Using Galaxy: RNA-seq

Stanford University
September 23, 2014

Jennifer Hillman-Jackson
Galaxy Team
Penn State University

http://galaxyproject.org/

Galaxy

amazon web services

The Agenda

Introduction

RNA-seq Example

- Data Prep: QC
  - Create, Edit, and Run a Workflow
  - Sharing Results
- Data Prep: Re-pairing (is nicer)
- Mapping with Tophat2
- Differential Expression using Cuffdiff

Resources

The Agenda

Goal is to demonstrate how Galaxy can help you explore and learn options, perform analysis, and then share, repeat, and reproduce your analyses.

Not The Agenda

This workshop will *not* cover

- details of how tools are implemented, or
- new algorithm designs, or
- which assembler or mapper or peak caller or ... is best for you.

While this workshop does cover RNA-seq, we are only using that specific example to learn how to best use and navigate the Galaxy Application …

to further your own analysis goals.

What is Galaxy?

A free (for everyone) web server

Open source software

These options result in several ways to use Galaxy

http://galaxyproject.org

Galaxy is available ...

As a free (for everyone) web server integrating a wealth of tools, compute resources, terabytes of reference data and permanent storage

http://usegalaxy.org

However, ***a centralized solution cannot support the different analysis needs of the entire world.***

Galaxy is available ...

- As a free (for everyone) web service

  http://usegalaxy.org

- As open source software

  http://getgalaxy.org

It is installed right here at Stanford on your *Galaxy SlipStream Appliance*

SLIPSTREAM APPLIANCE GALAXY EDITION
Galaxy made easy.

And locations around the world…

Galaxy is available ...

- As a free (for everyone) web ser
  http://usegalaxy.org
- As open source software
  http://getgalaxy.org
- *On the Cloud*

We are using this today.

http://aws.amazon.com/education
http://globus.org/

http://wiki.galaxyproject.org/Cloud

http://cloud1.galaxyproject.org/
http://cloud2.galaxyproject.org/
http://cloud3.galaxyproject.org/

Pair up into groups of TWO
Say HELLO
Assign cloud servers EVENLY - *important*!
Account on your cloud: User -> Register

The Agenda

Introduction

RNA-seq Example

- Data Prep: QC

- Create, Edit, and Run a Workflow

- Sharing Results

- Data Prep: Re-pairing (is nicer)

- Mapping with Tophat2

- Differential Expression using Cuffdiff

Resources

Know your DATA

RNA-seq (RNA Sequencing), also called "Whole Transcriptome Shotgun Sequencing" ("WTSS"), is a technology that uses the capabilities of next-generation sequencing to reveal a snapshot of RNA presence and quantity from a genome at a given moment in time.

http://en.wikipedia.org/wiki/RNA-Seq

Good for transcriptome **expression analysis**, just like **EST** sequencing in the 1800s and **Microarrays** 10 years ago.

RNA-Seq Analysis: QC - Get the Data

Create new history

 (cog) → Create New, then name your History

Import:

Shared Data → Data Libraries

→ RNA-Seq UCDavis 2013 Example Data*

→ Unfiltered Reads

→ MeOH_REP1_R1.fastq or
MeOH_REP1_R2.fastq

UCDAVIS Genome Center Bioinformatics Core

* RNA-Seq example datasets from the 2013 UC Davis Bioinformatics Short Course. http://bit.ly/ucdbsc2013

NGS Data Quality Control

- FASTQ format
- Examine quality in an RNA-seq dataset
- Trim/filter the data, many ways, we'll do a few
- Workflow the steps to avoid mistakes/tedium.
- Share your results with everyone
- Import the compliment - want paired R1 & R2

Quality Control is vital.

What is FASTQ?

- Specifies sequence (FASTA) and quality scores
  - » Sanger PHRED+33 is used in Galaxy as .fastqsanger
- Text format, 4 lines per entry

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
+
!''*((((***+))%%%++)(%%%%).1***-+*''))**55CCF>>>>>>CCCCCCC65
```

- FASTQ is such a cool standard, there are 3 (or 5) of them!

```
  SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
  ..............................IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
  ..........................XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  !"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_`abcdefghijklmnopqrstuvwxyz{|}~
  |                          |    |        |                              |                     |
 33                         59   64       73                            104                   126

