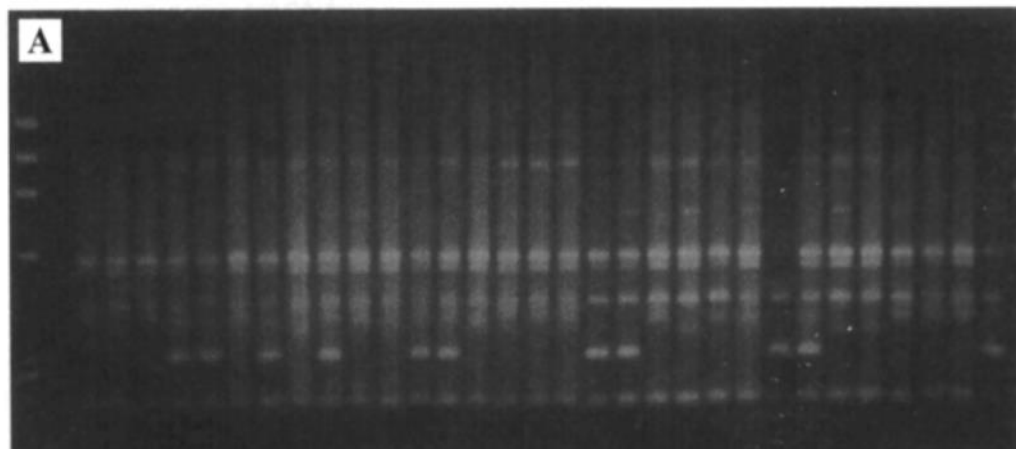
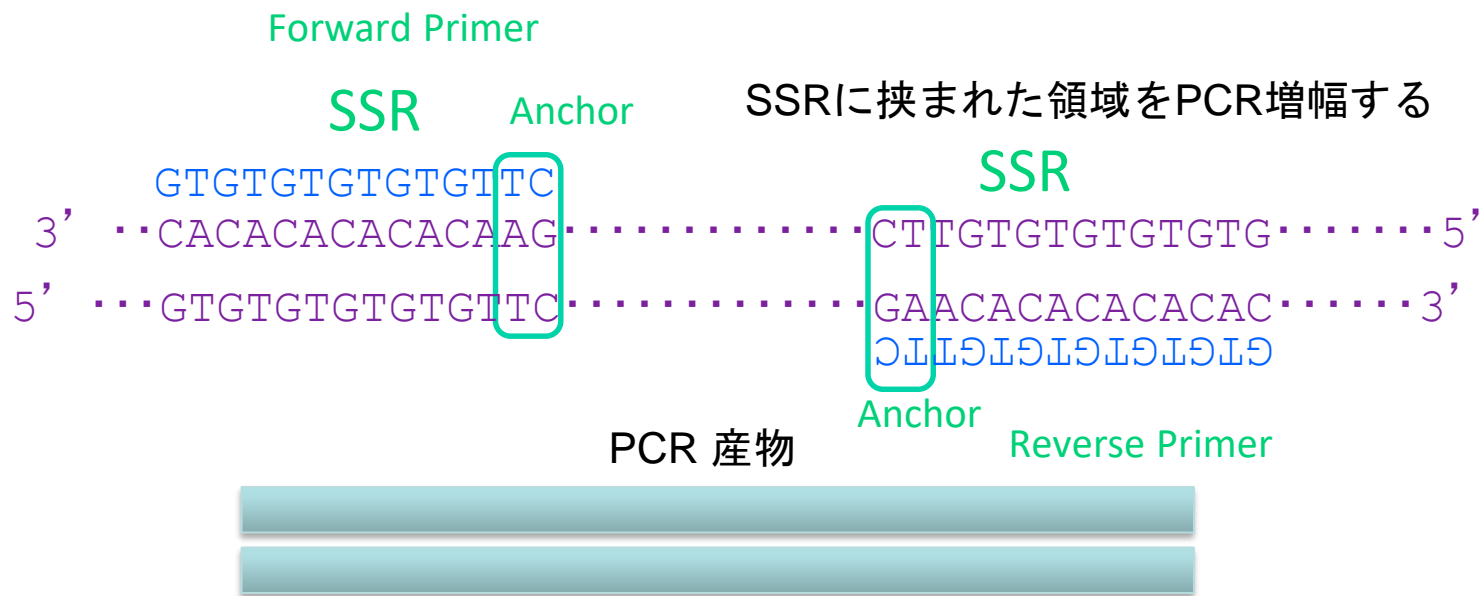


ISSR



(Inter-Simple Sequence Repeat)

背景 ISSR



Tsumura et al. 1996より

これをマルチプレックスにしてもっと多数の領域を増幅してライブラリーにすればよい

1. MIG-seq法の概要

- a) マルチプレックスPCRによる多数のSSR間領域のPCR増幅
- b) Tailed PCRによる配列付加
- c) サイズ選抜・定量(qPCR)、NGS (MiSeq)によるラン

