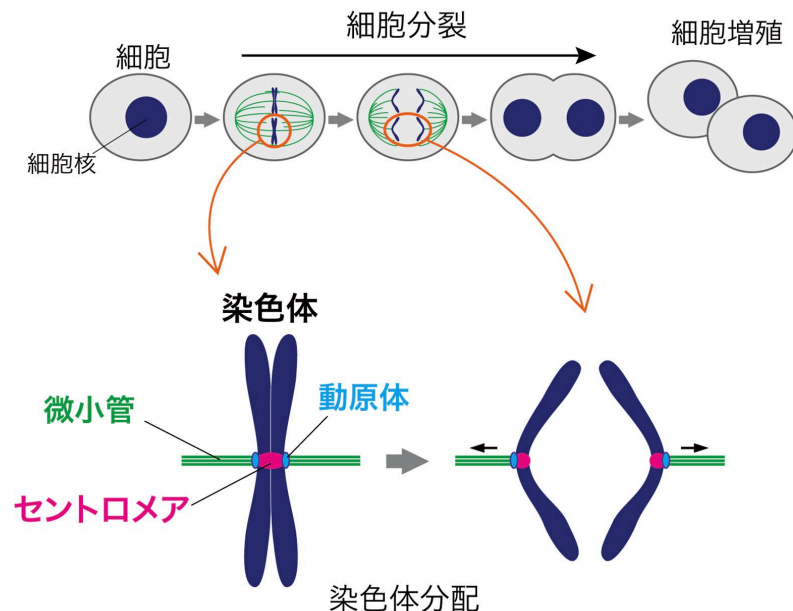


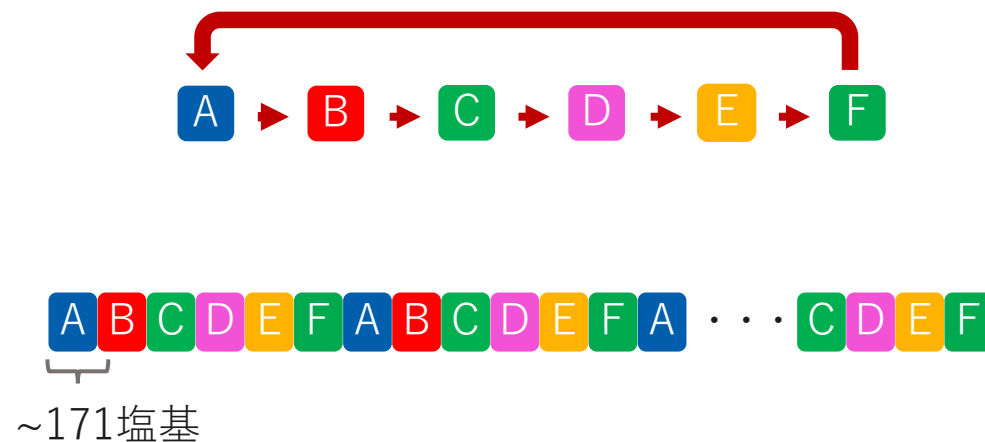
ヒトゲノムの未解明領域「セントロメア」の 大規模解析が実現

世界中に蓄積されたゲノムデータを活用し、がんなどの疾患研究を加速

国立がん研究センター研究所 ゲノム解析基盤開発分野 白石友一



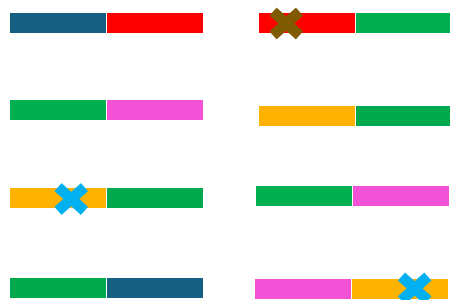
7番染色体のセントロメア配列の並び



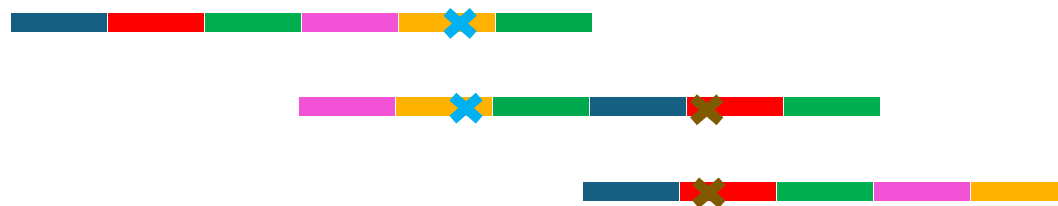
出典：大阪大学 深川竜郎研究室ウェブサイト（一部改変）

- 細胞分裂で、染色体を2つの細胞に正しく分けるために必要。
- ほぼ同じDNA配列が数百万塩基にわたって繰り返す。
- 従来のシーケンスデータでは解析困難。

従来のゲノムシーケンス (ショートリード)



