Mark Pinese

Contents

List of Figures

List of Tables

Acknowledgements

Abstract

Abbreviations

CPV	Clinico-pathological variable

Software versions

Unless otherwise specified, the following versions of software were used in all work.

bamtools	2.2.2
bedtools	2.18.2
cd-hit	4.6.1 TODO plus patch
FastQC	0.10.1
GATK	3.1-1
julia	0.3.2
mutect	1.1.6-4-g69b7a37
ncbi-blast	2.2.29
picard-tools	1.109
provean	1.1.5
Python	2.7.8 / 3.4.1
R	3.1.1
randomForest	4.6-10
samtools	1.0
SHRiMP	2.2.3
strelka	1.0.14
tabix	1.0
vcftools	0.1.10
VEP	76

Chapter 1

Introduction

Chapter 2

Prognostic Signatures

Outline ideas:

- Introduction / overview:
  - Overall thesis for this work:
    Sub-theses:
    *
- Methods
  1.
- Results
  1.
- Conclusion

Chapter 3

Comparative genomics

Outline ideas:

- Introduction / overview:
  - The use of models in PC (very brief)
  - Specific models used in PC, with strong focus on the most common (KPC), and derivates. Cover ease-of-use briefly.
  - Current knowledge re: how appropriate the models are. Consider histology, genetic features, disease progress (incl. metastatic potential), response to therapy. Highlight gap in genetic information, and relevance to response to therapy.
  - Brief overview of known genetic features of human disease. Raise possibility of subtypes.
  - Wrap-up with overview of project:
    1. Collect matched tumour-normal DNA from a range of GEMMs.
    2. Sequence and determine conserved model-specific and general patterns of somatic mutation.
    3. Compare observed patterns to human disease.
       * Are genetic features of human disease recapitulated generally in the models?
       * Does a single model match the genetic features of human disease much better than the others?
       * Do specific models serve as simulations of certain subtypes of human disease?
  - Overall thesis for this work:
    Matching patterns of genetic alterations in mouse models of pancreatic cancer to those seen in human disease can inform researchers as to which models are generally best, and which best match specific patient types.
    Sub-theses:
    * The patterns of mutations seen in common mouse models of pancreatic cancer match those consistently seen in human disease.

    * Different mouse models possess different mutation spectra, and models may be close fits to specific genetic subtypes of patients.

- Methods
    1. Models
    2. Sample origin and processing
    3. Sequencing
    4. QC, mapping, realignment, BQSR
    5. Somatic SNV and indel detection and analysis (pathway)
    6. CNV and LOH detection
- Results
    1. Somatic SNV and indels
    2. CNV and LOH
- Conclusion

Chapter 4

Conclusion