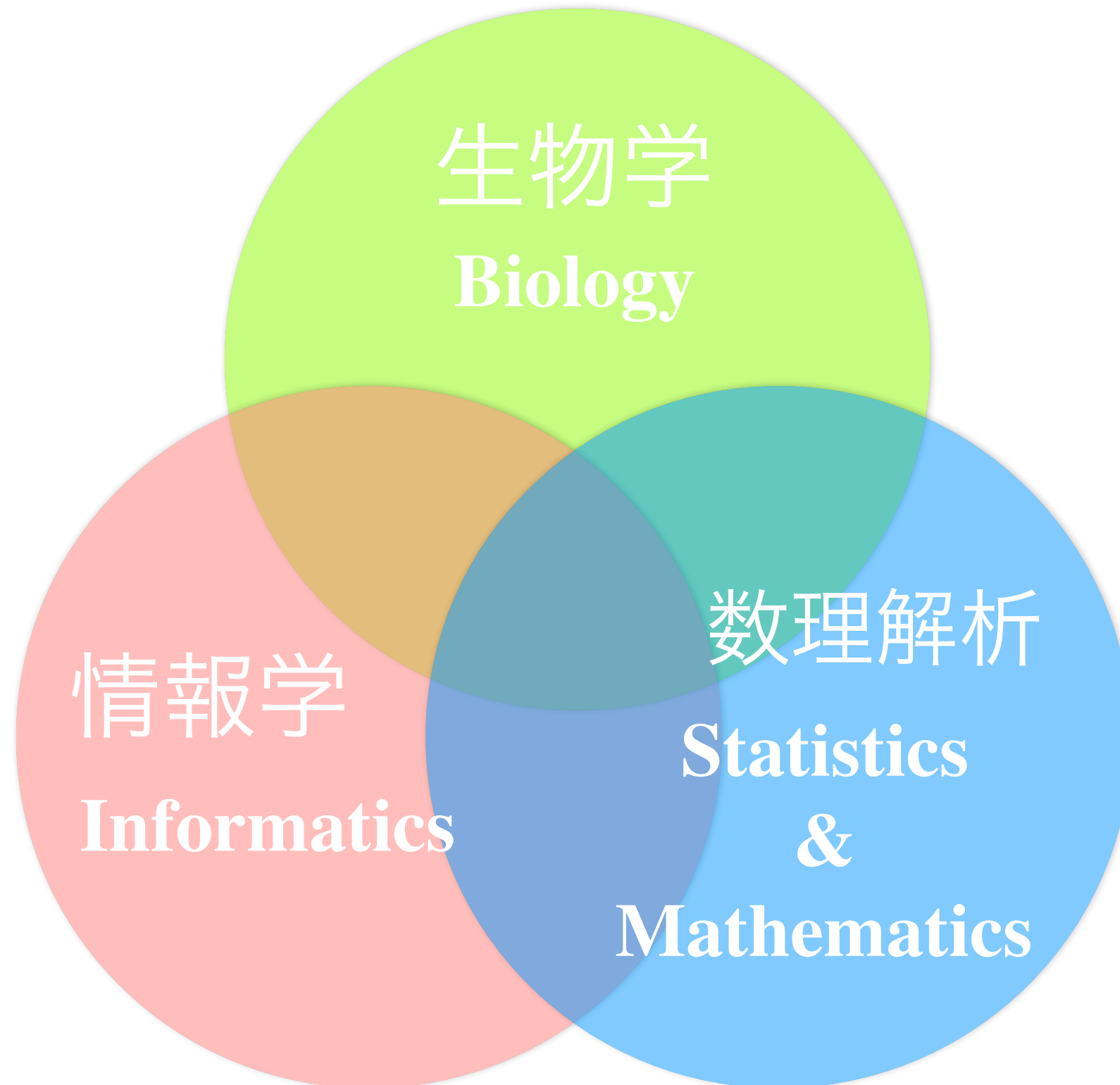


バイオインフォマティクスとは？

What is bioinformatics?



