Exercise III

Harry Freitas da Cruz
Data Management for Digital Health
Summer 2017
Exercise III

- Genome Browser
- Medical Knowledge Cockpit
- Use Case Heart Failure (Systems Medicine)
Exercise III Part 1
Genome Browser and Medical Knowledge Cockpit

Harry Freitas da Cruz, Milena Kraus
Hasso Plattner Institut
July 18, 2017
Assessment of Genetic Variants and Medical Knowledge Exploration

- **Genome Browser**
  - Interactive exploration of genetic variants
  - Identify SNP/InDels
  - Distinguish between functional and non-functional variants
  - Identify associated diseases harvesting int’l knowledge databases

- **Medical Knowledge Cockpit**
  - Interaction exploration of medical knowledge data sources
  - Patient-centered queries
  - Analysis medical data in real time
  - Links back to primary data sources
### Alignment tasks
- User-configured pipelines
- Parallelization of jobs
- Task scheduling

### List of variants
- Chromosome positions
- Associated diseases
- Affected genes

---

**Exercise III**
Data Management for Digital Health, Summer 2017
20
Test Exercise

Instructions:

These are the instructions for the test.

⚠️ This is an ungraded exercise. 🕳️ 0.0 points
<table>
<thead>
<tr>
<th>ID</th>
<th>Progress</th>
<th>Pipeline</th>
<th>Results</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>2276</td>
<td>Completed in 07h06m47s</td>
<td>BWA_s</td>
<td>1642578</td>
<td>Free</td>
</tr>
<tr>
<td>2285</td>
<td>Completed in 05h54m31s</td>
<td>BWA_s</td>
<td>327960</td>
<td>Free</td>
</tr>
<tr>
<td>2284</td>
<td>Completed in 02h25m56s</td>
<td>BWA_s</td>
<td>333180</td>
<td>Free</td>
</tr>
<tr>
<td>2283</td>
<td>Completed in 05h22m28s</td>
<td>BWA_s</td>
<td>276012</td>
<td>Free</td>
</tr>
<tr>
<td>2262</td>
<td>Completed in 05h17m58s</td>
<td>BWA_s</td>
<td>274805</td>
<td>Free</td>
</tr>
<tr>
<td>2261</td>
<td>Completed in 14h39m52s</td>
<td>BWA_s</td>
<td>1601151</td>
<td>Free</td>
</tr>
<tr>
<td>2260</td>
<td>Completed in 14h59m17s</td>
<td>BWA_s</td>
<td>1507363</td>
<td>Free</td>
</tr>
<tr>
<td>2259</td>
<td>Completed in 17h34m42s</td>
<td>BWA_s</td>
<td>1377275</td>
<td>Free</td>
</tr>
<tr>
<td>2258</td>
<td>Completed in 16h31m59s</td>
<td>BWA_s</td>
<td>1900591</td>
<td>Free</td>
</tr>
<tr>
<td>2257</td>
<td>Completed in 08h21m43s</td>
<td>BWA_s</td>
<td>1647173</td>
<td>Free</td>
</tr>
<tr>
<td>2256</td>
<td>Completed in 08h12m04s</td>
<td>BWA_s</td>
<td>1688232</td>
<td>Free</td>
</tr>
<tr>
<td>2255</td>
<td>Completed in 20h50m50s</td>
<td>BWA_s</td>
<td>187279</td>
<td>Free</td>
</tr>
<tr>
<td>2250</td>
<td>Completed in 10h34m64s</td>
<td>BWA_s</td>
<td>2044480</td>
<td>Free</td>
</tr>
<tr>
<td>Chr</td>
<td>Pos</td>
<td>Ref</td>
<td>Alt</td>
<td>Qual</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>5</td>
<td>35871190</td>
<td>G</td>
<td>A</td>
<td>82.30</td>
</tr>
<tr>
<td>5</td>
<td>35861068</td>
<td>T</td>
<td>C</td>
<td>7.80</td>
</tr>
<tr>
<td>22</td>
<td>36684354</td>
<td>T</td>
<td>C</td>
<td>6.98</td>
</tr>
<tr>
<td>3</td>
<td>12393125</td>
<td>C</td>
<td>G</td>
<td>7.80</td>
</tr>
<tr>
<td>10</td>
<td>8116241</td>
<td>GAAAAAAA</td>
<td>GAAAAAAA</td>
<td>4.58</td>
</tr>
<tr>
<td>3</td>
<td>12429385</td>
<td>attt</td>
<td>aTtt</td>
<td>11.50</td>
</tr>
<tr>
<td>4</td>
<td>55567888</td>
<td>A</td>
<td>C</td>
<td>7.80</td>
</tr>
<tr>
<td>10</td>
<td>12324057</td>
<td>T</td>
<td>C</td>
<td>4.13</td>
</tr>
<tr>
<td>10</td>
<td>43556521</td>
<td>T</td>
<td>G</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>69874515</td>
<td>A</td>
<td>G</td>
<td>6.20</td>
</tr>
<tr>
<td>14</td>
<td>81457257</td>
<td>A</td>
<td>C</td>
<td>3.54</td>
</tr>
<tr>
<td>3</td>
<td>68945553</td>
<td>A</td>
<td>T</td>
<td>58</td>
</tr>
<tr>
<td>7</td>
<td>140454973</td>
<td>G</td>
<td>T</td>
<td>14.20</td>
</tr>
</tbody>
</table>
Add a Cell Line

