

James Madison University

From the Selected Works of Ray Enke Ph.D.

June, 2016

Approaches to Teaching RNA-Seq in Undergraduate Lecture & Lab Courses

Ray A Enke

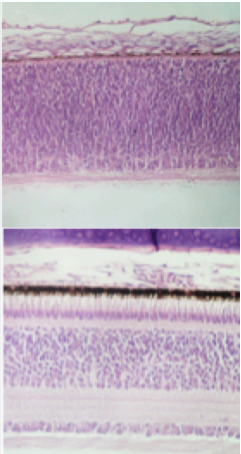
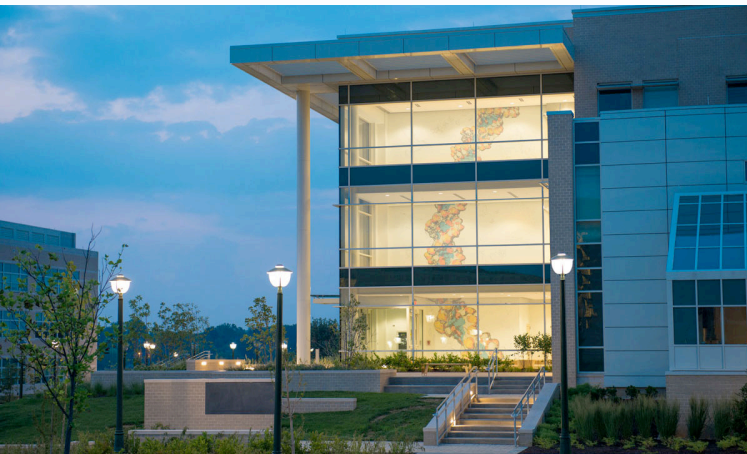


This work is licensed under a [Creative Commons CC BY-SA International License](https://creativecommons.org/licenses/by-sa/4.0/).



Available at: https://works.bepress.com/raymond_enke/S9/

Approaches to Teaching RNA-Seq in Undergraduate Lecture & Lab Courses



Ray Enke Ph.D.

James Madison University

Department of Biology, Harrisonburg VA

2016 RNA-Seq For the Next Generation Workshop

June 10th, 2016

Talk Overview

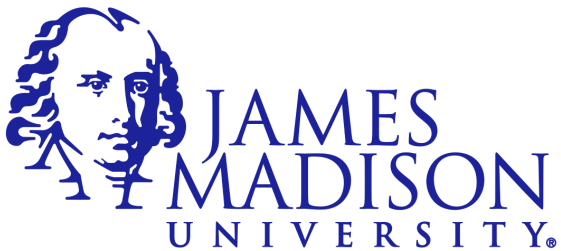
1. Background on JMU, my position & course and our research question
2. Overview of my 7 week course-embedded RNA-Seq module
3. Reflection from 3 semesters of developing RNA-Seq modules & future directions

About James Madison University



JMU Biosciences building

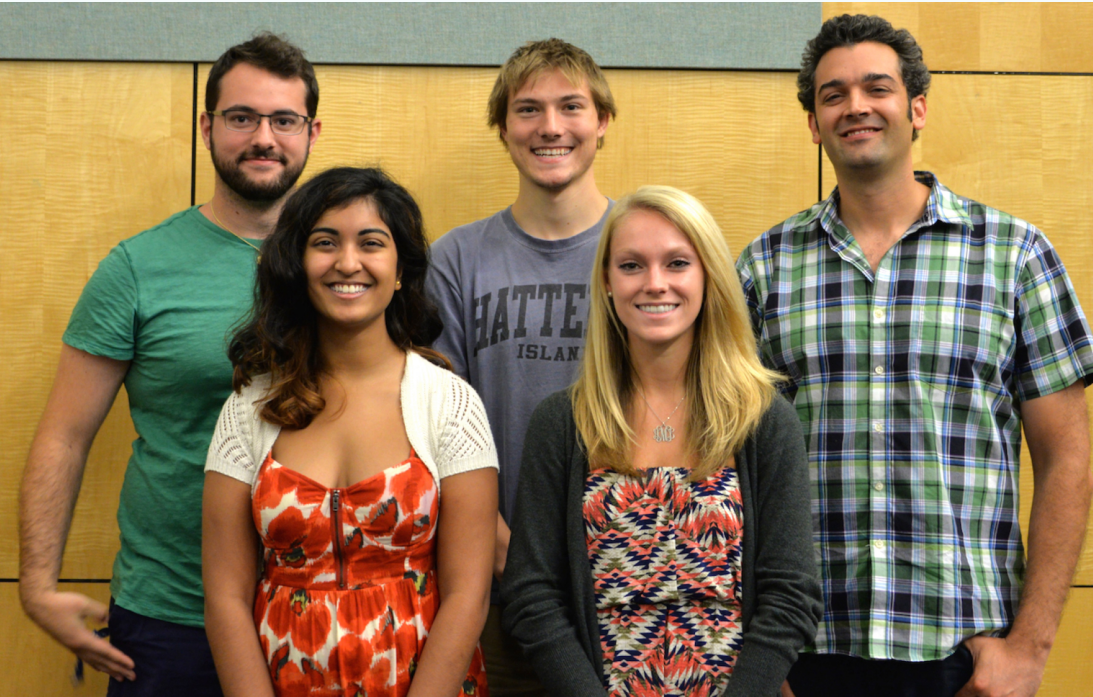
- Public PUI; Harrisonburg, VA
- ~19,000 undergraduates
- ~60% admission rate
- ~1,000 Biology/Biotech majors
- Emphasis on undergrad research & course based research



About My Lab & Position

Christophe

Nick



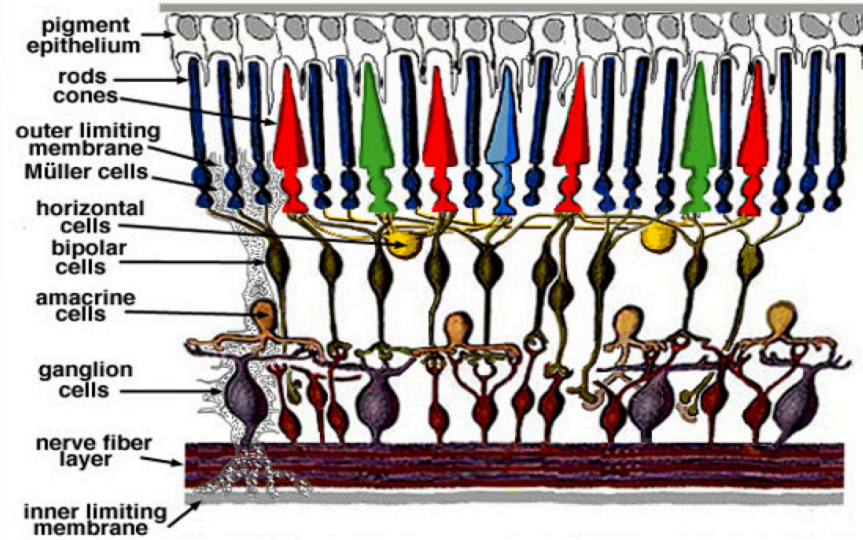
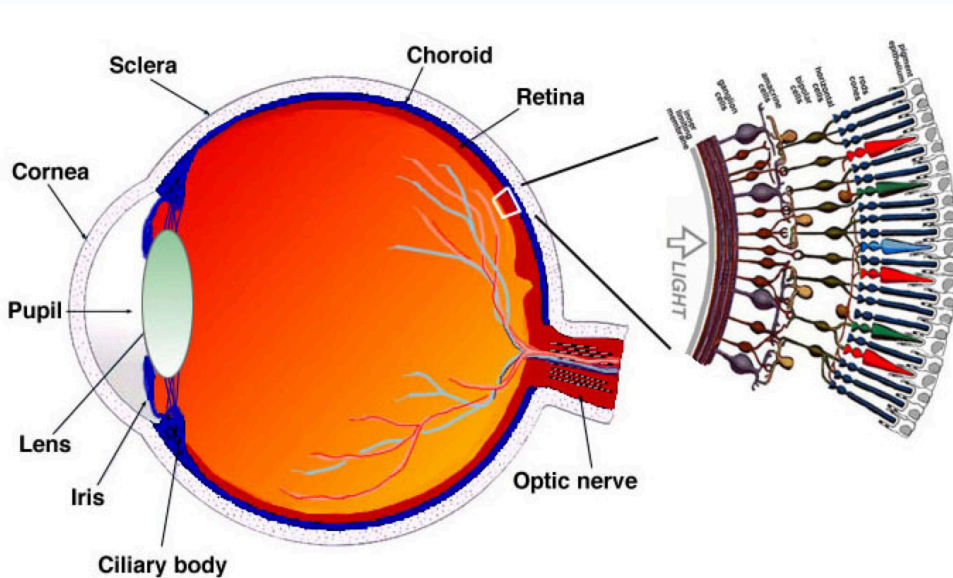
Sraavya

Annamarie

- Tenure track
 - just finished 3rd year
 - limited prior teaching experience
 - limited prior training in bioinformatics
- 50% : 40% : 10%
teaching : research : service
- 12 contact hours/semester
(combo of 4 lecture or labs/semester)
- Small undergraduate research lab (4-8 students)



The Vertebrate Retina



Images from <http://webvision.med.utah.edu/>

- Regulation of gene expression in the developing & diseased retina
 - chicken, mouse, human