ロングリードシーケンス



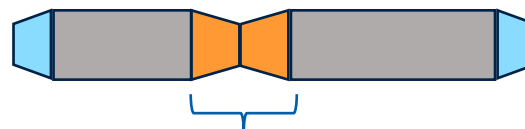
- ショートリード：DNAを約150塩基ずつ読み取る。
- ロングリード：DNAを数万～数十万塩基にわたって読み取る。
- 繰り返しの多いセントロメアも、詳しく解析できるように。

従来の参照ゲノム



- 2003年に発表。
- ヒトゲノム計画
- 現在最も活用されている。
- 未解明領域が多数残存。

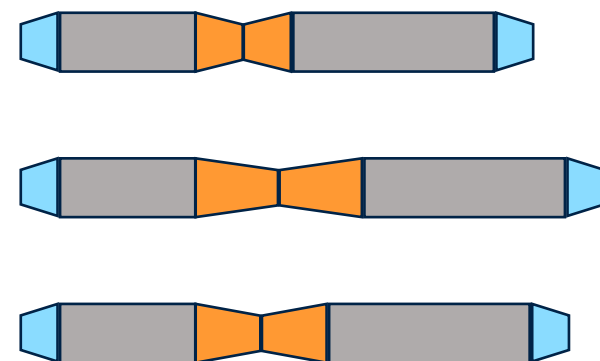
完全参照ゲノム



セントロメア領域

- 2022年発表。
- T2Tコンソーシアム
- 単一試料に由来する完全な参照ゲノム。
- 集団差・個人差を十分に表現できない。

パンゲノム参照配列



- 「多数」の参照ゲノム配列。
- 現在数百の参照配列が蓄積。
- 世界中で「パンゲノム」計画が進行中。

- ショートリードWGSデータは世界中に蓄積（2026年時点）。
 - UK Biobank（英国）：約50万人
 - All of Us（米国）：約54万人
 - 全ゲノム解析等実行計画（日本）：約4万症例
- 多くの臨床検体をロングリードで解析するには課題がある。
 - 多くのDNAが必要。
 - 解析費用が高い。
- パンゲノム研究で得られた「多数のセントロメア配列」を活用し、既存のショートリードWGSデータからセントロメア解析を可能に！