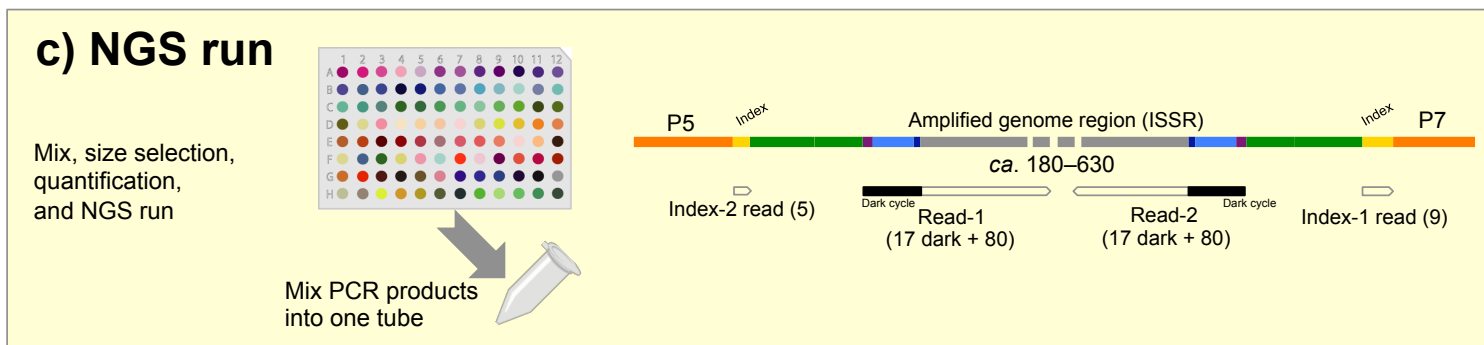
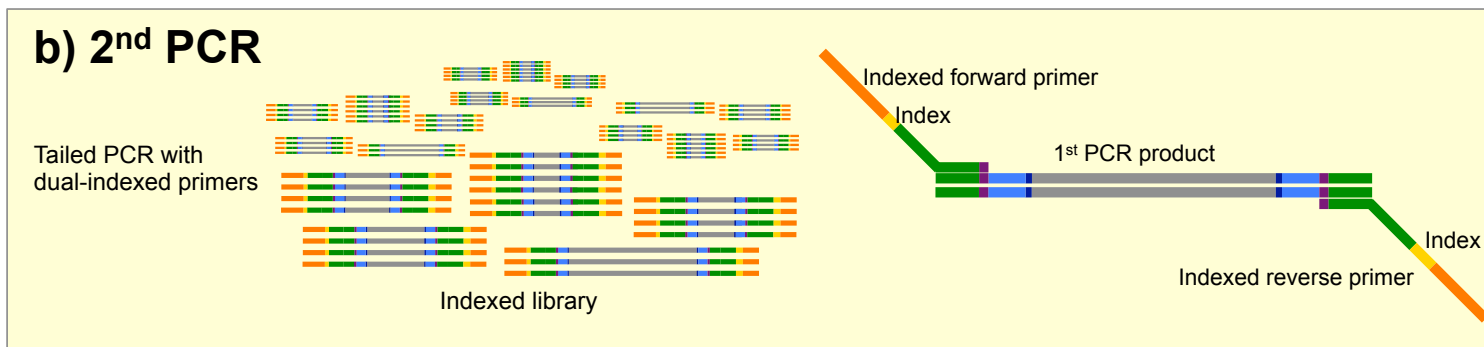
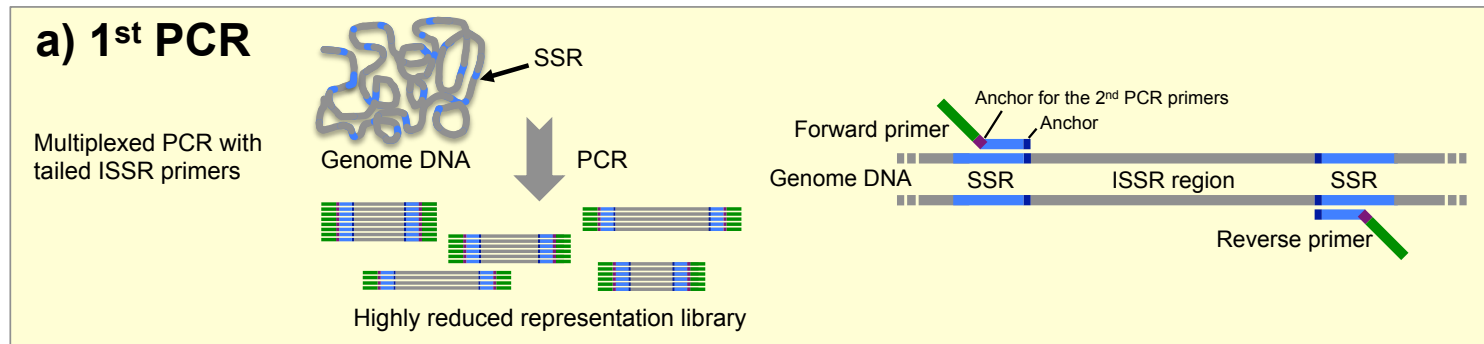
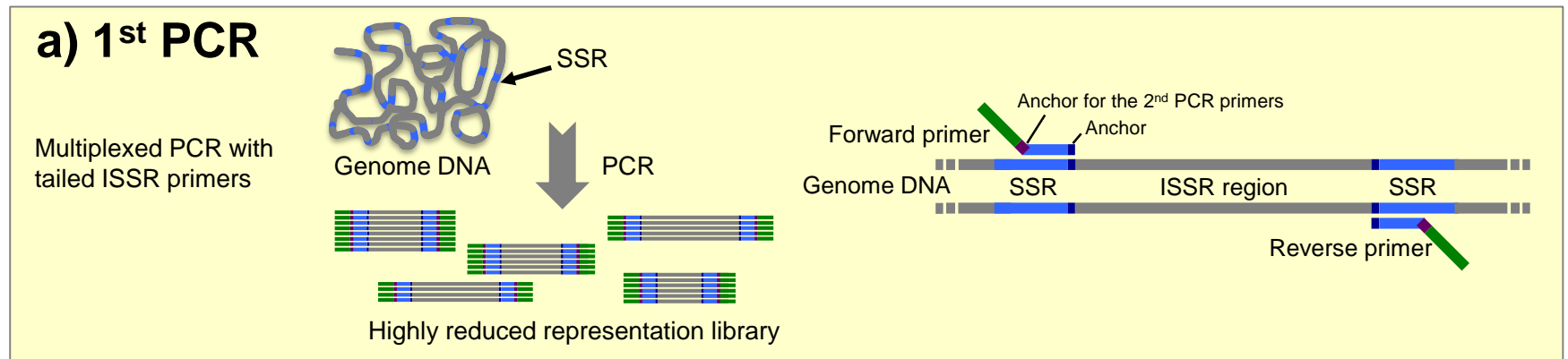


図 MIG-seq法の3ステップ

1. MIG-seq法の概要

Step #1: 1st PCR (多数のSSR間領域のPCR増幅)