BIO480 Advanced Molecular Biology

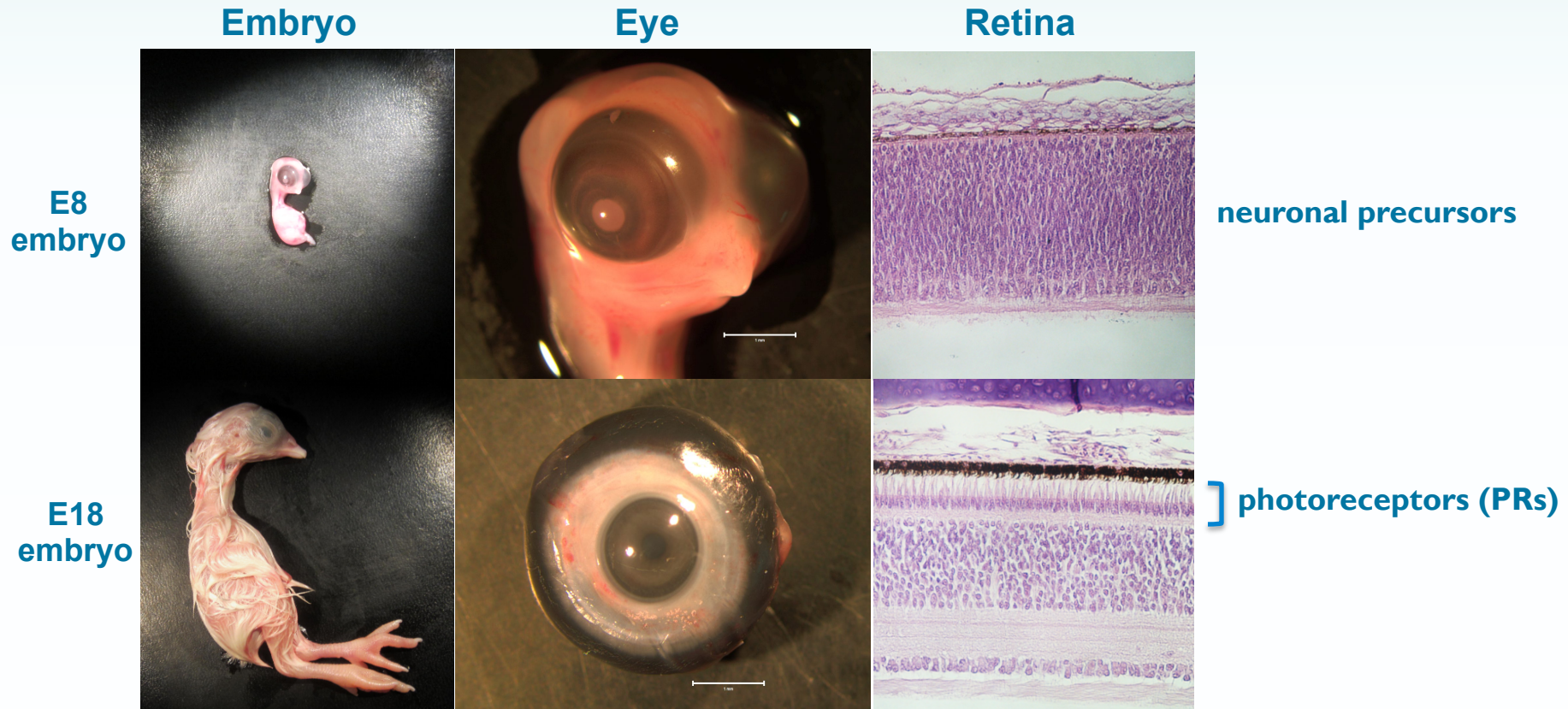
Spring 15 class



- Junior/senior level course
- 18-36 students in lecture
- 18 students/lab section
 - 6 groups of 3
- 4 credit lecture/lab course
 - 4 hours/week
- ~50% have lab research experience other than classes

- Developed & incorporated RNA-Seq analysis over the past 2 years over 3 semesters

Our Research Question: What genes are differentially regulated during development of the chicken retina?

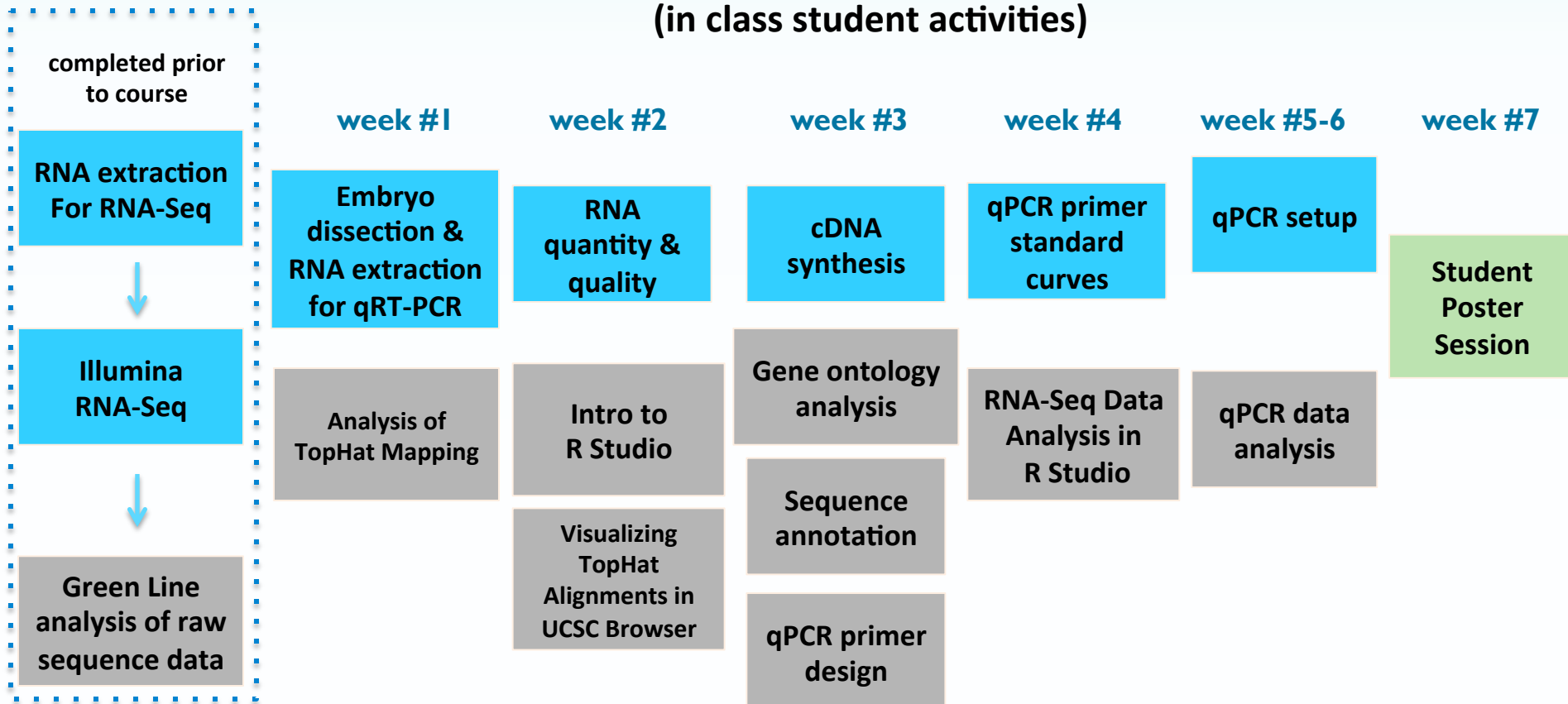


What genes are differentially expressed between:

- E8 retina vs E18 retina (developmental changes in gene expression)
- E18 retina vs E18 cornea (tissue-specific changes in gene expression)

BIO480 Course-embedded RNA-Seq Research Modules (4 hour lab/week)

(in class student activities)



*Green Line run prior to course activities

Key:

Wet lab
module

Computational
module

My Teaching Materials Website

<http://www.rnaseqforthenextgeneration.org/profiles/raymond-enke.html#teaching>

PROFILES: Working Group Faculty, 2014

[Previous](#) | [Next](#)

Ray Enke

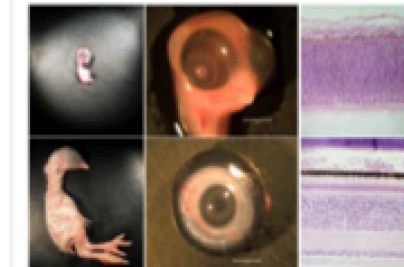
Assistant Professor of Biology, James Madison University
[Website](#)

Characterizing differential gene expression in the developing chicken retina

[Abstract](#)

[Research and Data](#)

[Teaching and Resources](#)



Wet lab protocols:

- Simultaneous DNA and RNA purification ([Download PDF](#))
- Spectrophotometric analysis of RNA on a Synergy H4 Hybrid plate reader ([Download PDF](#))
- Analysis of RNA quality by agarose gel electrophoresis ([Download PDF](#))
- 1st strand cDNA synthesis from total RNA ([Download PDF](#))

Computer lab tutorials:

- qRT-PCR primer design ([Download PDF](#))
- Obtaining and Importing data from the SRA database into DNA Subway ([Download PDF](#))

Wet lab combined with computational lab activities:

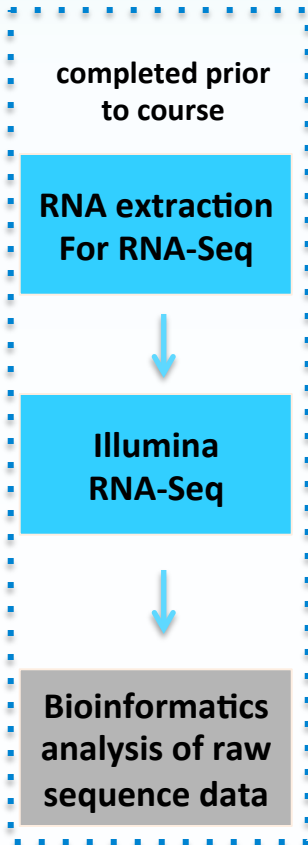
- 1st strand cDNA synthesis from total RNA, Gene Ontology (GO) analysis of CuffDiff output and sequence analysis/annotation ([Download PDF](#))
- qRT-PCR set up and analysis ([Download PDF](#))

Lectures:

- DNA sequencing technology ([Download PDF](#))
- Next Generation sequencing ([Download PDF](#))