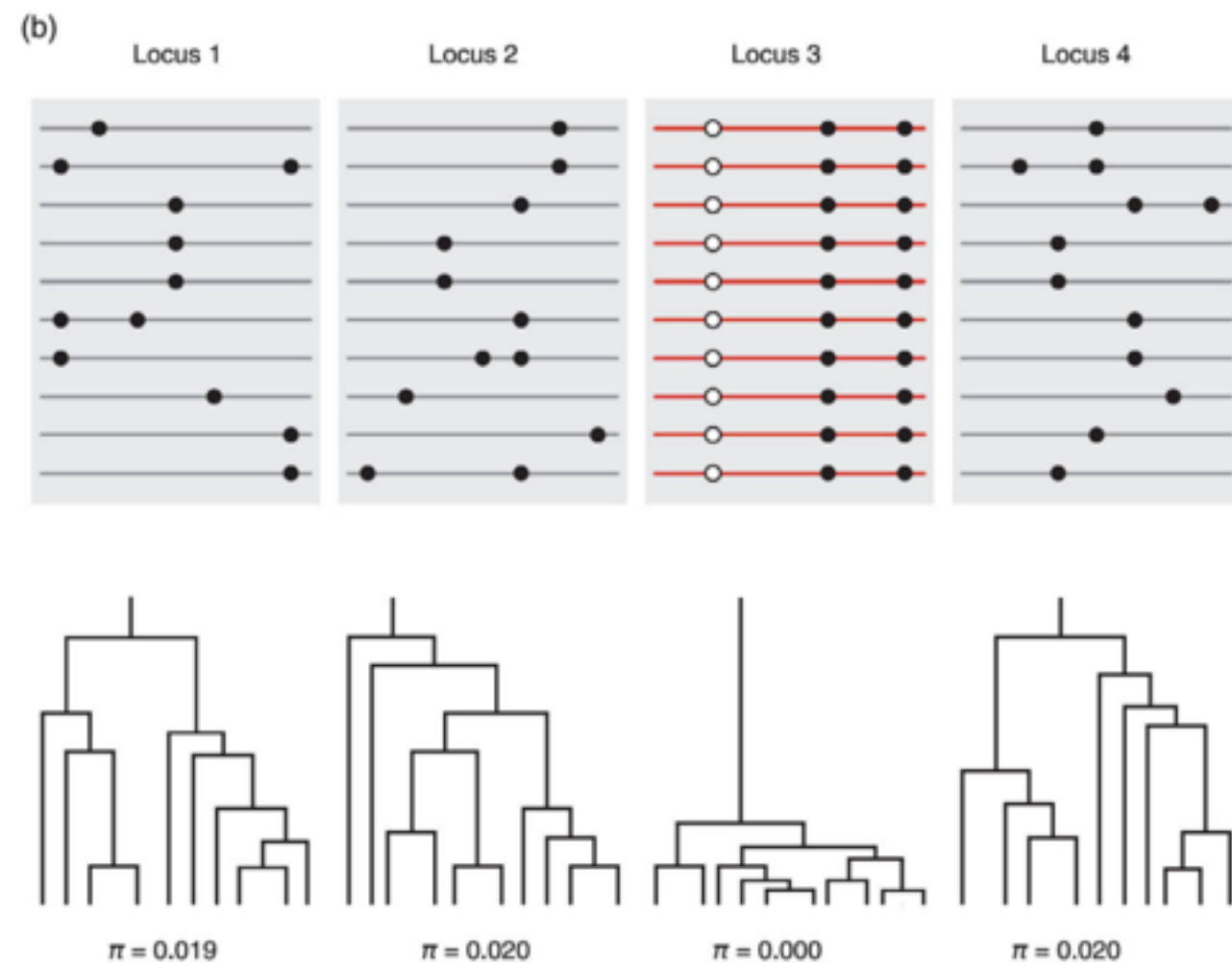