研究成果1

既存のショートリードWGSデータからセントロメアを解析する新手法を開発

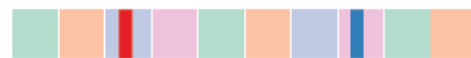
研究成果2

世界各地の3,332人のショートリードWGSデータの再解析から、地域・集団によるセントロメアの多様性を明らかにした

研究成果3

乏突起膠腫に特徴的な染色体異常の形成メカニズムの一端を解明

共通祖先の
セントロメア配列

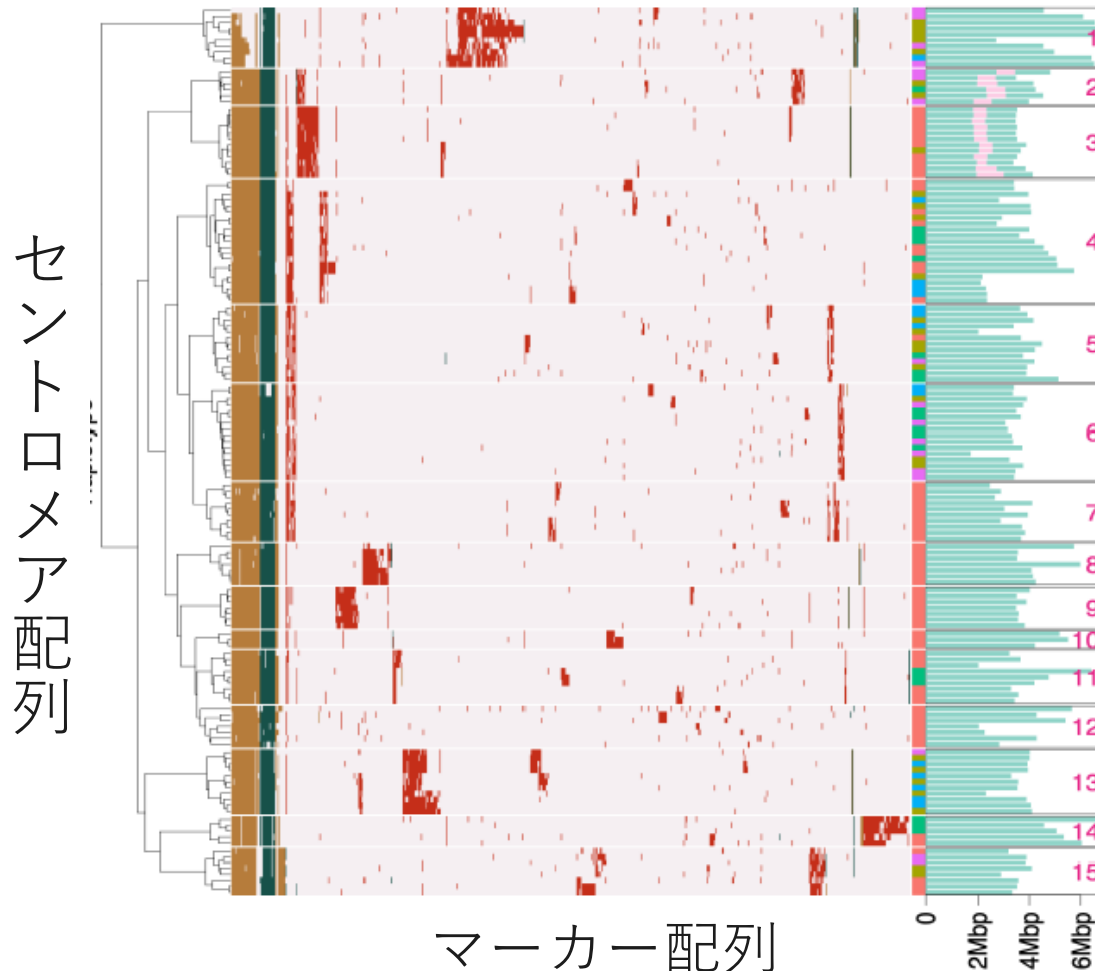


現代人の
セントロメア配列

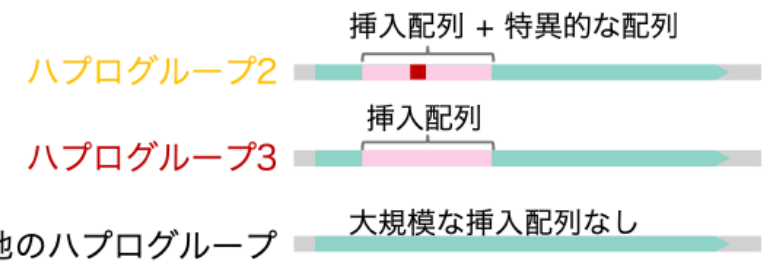


- セントロメアの繰り返し配列中には、ごく少数回しか現れない短いDNA配列（rare k-mer）がある。
- これらのDNA配列の種類や組み合わせは、セントロメアのタイプごとに異なる。
- この違いを『DNAの目印』として利用。