 S - Sanger       Phred+33,  93 values  (0, 93) (0 to 60 expected in raw reads)
 I - Illumina 1.3 Phred+64,  62 values  (0, 62) (0 to 40 expected in raw reads)
 X - Solexa       Solexa+64, 67 values (-5, 62) (-5 to 40 expected in raw reads)
```

http://en.wikipedia.org/wiki/FASTQ_format

NGS Data Quality: Assessment tools

NGS QC and Manipulation → FastQC

Gives you a lot of information but little control over how it is calculated or presented.

Still, it is almost always what I use, every time

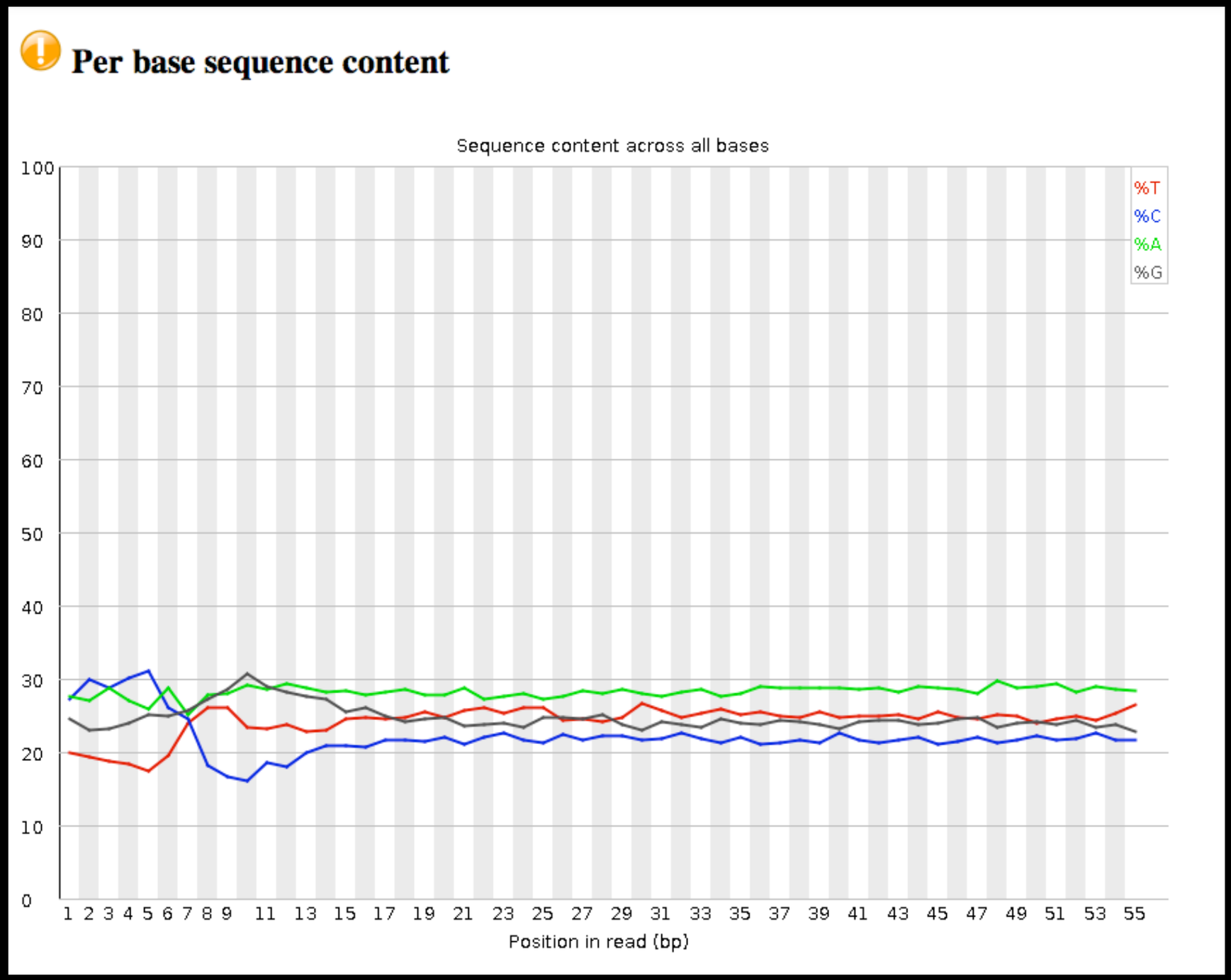
» verify quality score format (wrong == start over)

» baseline statistics before manipulations (if any)

» final statistics to aid with tool parameter choices

http://bit.ly/FastQCBoxPlot

NGS Data Quality: Sequence bias at front of reads?



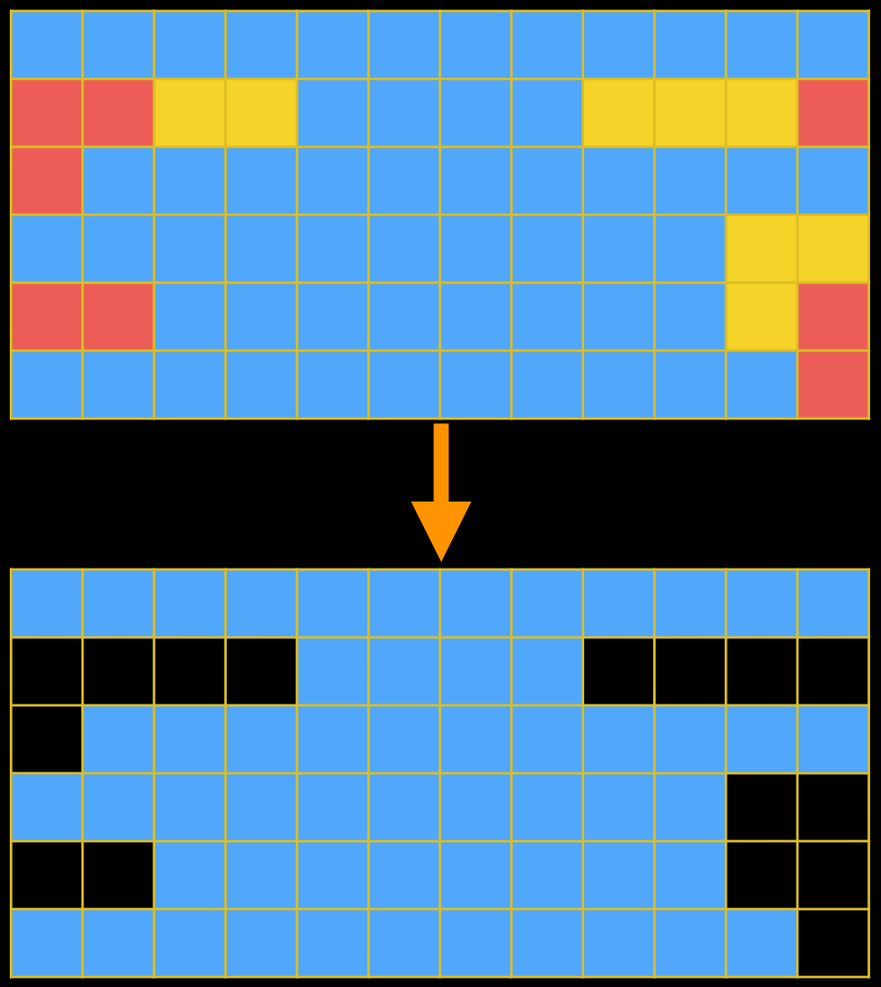
From a sequence specific bias that is caused by use of random hexamers in library preparation.

Hansen, *et al.*, "Biases in Illumina transcriptome sequencing caused by random hexamer priming" *Nucleic Acids Research*, Volume 38, Issue 12 (2010)

Trim? Filter? *How? When?*

- **Choice depends on downstream tools**
- **Find out assumptions & requirements for downstream tools and make appropriate choice(s) now.**
- **How to do that?**
  - **Read the tool documentation**
  - **http://biostars.org/**
  - **http://seqanswers.com/**
  - **http://galaxyproject.org/search**

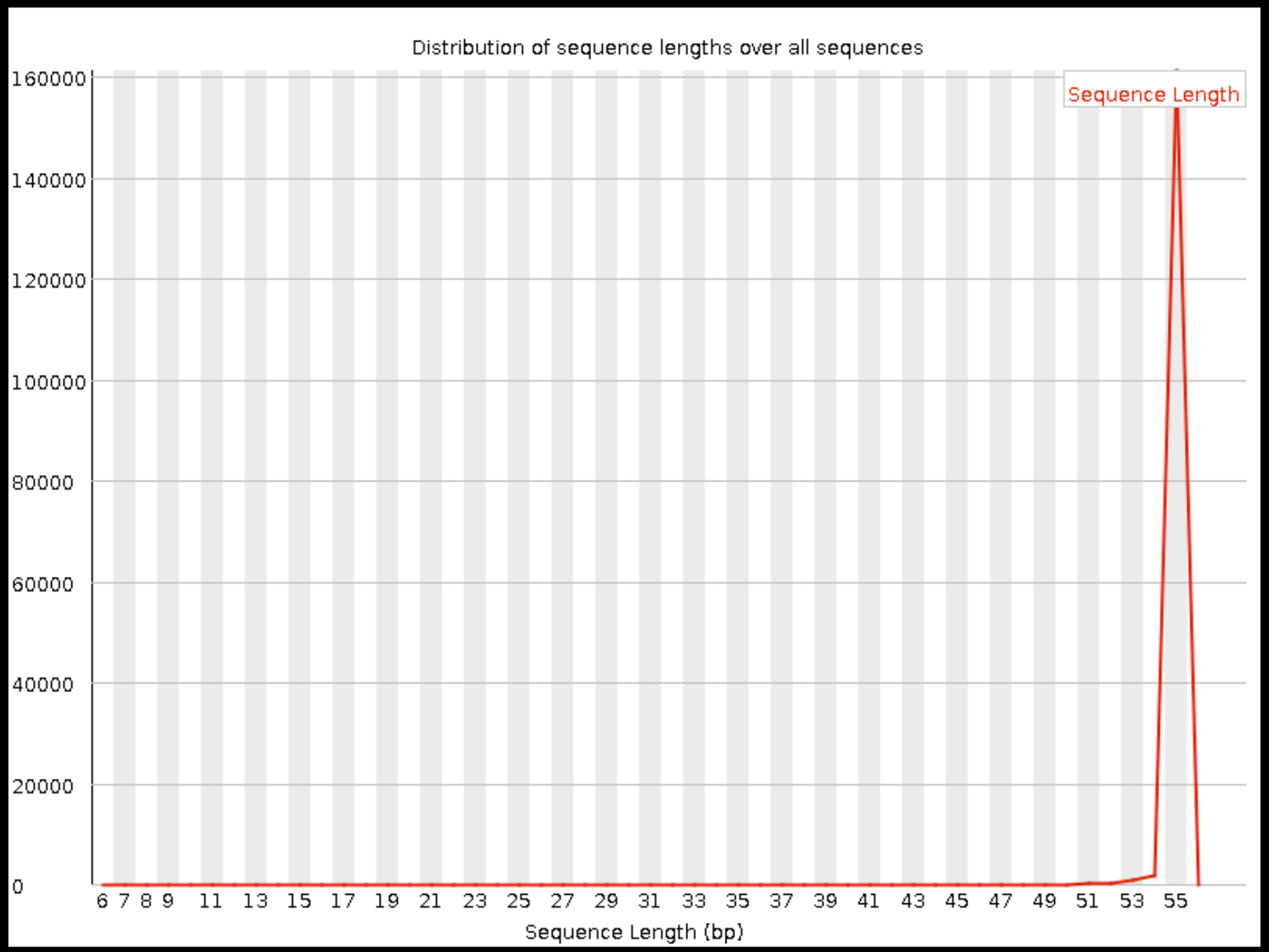
NGS Data Quality: Base Quality Trimming



We will trim by sliding windows.

- NGS QC and Manipulation → **FASTQ Quality Trimmer by sliding window**
- Trim from both ends, using sliding windows, until you hit a high-quality section.
- Produces variable length reads
- Enhances mapability, but remember, with expression data: less is often more
- Remove *just the bases that would prevent the rest from mapping*

NGS Data Quality: Base Quality Trimming



New Problem?

Now some reads are so short they are just noise and can't be meaningfully mapped. Have potential to bog down mapping.

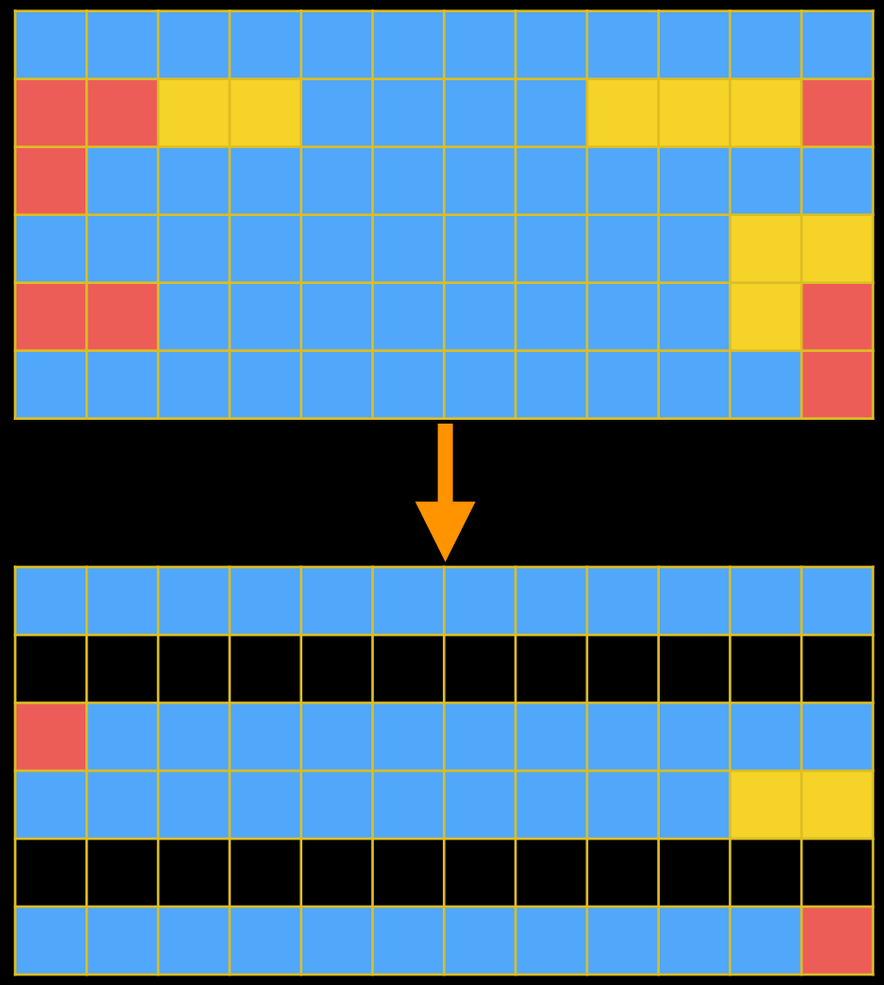
Filter to fix this (breaks pairings, but that can be remedied).

Or, your mapper may have an option to ignore shorter reads.

NGS Data Quality: Filtering by length

We will filter for a minimum length, 35

- »Empty« reads cause problems
- Too short, won't meet map criteria (false re-pairing, if done)
- NGS QC and Manipulation → **Filter FASTQ reads by quality score and length.**
  - Keep or discard whole reads
  - Can have different thresholds for different regions of the reads.
  - Keeps original read length.



NGS Data Quality: Sequencing Artifacts