Please Choose a Cell Line

Add

Reading frames
MUTATION DETAILS FOR
CHR5 AT 35,861,068

DGV

- Pubmed ID: 15286789
- Short description: Twenty samples were obtained from normal individuals whose DNA are stored at the Coriell Cell Repository (Camden, New Jersey) with the following ethnic representation: 10 European/Caucasian, 4 Native American, 2 Chinese, 2 Indo-Pakistani, and 2 Sub-Saharan African. Twelve samples were from fixed cell pellets of cytogenetically normal individuals from the Brigham and Women's Cytogenetics Laboratory and 7 samples from cytogenetically normal individuals were obtained from the Hospital for Sick Children.

Exercise III
Data Management for Digital Health, Summer 2017
26
<table>
<thead>
<tr>
<th>ID</th>
<th>Progress</th>
<th>Pipeline</th>
<th>Results</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>2276</td>
<td>Completed in 07h06m47s</td>
<td>BWA_s</td>
<td>1642578</td>
<td>Free</td>
</tr>
<tr>
<td>2285</td>
<td>Completed in 05h54m31s</td>
<td>BWA_s</td>
<td>327960</td>
<td>Free</td>
</tr>
<tr>
<td>2284</td>
<td>Completed in 02h25m56s</td>
<td>BWA_s</td>
<td>333180</td>
<td>Free</td>
</tr>
<tr>
<td>2283</td>
<td>Completed in 05h22m28s</td>
<td>BWA_s</td>
<td>276012</td>
<td>Free</td>
</tr>
<tr>
<td>2282</td>
<td>Completed in 05h17m58s</td>
<td>BWA_s</td>
<td>274805</td>
<td>Free</td>
</tr>
<tr>
<td>2281</td>
<td>Completed in 14h39m52s</td>
<td>BWA_s</td>
<td>1601151</td>
<td>Free</td>
</tr>
<tr>
<td>2280</td>
<td>Completed in 14h59m17s</td>
<td>BWA_s</td>
<td>1507363</td>
<td>Free</td>
</tr>
<tr>
<td>2259</td>
<td>Completed in 17h34m42s</td>
<td>BWA_s</td>
<td>1377275</td>
<td>Free</td>
</tr>
<tr>
<td>2258</td>
<td>Completed in 16h31m59s</td>
<td>BWA_s</td>
<td>1900591</td>
<td>Free</td>
</tr>
<tr>
<td>2257</td>
<td>Completed in 08h21m43s</td>
<td>BWA_s</td>
<td>1647173</td>
<td>Free</td>
</tr>
<tr>
<td>2256</td>
<td>Completed in 08h12m04s</td>
<td>BWA_s</td>
<td>1688232</td>
<td>Free</td>
</tr>
<tr>
<td>2255</td>
<td>Completed in 20h50m50s</td>
<td>BWA_s</td>
<td>187279</td>
<td>Free</td>
</tr>
<tr>
<td>2250</td>
<td>Completed in 10h34m54s</td>
<td>BWA_s</td>
<td>2044480</td>
<td>Free</td>
</tr>
</tbody>
</table>
Medical Knowledge Cockpit

Enter search terms

How about A2M, AADAC or PIK3CA?

Search terms
Medical Knowledge Cockpit

Enter search terms
How about A2M, AADAC or PIK3CA?

A2M  ×  AADAC  ×

Cancer Gene Census

Definition

AADAC
Microsomal arylacetamide deacetylase competes against the activity of cytosolic arylamine N-acetyltransferase, which catalyzes one of the initial biotransformation pathways for arylamine and heterocyclic amine carcinogens (provided by RefSeq, Jul 2008).

A2M
Alpha-2-macroglobulin is a protease inhibitor and critical for protein homeostasis.