19番染色体セントロメアの分類

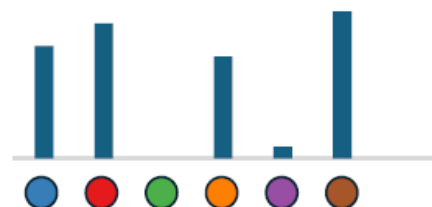


挿入配列をもつハプログループ: 2, 3

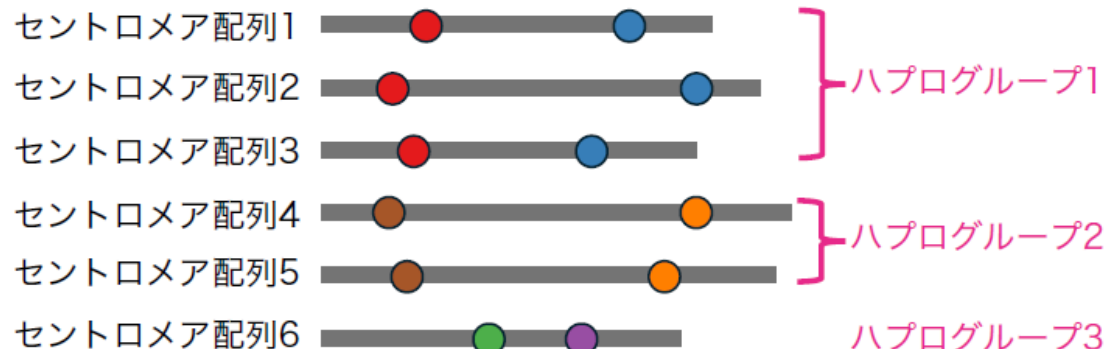


- 各染色体で3~16、全体で214のタイプ（ハプログループ）に分類。
- ハプログループはセントロメアの系統関係を反映。

通常的全ゲノムデータから得られる「DNAの目印」のパターン