Take another look at FastQC Report .. what's it is?

Anyone with the MeOH Rep1 R2 (the reverse reads) will see this:

Overrepresented sequences

Sequence	Count	Percentage	Possible Source
CTGTGTATTTGTCAATTTTCTTCTCCACGTTCTTCTCGGCCTGTTTCCGTAGCCT	590	0.3541692929220167	No Hit
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	342	0.2052981325073385	No Hit
CGGCCACAAATAAACACAGAAATAGTCCAGAATGTCACAGGTCCAGGGCAGAGGA	325	0.1950932545756871	No Hit
CTGCATTATAAAAAGGACAGCCAGATATCAACTGTTACAGAAATGAAATAAGACG	230	0.1380659955458709	No Hit
CGGCCGCAAATAAACACAGAAATAGTCCAGAATGTCACAGGTCCAGGGCAGAGGA	199	0.1194571004940361	No Hit
GTCAGCTCAACTTGTAGGCCCCAAAAGAAAACAGCGTCTTACTGGGGAGGGATAT	197	0.1182565266197242	No Hit

For some pipelines, like assembly, you want to remove, use:

NGS QC and Manipulation → **Remove sequencing artifacts**

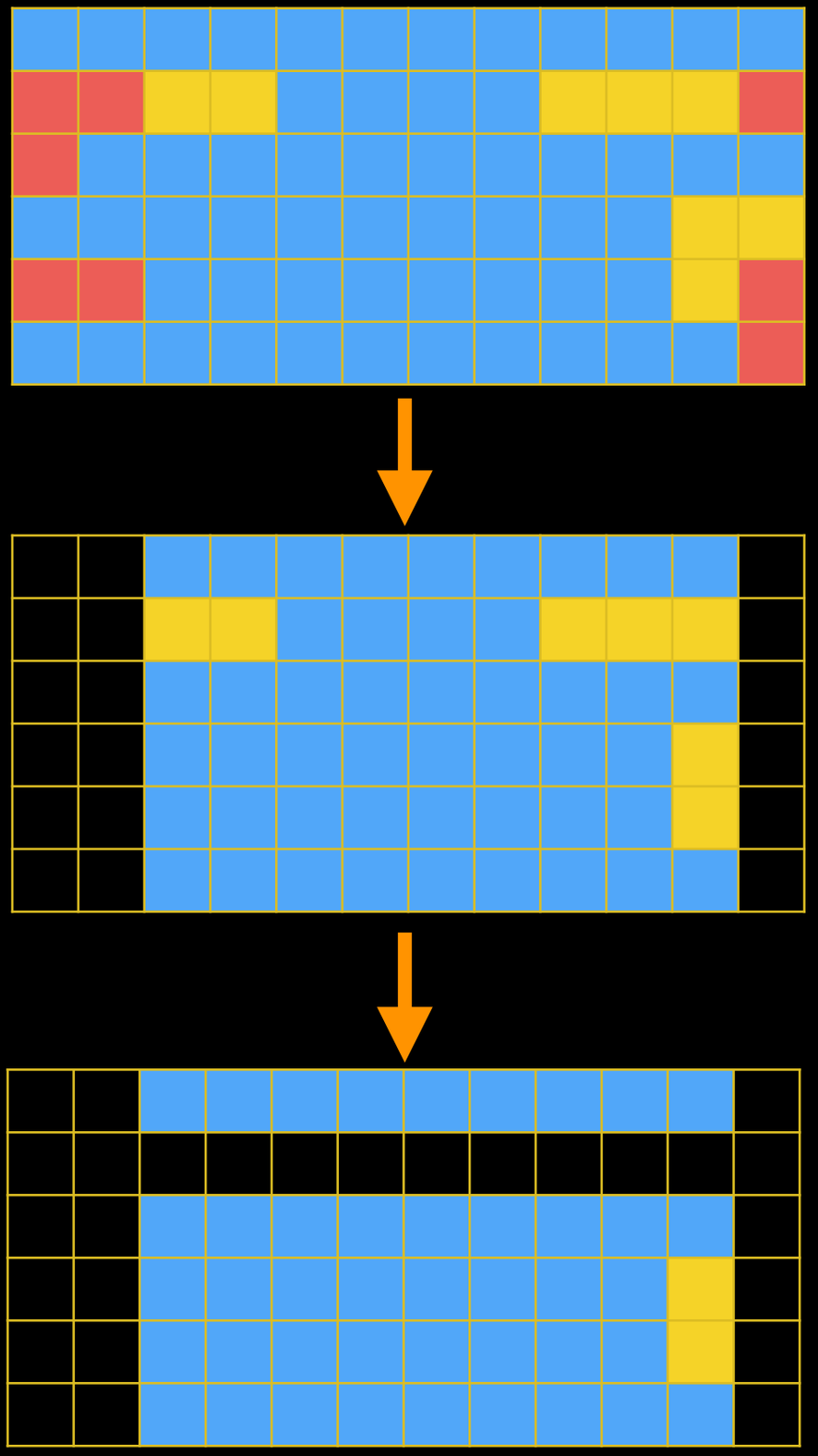
We will use, too, to cull artifacts from potential contamination.

Mapper won't map either, but may want to investigate contam.

Sometimes it helps to mix it up.

With your own RNA data, experiment, and test with your mapper and downstream tools.

…and read-the-manual, *reallly*



Option 1
(by column)

+

Option 2
(by row)

NGS Data Quality: Check your data *after*

NGS QC and Manipulation → FastQC

The original FastQC report no longer represents your data.

Run it again.

NGS Data Quality: Make a *Workflow*

History menu → Extract Workflow

Now let's go take look, edit it, and run it. Then try:

Workflow → Published Workflow → RNA-Seq Step 1: QC

Sharing in Galaxy

Share:

Make something available to someone else or everyone else who has the link

Publish:

Make something available to everyone

1. Add »R1« or »R2« to your History name and Publish

History menu → Share or Publish → Publish

2. *Find a published History of opposite type and Import*

Shared Data → Published Histories → Import (green + icon)

3. *Copy the result datasets into your own History*

History menu → Copy Datasets → Pick Original History

Keeping paired ends paired: Pretty or Required?

- Important for many **Variant Analysis** pipelines, but not required for the RNA-seq Tuxedo Suite (Tophat, Cuff*)
  - » pairs not properly mapped are ignored

- Run the Picard Paired Read Mate Fixer after mapping reads. You don't need this now, but is required for Variant.

- Do you really need to do anything? Test it. (later)

- Run a workflow (like "RNA-seq Step 2: Re-pair") that removes any unpaired reads before mapping.
  We'll do this. *Where are Published Workflows again?*

Re-Pair Paired ends. Not required for RNA-seq,
but it is good to ***learn how.…***

Workflow takes 4 inputs

- Forward Reads, before QC
