Aim 1

- Build and calibrate an integrated **pipeline** of tools for discovering complex SVs based on a 'pilot' set of samples obtained from the **sequencing centers**
- Discover and validate SVs in a large (comprising 20K individuals) sample set yielding a **reference SV database** for use in the research community

Aim 2

- Develop the **functional annotation pipeline (SVIM)**:
- Prioritize SVs with respect to their impact to determine **high-impact variants**

Aim 3

- **Genotype** the reference set of variants in the ~200K full-cohort set of individuals
- Develop appropriate statistical framework and perform **genome-wide association studies**