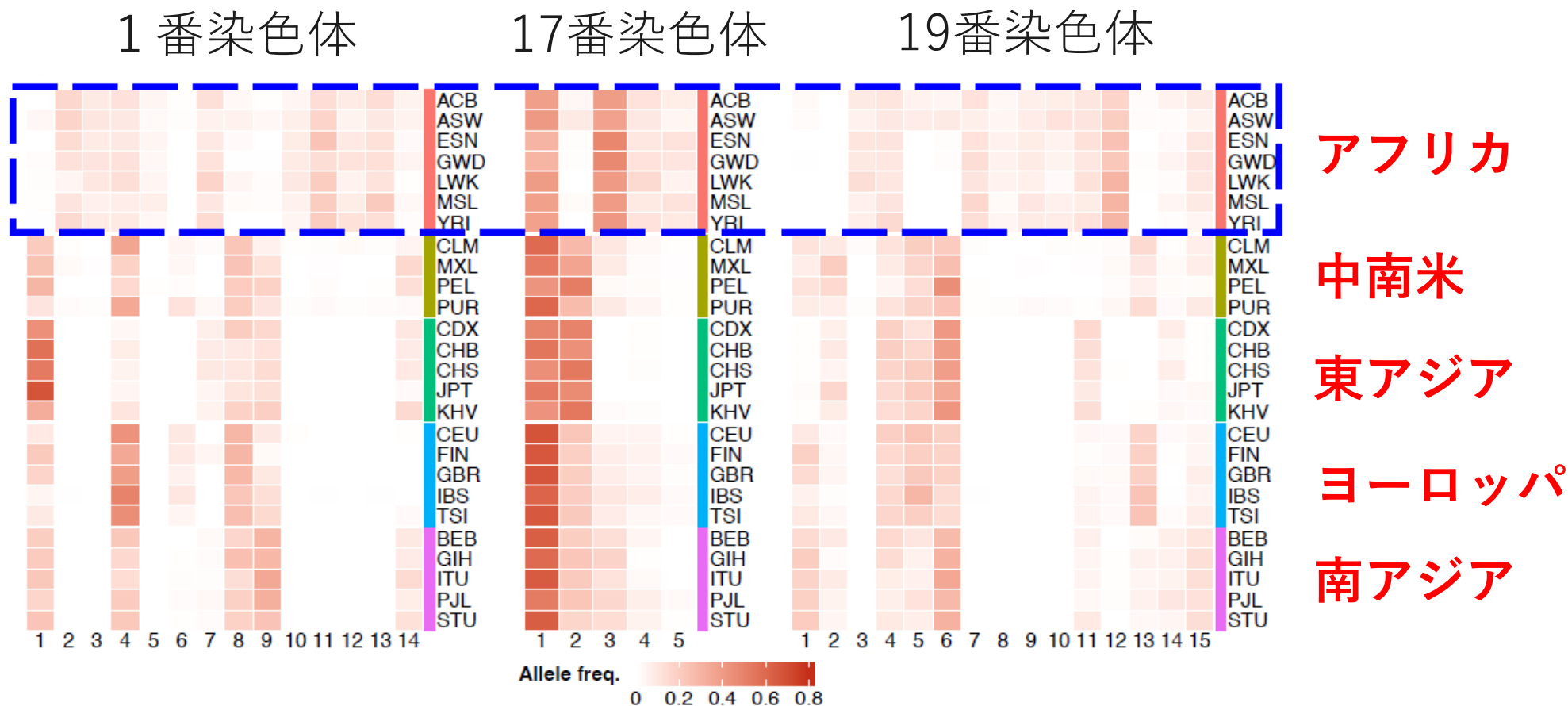
パンゲノムから作成したセントロメア参照パネル



パターンが最もよく合うセントロメア配列・ハプログループを推定

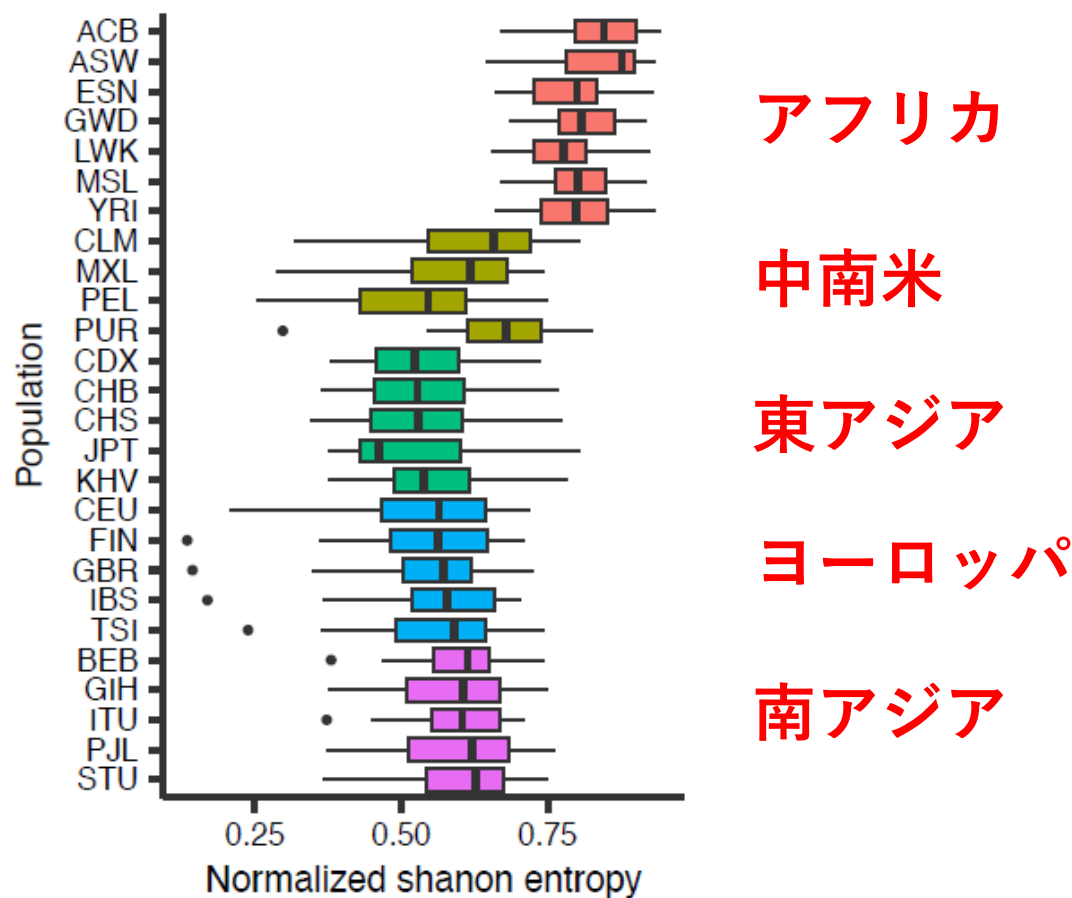


- 染色体ごとの推定精度：80.7～100%（中央値92.7%）
- プログラムはGitHub上で公開
(<https://github.com/friend1ws/asclairn>)