- Reverse Reads, before QC
- Forward Reads, after QC
- Reverse Reads, after QC

And produces 2 outputs

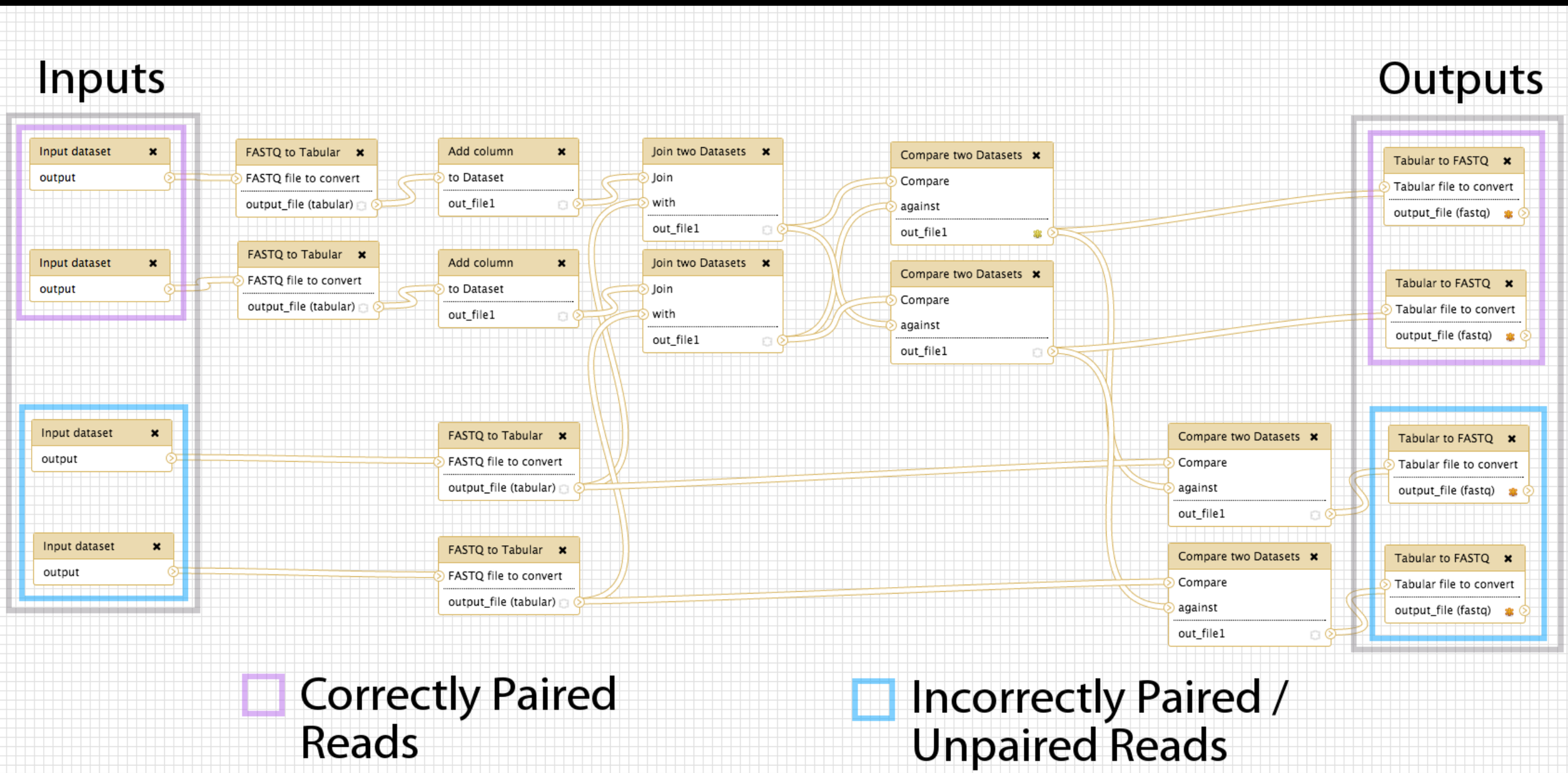
- Forward reads, re-paired
- Reverse reads, re-paired

Workflow assumes pre-QC reads are correctly paired.

Can be edited later to also output any strays or
»singletons«

Is just a collection of text manipulation tools.
Workflows are like making your OWN tools. On demand.

Open in workflow editor: RNA-seq Step 2: Re-pair



Who can tell me what is different about this workflow and the one actually used today?

NGS Data Quality: Further reading & Resources

FastQC Documenation

Read Quality Assessment & Improvement

by Joe Fass

From the UC Davis 2013 Bioinformatics Short Course

Manipulation of FASTQ data with Galaxy

by Blankenberg, *et al*.

Mapping with Tophat

RNA-Seq: Mapping with Tophat

Create new history or Use what you have

(cog) → Create New

Get filtered reads if you want to start over, or skip

Shared Data → Data Libraries

→ RNA-Seq UCDavis 2013 Example Data*

→ **Reads, Post-QC**

→ Select **MeOH_REP1_R1, MeOH_REP1_R2**

Everyone get the Reference Annotation **genes_chr12.gtf**

And then Import to current history what you selected

UCDAVIS Genome Center Bioinformatics Core

* RNA-Seq example datasets from the 2013 UC Davis Bioinformatics Short Course. http://bit.ly/ucdbsc2013

RNA-seq Exercise: Mapping with Tophat

- Tophat looks for best place(s) to map reads, and best places to insert introns
- *Imagine pages and pages of discussion on the intricacies and pitfalls of RNA-seq mapping here.*

Mapping with Tophat: mean inner distance

Expected distance between paired end reads

- Determined by sample prep
- We'll use 90* for mean inner distance
- We'll use 50 for standard deviation

✳ The library was constructed with the typical Illumina TruSeq protocol, which is supposed to have an average insert size of 200 bases. Our reads are 55 bases (R1) plus 55 bases (R2). So, the Inner Distance is estimated to be 200 - 55 - 55 = 90

From the 2013 UC Davis Bioinformatics Short Course

Mapping with Tophat: Use Existing Annotations?

You can bias Tophat towards known annotations

- Use Own Junctions → Yes
  - Use Gene Annotation → Yes
    - Gene Model Annotation → genes_chr12.gtf
- Use Raw Junctions → Yes (tab delimited file)
- Only look for supplied junctions → Yes

Mapping with Tophat: Make it quicker?

Warning: Here be dragons!

- Allow indel search → No
- Use Coverage Search → No (wee dragons)