「**ほどほどに**」多数の領域をゲノム内から引っ張り出す。



ユニバーサルなプライマーセットによる超マルチプレックスPCRによって多数のISSR領域を増幅し、**ゲノムの縮約ライブラリー**を作る。

1. MIG-seq法の概要

Step #1: 1st PCR

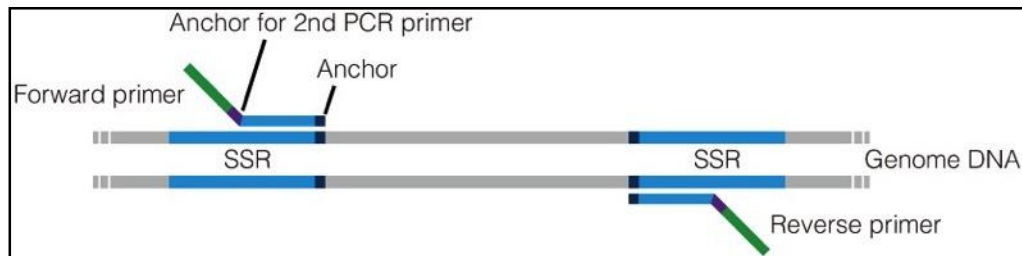


Table 1. Sequences of MIG-seq primer set-1 for the 1st PCR

Name	Sequences (5′–3′)
Forward primers:	(Tail + anchor: CTG) + SSR + anchor
(ACT) ₄ TG-f	<u>CGCTCTTCCGATCT</u> CTG ACTACTACTACT TTG
(CTA) ₄ TG-f	CGCTCTTCCGATCT CTG CTACTACTACTAT G
(TTG) ₄ AC-f	CGCTCTTCCGATCT CTG TTGTTGTTGTT GAC
(GTT) ₄ CC-f	CGCTCTTCCGATCT CTG GTTGTTGTTGTT CC
(GTT) ₄ TC-f	CGCTCTTCCGATCT CTG GTTGTTGTTGTT TC
(GTG) ₄ AC-f	CGCTCTTCCGATCT CTG GGTGGTGGTGGT GAC
(GT) ₆ TC-f	CGCTCTTCCGATCT CTG GTGTGTGTGTGT TTC
(TG) ₆ AC-f	CGCTCTTCCGATCT CTG TGTGTGTGTGT GAC
Reverse primers:	(Tail + anchor: GAC) + SSR + anchor
(ACT) ₄ TG-r	TGCTCTTCCGATCT GAC ACTACTACTACT TTG
(CTA) ₄ TG-r	TGCTCTTCCGATCT GAC CTACTACTACTAT G
(TTG) ₄ AC-r	TGCTCTTCCGATCT GAC TTGTTGTTGTT GAC
(GTT) ₄ CC-r	TGCTCTTCCGATCT GAC GTTGTTGTTGTT CC
(GTT) ₄ TC-r	TGCTCTTCCGATCT GAC GTTGTTGTTGTT TC
(GTG) ₄ AC-r	TGCTCTTCCGATCT GAC GTGGTGGTGGT GAC
(GT) ₆ TC-r	TGCTCTTCCGATCT GAC GTGTGTGTGT TTC
(TG) ₆ AC-r	TGCTCTTCCGATCT GAC TGTGTGTGT GAC

すべてのISSRとアンカーの配列組み合わせの中から問題のない組み合わせを選んだ。

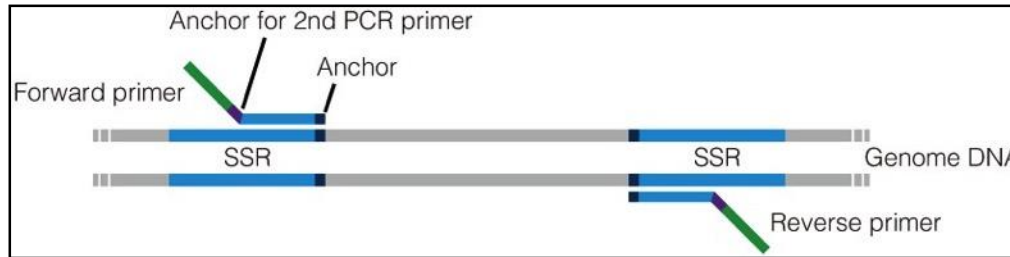
Underlined and boldface nucleotides denote tail and anchor sequences, respectively. The difference between forward and reverse primer sets is only in their tail sequences.

1. MIG-seq法の概要

Step #1: 1st PCR

一応、2つのセットを作った。
座数が足りない場合等に
使うことができる。

Tailed ISSR primers set-1 & 2



Primer set-1

Primer set-2

Table 1. Sequences of MIG-seq primer set-1 for the 1st PCR

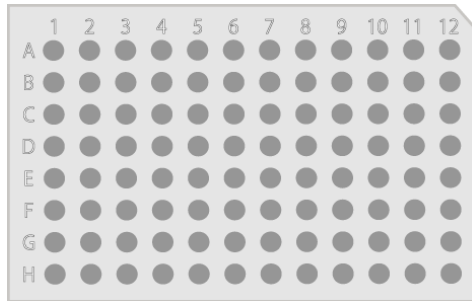
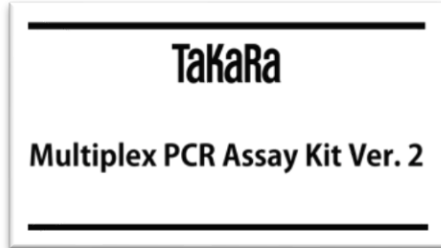
Name	Sequences (5'–3')
Forward primers:	(Tail + anchor: CTG) + SSR + anchor
(ACT) ₄ TG-f	CGCTCTTCCGATCTCTGACTACTACTACTTG
(CTA) ₄ TG-f	CGCTCTTCCGATCTCTGCTACTACTACTATG
(TTG) ₄ AC-f	CGCTCTTCCGATCTCTGTTGTTGTTGTTGAC
(GTT) ₄ CC-f	CGCTCTTCCGATCTCTGGTTGTTGTTGTTCC
(GTT) ₄ TC-f	CGCTCTTCCGATCTCTGGTTGTTGTTGTTTC
(GTG) ₄ AC-f	CGCTCTTCCGATCTCTGGTGGTGGTGGTGAC
(GT) ₆ TC-f	CGCTCTTCCGATCTCTGGTGTGTGTGTGTTTC
(TG) ₆ AC-f	CGCTCTTCCGATCTCTGTGTGTGTGTGTGAC
Reverse primers:	(Tail + anchor: GAC) + SSR + anchor
(ACT) ₄ TG-r	TGCTCTTCCGATCTGACACTACTACTACTTG
(CTA) ₄ TG-r	TGCTCTTCCGATCTGACTACTACTACTATG
(TTG) ₄ AC-r	TGCTCTTCCGATCTGACTTGTGTTGTTGTTGAC
(GTT) ₄ CC-r	TGCTCTTCCGATCTGACGTTGTTGTTGTTCC
(GTT) ₄ TC-r	TGCTCTTCCGATCTGACGTTGTTGTTGTTTC
(GTG) ₄ AC-r	TGCTCTTCCGATCTGACGTGGTGGTGGTGAC
(GT) ₆ TC-r	TGCTCTTCCGATCTGACGTGTGTGTGTGTTTC
(TG) ₆ AC-r	TGCTCTTCCGATCTGACTGTGTGTGTGTGAC

Supplementary Table 1. Sequences of MIG-seq primer set-2 for the 1st PCR

Name	Sequences (5'–3')
Forward primers:	(Tail + anchor: CTG) + SSR + anchor
(CTG) ₄ AC-f	CGCTCTTCCGATCTCTGCTGCTGCTGCTGAC
(GCT) ₄ TG-f	CGCTCTTCCGATCTCTGGCTGCTGCTGCTTG
(GTA) ₄ TG-f	CGCTCTTCCGATCTCTGGTAGTAGTAGTATG
(GTT) ₄ AG-f	CGCTCTTCCGATCTCTGTTGTTGTTGTTAG
Reverse primers:	(Tail + anchor: GAC) + SSR + anchor
(CTG) ₄ AC-r	TGCTCTTCCGATCTGACCTGCTGCTGCTGAC
(GCT) ₄ TG-r	TGCTCTTCCGATCTGACGCTGCTGCTGCTTG
(GTA) ₄ TG-r	TGCTCTTCCGATCTGACGTAGTAGTAGTATG
(GTT) ₄ AG-r	TGCTCTTCCGATCTGACGTTGTTGTTGTTAG