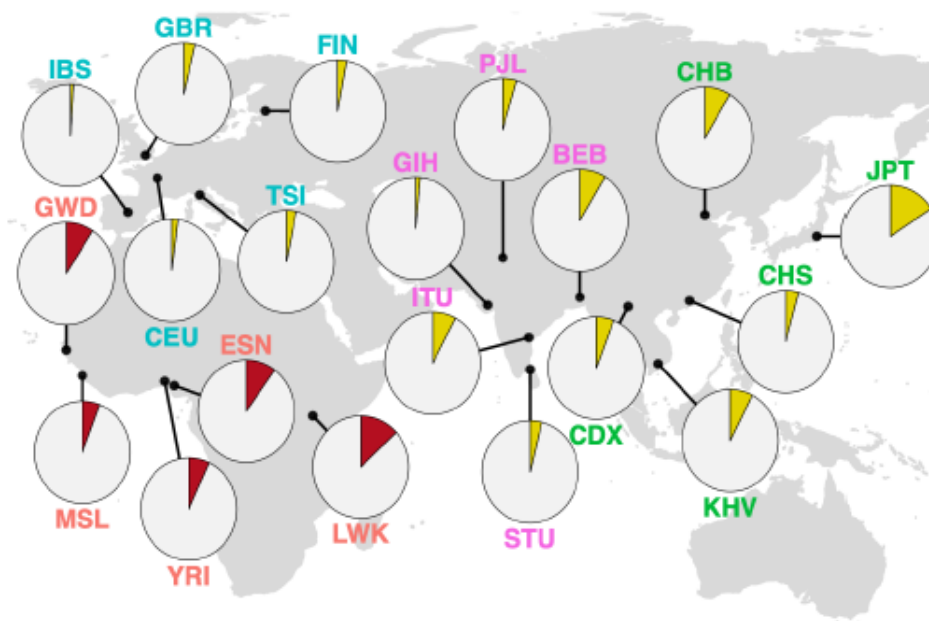
* 掲載図は1000 Genomes Projectの2,504人の解析結果

- 世界規模のWGSデータから、セントロメアのタイプを推定。
- アフリカ由来の集団で多様性が最も高い。

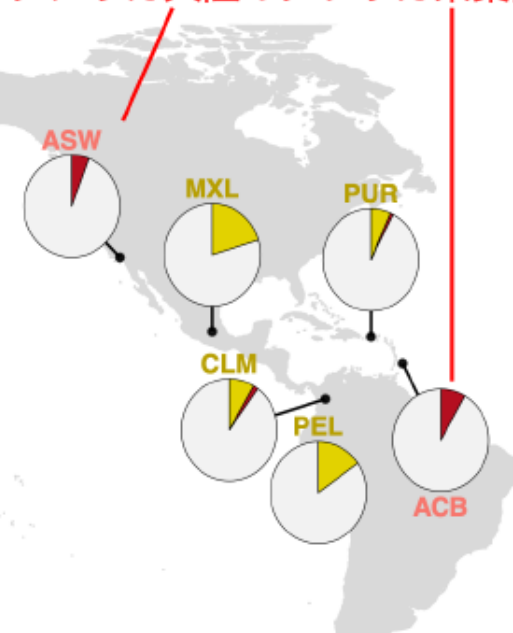


* 掲載図は1000 Genomes Projectの2,504人の解析結果

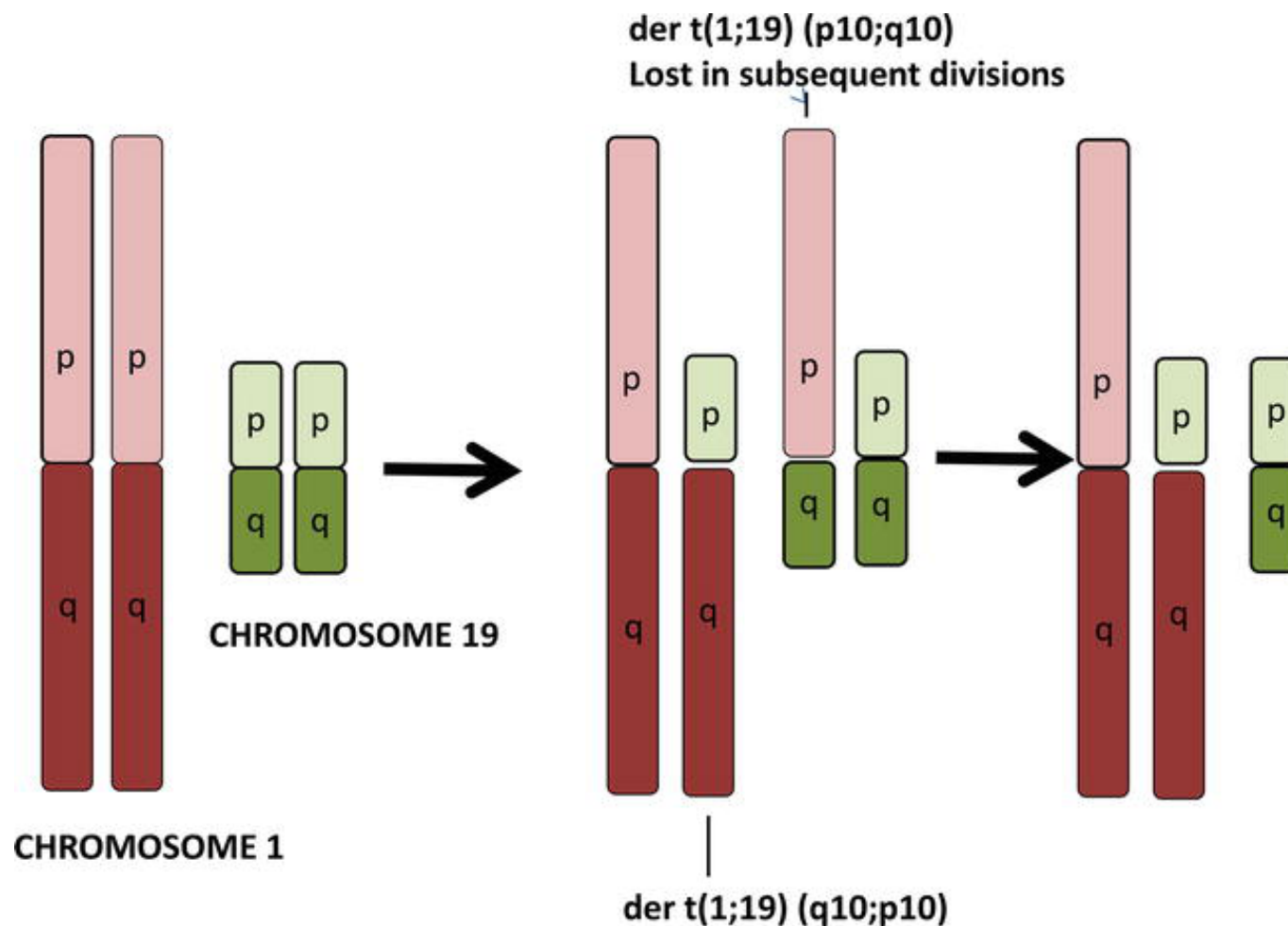
- セントロメアの多様性を定量的に評価した結果、アフリカ由来の集団で最も高かった。
- 一方、東アジア由来の集団では比較的低い傾向が見られた。