TopHat generates its database of possible splice junctions from two sources of evidence. The first and strongest source of evidence for a splice junction is when two segments from the same read (for reads of at least 45bp) are mapped at a certain distance on the same genomic sequence or when an internal segment fails to map - again suggesting that such reads are spanning multiple exons. With this approach, "GT-AG", "GC-AG" and "AT-AC" introns will be found *ab initio*. The second source is pairings of "coverage islands", which are distinct regions of piled up reads in the initial mapping. Neighboring islands are often spliced together in the transcriptome, so TopHat looks for ways to join these with an intron. We only suggest users use this second option (--coverage-search) for short reads (< 45bp) and with a small number of reads (<= 10 million). This latter option will only report alignments across "GT-AG" introns

TopHat Manual

Mapping with Tophat: Max # of Alignments Allowed

Some reads align to more than one place equally well.

For such reads, how many should Tophat include?

If more than the specified number, Tophat will pick those with the best mapping score.

Tophat breaks ties randomly.

Tophat assigns equal fractional credit to all *n* mappings

Instructs TopHat to allow up to this many alignments to the reference for a given read, and choose the alignments based on their alignment scores if there are more than this number. The default is 20 for read mapping. Unless you use --report-secondary-alignments, TopHat will report the alignments with the best alignment score. If there are more alignments with the same score than this number, TopHat will randomly report only this many alignments. In case of using --report-secondary-alignments, TopHat will try to report alignments up to this option value, and TopHat may randomly output some of the alignments with the same score to meet this number.

TopHat Manual

Mapping with Tophat: Execute or Use Workflow

- Run the tool directly from the Tool Panel
- Or use the Shared Workflow: **RNA-seq Step 3: Mapping**

Please just do one or the other. We want everyone to have a chance to run the job. *This may take a while …*

While we are waiting, why not look at one that is already done? Which Published History is it?

(hint, we are at Step 3. View it. Then Import it.)

There five output files. Where? Reveal hidden ones.
Delete them. *Permenently* Delete them.

RNA-Seq Mapping With Tophat: Resources

RNA-Seq Concepts, Terminology, and Work Flows
by Monica Britton

Aligning PE RNA-Seq Reads to a Genome
by Monica Britton

both from the UC Davis 2013 Bioinformatics Short Course

RNA-Seq Analysis with Galaxy
by Jeroen F.J. Laros, Wibowo Arindrarto, Leon Mei

from the GCC2013 Training Day

RNA-Seq Analysis with Galaxy
