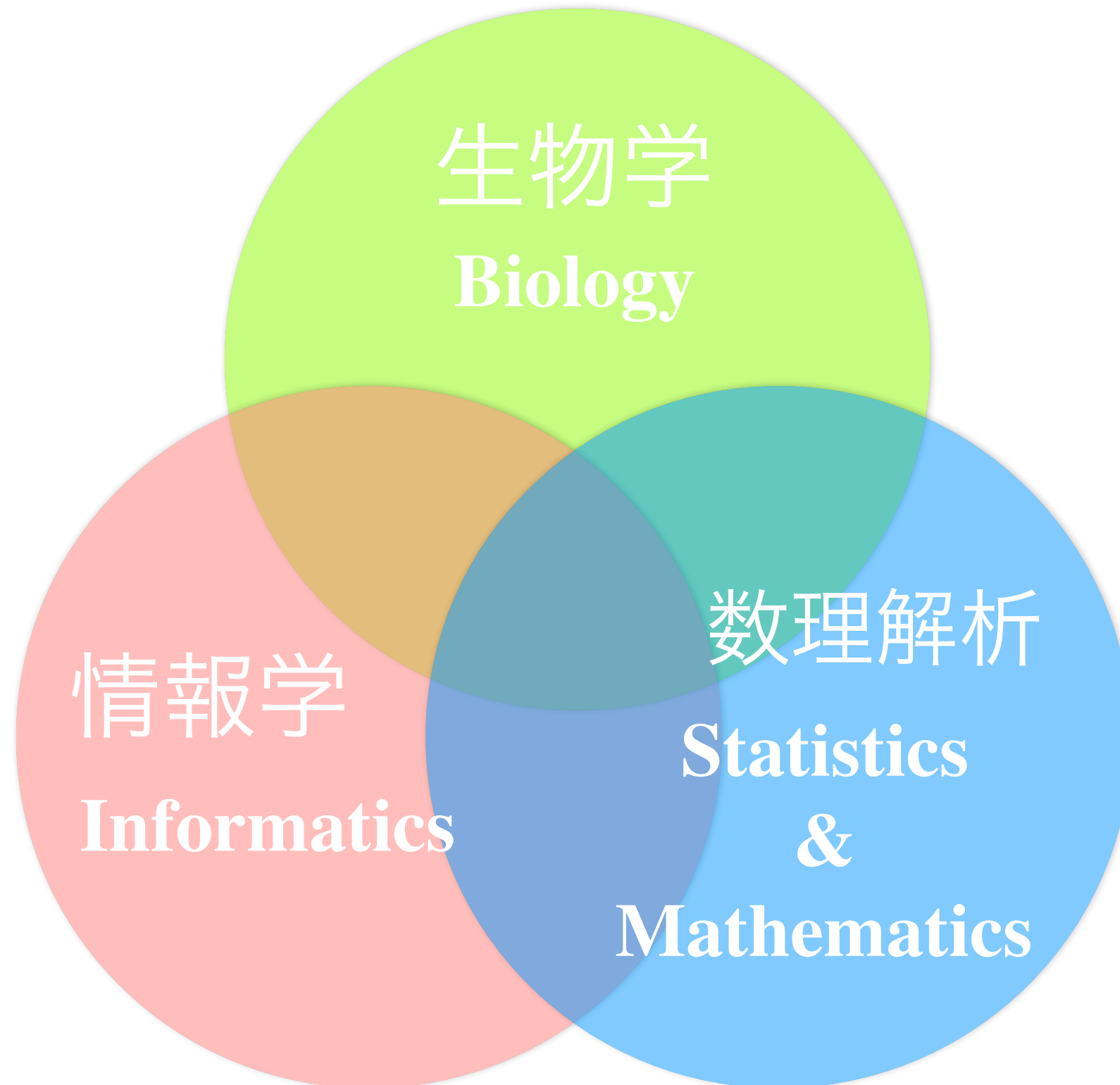


バイオインフォマティクスとは？

What is bioinformatics?