1. Brief introduction of MIG-seq

Step #1: 1st PCR

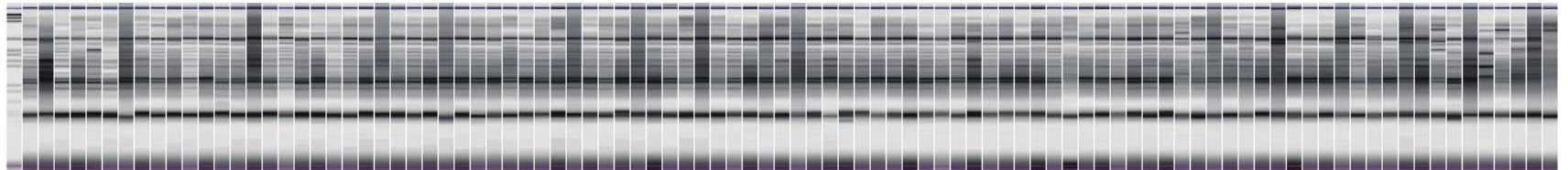


Forward primers: (Tail + anchor: CTG) + SSR + anchor
 (ACT)₄TG-f CGCTCTTCCGATCTCTGACTACTACTACTTG
 (CTA)₄TG-f CGCTCTTCCGATCTCTGCTACTACTACTATG
 (TTG)₄AC-f CGCTCTTCCGATCTCTGTTGTTGTTGTTGAC
 (GTT)₄CC-f CGCTCTTCCGATCTCTGGTTGTTGTTGTTCC
 (GTT)₄TC-f CGCTCTTCCGATCTCTGGTTGTTGTTGTTTC
 (GTG)₄AC-f CGCTCTTCCGATCTCTGGTGGTGGTGGTGAC
 (GT)₆TC-f CGCTCTTCCGATCTCTGGTGTGTGTGTGTTCC
 (TG)₆AC-f CGCTCTTCCGATCTCTGTGTGTGTGTGTGAC

Reverse primers: (Tail + anchor: GAC) + SSR + anchor
 (ACT)₄TG-r TGCTCTTCCGATCTGACACTACTACTACTTG
 (CTA)₄TG-r TGCTCTTCCGATCTGACCTACTACTACTATG
 (TTG)₄AC-r TGCTCTTCCGATCTGACTTGTGTTGTTGAC
 (GTT)₄CC-r TGCTCTTCCGATCTGACGTTGTTGTTGTTCC
 (GTT)₄TC-r TGCTCTTCCGATCTGACGTTGTTGTTGTTTC
 (GTG)₄AC-r TGCTCTTCCGATCTGACGTGGTGGTGGTGAC
 (GT)₆TC-r TGCTCTTCCGATCTGACGTGTGTGTGTGTTCC
 (TG)₆AC-r TGCTCTTCCGATCTGACTGTGTGTGTGTGAC

Template DNA	1.0 μL
[Mixture]	
2x Multiplex PCR Buffer	10.0 μL
1 st PCR primers	3.2 μL each @20 μM
Multiplex PCR Enzyme mix	0.1 μL
Water	5.7 μL
	<hr/>
	20.0 μL

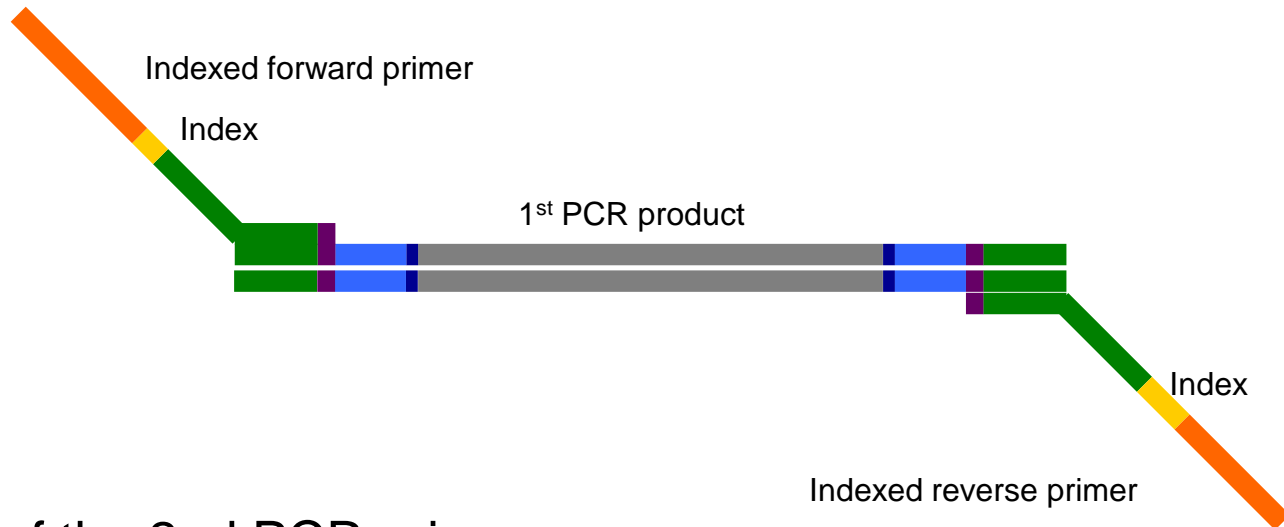
94°C	1 min	
94°C	30 sec	
38°C	1 min	18–25 cycles
72°C	1 min	
72°C	10 min	



1. MIG-seq法の概要

Step #2: 2nd PCR (Tailed PCR)

Tailed PCRによって、1st PCR産物の両端にシーケンシングに必要な配列とインデックス配列を付加する。



New version of the 2nd PCR primers

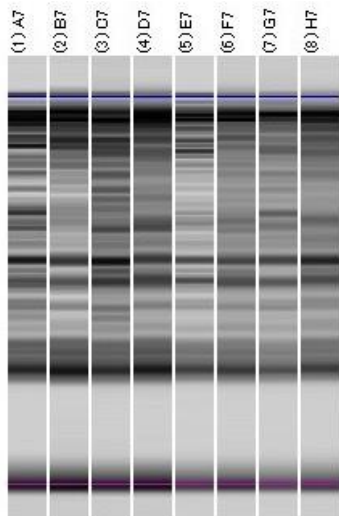
	Index
Forward (index)	AATGATACGGCGACCACCGAGATCTACAC xxxxxx ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTG
Reverse (index)	CAAGCAGAAGACGGCATACGAGAT xxxxxxxxxx GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAC
	Index

1. MIG-seq法の概要

Step #3-1,2 : サイズ選択、定量

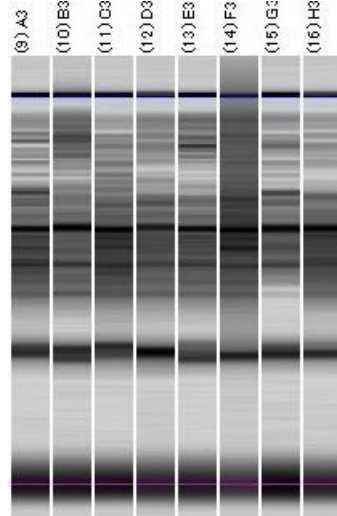
1st PCR

(Multiplexed Tailed PCR)