*updated materials will posted after the workshop has concluded

Wk#1 (computational): Analysis of TopHat Mapping

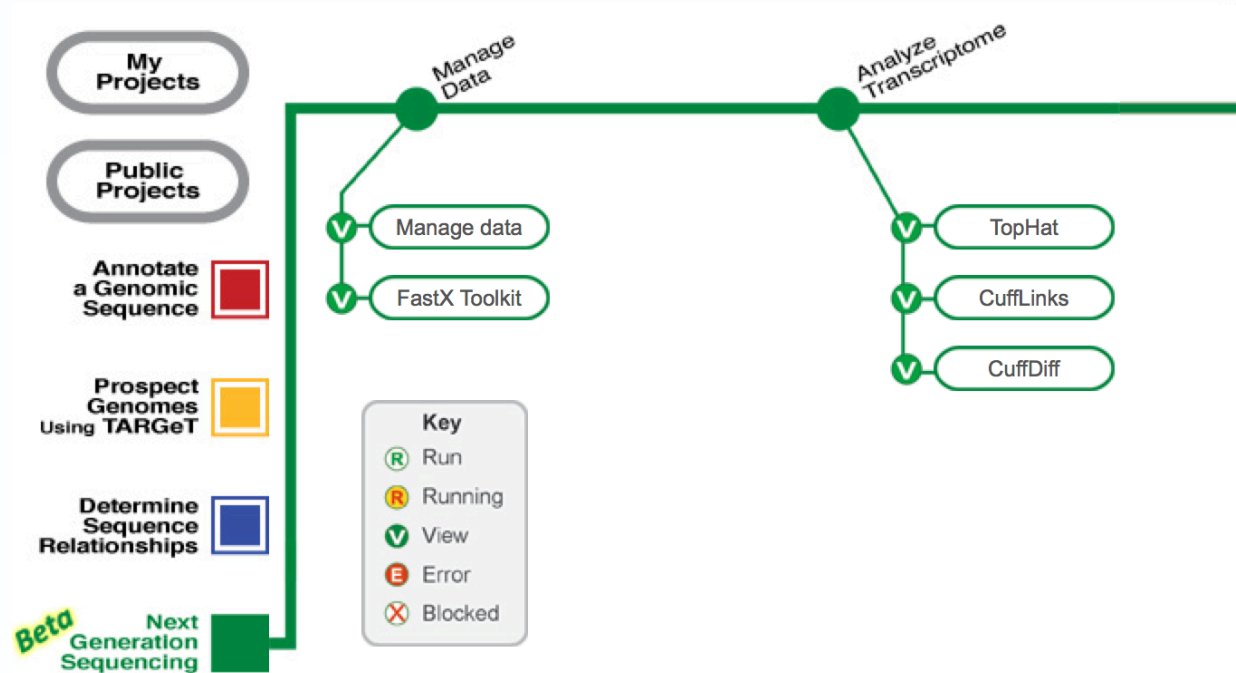


week #1

Embryo dissection & RNA extraction for qRT-PCR

Analysis of TopHat Mapping

- Students create a DNA Subway account and navigate to a pre run public project

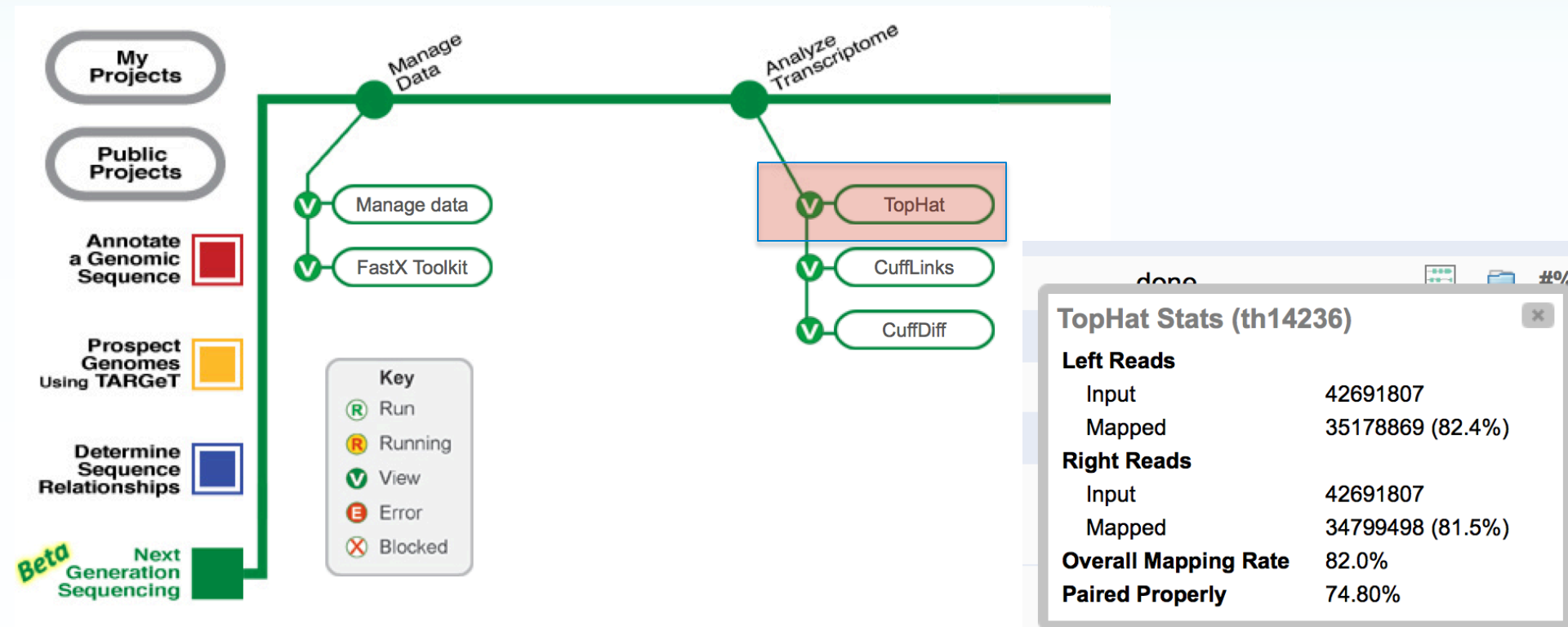


Key:

Wet lab module

Computational module

Wk#1 (computational): Analysis of TopHat Mapping



- Students navigate to & view individual TopHat alignment files to gather info for each sample

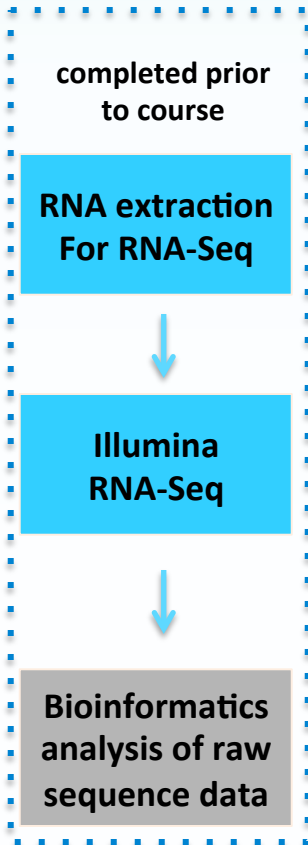
Wk#1 (computational): Analysis of TopHat Mapping

Students fill in TopHat alignment info on a GoogleSheet I provided to them

FASTQ Filename	TopHat Job#	Sample	Read End	# Input Reads	# Mapped Reads	% Mapped Reads	% Overall Mapping Rate	% Properly Paired Reads	# of mapped paired end reads (in millions)***
RNA5_S5_R1.fastq.gz	th14236	E8 retina replicate #1	1-left	42691807	35178869	82.40192082	81.95760723	74.8	26.17
RNA5_S5_R2.fastq.gz	th14236	E8 retina replicate #1	2-right	42691807	34799498	81.51329364	N/A	N/A	N/A
RNA6_S6_R1.fastq.gz		E8 retina replicate #2	1-left						0.00
RNA6_S6_R2.fastq.gz		E8 retina replicate #2	2-right				N/A	N/A	N/A
RNA7_S7_R1.fastq.gz		E18 retina replicate #1	1-left						0.00
RNA7_S7_R2.fastq.gz		E18 retina replicate #1	2-right				N/A	N/A	N/A
RNA8_S8_R1.fastq.gz		E18 retina replicate #2	1-left						0.00
RNA8_S8_R2.fastq.gz		E18 retina replicate #2	2-right				N/A	N/A	N/A
292652_S1_R1.fastq.gz		E18 cornea replicate #1	1-left						0.00
292652_S1_R2.fastq.gz		E18 cornea replicate #1	2-right				N/A	N/A	N/A
293205_S8_R1.fastq.gz		E18 cornea replicate #2	1-left						0.00
293205_S8_R2.fastq.gz		E18 cornea replicate #2	2-right				N/A	N/A	N/A
***this is the number of paired end sequencing reads available to calculate differential expression compared to another sample group									
Note: this GoogleSheet is set to view only and not edit. Download and editable version of this spreadsheet to complete the lab assignment									

- The activity gives students an introductory experience with Green Line
- Also demonstrates the huge amount of sequence data in analysis pipeline

Wk#:1 (wet lab): Chicken Embryo Dissections

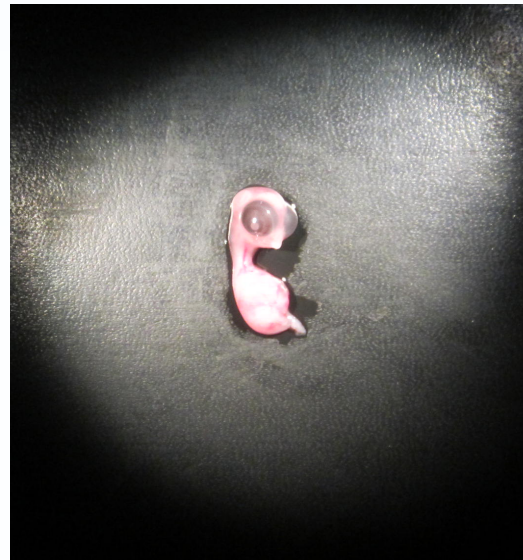


week #1

Embryo dissection & RNA extraction for qRT-PCR

Analysis of TopHat Mapping

- E8 & E18 embryo dissections
- Silica column-based RNA extractions



Required equipment/associated costs:

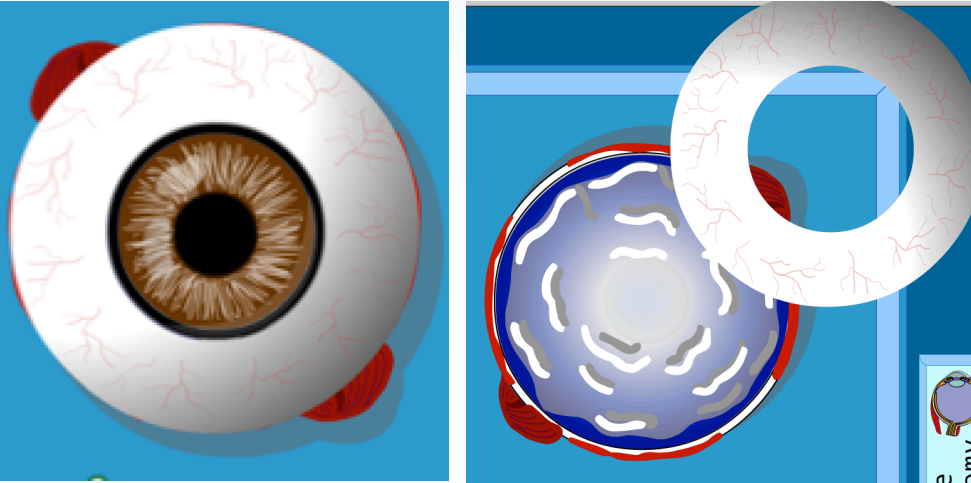
- Eggs
- Dissection tools & supplies

Key:

Wet lab module

Computational module

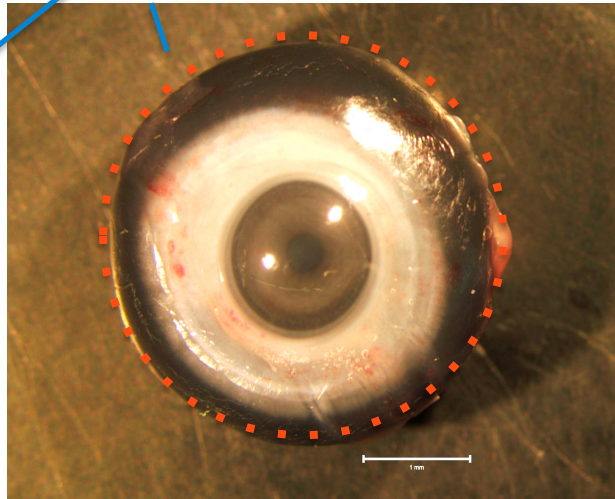
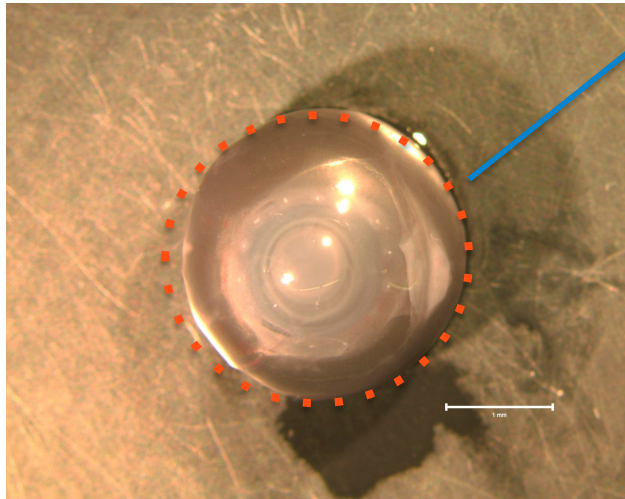
Wk#:1 (wet lab): Chicken Embryo Dissections



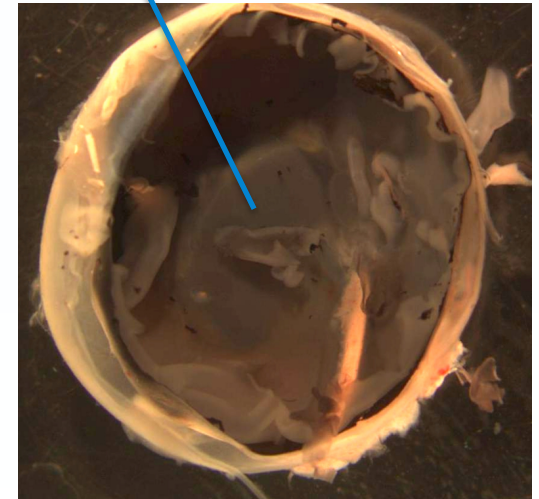
6. Scalpel & scissors to cut around equator of eye

7. Remove anterior portion (front half)
-save for cornea

Cut front half off



Retina



Wk#1 (wet lab): Total RNA Extractions

completed prior
to course

RNA extraction
For RNA-Seq



Illumina
RNA-Seq



Bioinformatics
analysis of raw
sequence data

week #1

Embryo
dissection &
RNA extraction
for qRT-PCR

Analysis of
TopHat Mapping

- RNeasy Mini Kit
- Silica column-based RNA extractions



Required equipment/associated costs:

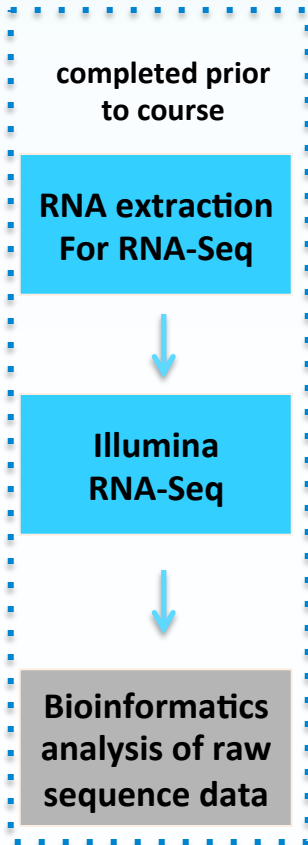
- RNeasy kit- ~\$260
- Qias shredders- ~\$75
- DNaseI kit ~\$90

Key:

Wet lab
module

Computational
module

Wk#2 (wet lab): RNA Quality & Quantity Analysis



week #2

RNA
quantity &
quality

Intro to
R Studio

Visualizing
TopHat
Alignments

- UV Spectrophotometer to obtain concentration & 260/280 ratio
 - optimally ~2.0 for pure RNA
- Agarose gel electrophoresis of total RNA
 - no formaldehyde needed!
 - do need a dedicated RNA only gel box

Required equipment/associated costs:

- UV spectrophotometer

Key:

Wet lab
module

Computational
module

Wk#2 (wet lab): RNA Quality & Quantity Analysis

completed prior
to course

RNA extraction
For RNA-Seq



Illumina
RNA-Seq



Bioinformatics
analysis of raw
sequence data

week #2

RNA
quantity &
quality

Intro to
R Studio

Visualizing
TopHat
Alignments

- Students input spec data into a GoogleSheet
- Calculate 1 μ g mixture for gel analysis

example: 1 blue	Gg E18 retina	83	from spec	use equation	use equation	3	
Group #/color	sample	ng/ μ L	260/280	μ L for 1 μ g	μ L H ₂ O to bring up to 15 μ L	μ L of 6X gel dye	notes
#3 Green	Gg E8.1 retina	299.172	2.046	3.342559	11.657441	3	
#3 Green	Gg E8.2 retina	397	2.059	2.518892	12.481108	3	
#3 Green	Gg E18.1 retina	86.388	2.029	11.57568	3.42432	3	
#3 Green	Gg E18.2 retina	78.547	2.117	12.73123	2.26877	3	
#4 Orange	Gg E8.1 retina	459.778	2.067	2.174962699	12.8250373	3	
#4 Orange	Gg E8.2 retina	320.299	2.065	3.122082804	11.8779172	3	
#4 Orange	Gg E18.1 retina	29.672	2.113	33.70180642	-18.70180642	5	25 μ L RNA; low quality
#4 Orange	Gg E18.2 retina	8.328	2.708	120.0768492	-105.0768492	5	25 μ L RNA; low quality
#1 Blue	Gg E8 1 retina	111.056	2.065	9.004466215	5.995533785	3	18
#1 Blue	Gg E8 2 retina	495.191	2.072	2.019422809	12.98057719	3	18
#1 Blue	Gg E18 1 retina	223.69	2.06	4.470472529	10.52952747	3	18
#1 Blue	Gg E18 2 retina	265.155	2.039	3.771379005	11.228621	3	18
Pink #5	Gg E18 retina	11.346	2.446	88.13679	-73.13679	6	25 μ L RNA; low quality
Pink #5	Gg E18 retina	14.057	2.375	71.13893	-56.13893	6	25 μ L RNA; low quality
Pink #5	Gg E18 retina	7.153	3.036	139.8015	-124.8015	6	25 μ L RNA; low quality

Required equipment/associated costs:

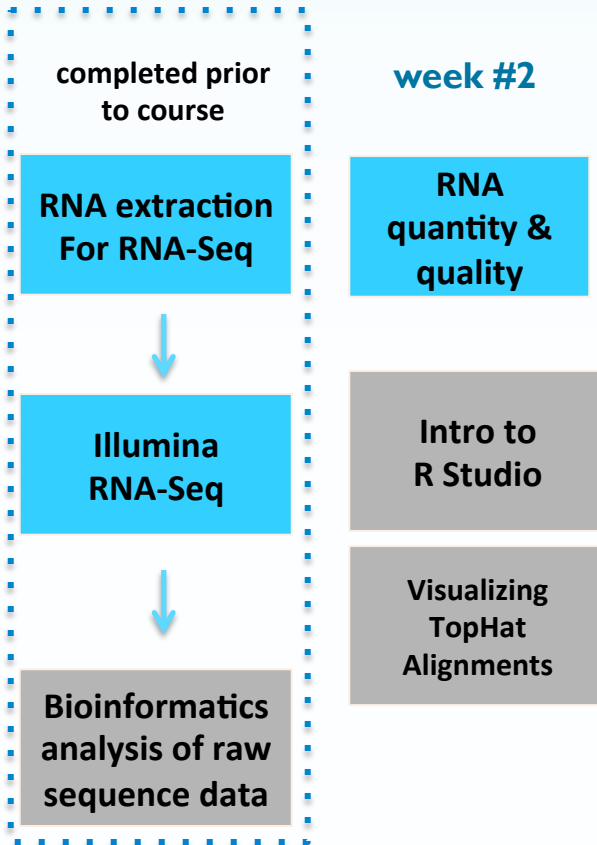
- UV spectrophotometer

Key:

