# (1) Genotype-Tissue Expression (GTEx)

- Largest systematic study of genetic regulation in multiple tissues to date
  - 53 tissues, 500+ donors, 9K samples, 180M gene RNA-Seq, 14M eQTL, 2M+ gene interactions, 30K+ gene annotations
- Data
  - Donor
    - Age, Gender, Death Type
    - Tissue
      - RNA-Seq expression values
  - Gene
    - Tissue
      - Genetic SNP associations
    - Interaction partners
    - Biological processes, Molecular functions, Diseases
  - Tissue Hierarchy
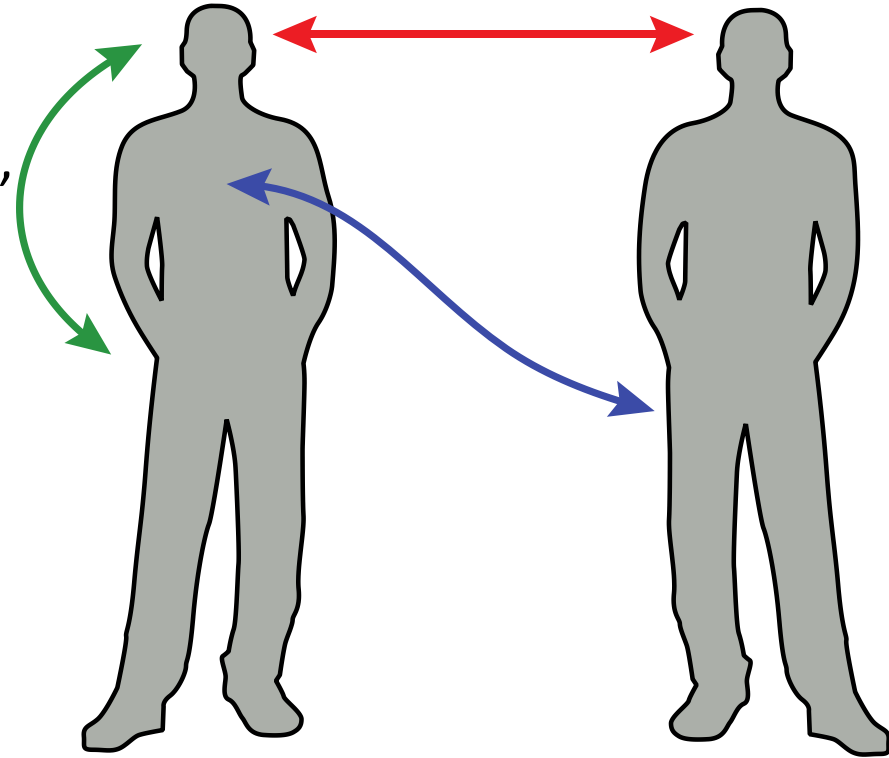
■ Relations between tissues and within donors
■ Relations between donors and within tissues
■ Relations between donors and between tissues

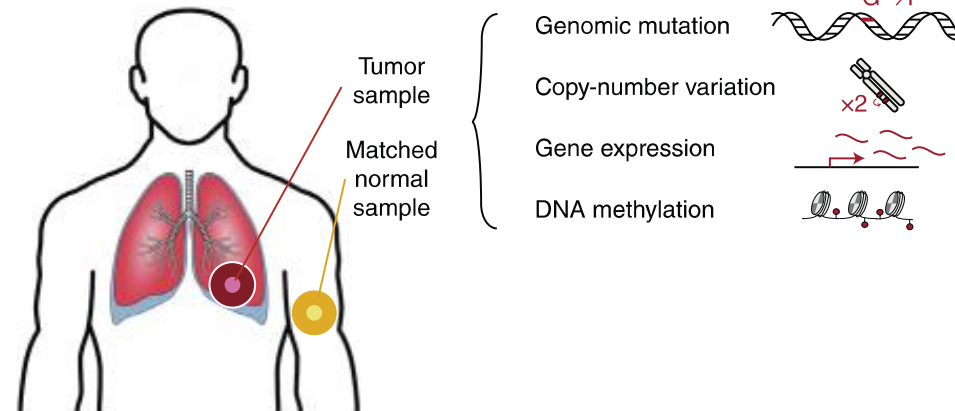Mele et al *Science* 2015, Ardi et al *Science* 2015