by Curtis Hendrickson, David Crossman, Jeremy Goecks

from the GCC2012 Training Day

RNA-Seq: Differential Expression with Cuffdiff

Is your Tophat job still running?
Doesn't matter.
We need way more inputs now.

Back to the Shared Data Library ...

RNA-Seq Differential Expression: Get the Data

Create new history. Really. Everyone.

(cog) → Create New

Import:

Shared Data → Data Libraries

→ RNA-Seq UCDavis 2013 Example Data*

→ Tophat Outputs

→ Select all accepted_hits datasets

Also select **genes_chr12.gtf**

And then Import to current history

UCDAVIS Genome Center Bioinformatics Core

* RNA-Seq example datasets from the 2013 UC Davis Bioinformatics Short Course. http://bit.ly/ucdbsc2013

Cuffdiff

- Part of the Tuxedo RNA-Seq Suite (as are Tophat and Bowtie)
- Identifies differential expression between multiple datasets
- Widely used and widely installed on Galaxy instances

NGS: RNA Analysis → Cuffdiff

Cuffdiff

Cuffdiff uses FPKM/RPKM as a central statistic.
Total # mapped reads heavily influences FPKM/RPKM.
Can lead to challenges when you have very highly expressed genes in the mix.

What's the difference?

They're *almost* the same thing.

RPKM stands for Reads Per Kilobase of transcript per Million mapped reads.
FPKM stands for Fragments Per Kilobase of transcript per Million mapped reads.
In RNA-Seq, the relative expression of a transcript is proportional to the number of cDNA fragments that originate from it. **Paired-end RNA-Seq experiments produce two reads per fragment, but that doesn't necessarily mean that both reads will be mappable.** For example, the second read is of poor quality. If we were to count reads rather than fragments, we might double-count some fragments but not others, leading to a skewed expression value.

Thus, FPKM is calculated by counting fragments, not reads.
(this is why re-pairing doesn't matter, FRKM is only counted if both ends are there)

The Manaul
http://cufflinks.cbcb.umd.edu/faq.html

Cuffdiff

- Running with 2 Groups: MeOH and R3G
- Each group has 3 replicates each

Cuffdiff

- Which Transcript definitions to use?
  - Official (genes_chr12.gtf in our case)
  - MeOH or R3G Cufflinks transcripts
  - Results of Cuffmerge on MeOH & R3G Cufflinks transcripts
- Depends on what you care about

NGS: RNA Analysis → Cuffdiff

Cuffdiff

Produces many output files, all explained in doc

We'll focus on gene differential expression testing

test_id	gene_id	gene	locus	sample_1	sample_2	status	value_1	value_2	log2(fold_change)	test_stat	p_value	q_value	significant
A2M	A2M	A2M	chr12:9217772-9268558	MeOH	R3G	NOTEST	3.32147	3.13694	-0.0824644	0	1	1	no