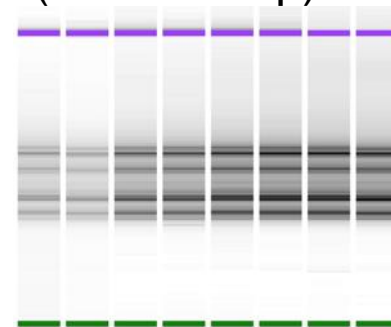


2nd PCR

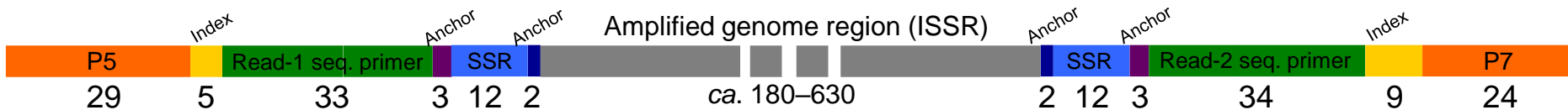
(Tailed PCR + Index)



Pool, Size selection (350–800 bp)



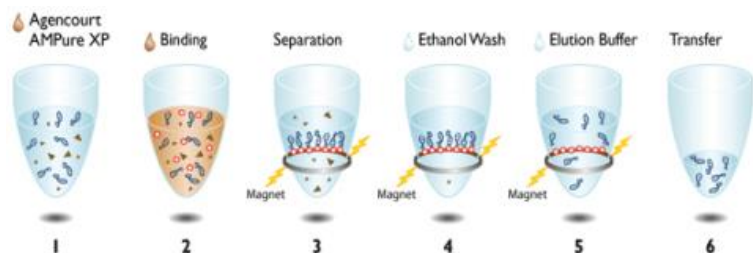
Pippin Prep
(Sage Science)



ライブラリーの構造

1. MIG-seq法の概要

Step #3-1: サイズ選択 (350–800 bp)



あるいは



ゲルカセット



Pippin Prep
(Sage Science)

磁性ビーズ法

(SPRIselect or AMPure XP, Beckman Coulter)

Step #3-2: 定量



定量用キット



qPCR

Step #3-3: NGSラン



MiSeq用キット



NGS
(MiSeq等)

1. MIG-seq法の概要

Step #3-3: MiSeq Run

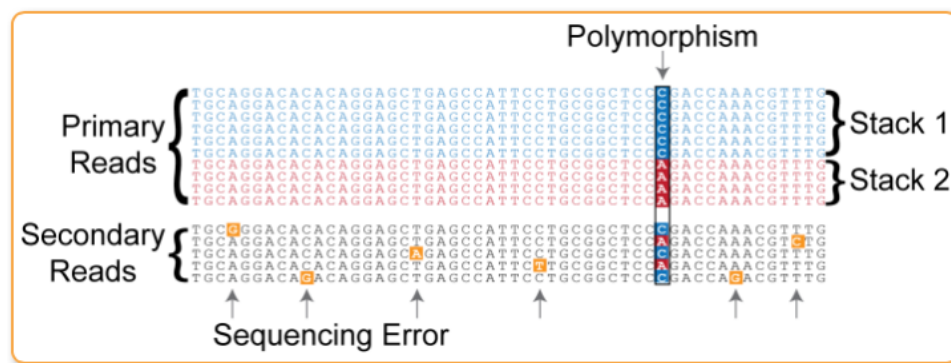
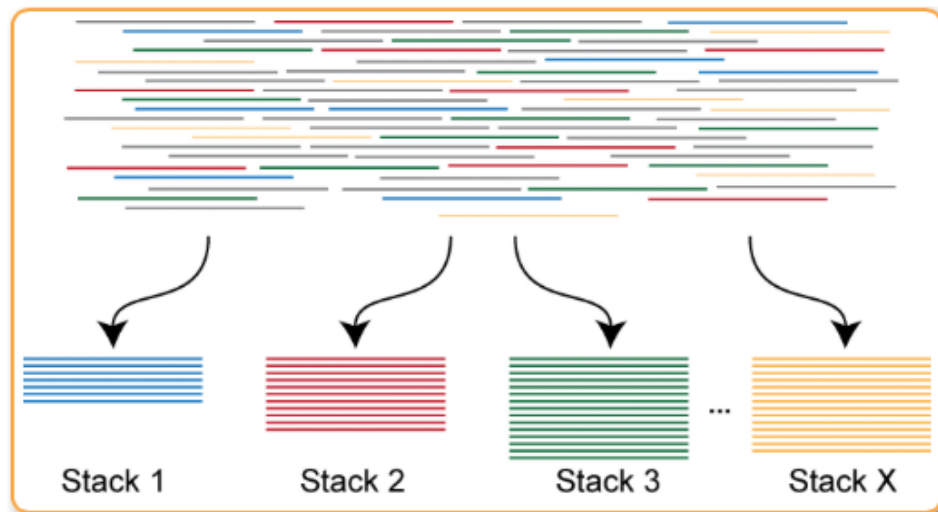


MiSeqを用いた1ラン
(96~192サンプル)の平均的データ量

生リード数: ~6000万リード

SNP座数: ~10,000 SNPs

Step #4: データ解析



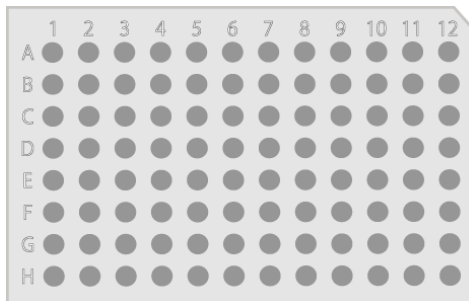
データ解析ソフト: Stacks (Catchen et al 2011)

1. MIG-seq法の概要

時間と費用（MiSeqを用いた標準的な分析例）

1日目

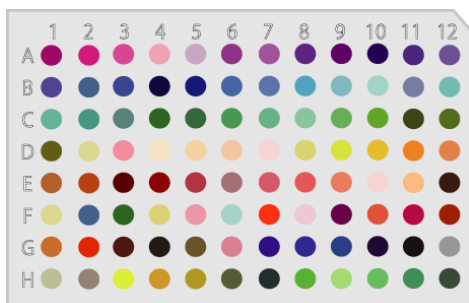
- 1) 1st PCR
マルチプレックスPCR
- 2) 2nd PCR
必要な配列を付加



96-192 サンプル
(前情報不要)

2日目

- 1) 全サンプルを混合
サイズ選択、定性・定量
- 2) 一度にNGSでラン



時間：3日間

ライブラリ構築: 1日
NGSラン: 1日
データ解析: 1日

3日目

データ解析



費用：約15万円

1. MIG-seq法の概要
- 2. 研究例の紹介**
3. 手法の改良
4. iSeq100の利用
5. その他の応用

- 1) 個体間レベル
ジェネット（クローン） 識別
- 2) 集団間レベル
集団遺伝・分子系統地理
- 3) 近縁種間レベル
雑種識別
- 4) 種（あるいは属）間レベル
系統関係

1. MIG-seq法の概要
2. 研究例の紹介
- 3. 手法の改良**
4. iSeq100の利用
5. その他の応用

3. 手法の改良 (1)

