

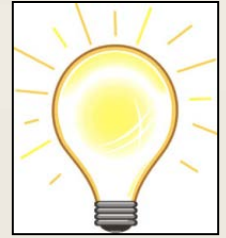
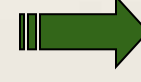
Machine Learning/Stats for Biology

GWAS:



+

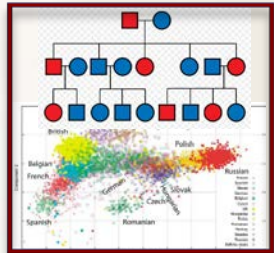
```
ATGTTGAATCTG
AAAGTGAAATGT
TATTATACGAAG
AAGTATTTGCTA
GACCTCAAACC
CTTCATCATAAC
```



DNA from set of individuals

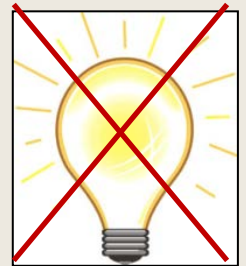
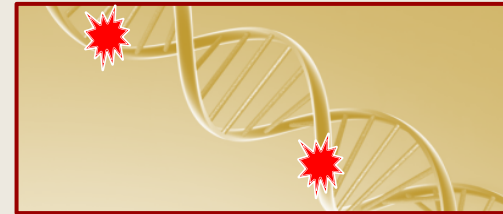
biological & clinical insights

Problem:



+

```
ATGTTGAATCTG
AAAGTGAAATGT
TATTATACGAAG
AAGTATTTGCTA
GACCTCAAACC
CTTCATCATAAC
```



related individuals and/or ethnic diversity

Solution:

- linear mixed models (LMM)
- **FaST-LMM** :
 - Algebraic tricks to make it **FAST (cubic to linear)**.
 - Statistical insights to give **MORE POWER**.

Other statistical problems:

- Time series alignment.
- Imputing immunological data.
- Assigning latent variables to understand experiments.