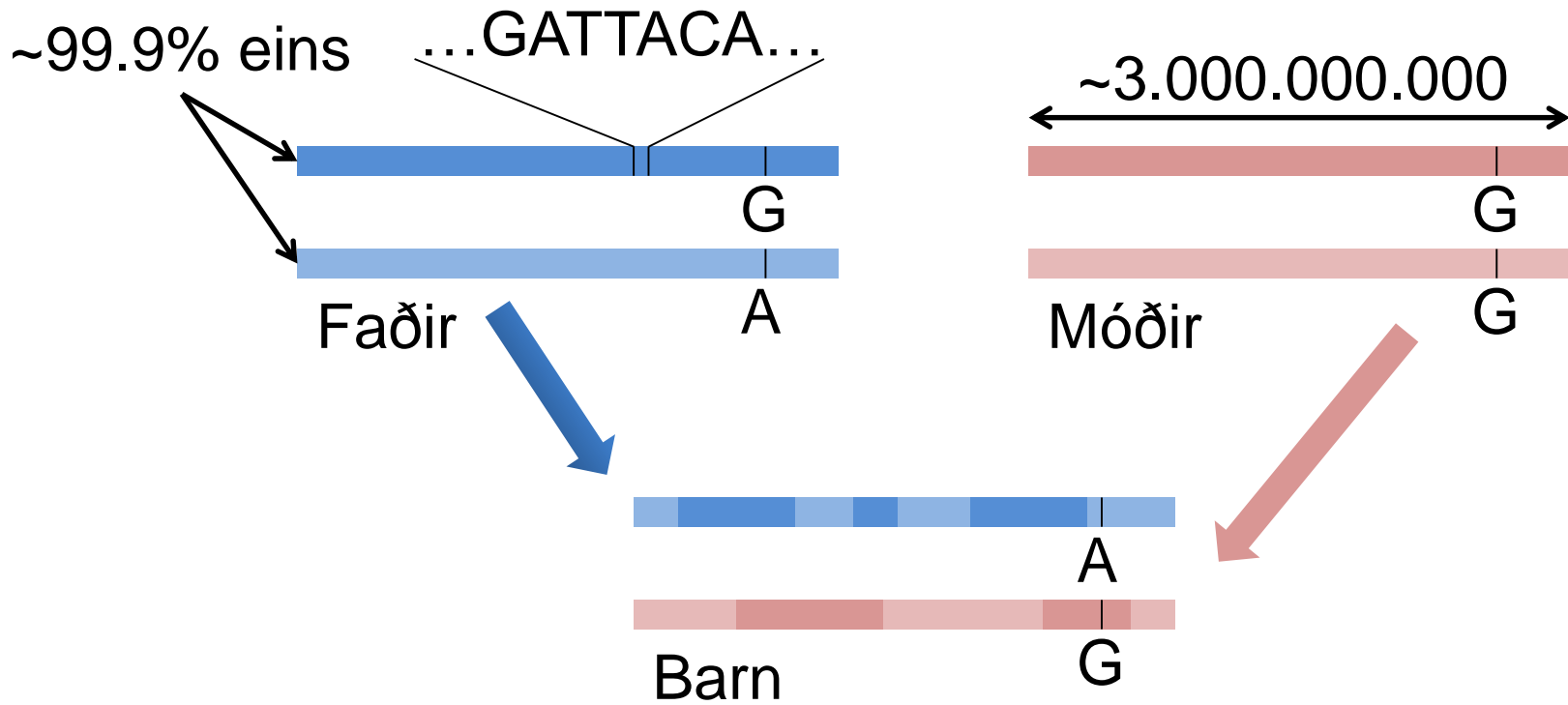




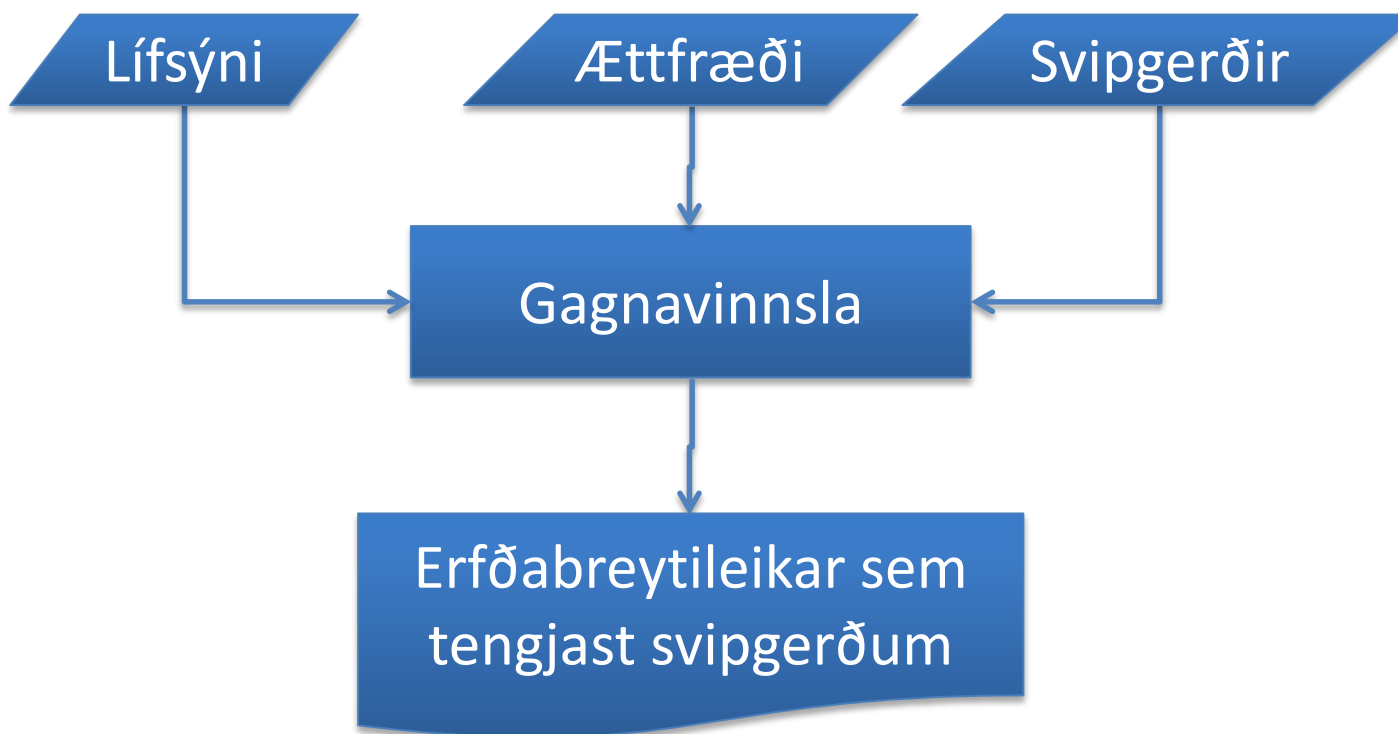
Erfðarannsóknir með stórvirkum vinnuvélum  
- gögnum breytt í uppgötvanir

Gísli Másson

# Örstutt um erfðafræði



# Verkefni í stórum dráttum



# Yfirlit gagnavinnslu



## Stórvirk vinnsla

- **Vinnsla fer öll fram miðlægt**
  - Skiptist í frumvinnslu og eftirvinnslu
- **Vinnslulína þróuð innanhúss**
  - Sérhæfð fyrir okkar þarfir
- **Rekjanleiki er grundvallarkrafa**
  - Allur hugbúnaður undir útgáfustjórn
  - Sama gildir um öll gögn

# Frumvinnsla

- **Keyrir á öllum hrágögnum**
- **Hrágögn af mælitækjum: ~5TB á dag**
- **Frumvinnsla: ~4 kjarna-klst/GB**
  - ~20.000 kjarna-klst á dag
  - ~1.000 kjarnar í stöðugri notkun
- **Erfitt að breyta eftirá**
  - Breytingar kalla á að allt sé keyrt aftur
  - Annars er hættu á kerfisbundinni skekkju

## Eftirvinnsla

- **Nýjar aðferðir bætast stöðugt við**
  - Aðferðir þróaðar af vísindamönnum
  - Sérfræðingar aðlaga að gagnavinnsulínu
- **Ný gögn bætast stöðugt við**
  - Allar aðferðir þurfa að vaka yfir viðbótum
  - Nýjar aðferðir þurfa að vinna öll gömul gögn

# Áskoranir

- **Mikið gagnamagn**
  - Vanda þarf val á því sem er geymt
- **Mikið gagnaflæði**
  - Flutningsgeta er takmörkuð auðlind
- **Miklir reikningar**
  - Forgangsröðun er nauðsynleg
- **Mikill vélbúnaður**
  - Bilanir mega ekki trufla gagnavinnslu