



# 遺伝子系図の推定と集団遺伝学への応用

シュパイデル玲雄  
れお

Department of Statistics, University  
of Oxford, Oxford UK

近年、人を含む数々の種において、数千から数万の全ゲノムシーケンスが行われてきた。我々は今、前代未聞の規模で遺伝的多様性を記録している。遺伝子系図 (gene genealogy) は、このような DNA サンプルの遺伝的関係を記述する (図1)。人類の遺伝子系図は、数百万年の進化史を記録する。それは、例えばホモ・サピエンスの進化と移住、および様々な環境変化への適応状況を含むものだ。

我々は、遺伝子系図の正確な推定を目的とした手法「Relate」を開発した<sup>1)</sup>。従来の手法は解析可能なサンプル数が百以下に留まっていたところ、Relateは数千のDNAサンプルに適用可能で、初めて大規模な遺伝子系図の推定を可能にした。

従来の手法では遺伝子系図の推定が困難であったことから、データの要約統計量を用いた解析が主流であった。主成分分析や Tajima's D などがその例である。このようなアプローチと比べて、遺伝子系図を用いた本解析は、検出力や精度の改善が大いに期待される。なぜなら、具体的な遺伝的過去を再現することにより、DNAデータに含まれる進化史の情報を最大限に引き出し、直接的な推定が可能になるからである。

幅広い応用性を示すべく、Relateを千人ゲノム・プロジェクトのデータに適用し、アフリカ・東アジア・南アジア・ヨーロッパ・アメリカに及ぶ26の集団、全2,478人の遺伝子系図を推定した。推定した遺伝子系図を利用し、集団有効サイズの変化と集団の分離時期を明らかにし、ユーラシア人とネアンデルタール人・デニソワ人とが浸透交雑していた様子を再現した (図2)。

さらに、正の自然選択を突然変異の拡散速度に基づいて定量化した。この統計量は、Tajima's D や iHS より検出力が高く、ヨーロッパ人の乳糖耐性の原因の一つとされてきた LCT 遺伝子の突然変異の正の自然選択など、様々なシグ

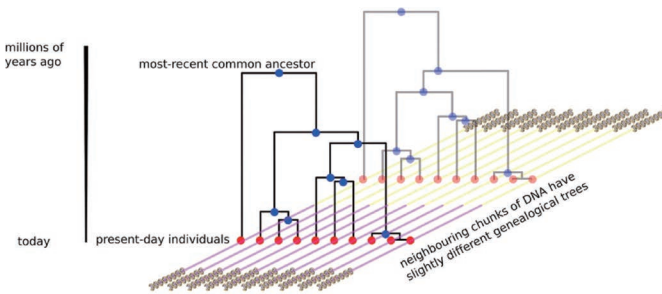
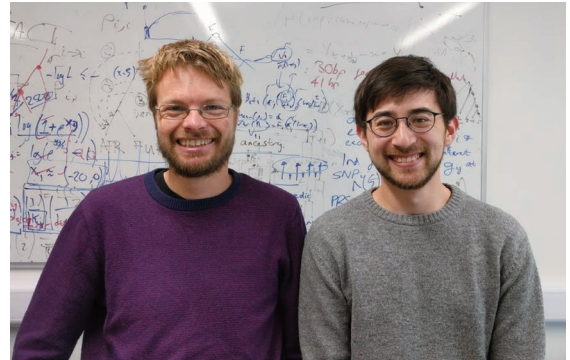


図1 遺伝子系図の模式図。ゲノムの特定なポジションにおける遺伝子系図は二分木構造によって表すことができ、現代のDNAサンプルを最も近い共通祖先に繋げている。この二分木構造は、ゲノムを辿っていくと、遺伝的組み換えにより変化していく。また、二分木構造に突然変異を置くことにより、現代人の遺伝的変異が生成される。



Simon Myers

シュパイデル玲雄

ナルを再現・明らかにした。加えて、多数の遺伝子によって影響される形質 (polygenic trait) の自然選択を検出することにも応用し、髪の色・背丈・BMIなどの適応を定量化した。

Relateは人類以外の様々な種にも応用可能である。特に、突然変異の年齢の特定や、集団の遺伝的構造の検出に優れている。Relateが今後、集団遺伝学の様々な応用に役立つことを強く願っている。

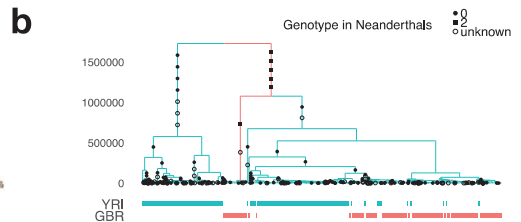
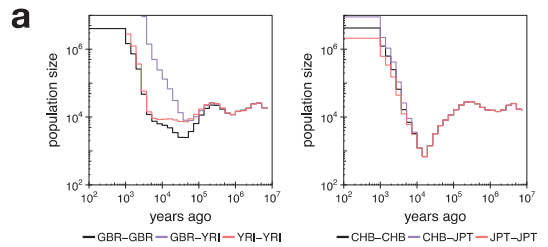


図2

a. 千人ゲノム・プロジェクトにおける、イギリス人 (GBR)、ヨルバ人 (YRI)、日本人 (JPT)、中国人 (CHB) の有効集団サイズと、これらの集団の分離時期の推定値。b. とあるゲノムのポジションにおける、イギリス人 (GBR) とヨルバ人 (YRI) の遺伝子系図。ネアンデルタール人が約80万年前にホモ・サピエンスの祖先から分離し、その後約4万年前にユーラシア人と浸透交雑した仮説と一致している。

1) L. Speidel, M. Forest, S. Shi, S. R. Myers. Nature Genetics 51: 1321-1329, 2019.