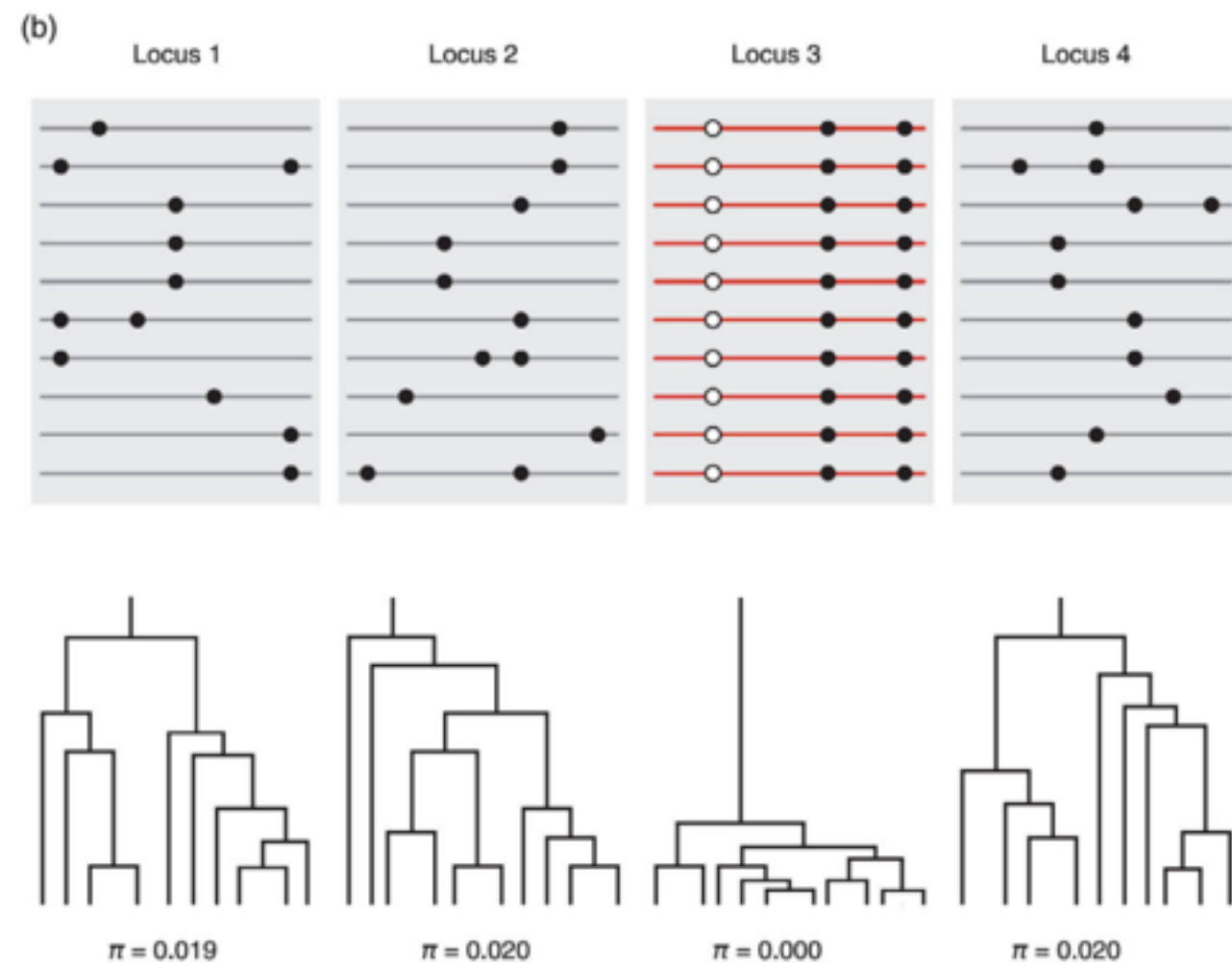