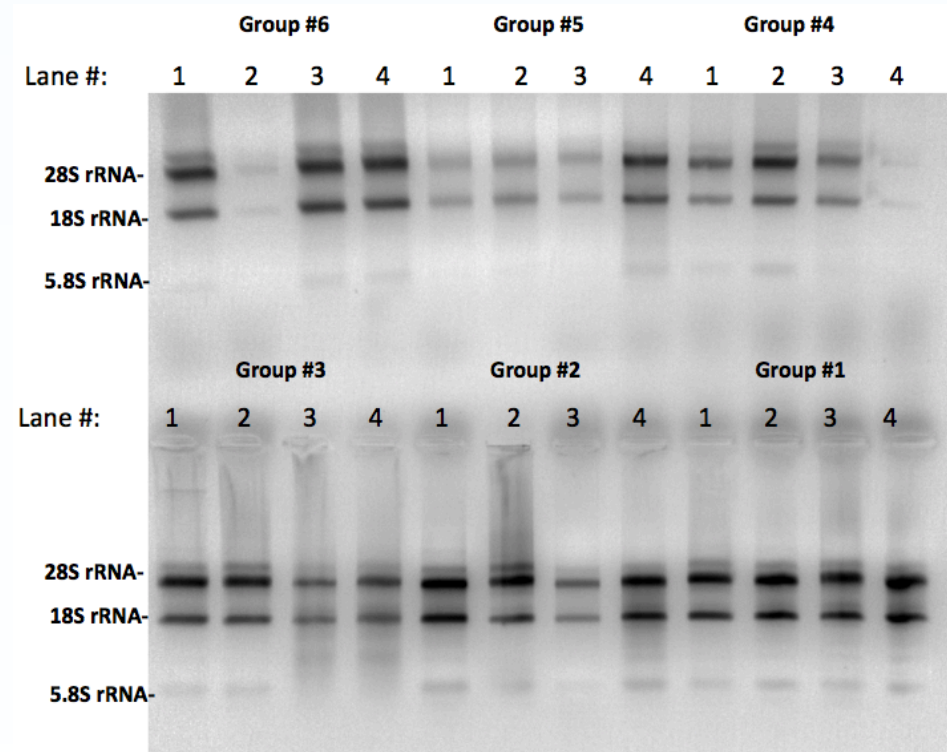
Wet lab
module

Computational
module

Wk#2 (wet lab): RNA Quality & Quantity Analysis



- Regular agarose w/ no formaldehyde!
- Students RNA quality generally looks good!



Required equipment/associated costs:

- Agarose ~\$100
- RNA only gel rig

Key:

Wet lab
module

Computational
module

Wk#2 (computational): Intro to R Studio

completed prior
to course

RNA extraction
For RNA-Seq



Illumina
RNA-Seq



Bioinformatics
analysis of raw
sequence data

week #2

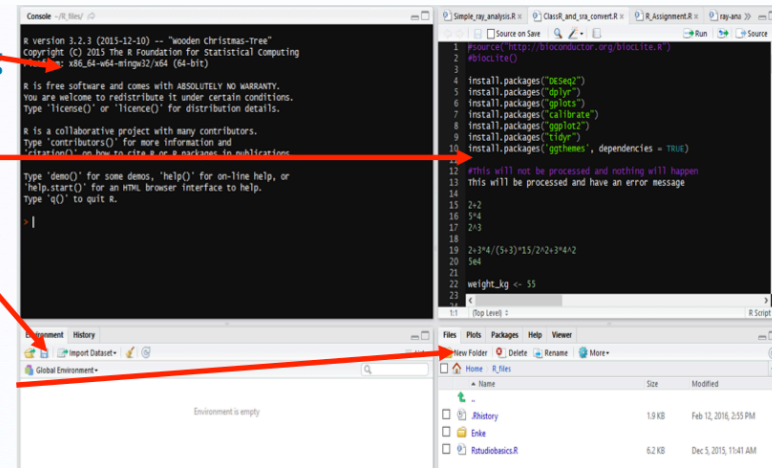
RNA
quantity &
quality

Intro to
R Studio

Visualizing
TopHat
Alignments



1. **Console:** where the processing and computation occurs
2. **Source:** Where you type your code.
3. **Environnment:** list the variables, tables, etc.
4. **Files, Plots, Packages, help, View.** Lists all the files, plots, packages, and help.

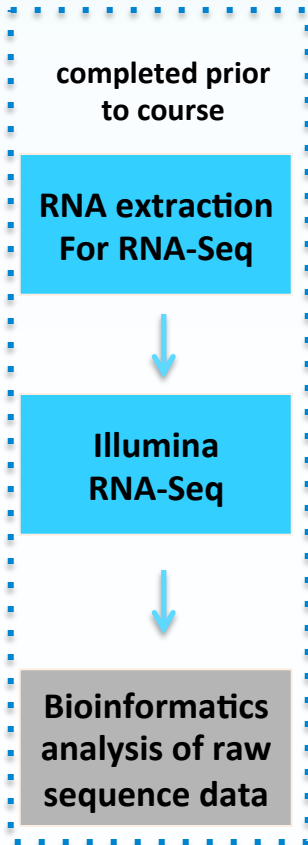


Key:

Wet lab
module

Computational
module

Wk#2 (computational): Visualizing TopHat Alignments in the UCSC Genome Browser (dedicated lecture 6/15 @ 1pm)



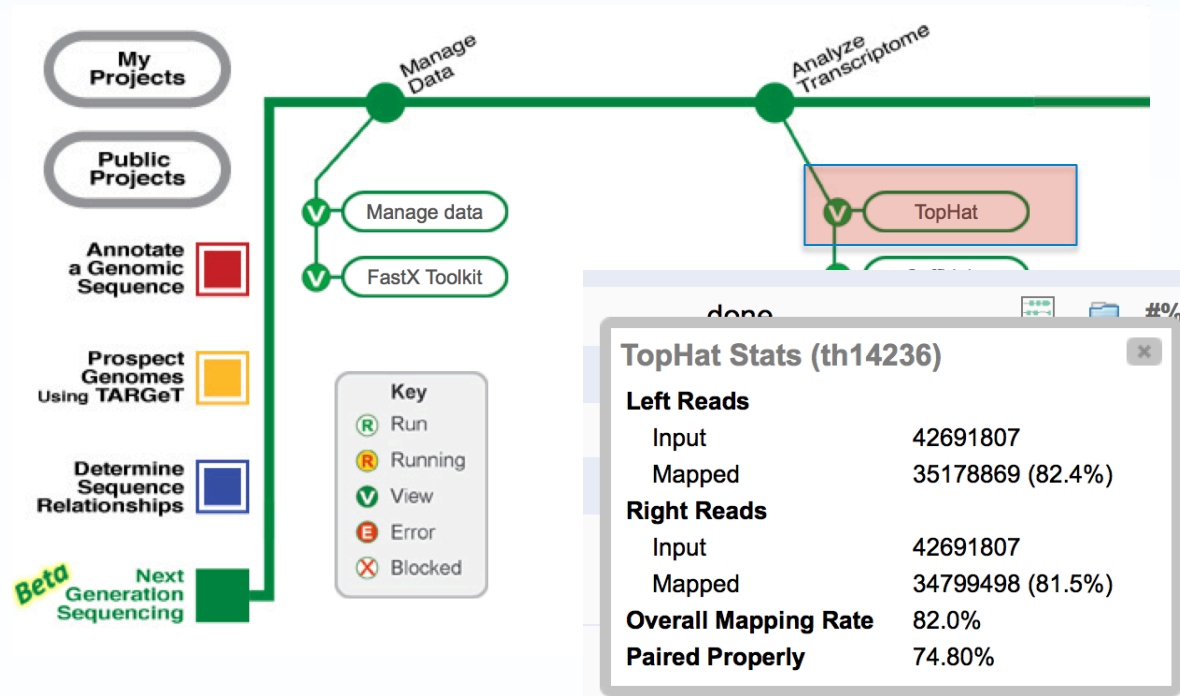
week #2

RNA quantity & quality

Intro to R Studio

Visualizing TopHat Alignments

- TopHat alignments (BAM files) can be visualized in a genome browser (IGV, UCSC, etc)

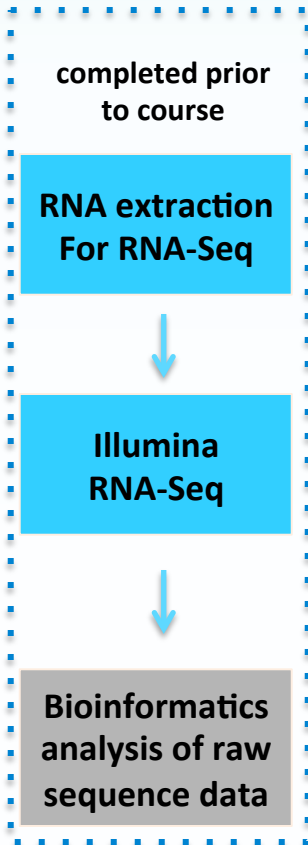


Key:

Wet lab module

Computational module

Wk#2 (computational): Visualizing TopHat Alignments in the UCSC Genome Browser (dedicated lecture 6/15 @ 1pm)



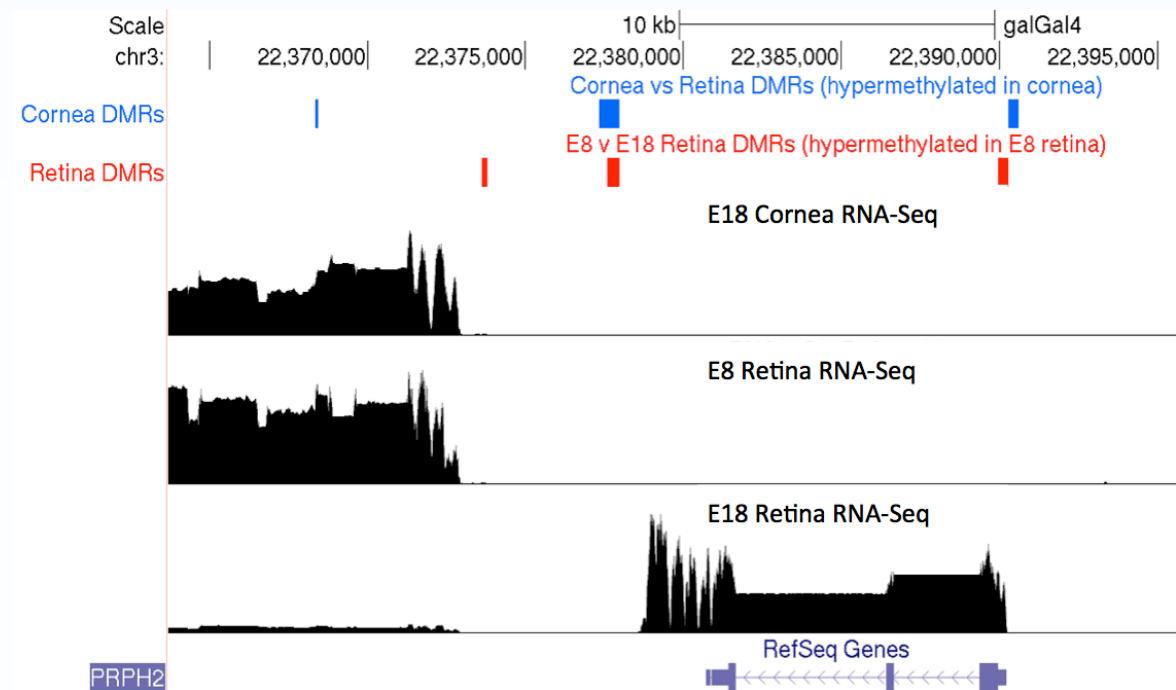
week #2

RNA quantity & quality

Intro to R Studio

Visualizing TopHat Alignments

- TopHat alignments (BAM files) can be visualized in a genome browser (IGV, UCSC, etc)