アメリカ大陸のアフリカ系集団

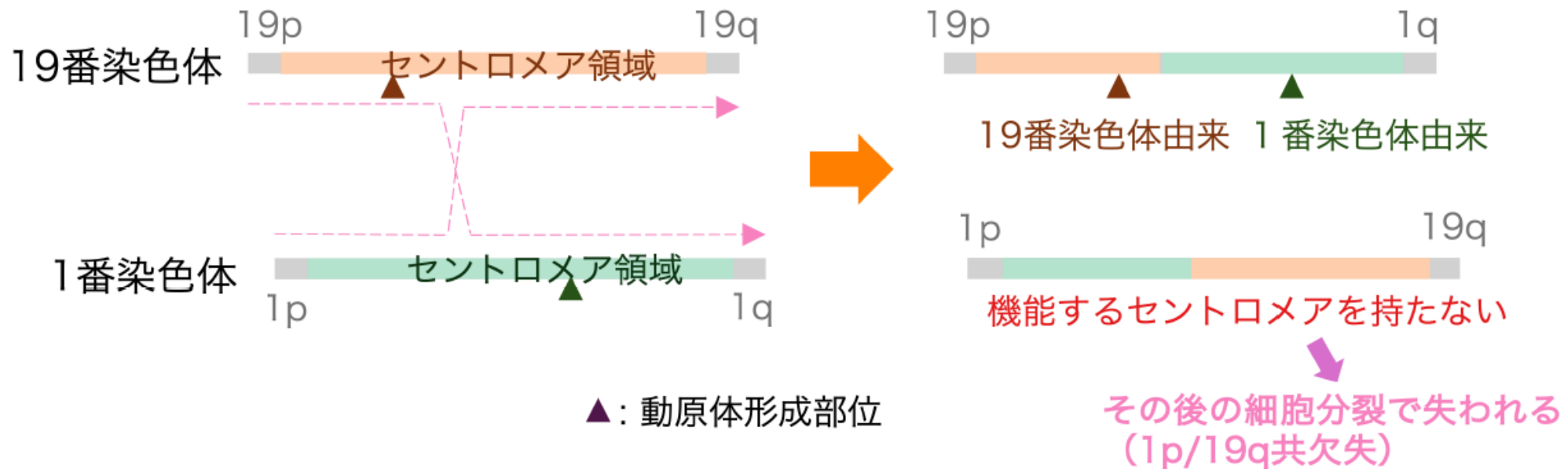


特徴的なセントロメア・ハプログループの世界的な分布を明らかにし、人類史を反映する地域差を発見。

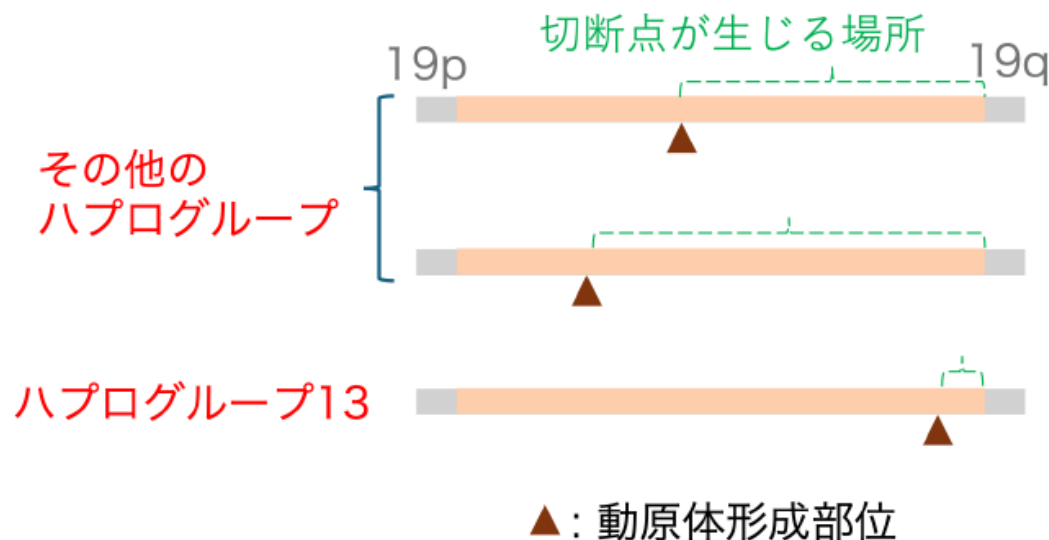
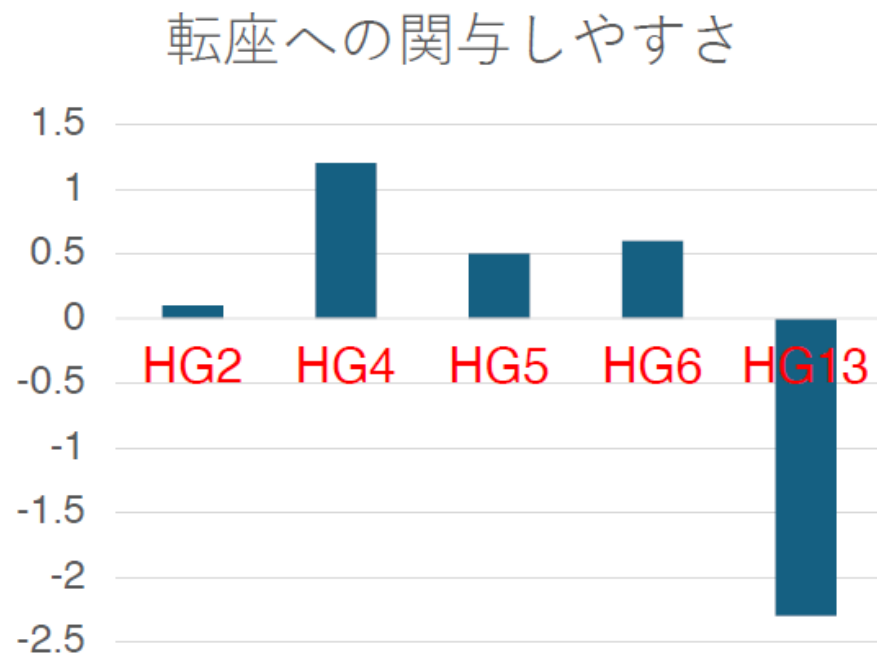


1p/19q共欠失の形成過程を、セントロメアの切断位置から解析。

研究成果3-1: 切断位置の偏りが、1p/19q共欠失の形成モデルと整合



- がんゲノムデータベース、TCGAに登録された乏突起膠腫142例のショートリードWGSを再解析し、染色体の切断位置を推定。
- 切断点の位置
 - 切断点は、動原体形成部位を境に1番染色体では1p側、19番では19q側に偏る。
 - 1p/19q共欠失の形成過程と整合。
- さらに2症例をロングリードで解析し、転座構造を直接確認。



- 転座が起こり得る範囲が長いタイプほど転座に関与しやすい。
- 19番染色体のハプログループ13は転座に関与しにくい。

- 新手法の開発：
既存のショートリードWGSデータから、セントロメアのタイプを推定する手法を開発。
- 世界規模の解析：世界各地の3,332人のWGSデータを再解析し、地域・集団による多様性を確認。
- がん研究への応用：セントロメア内の切断位置の偏りから、1p/19q共欠失の形成メカニズムの一端を解明。
- 今後の展開：
 - 世界中に蓄積された既存のゲノムデータを活用し、がんなどの疾患とセントロメアとの関係で大規模に調べる。
 - 古代人ゲノムにも解析を広げ、セントロメアの多様性から人類の分岐・移動史を探る。