異なるPCR条件下における生産性・再現性試験

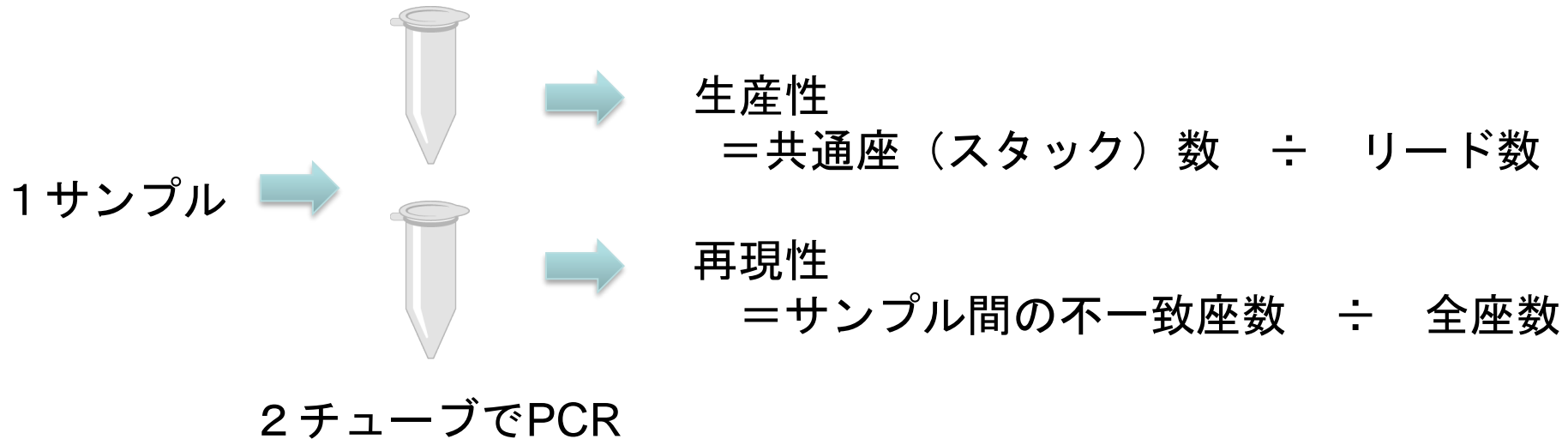
Step #1: 1st PCRの条件検討

- ・ PCRキット
(original: TaKaRa Multiplex PCR Assay Kit Ver. 2)
- ・ プライマー濃度
(original: 0.2 μ M)
- ・ アニール温度
(original: 48° C)

3. 手法の改良（1）

異なるPCR条件下における生産性・再現性試験

材料: 異なる8種



3. 手法の改良 (1)

異なるPCR条件下における生産性・再現性試験

1) 1st PCRのアニーリング温度

Step #1: 1st PCR

48° C

VS.

28, 33, 38, 43, 53° C

Profile of the 1st PCR

94° C 1 min.

(18–25 cycles)

94° C 30 sec.

XX° C 1 min.

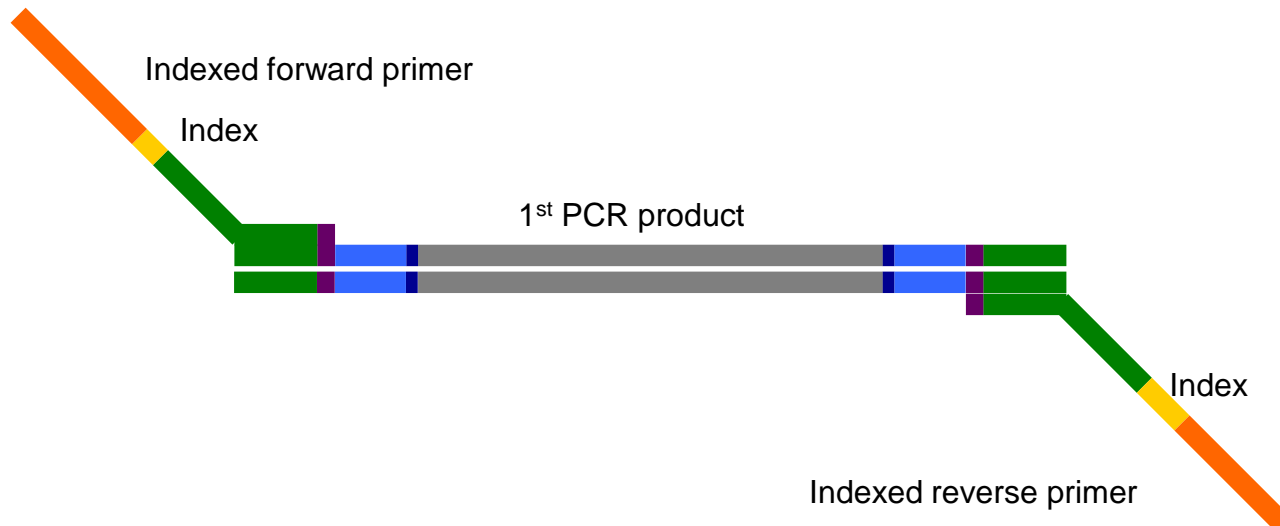
72° C 1 min.

72° C 10 min.

3. 手法の改良（2）

Step #2: 2nd PCRの条件等検討

- ・ PCRサイクル
- ・ インデックス
（シングル→デュアル）
- ・ その他



3. 手法の改良

2) インデックス

(シングル→デュアル)

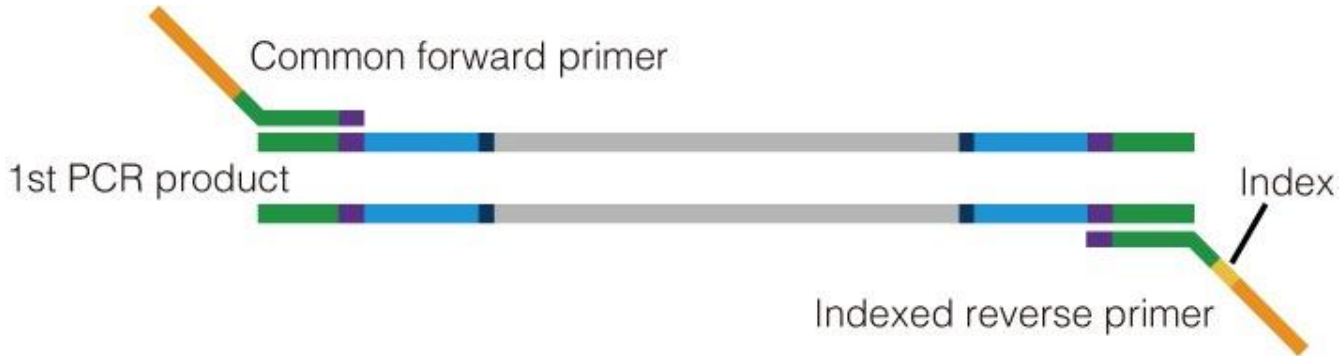


Fig. Schematic diagram of MIG-seq 2nd PCR

2nd PCR primers

Forward
(Common)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTG

Reverse

CAAGCAGAAGACGGCATAACGAGATxxxxxxGTGACTGGAGTTCAGACGTGCGTGCTCTTCCGATCTGAC

Index

シングルインデックスによる解析で実際に検出された振り分け間違いの例

: 358 ± 432 reads / index

その他の改変

2) インデックス

(シングル→デュアル)