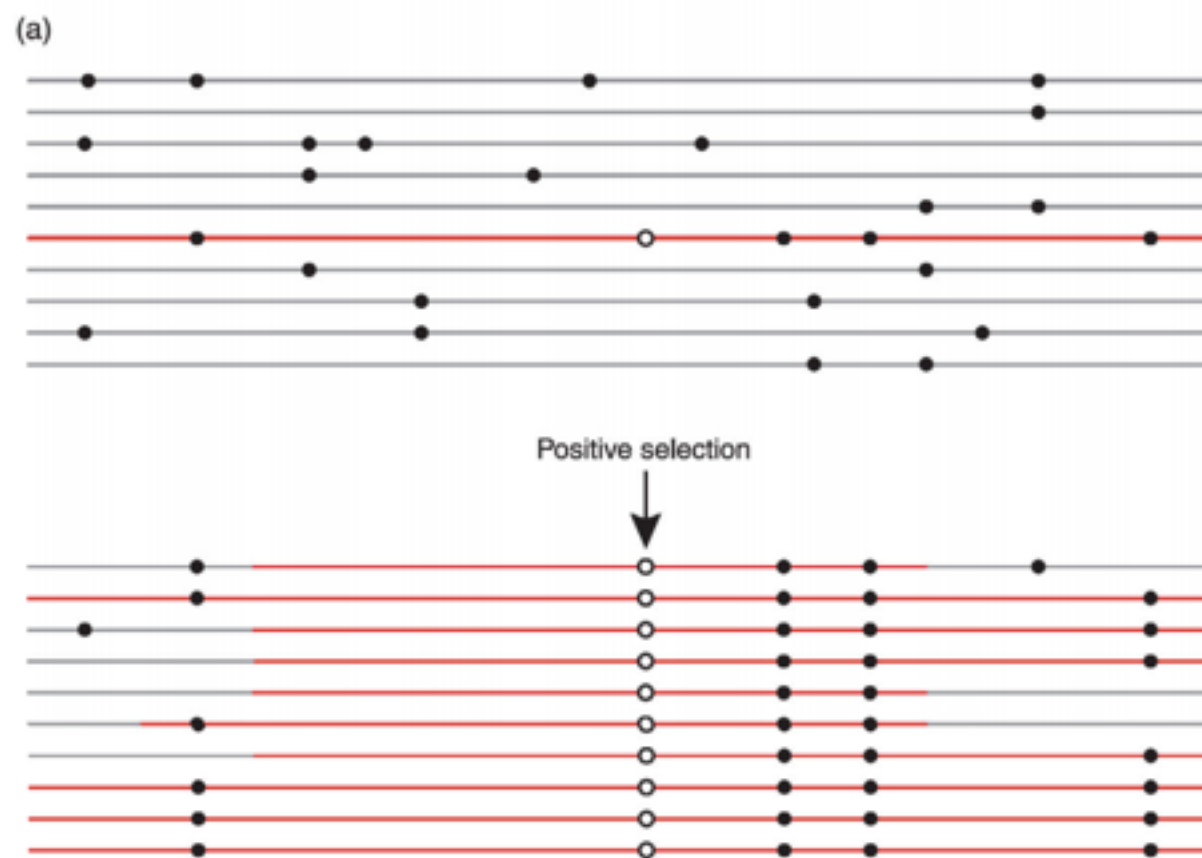
バイオインフォマティクスとは？