(1) GTEx: Personal transcriptomics

- Many key challenges in medicine involve tissue specificity

- Generic tasks:
  - Find top 100 active interaction modules in a tissue
  - Find informative tissue features

- Specialized tasks:
  - Predict tissue-specific processes
  - Predict gene-disease associations
  - Predict cross-tissue interactions
  - Predict tissue-specific interactions

Mele et al *Science* 2015
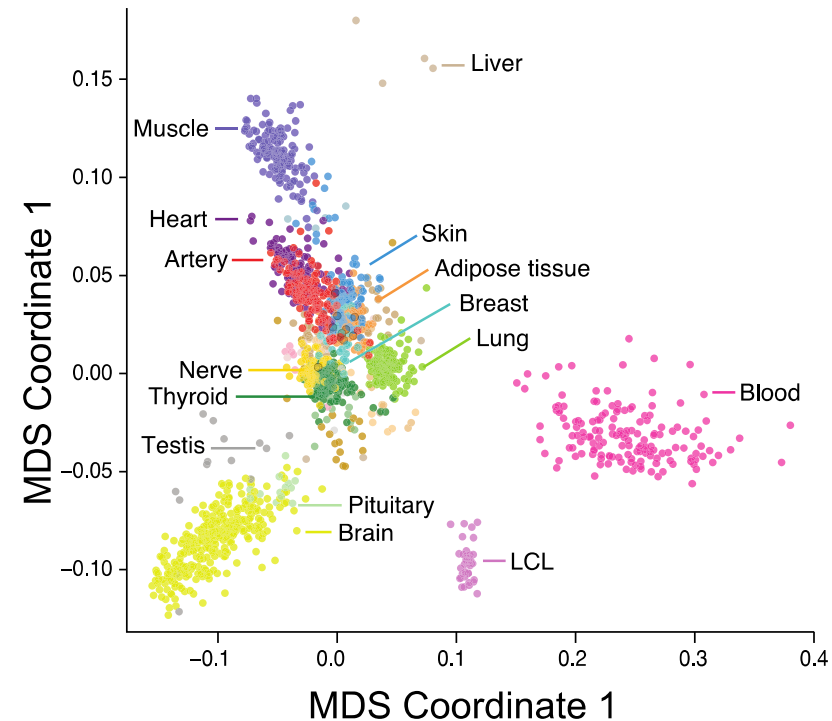
(2) Cancer Genome Consortium (ICGC)

- International data collection of cancer data:
  - 60+ cancer types, 18K+ donors, 37M+ mutations

- Data
  - Cancer type
    - Donor
      - Tumor stage, Vital status, Disease status, Age, Relapse, Gender
      - Healthy samples: Microarray expressions, RNA-seq expressions, Mutations, Methylation, Copy-number variation
      - Tumor samples: Microarray expressions, RNA-seq expressions, Mutations, Methylation, Copy-number variation
  - Gene-gene interactions

Tumor sample

Matched normal sample

Genomic mutation   G→T

Copy-number variation   ×2

Gene expression

DNA methylation

Lawrence et al *Nature* 2013, Creixell et al. *Nature Methods* 2015

(2) Cancer Genome Consortium (ICGC)

- Goal:
  - Integrate multiple layers of information across all cancers
- Tasks:
  - Predict patients' survival time
  - Identify cancer subtypes with clinically relevant differences
  - Identify abrupt changes between tumor stages
    progression -> remission -> relapse
- Relevant papers:
  - *Network-based stratification of tumor mutations*, M. Hofree et al., Nature Methods 2013
  - *Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes*, MDM Leiserson et al., Nature Genetics 2015
  - *Pathway and network analysis of cancer genomes*, P. Creixell et al., Nature Methods 2015