デュアルインデックスによる解析で実際に検出された振り分け間違い
: 1.6 ± 12.5 reads / index

New version of the 2nd PCR primers

Index

Forward (index) **AATGATACGGCGACCACCGAGATCTACAC**xxxxx**ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTG**

Reverse (index) **CAAGCAGAAGACGGCATACGAGAT**xxxxxxxxxx**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAC**

Index

3. 手法の改良 (3)

データ解析手法の改善

<これまで>

- ・ 解析ソフトStacksを使った「厳しめ」の解析

Stacks



<改良法 1>

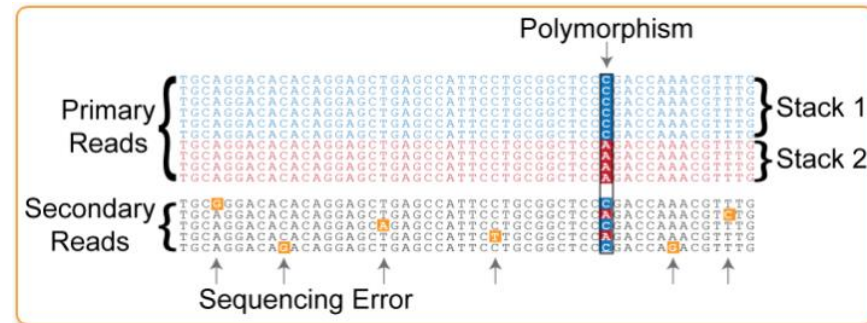
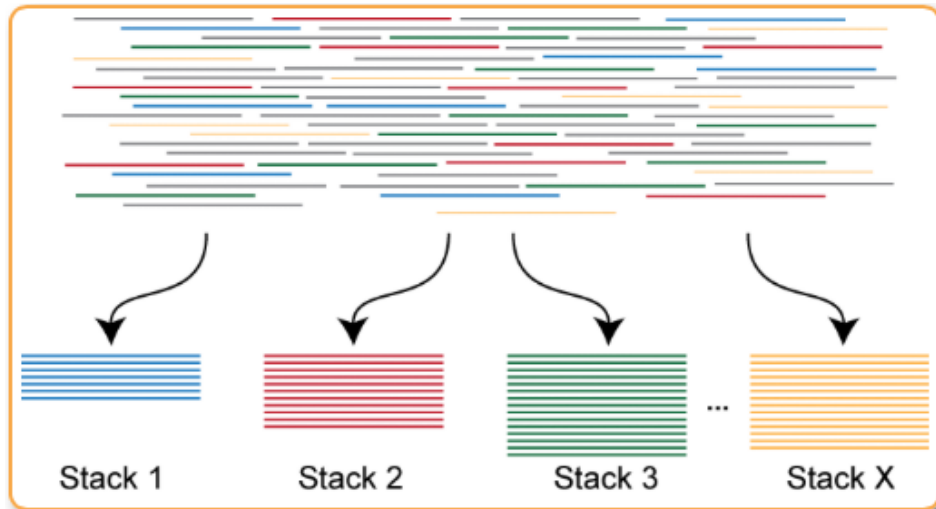
- ・ 最小デプスを下げた (m=3) 解析 = Stacksのデフォルト設定

<改良法 2>

- 1) Stacksで各サンプルの座を抽出
- 2) 全サンプルでリファレンスを作製
- 3) リファレンスに各サンプルの座を貼り付ける

<その効果>

- ・ 座数と解析可能サンプル数の増加
- ・ 正確性の向上



1. MIG-seq法の概要
2. 研究例の紹介
3. 手法の改良
- 4. iSeq100の利用**
5. その他の応用

4. iSeq100の利用

iSeq 100システム

400万リード、1.2GB/ランのNGSが驚きの300万円台



CCGA
TAGAT
AGA
AC

400万
リード



9-17 hr
ラン時間

10101
01100
11010

1.2 Gb
アウトプット



2x150 bp
ペアエンドリード

iSeq 100 システム本体

3,900,000円

iSeq 100 i1 試薬

400万リードのシーケンス試薬

1キットパッケージ or 4キットパッケージ

110,000円

415,000円

価格は税抜です。
2018年4月時点の価格です。

For Research Use Only. Not for use in
diagnostic procedures.