What is bioinformatics?

- ・ バイオインフォマティクスとは、生物学、情報学、数理解析の統合分野です。
- ・ 生物学がマーケティング情報に変わると、最近流行っているデータサイエンスになります。
- ・ 数理解析は様々な方法がありますが、これから3回ほど、集団遺伝学を用いたゲノム解析について勉強します。

選択的一掃とは？

What is selective sweep?



Jay F Storz (2005). Molecular Ecology

INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence

選択的一掃とは？

What is selective sweep?

(a) Effects of genetic hitch-hiking along a recombining chromosome. Horizontal lines depict a population sample of homologous chromosomes, and filled symbols depict neutral mutations. In this example an advantageous mutation (open symbol) arises and is rapidly driven to fixation by positive selection. Although the mutation is recombined against new genetic backgrounds during the course of the selective sweep, a sizable fraction of the ancestral haplotype (shown in red) also becomes fixed. Consequently, neutral variants that were initially linked to the advantageous mutation undergo a dramatic increase in frequency as a result of hitch-hiking. (b) In this example, locus 3 has been rendered monomorphic by a selective sweep. Sampled gene copies (denoted by tips of the gene tree) share a very recent common ancestor, and $\pi = 0$ (where π = nucleotide diversity; Nei & Li 1979). By comparison, unlinked, neutrally evolving regions of the genome (loci 1, 2, and 4) are characterized by deeper genealogies, and higher levels of nucleotide diversity ($\pi = 0.019$ – 0.020). Note that the gene trees depict the true genealogies of the samples, not the genealogies inferred from observed variation.