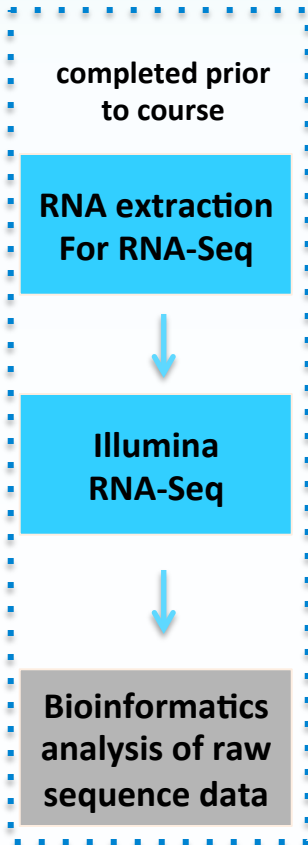
*similar to IGV view of BAM files

Key:

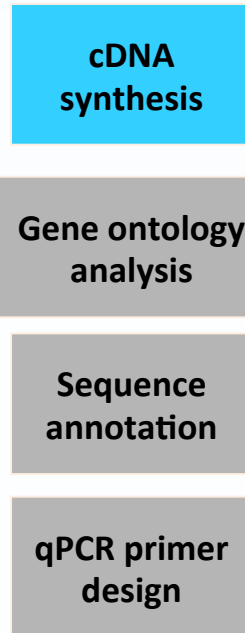
Wet lab module

Computational module

Wk#3 (wet lab): cDNA Synthesis from Total RNA



week #3



- cDNA synthesis using RNAs from wk#1

Groups #1+2			
Tube #	RNA template	ng/μL	for 400 ng
1	E8 retina #1	26.67	15.00
2	E8 retina #2	26.67	15.00
3	E8 retina #3	26.67	15.00
*4	E8 retina #3 (no RT)	26.67	15.00
5	E18 retina #1	26.67	15.00
6	E18 retina #2	26.67	15.00
7	E18 retina #3	26.67	15.00
*8	E18 retina #3 (no RT)	26.67	15.00

RT reaction master mix:

per reaction	ingredient	X7
4 uL	5X buffer	28 uL
1 uL	RT	7 uL

no RT mix:

per reaction	ingredient	X3
4 uL	5X buffer	12 uL
1 uL	H2O	3 uL

Add 5 uL of RT mix to 15 uL RNA in tubes #1-3 & #5-7 and keep on ice. Add 5 uL of no RT mix to 15 uL RNA in tubes #4 & 8 and keep on ice.

Required equipment/associated costs:

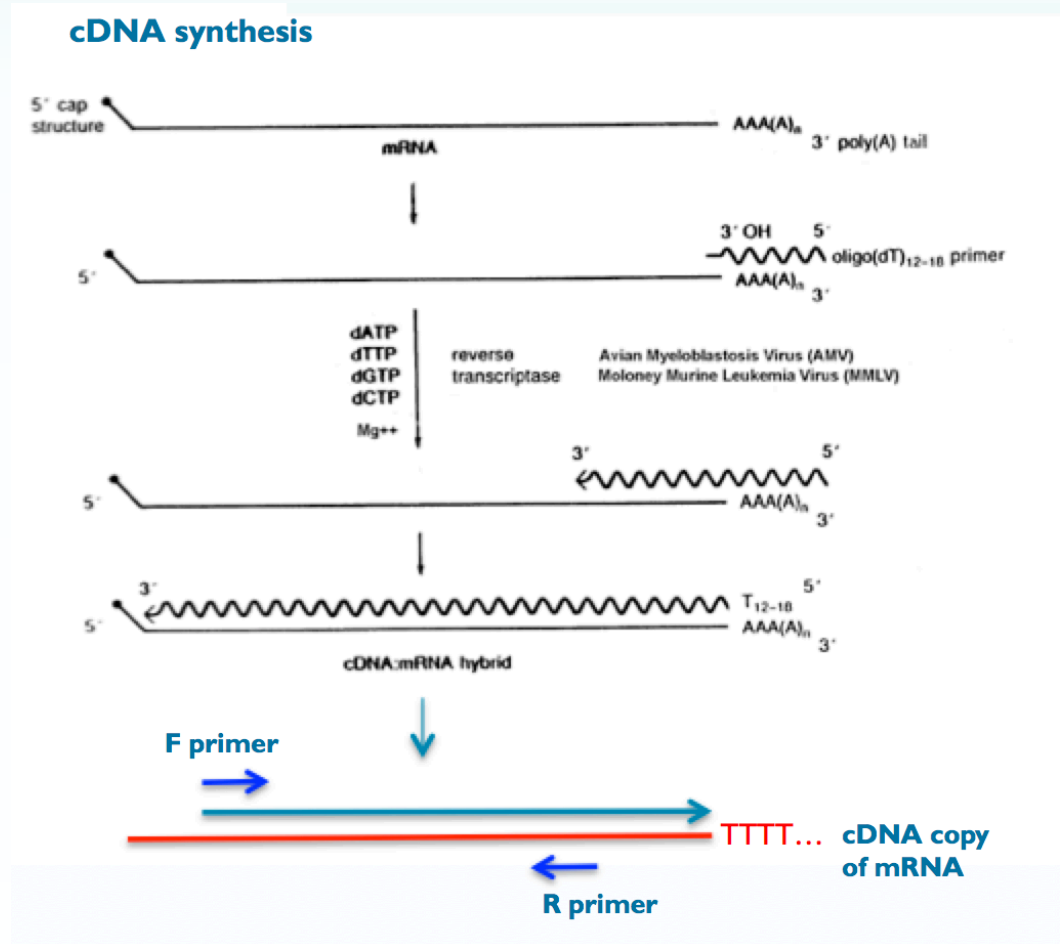
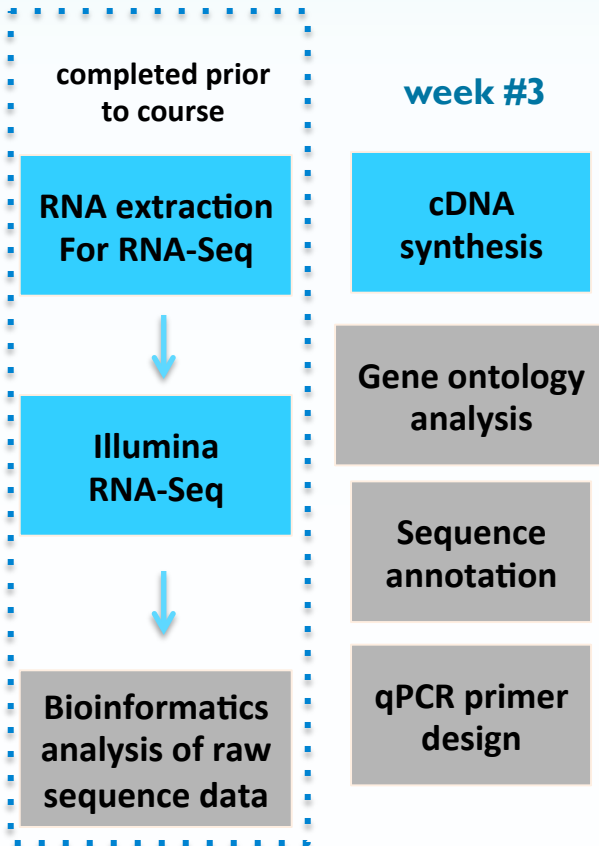
- iScript cDNA synthesis kit ~\$340/100 rctns
- Conventional PCR cycler

Key:

Wet lab module

Computational module

Wk#3 (wet lab): cDNA Synthesis from Total RNA



Required equipment/associated costs:

- iScript cDNA synthesis kit ~\$340/100 rxns
- Conventional PCR cycler

Key:

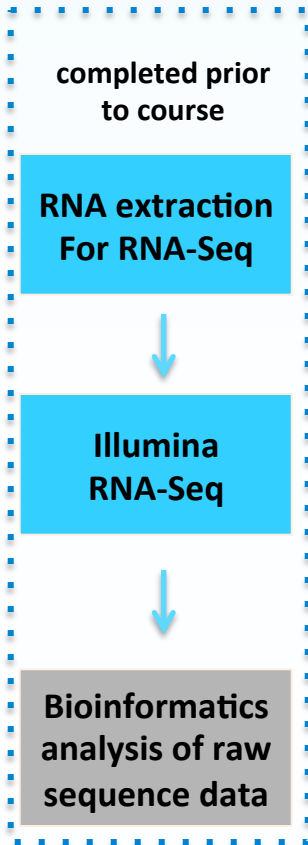
Wet lab
module

Computational
module

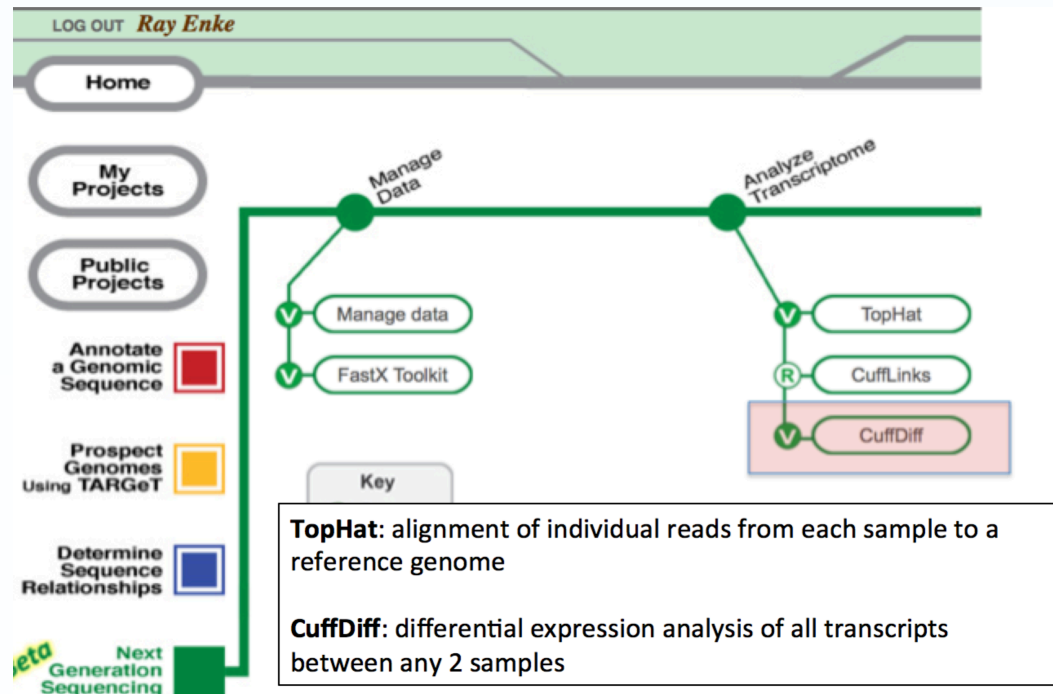
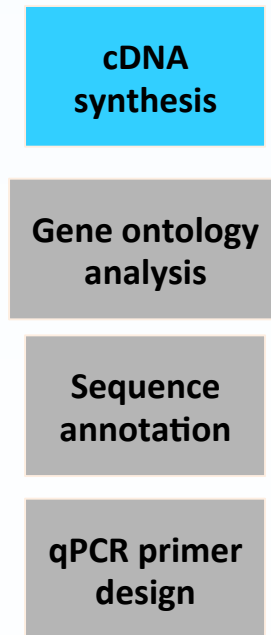
Wk#3 (computational): GO Analysis of DEGs

GO=gene ontology; DEG= differentially expressed genes

- Students navigate to and download CuffDiff data from Green Line



week #3



Key:

Wet lab module

Computational module

Wk#3 (computational): GO Analysis of DEGs

GO=gene ontology; DEG= differentially expressed genes

- Students navigate to and download CuffDiff data from Green Line

Cuffdiff

cd14319= E8 retina vs E18 retina DEGs (groups #3-6)
cd14319= E18 cornea vs E18 retina DEGs (groups #1-2)

Gene summary tables

#	Job Name	Start Time	End Time	Status	View Results
1	cd14319	2016-02-18T11:46:55.394-06:00	2016-02-18T13:05:58.000-06:00	done	
2	cd14321	2016-02-18T11:51:06.094-06:00		error	
3	cd14322	2016-02-18T21:04:06.978-06:00	2016-02-18T22:42:06.000-06:00	done	

Cuffdiff

View: Genes | EB_retina | E18_retina | Go

Gene summary table:
cd14319= E8 retina vs E18 retina DEGs

Export data to spreadsheet

Transcripts sorted by Q-Value

Search:

Transcript	Nearest Ref Id	Gene	Alias	Fold Change	Direction	Sample 1 FPKM	Sample 2 FPKM	Q-Value	Description
ENSGALT00000000003	-	ENSGALT00000000003	PANX2	2.16	UP	9.25776	19.9978	0.0001	
ENSGALT00000000026	-	ENSGALT00000000019	AAK1	3.89	UP	18.9927	73.8016	0.0001	
ENSGALT00000000064	-	ENSGALT00000000049	CSDC2	3.10	UP	44.2807	137.052	0.0001	
ENSGALT00000000070	-	ENSGALT00000000055	LAMTOR3	1.74	UP	80.5424	140.282	0.0001	
ENSGALT00000000071	-	ENSGALT00000000056	DAPP1	8.58	DOWN	10.8211	1.2606	0.0001	
ENSGALT00000000074	-	ENSGALT00000000059	TUBB3	3.10	DOWN	869.907	280.966	0.0001	
ENSGALT00000000104	-	ENSGALT00000000076	NRBP2	3.23	UP	2.94817	9.53326	0.0001	
ENSGALT00000000141	-	ENSGALT00000000102	UBE2T	9.17	DOWN	65.529	7.14856	0.0001	

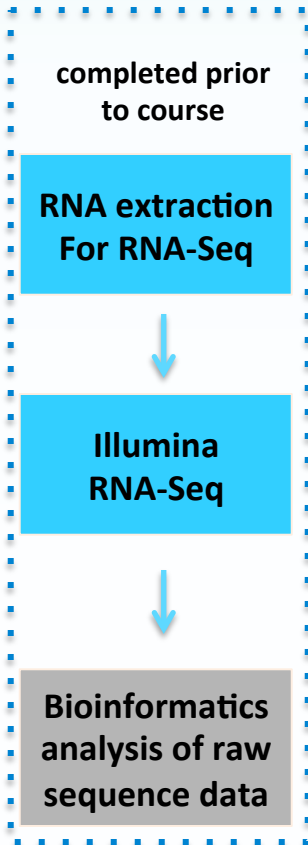
Key:

Wet lab
module

Computational
module

Wk#3 (computational): GO Analysis of DEGs

GO=gene ontology; DEG= differentially expressed genes



week #3

cDNA synthesis

Gene ontology analysis

Sequence annotation

qPCR primer design

- Students sort the DEG data by biological & statistical significance

Assignment

1. How many total DEGs in the unfiltered CuffDiff data?
2. How many biologically & statistically significant up regulated DEGs are there in E18 retina?
3. What's the highest biologically & statistically significant up regulated gene in E18 retina? How much up regulated?
4. How many biologically & statistically significant down regulated DEGs are there?
5. What's the most decreased biologically & statistically significant down regulated gene in E18 retina? How much down regulated?
6. Add data table to your detailed notebook page

Cuffdiff

View: Genes | E8_retina | E18_retina | Go

Gene summary table:
cd14319= E8 retina vs E18 retina DEGs

Export data to spreadsheet

Transcripts sorted by Q-Value

Search:

Transcript	Nearest Ref Id	Gene	Alias	Fold Change	Direction	Sample 1 FPKM	Sample 2 FPKM	Q-Value	Description
ENSGALT00000000003	-	ENSGALG00000000003	PANX2	2.16	UP	9.25776	19.9978	0.0001	
ENSGALT00000000026	-	ENSGALG00000000019	AAK1	3.89	UP	18.9927	73.8016	0.0001	
ENSGALT00000000064	-	ENSGALG00000000049	CSDC2	3.10	UP	44.2807	137.052	0.0001	
ENSGALT00000000070	-	ENSGALG00000000055	LAMTOR3	1.74	UP	80.5424	140.282	0.0001	
ENSGALT00000000071	-	ENSGALG00000000056	DAPP1	8.58	DOWN	10.8211	1.2606	0.0001	
ENSGALT00000000074	-	ENSGALG00000000059	TUBB3	3.10	DOWN	869.907	280.966	0.0001	
ENSGALT00000000104	-	ENSGALG00000000076	NRBP2	3.23	UP	2.94817	9.53326	0.0001	
ENSGALT00000000141	-	ENSGALG00000000102	UBE2T	9.17	DOWN	65.529	7.14856	0.0001	

Key:

Wet lab
module

Computational
module

RNA-Seq: Cuffdiff output

- 1,416 significant (<0.05) downregulated genes sorted by log2 fold change

	A	B	C	D	E	F	G
1	gene_id	associated gene name	locus	E8 retina FPKM	E18 retina FPKM	log2 (fold change E8:E18)	p_value
2	ENSGALT00000037609	SMC2	Z:66288452-66312320	138.159	0.144311	-9.90293	5.00E-05
3	ENSGALT00000025338	SLBP	4:83495205-83502378	42.2289	0.0659106	-9.32351	0.0005
4	ENSGALT00000013400	BUB1	3:3037096-3057283	31.3715	0.0493455	-9.31232	5.00E-05
5	ENSGALT00000016208	CDC20	8:18577131-18584614	104.17	0.176012	-9.20906	0.0007
6	ENSGALT00000023752	DEPDC1B	Z:18712424-18733642	14.4097	0.0265575	-9.08371	0.00035
7	ENSGALT00000021577	CENPE	4:60597387-60634806	20.5329	0.0399888	-9.00413	5.00E-05
8	ENSGALT00000009864	PLK1	14:6884684-6889557	138.655	0.279497	-8.95446	5.00E-05
9	ENSGALT00000043007	UBE2C	20:10853788-10854730	318.565	0.959296	-8.37539	5.00E-05
10	ENSGALT00000026525	RRM2	3:96086207-96107950	175.158	0.548728	-8.31835	5.00E-05
11	ENSGALT00000027509	CKAP2	1:169701517-169710954	167.704	0.965468	-7.44048	5.00E-05
12	ENSGALT00000025333	TACC3	4:83447127-83469304	111.934	0.675251	-7.37301	5.00E-05
13	ENSGALT00000013978	NUSAP1	5:23859403-23873313	92.6816	0.585096	-7.30746	5.00E-05
14	ENSGALT00000043356	MKI67	6:32889675-32910556	16.1344	0.112181	-7.16817	5.00E-05
15	ENSGALT00000032863	CENPH	Z:21769912-21776084	105.188	0.763317	-7.10647	5.00E-05
16	ENSGALT00000044727	ZIC2	1:143534165-143535092	59.1463	0.446545	-7.04934	5.00E-05
17	ENSGALT00000010135	RACGAP1	22C19W28_E50C23:335215-342	57.6397	0.452772	-6.99214	5.00E-05
18	XLOC_018967	FANCF	5:2945015-2947340	7.29388	0.0618329	-6.88217	0.0009
19	ENSGALT00000006667	KIF4A	4:1248685-1266250	83.2927	0.715993	-6.8621	5.00E-05
20	ENSGALT00000027693	CENPJ	1:178533779-178551434	2.99984	0.0284654	-6.71953	0.00025

Lots of DEGs, what's next?

