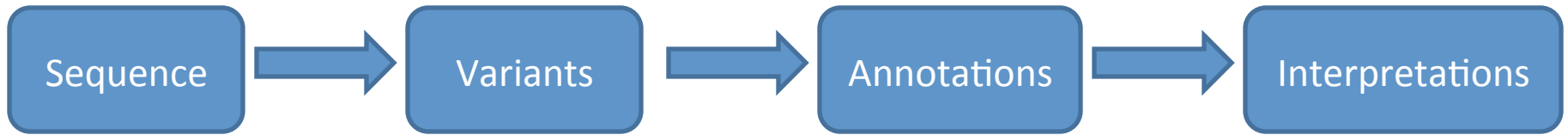


# Module Review

MED676 Bioinformatics Module

Dennis Wang

# Main Take Aways



**Genomic features**

**Knowledge**

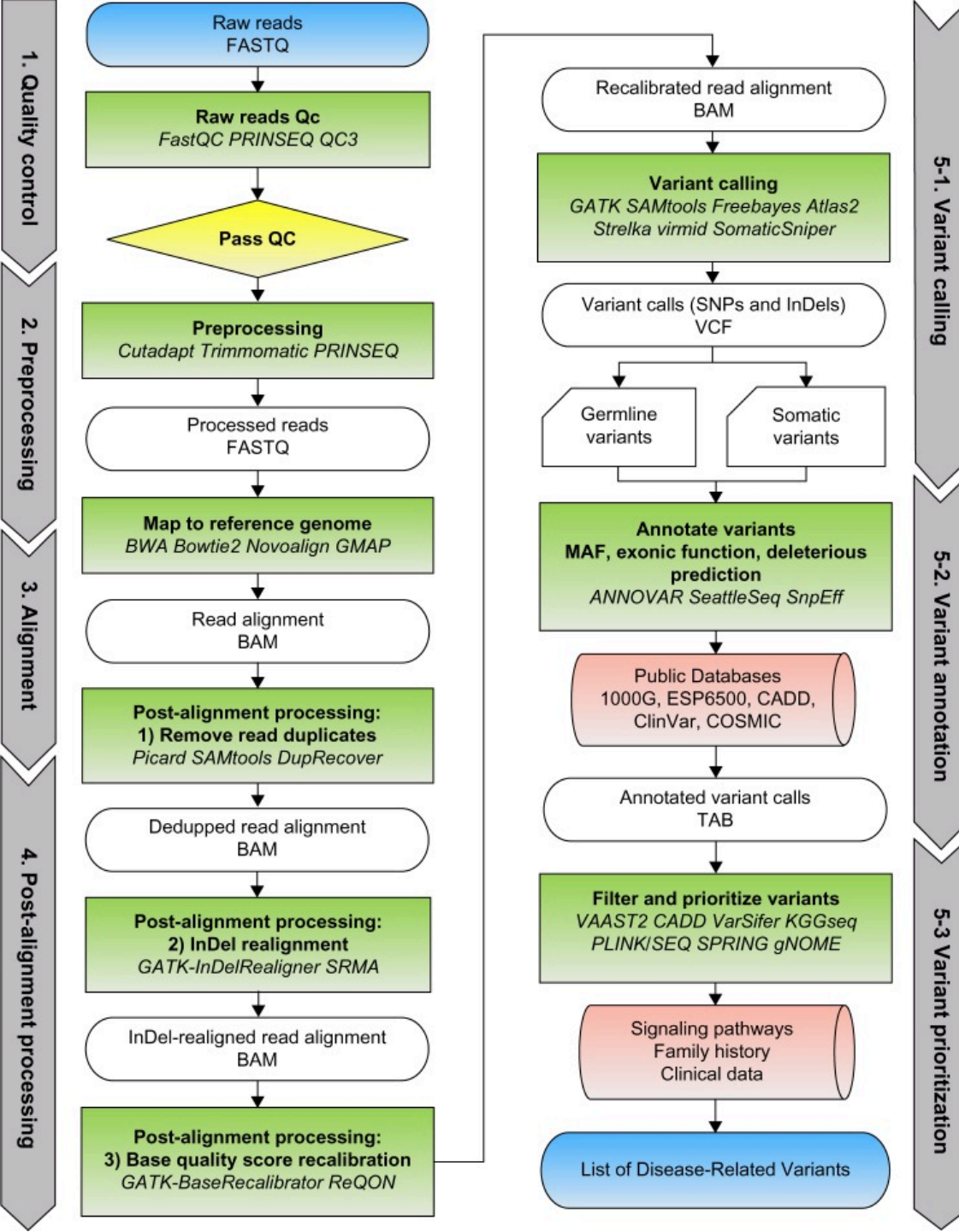
**Automation**

**Files & Data types**

# Where to start on a problem?

## **Data driven approach:**

1. Define a specific disease area and population.
2. What datasets look interesting and may be useful.
3. Propose a filtering or pattern recognition strategy.



**Complete example from:**  
[Cancer Inform. 2014; 13\(Suppl 2\): 67–82.](#)

Do not just list a bunch of tools in your proposal! Need descriptions and justifications.

**Also:**

ExAC, VEP, dbSNP,  
GGV browser,  
Polyphen2, PROVEAN


OMIM, HGMD,  
SNOMED, HPO,  
PhenIX, GO

STRING,  
PathwayCommons,  
DAVID, mSIGdb,  
SFARI

# Where to start on a problem?

## **Hypothesis driven approach:**

1. Define success (impact) for your proposal and how it is measured.
2. What do you expect the result (hypothesis) to be?
3. Identify what you need to support the hypothesis

[illegible]

A diagram illustrating the concept of a hypothesis. It features a central yellow oval with the word "Hypothesis" written inside. Four red arrows point towards this central oval from the top-left, top, top-right, and bottom-right directions, symbolizing multiple lines of evidence or different perspectives converging on a single hypothesis.

For each resource/tool, think about their advantages and **limitations** (eg. EPRs, protein positions)