Jay F Storz (2005). Molecular Ecology

INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence

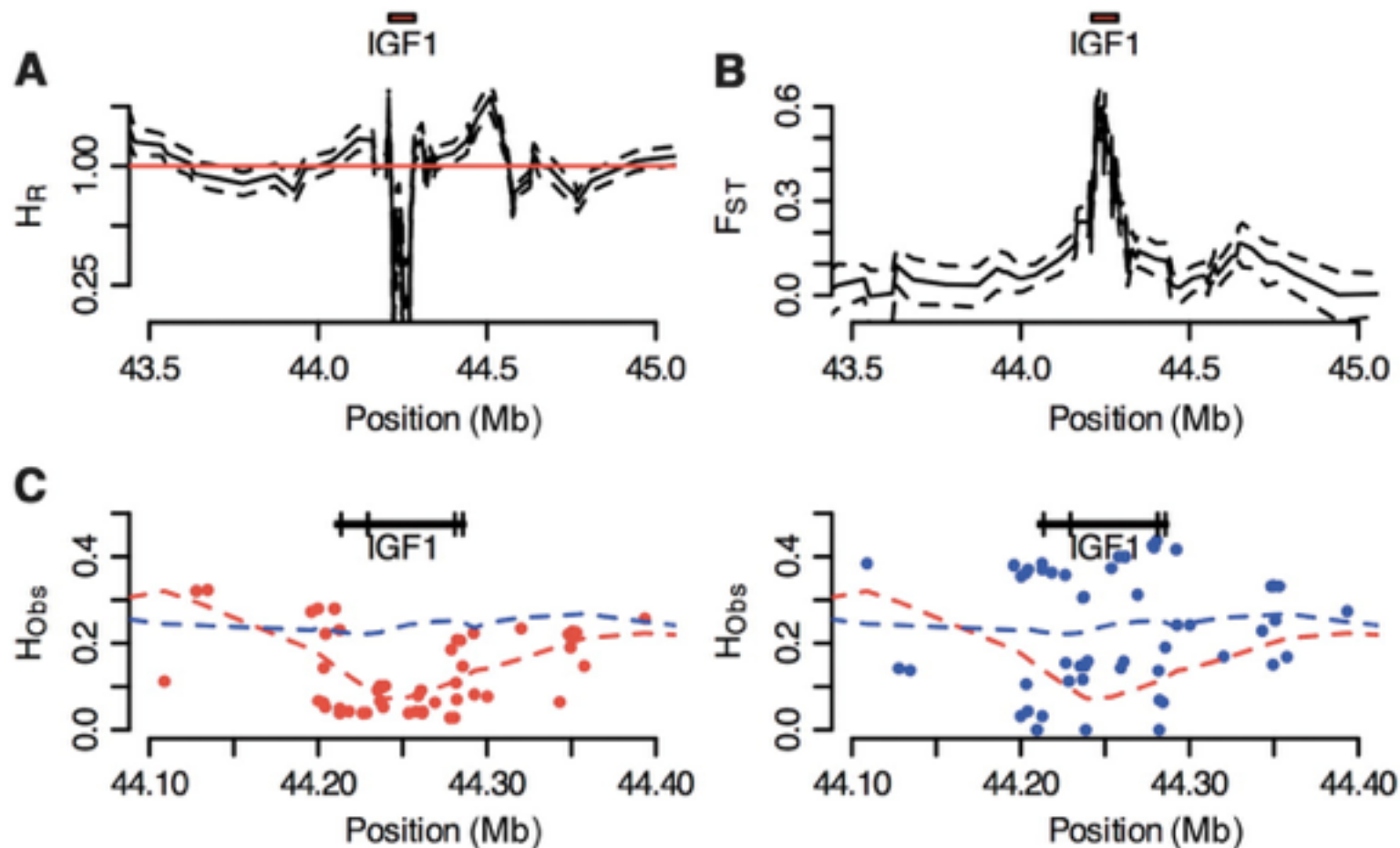
選択的一掃とは？

What is selective sweep?

- 正の自然選択が働いた領域は多様性を失います。そのような領域を英語でselective sweepと言います。
- 多様性が極端に低い領域を探すことで、自然選択の影響を受けた領域を特定できます。それをゲノムスキャンと呼び、仮説を検証するだけでなく、仮説を立てることができます。
- π とは多様性を表すための指標で、今回は、皆さんに π を計算するプログラムを作ってもらいます。

集団遺伝学によるゲノム解析の例

The example of genome analysis with population genetics



Nathan B. Sutter et al. (2007). science

A Single *IGF1* Allele Is a Major Determinant of Small Size in Dogs

集団遺伝学によるゲノム解析の例

The example of genome analysis with population genetics

Fig. 2. Signatures of recent selection on the IGF1 locus across 22 small and giant dog breeds. (A) Heterozygosity ratio (HR) for small versus giant dogs. (B) Genetic differentiation (F_{ST}) for small versus giant dogs. For both (A) and (B), a sliding 10-SNP window across IGF1 was used. Dashed lines delimit the 95% confidence intervals based on nonparametric bootstrap resampling. The IGF1 gene interval is indicated above the graphs as a red box drawn to scale. (C) Observed heterozygosity (H_{Obs}) of SNPs near IGF1 typed in small breeds (<9 kg) and giant breeds (>30 kg). Small breeds have a reduction in observed heterozygosity compared with that of giant breeds. Red and blue points are average observed heterozygosity in small and giant breeds, respectively. Dashed lines are locally weighted scatterplot smoothing (LOWESS) best fit to the data. The IGF1 gene is shown as a black bar with exons indicated by vertical lines.

Nathan B. Sutter et al. (2007). science

A Single *IGF1* Allele Is a Major Determinant of Small Size in Dogs

集団遺伝学によるゲノム解析の例

The example of genome analysis with population genetics

- *IGF1*が犬の体のサイズに関わる遺伝子であるということがゲノムスキャンによってわかりました。
- 小さい犬の*IGF1*は多様性を失っています。一方で、大きい犬の*IGF1*は多様性を維持しています。これは小さい犬の*IGF1*が何らかの自然選択を受けたことを示唆しています。
- F_{ST} とは集団間の遺伝的距離（今回では、赤色と青色の線がいかに離れているのか）を表すための指標で、今回は、皆さんに π と同様、 F_{ST} を計算するプログラムを作ってもらいます。
- 起源種（今回では、大きな犬）を知っているからこそ、多様性を”失った”と分かります。すなわち、起源種は研究する上で非常に重要です。
- ゲノム解析は生き物を選びません。これから皆さんがどの生物の分野行こうとも、役に立つと思います。