PubMed Links

AADAC
- Human liver arylacetamide deacetylase. Molecular cloning of a novel esterase involved in the metabolic activation of arylamine carcinogens with high sequence similarity to hormone-sensitive lipase.

Clinical Trials

Internal

External
- Long-Term Extension Study in Subjects With Multiple Sclerosis Who Have Completed Study 205MS01 (NCT01864481) to Evaluate the Safety and Efficacy of DAS-HYD (EXTEND)

Data Management for Digital Health, Summer 2017

Exercise III
29
Medical Knowledge Cockpit
Context-sensitive definitions for search terms

- Enter search terms
  - How about A2M, AADAC or PIK3CA?

- BRAF, lung, cancer

- Jane Dough
  - female, 51 years, non-smoker
  - Markers: KRAS, EGFR, BRAF, NRAS
  - Diagnosis: non-small cell lung cancer, stage IV

- Definition:
  - **BRAF**
    - This gene encodes a protein belonging to the raf/mil family of serine/threonine protein kinases.

- **AADAC**
  - Microsomal arylacetamide deacetylase competes against the activity of cytosolic arylamine N-acetyltransferase, which catalyzes one of the initial biotransformation pathways for arylamine and heterocyclic amine carcinogens (provided by RefSeq, Jul 2008).

- Cancer Gene Census
  - **BRAF**
    - **Somatic tumors**
      - melanoma, colorectal, papillary thyroid, borderline ov, Non small-cell lung cancer (NSCLC), cholangiocarcinoma, pilocytic astrocytoma

**Exercise III**
Data Management for Digital Health, Summer 2017
30
Medical Knowledge Cockpit
Latest research papers and clinical trials

**PubMed Links**
- Complete coding sequence of a human B-raf cDNA and detection of B-raf protein kinase with isozyme specific antibodies
- B-raf and a B-raf pseudogene are located on 7q in man
- Mutations of the BRAF gene in human cancer
- The ins and outs of Raf kinases
- Autoregulation of the Raf-1 serine/threonine kinase

**Clinical Trials**

**Internal**

**External**
- CEP-32496 in Patients With Advanced Solid Tumors in Phase 1 and Advanced Melanoma and Metastatic Colorectal Cancer in Phase 2
- GSK2141795 and Dabrafenib in Treating Patients With Stage IIIIC-IV Cancer
- Selumetinib and Akt Inhibitor MK2206 in Treating Patients With Stage III or Stage IV Melanoma Who Failed Prior Therapy With Vemurafenib or Dabrafenib
- Dabrafenib/Trametinib/Navitoclax in Braf Mutant Melanoma
- Vemurafenib, Cetuximab, and Irinotecan in Advanced Solid Cancers

**ClinicalTrials.gov**

**Exercise III**
Data Management for Digital Health, Summer 2017
31
Coagulation and complement cascades

**Coagulation cascade**
- **Extrinsic pathway**: Tissue damage
  - F3
  - F7
- **Intrinsic pathway**: Contact with damaged vessel
  - F10
  - TPPI
- **Activated thrombin**
  - F12
  - F10
  - TPPI
- **Inflammatory mediator** regulation of TFPI channel
  - Inflammation, vasodilation
- **Platelet activation**
- **Plasmin**

**Complement cascade**
- **Alternative pathway**
  - Microbes
  - C3b, C5b
- **Classical pathway**
  - Antigen-antibody complex
  - C1, C4, C2
- **Lectin pathway**
  - Carbohydrates
  - C1r, C1s, C2, C4

**Platelets**
- Fibronectin
- Cross-linked fibrin polymer
- Fibronectin

**Fibrinolytic system**
- Fibrinolysis
- Platelet secretion
- Metalloproteinase

**Exercise III**
- Data Management for Digital Health, Summer 2017
Exercise III Part 2
Systems Medicine Approach for Heart Failure

Harry Freitas da Cruz, Milena Kraus
Hasso Plattner Institut
July 18, 2017
Systems Medicine Approach for Heart Failure

- Systems medicine concept
- Biomedical and clinical facts regarding heart failure
- RNA sequencing and differential gene expression analysis
- Clustering of mixed-type data

Biomedical and Clinical Facts
Heart Failure


https://www.youtube.com/watch?v=B93TsbJXnMc
RNA Sequencing and Differential Gene Expression Analysis

Clustering of Mixed-type Data


Original index: \(-2, -1, 0, 1, 2\) = \(i\)

Converted to rank: \(1, 2, 3, 4, 5\) = \(r\)

Normalized rank: \(0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\)

\[ x = \frac{r - 1}{R - 1} \]