イルミナ社のHPより

4. iSeq100の利用

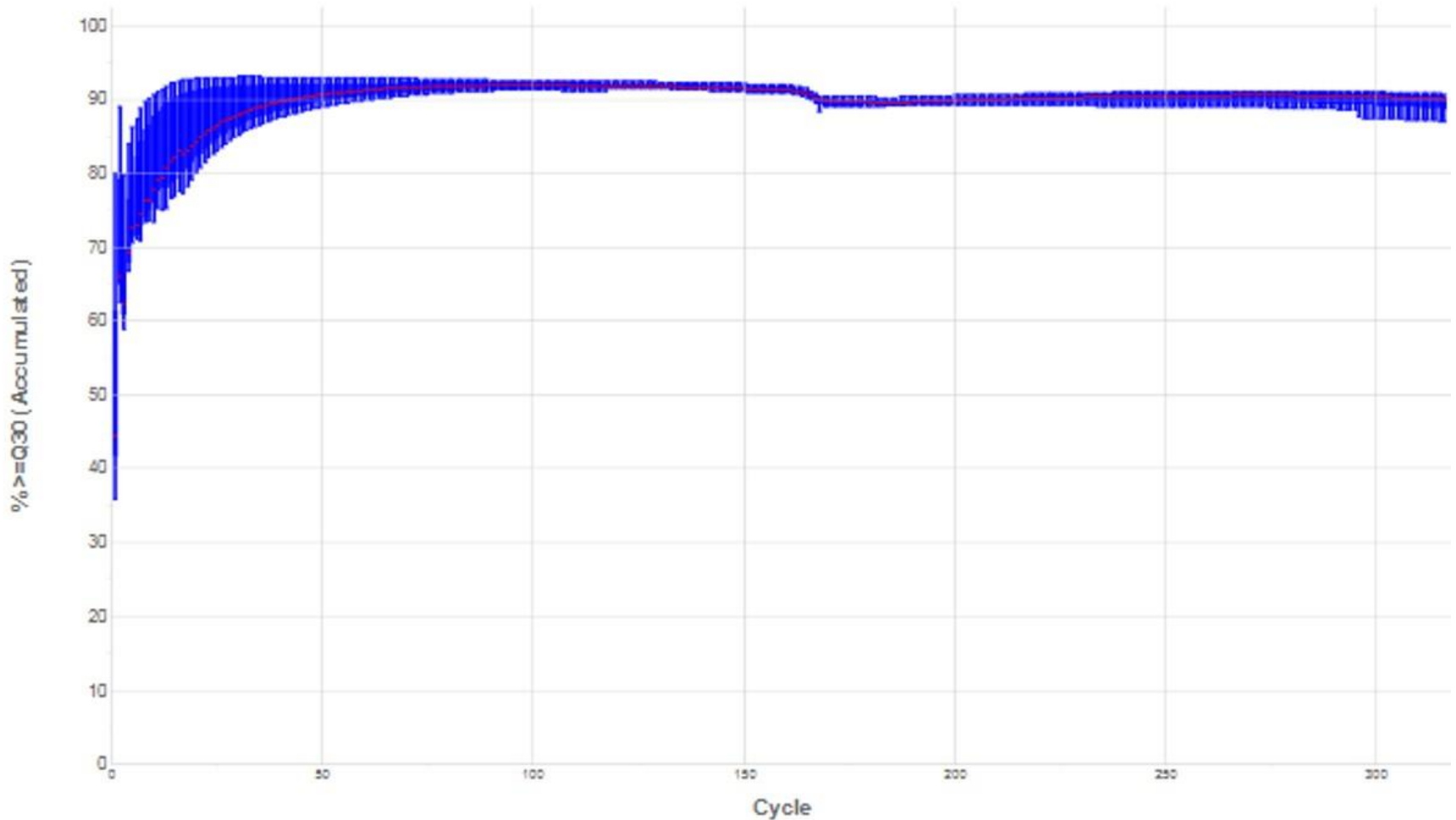


図 iSeq100を用いたMIG-seqライブラリーのランにおける
累計Q30（以上）平均値の変化

iSeq100i1 Reagent (300 Cycles)による 1 ランにより48サンプルを分析。

4. iSeq100の利用

表 iSeq100を用いたMIG-seq分析により7分類群の生物から得られたSNP数

材料	サンプル数	集団等	総リード数	最小リード数	最大リード数	SNP数
動物プランクトン	8	2地域	1,151,360	120,052	174,201	2,974
アノール	8	2地域	835,232	79,080	128,729	1,222
スギ	8	4地域	1,008,344	98,999	143,518	5,489
アツモリソウ属	8	2地域	944,908	101,985	131,681	2,338
キスゲ類	8	2近縁種	1,101,455	114,989	184,300	3,905
ササ	4	1集団	411,815	91,207	107,897	3,371
米	4	4品種	947,689	202,002	256,736	850

iSeq100i1 Reagent (300 Cycles) を用いた1ランにより48サンプルの分析を行なった例。

4. iSeq100の利用

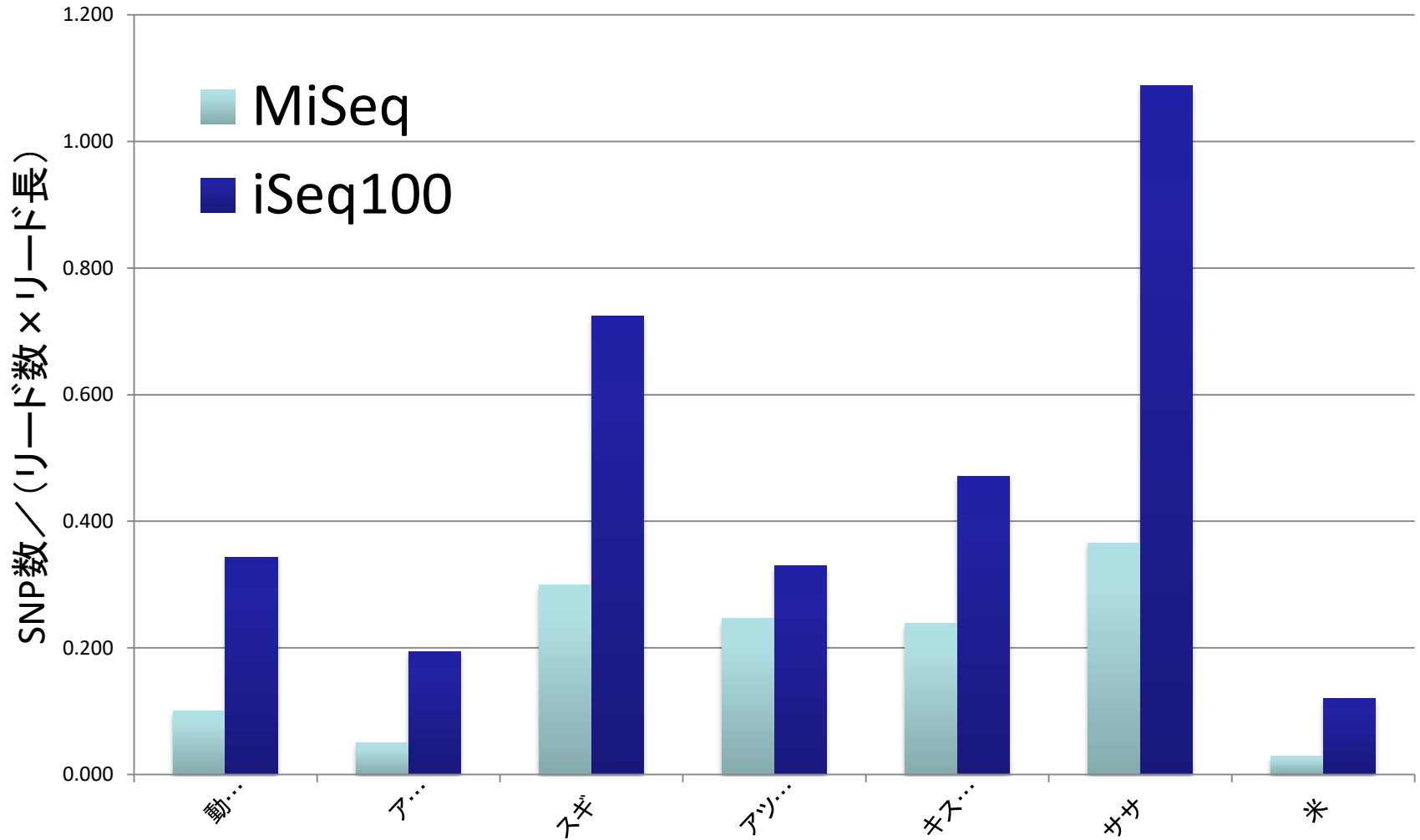


図 iSeq100およびMiSeqを用いたMIG-seq分析により7分類群の生物から得られたSNP数の比較

iSeq100 : iSeq100i1 Reagent (300 Cycles)による1ランにより48サンプルを分析。
MiSeq : MiSeq Reagent Kit v3 (150 Cycles)による1ランにより48サンプルを分析。

1. MIG-seq法の概要
2. 研究例の紹介
3. 手法の改良
4. iSeq100の利用
5. **その他の応用**

まとめ: MIG-seqの特徴

長所

- ・ 低品質／微量DNAの利用が可能
- ・ 早い（3日間）
- ・ 簡単（2回のPCRとNGSランのみ）
- ・ 安い（1サンプル750～1500円程度）

短所

- ・ 数万座以上の解析には向かない
- ・ SSR近傍の配列解析に限定される

向いている研究

- ・ 非モデル野外生物の集団遺伝
- ・ 種・品種・雑種・クローン識別
- ・ 種間比較

MIG-seq法により、大まかな遺伝的情報の取得が可能

→クローン識別

→遺伝的集団構造解析

→遺伝的多様性評価

→種間関係解析

よりよい研究にお役立ていただければ幸いです。