A2M-AS1	A2M-AS1	A2M-AS1	chr12:9217772-9268558	MeOH	R3G	NOTEST	7.45797	13.9413	0.902515	0	1	1	no
A2ML1	A2ML1	A2ML1	chr12:8975149-9029381	MeOH	R3G	NOTEST	4.83055	7.79884	0.691072	0	1	1	no
A2MP1	A2MP1	A2MP1	chr12:9381128-9386803	MeOH	R3G	NOTEST	2.49656	0	-inf	0	1	1	no
AAAS	AAAS	AAAS	chr12:53701239-53715412	MeOH	R3G	OK	269.035	159.23	-0.756683	-2.22857	0.0005	0.00194017	yes
AACS	AACS	AACS	chr12:125549924-125627871	MeOH	R3G	NOTEST	29.2933	35.0339	0.258178	0	1	1	no
ABCB9	ABCB9	ABCB9	chr12:123405497-123451056	MeOH	R3G	NOTEST	4.68869	1.7732	-1.40283	0	1	1	no
ABCC9	ABCC9	ABCC9	chr12:21950323-22089628	MeOH	R3G	OK	553.247	487.261	-0.18323	-2.02806	0.0004	0.00162143	yes
ABCD2	ABCD2	ABCD2	chr12:39945021-40013843	MeOH	R3G	OK	86.1377	172.795	1.00435	4.3436	5e-05	0.000246739	yes
ACACB	ACACB	ACACB	chr12:109577201-109706030	MeOH	R3G	NOTEST	8.45306	15.5772	0.881885	0	1	1	no
ACAD10	ACAD10	ACAD10	chr12:112123856-112194911	MeOH	R3G	NOTEST	21.8237	27.8326	0.350882	0	1	1	no
ACADS	ACADS	ACADS	chr12:121163570-121177811	MeOH	R3G	NOTEST	38.644	16.1739	-1.25658	0	1	1	no
ACRBP	ACRBP	ACRBP	chr12:6747241-6756580	MeOH	R3G	NOTEST	2.96987	3.26939	0.138621	0	1	1	no
ACSM4	ACSM4	ACSM4	chr12:7456927-7480969	MeOH	R3G	NOTEST	0	0	0	0	1	1	no
ACSS3	ACSS3	ACSS3	chr12:81471808-81649582	MeOH	R3G	NOTEST	0	0	0	0	1	1	no
ACTR6	ACTR6	ACTR6	chr12:100593864-100618202	MeOH	R3G	OK	475.594	421.324	-0.174799	-0.797581	0.1588	0.258406	no
ACVR1B	ACVR1B	ACVR1B	chr12:52345450-52390863	MeOH	R3G	NOTEST	32.5737	38.3075	0.233922	0	1	1	no
ACVRL1	ACVRL1	ACVRL1	chr12:52301201-52317145	MeOH	R3G	NOTEST	1.27713	2.16161	0.759201	0	1	1	no
ADAM1A	ADAM1A	ADAM1A	chr12:112336866-112339706	MeOH	R3G	NOTEST	30.0162	55.2154	0.879331	0	1	1	no
ADAMTS20	ADAMTS20	ADAMTS20	chr12:43748011-43945724	MeOH	R3G	NOTEST	0.453322	0.502067	0.147346	0	1	1	no
ADCY6	ADCY6	ADCY6	chr12:49159974-49182820	MeOH	R3G	NOTEST	9.32722	17.6743	0.922135	0	1	1	no
ADIPOR2	ADIPOR2	ADIPOR2	chr12:1800246-1897845	MeOH	R3G	OK	207.468	179.333	-0.210248	-1.02392	0.09	0.158988	no
AEBP2	AEBP2	AEBP2	chr12:19592607-19675173	MeOH	R3G	OK	143.039	128.293	-0.156957	-0.688267	0.2254	0.344537	no
AGAP2	AGAP2	AGAP2	chr12:58118075-58135944	MeOH	R3G	OK	98.2385	116.302	0.243511	0.935119	0.11475	0.198086	no
AICDA	AICDA	AICDA	chr12:8754761-8765442	MeOH	R3G	NOTEST	78.1514	63.4313	-0.301077	0	1	1	no
AKAP3	AKAP3	AKAP3	chr12:4724675-4754343	MeOH	R3G	NOTEST	6.12385	7.89626	0.366731	0	1	1	no
ALDH1L2	ALDH1L2	ALDH1L2	chr12:105413561-105478341	MeOH	R3G	NOTEST	7.11374	8.11722	0.190377	0	1	1	no
ALDH2	ALDH2	ALDH2	chr12:112204690-112247789	MeOH	R3G	NOTEST	12.8033	8.05635	-0.668321	0	1	1	no
ALG10	ALG10	ALG10	chr12:34175215-34181236	MeOH	R3G	NOTEST	54.8575	59.3459	0.11346	0	1	1	no