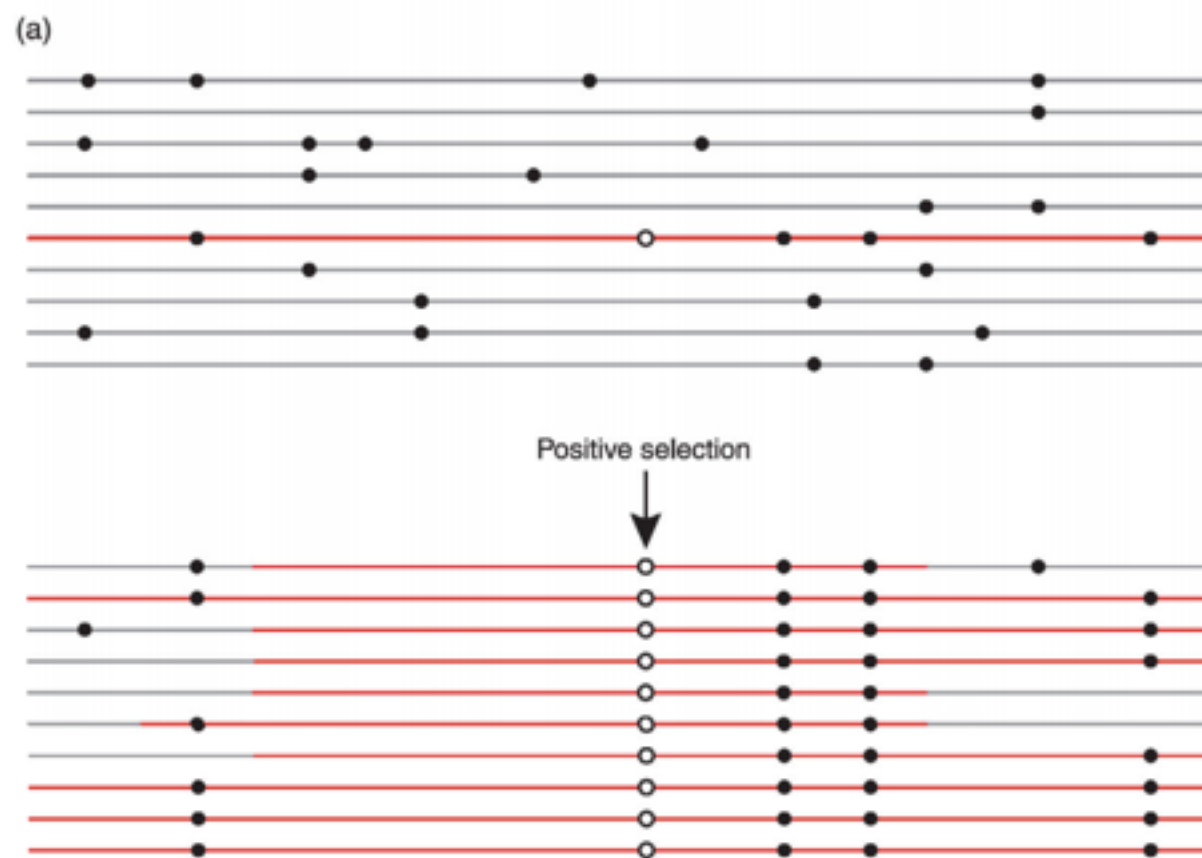
バイオインフォマティクスとは？