ALG10B	ALG10B	ALG10B	chr12:38710556-38723528	MeOH	R3G	NOTEST	43.8157	63.0457	0.524952	0	1	1	no
ALKBH2	ALKBH2	ALKBH2	chr12:109525992-109531293	MeOH	R3G	OK	679.517	297.183	-1.19316	-3.34255	5e-05	0.000246739	yes
ALX1	ALX1	ALX1	chr12:85674035-85695561	MeOH	R3G	NOTEST	0	0	0	0	1	1	no

Cuffdiff: differentially expressed genes

Column	Contents
test_stat	value of the test statistic used to compute significance of the observed change in FPKM
p_value	Uncorrected P value for test statistic
q_value	FDR-adjusted p-value for the test statistic
status	Was there enough data to run the test?
significant	and, was the gene differentially expressed?

Cuffdiff

- Column 7 (“status”) can be FAIL, NOTEST, LOWDATA or OK
  - Filter and Sort → Filter
    - c7 == 'OK'
- Column 14 (“significant”) can be yes or no
  - Filter and Sort → Filter
    - c14 == 'yes'

Returns the list of genes with
1) enough data to make a call, and
2) that are called as differentially expressed.

Cuffdiff: Next Steps

Try running Cuffdiff with different normalization and dispersion estimation methods.

Later, please. On the Slipstream. But do. Understand it.

Compare the differentially expressed gene lists.
Which settings have what type of impacts on the results?

RNA-Seq Differential Expression with Cuffdiff: Resources

RNA-Seq Concepts, Terminology, and Work Flows

by Monica Britton

from the UC Davis 2013 Bioinformatics Short Course

RNA-Seq Analysis with Galaxy

by Jeroen F.J. Laros, Wibowo Arindrarto, Leon Mei

from the GCC2013 Training Day

RNA-Seq Analysis with Galaxy

by Curtis Hendrickson, David Crossman, Jeremy Goecks

from the GCC2012 Training Day

Galaxy Project: Resources

http://galaxyproject.org

http://usegalaxy.org

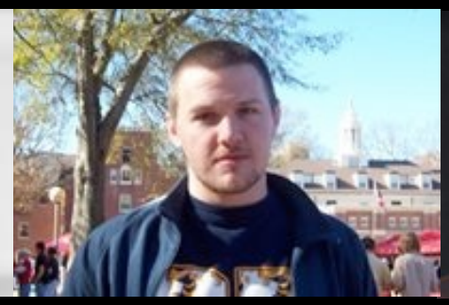
http://getgalaxy.org

http://wiki.galaxyproject.org/Cloud

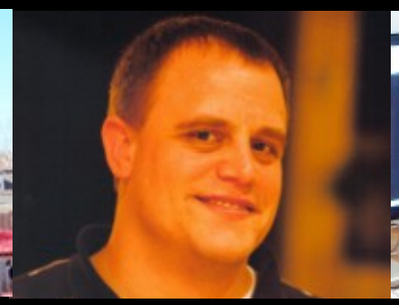
http://bit.ly/gxychoices

The Galaxy Team

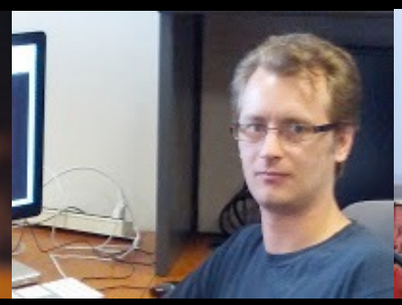
Enis Afgan



Dannon Baker



Dan Blankenberg



Dave Bouvier

Marten Cech

John Chilton

Dave Clements

Nate Coraor

Carl Eberhard

Dorine Francheteau

Jeremy Goecks

Sam Guerler

Jen Jackson

Greg von Kuster

Ross Lazarus

Anton Nekrutenko

Nick Stoler

James Taylor

http://wiki.galaxyproject.org/GalaxyTeam

Thanks

Dave Clements, Dannon Baker, and Enis Afgan

Galaxy Project

Anushka Brownley

Bioteam

Yes, I live right by here and took this picture all by myself.

But I had much help with this tutorial from those listed here.

outreach@galaxyproject.org

Galaxy Project: Resources

http://galaxyproject.org

http://usegalaxy.org

http://getgalaxy.org

http://wiki.galaxyproject.org/Cloud

http://bit.ly/gxychoices