What is bioinformatics?

- Bioinformatics is a interdisciplinary field which combines biology, informatics, statistics.
- If you analyze commercial information, it is so-called “data science”. (so they are very close to each other.)
- There are so many analysis for bioinformatics. In this section, we will use population genetics.

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

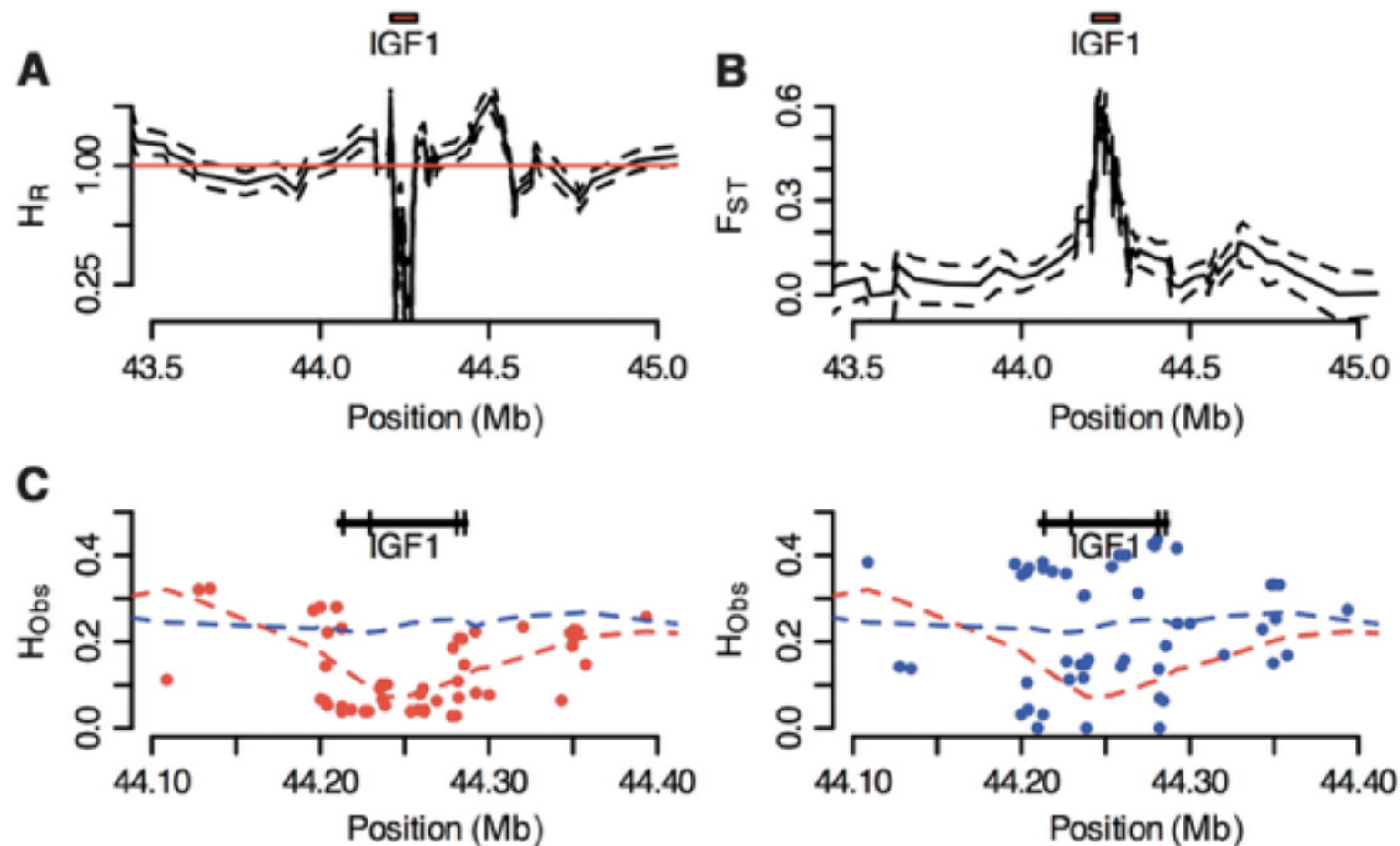
選択的一掃とは？

What is selective sweep?

- The locus which is affected by positive selection loss their diversity. That locus is referred to as “selective sweep”.
- If you find the locus which has extremely low diversity, maybe, that locus affected by the natural selection. That method to detect the natural selection is referred to as “genome-scan”. Genome-scan not only proof a hypothesis, but also make a hypothesis.
- π is the measure for genetic diversity. In this section, you will learn about how to make a program to calculate π .

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. 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* is related to the size of dog, which is found by genome-scan in dog.
- In small dog, *IGF1* lost their diversity. On the other hand, in the big (or normal) dog, *IGF1* maintain their diversity, which shows the evidence for the natural selection.
- F_{ST} is the measure for genetic distance between populations. In this section, you will learn about how to make a program to calculate F_{ST} as well as π .
- It is important for genome-scan to know the origin of your material; for example, big dog in this case. If you don't know the origin, you cannot distinguish “they lost diversity” from “initially, they don't have diversity”.
- You can analyze any material (crop, wild plant, livestock and insect etc..) by genome analysis!