- Students parse out CuffDiff list into separate list of significant up & down regulated genes

RNA-Seq: Cuffdiff output

- 1,416 significant (<0.05) downregulated genes sorted by log2 fold change

	A	B	C	D	E	F	G
1	gene_id	associated gene name	locus	E8 retina FPKM	E18 retina FPKM	log2 (fold change E8:E18)	p_value
2	ENSGALT00000037609	SMC2	Z:66288452-66312320	138.159	0.144311	-9.90293	5.00E-05
3	ENSGALT00000025338	SLBP	4:83495205-83502378	42.2289	0.0659106	-9.32351	0.0005
4	ENSGALT00000013400	BUB1	3:3037096-3057283	31.3715	0.0493455	-9.31232	5.00E-05
5	ENSGALT00000016208	CDC20	8:18577131-18584614	104.17	0.176012	-9.20906	0.0007
6	ENSGALT00000023752	DEPDC1B	Z:18712424-18733642	14.4097	0.0265575	-9.08371	0.00035
7	ENSGALT00000021577	CENPE	4:60597387-60634806	20.5329	0.0399888	-9.00413	5.00E-05
8	ENSGALT00000009864	PLK1	14:6884684-6889557	138.655	0.279497	-8.95446	5.00E-05
9	ENSGALT00000043007	UBE2C	20:10853788-10854730	318.565	0.959296	-8.37539	5.00E-05
10	ENSGALT00000026525	RRM2	3:96086207-96107950	175.158	0.548728	-8.31835	5.00E-05
11	ENSGALT00000027509	CKAP2	1:169701517-169710954	167.704	0.965468	-7.44048	5.00E-05
12	ENSGALT00000025333	TACC3	4:83447127-83469304	111.934	0.675251	-7.37301	5.00E-05
13	ENSGALT00000013978	NUSAP1	5:23859403-23873313	92.6816	0.585096	-7.30746	5.00E-05
14	ENSGALT00000043356	MKI67	6:32889675-32910556	16.1344	0.112181	-7.16817	5.00E-05
15	ENSGALT00000032863	CENPH	Z:21769912-21776084	105.188	0.763317	-7.10647	5.00E-05
16	ENSGALT00000044727	ZIC2	1:143534165-143535092	59.1463	0.446545	-7.04934	5.00E-05
17	ENSGALT00000010135	RACGAP1	22C19W28_E50C23:335215-342	57.6397	0.452772	-6.99214	5.00E-05
18	XLOC_018967	FANCF	5:2945015-2947340	7.29388	0.0618329	-6.88217	0.0009
19	ENSGALT00000006667	KIF4A	4:1248685-1266250	83.2927	0.715993	-6.8621	5.00E-05
20	ENSGALT00000027693	CENPJ	1:178533779-178551434	2.99984	0.0284654	-6.71953	0.00025

Lots of candidates, what next?

- What's the full gene name?
- What's the associated function of the gene?

GO analysis using Ensembl Biomart

1	gene_id	gene	locus	E8 retina FPKM	E18 retina FPKM	log2 (fold_change E8:E18)	p_value
954	ENSGALT0000001637	NFKBIE	3:29611731-29620288	6.47794	12.2766	0.922305	0.0007
955	ENSGALT0000003803	GARNL3	17:10400898-10448322	7.83474	14.8297	0.920531	0.01405
956	ENSGALT0000000003	RHOBTB2	22:1277203-1295625	8.432	15.873	0.913204	0.00045
957	ENSGALT0000000462	TMEM55A	2:124512881-124538359	12.4378	23.4139	0.912631	0.0007
958	ENSGALT0000000282	ENSGALT00000017553	27:3836211-3842176	12.6749	23.8305	0.910833	0.0048
959	ENSGALT0000000002	THRA	27:4342651-4348356	23.2756	43.7479	0.910392	0.0003
960	ENSGALT0000000391	CD151	5:15037217-15068612	9.37814	17.5764	0.906269	0.00015
961	ENSGALT0000000297	ENSGALT00000018781	ADN030039360:16747-7970	5.68679	10.6397	0.903774	0.01395
962	XLOC_022183	SLC35G2	9:361086-373502	17.3758	32.4342	0.900437	0.0017
963	ENSGALT0000000259	RNF103	4:85175054-85192537	13.1117	24.2797	0.888897	0.00025
964	ENSGALT0000000397	SPTAN1	17:5139816-5193697	26.0995	48.3296	0.888882	5.00E-05
965	ENSGALT0000000008	C2H3ORF75	2:851260-867494	7.30864	13.5215	0.887583	0.0027
966	ENSGALT0000000072	DRAXIN	21:5645769-5657879	7.54955	13.9259	0.883309	0.00155
967	ENSGALT0000000318	HIGD1A	2:2061654-2066620	43.0516	79.3184	0.88159	5.00E-05
968	ENSGALT0000000317	FBXO48	3:11240339-11243704	4.26504	7.84938	0.880019	0.0102
969	ENSGALT0000000438	SLC35F1	3:62486815-62723122	18.9053	34.7759	0.879297	5.00E-05
970	ENSGALT0000000123	EFEMP1	3:225183-277692	5.22139	9.59035	0.87715	0.00105
971	XLOC_003103	ENSGALT00000016249	1:112077952-112079696	11.2323	20.5953	0.874665	0.00125
972	ENSGALT0000000169	GLO1	3:28858537-28865172	123.985	227.285	0.874341	5.00E-05
973	ENSGALT0000000436	MEI30	2:134983353-135002167	4.0047	7.33202	0.872517	0.00095
974	ENSGALT0000000050	CLTC	19:7283352-7310413	38.0908	69.7291	0.872319	5.00E-05
975	ENSGALT0000000452	DNAIC27	3:105062346-105072565	6.6868	12.2269	0.870674	0.00375
976	ENSGALT0000000001	STX12	23:1410226-1422896	44.7438	81.1303	0.858552	5.00E-05
977	ENSGALT0000000008	SLC04A1	20:8527769-8555187	3.03701	5.50518	0.858135	0.0063
978	XLOC_000291	gga-let-7i	1:33018777-33022900	1.81779	3.29149	0.856554	0.00425
979	ENSGALT0000000174	GOLGA5	5:44104939-44121283	5.80064	10.4819	0.853615	0.00085
980	ENSGALT0000000443	ENSGALT00000026654	13:16565757-16577228	7.8578	14.1768	0.851335	0.0087
981	XLOC_015476	CDC42EP3	3:31127956-31154514	8.85959	15.9356	0.846941	0.00125
982	XLOC_000221	GPR85	1:26614541-26640975	13.9272	25.0222	0.845305	0.0006
983	ENSGALT0000000121	SLC9A3R1	18:10711116-10718410	88.2415	158.102	0.841329	5.00E-05
984	ENSGALT0000000396	PGAM1	6:21893887-21895347	478.442	856.169	0.839551	5.00E-05
985	ENSGALT0000000009	XBP1	15:7713030-7716156	49.3108	88.2253	0.839288	0.0026
986	ENSGALT0000000009	ABTB1	12:9857434-9865180	7.01307	12.5255	0.836752	0.0015
987	ENSGALT0000000005	EGFLAM	2:11778981-11853316	51.758	92.367	0.835596	5.00E-05
988	ENSGALT0000000019	TMEM222	23:1913664-1920499	11.3426	20.2143	0.833631	0.0012
989	ENSGALT0000000324	LRRCS7	5:2502022-25024747	11.2078	19.9563	0.832344	0.0009
990	XLOC_017582	GPRIN3	4:34687587-34723721	5.59993	9.96474	0.831422	0.00015
991	XLOC_022582	ENSGALT00000024369	9:3031932-3037191	13.9929	24.8781	0.830175	0.0001
992	ENSGALT0000000147	PIGK	8:18082727-18148014	25.0446	44.4455	0.827539	0.00025
993	ENSGALT0000000201	BOH2	4:60581710-60595813	2.25257	3.9958	0.826912	0.00775
994	ENSGALT0000000202	SYN3	1:52867084-53186826	3.36118	5.94407	0.822482	0.0099
995	ENSGALT0000000190	BBS9	2:48355516-48554134	3.23873	5.70227	0.816108	0.01315
996	ENSGALT0000000006	CHST12	14:3136282-3137952	3.61126	6.32278	0.808055	0.01125
997	ENSGALT0000000446	AMPD2	ADN03013667:110-3184	14.4499	25.2649	0.806069	0.0031
998	XLOC_023904	masRNA-menRNA	ADN03016580:11101-6473	455.562	795.798	0.804756	5.00E-05
999	ENSGALT0000000012	MLLT11	25:1746787-1963328	76.9278	134.108	0.801819	0.01355
1000	ENSGALT0000000006	ENOX2	4:1508589-1532529	11.568	20.045	0.793106	0.0115
1001	ENSGALT0000000126	FAM177A1	5:35371255-35397675	25.3797	43.969	0.792812	0.00145
1002	XLOC_017857	SNORA26	4:65582744-65589564	7.90798	13.698	0.792581	0.0141
1003	ENSGALT0000000280	ENSGALT00000017589	25:475092-478012	11.6658	20.1761	0.790366	0.0025
1004	ENSGALT0000000181	TAX1BP1	2:33445896-33496509	30.0595	51.8669	0.786996	0.0003
1005	ENSGALT0000000408	ULK3	10:1911381-1917110	5.47033	9.42648	0.785091	0.01235

Animals:

<http://www.ensembl.org/biomart/>

Plants:

<http://plants.ensembl.org/biomart/martview/b95ebad3df7cde39d430b6b8efedd2ae>

Protists:

<http://protists.ensembl.org/biomart/martview/ac3930f7a0a4a426957b0afd6375af5c>

Fungi:

<http://fungi.ensembl.org/biomart/martview/aa9629f5dcf60a90a57c15a527c304cc>

- Filter associated gene names in cuffdiff output through Biomart to obtain full gene names and GO terms

GO analysis using Ensembl Biomart

The screenshot displays the Ensembl Biomart interface. On the left, a table lists genes with columns for 'gene_id' and 'gene'. A red box highlights the 'gene' column, which contains gene names like NFKBIE, GARNL3, RHOBTB2, TMEM55A, etc. The main panel shows the 'Dataset' as 'Gallus gallus genes (Galgal4)' and 'Filters' as 'Associated Gene Name(s) [e.g. BRCA2]: [ID-list specified]'. The 'Attributes' section is set to 'Ensembl Gene ID' and 'Ensembl Transcript ID'. The 'Dataset' is currently '[None Selected]'. The 'REGION' and 'GENE' sections are empty. The 'Limit to genes (external references)...' checkbox is unchecked. The 'Input external references ID list [Max 500 advised]' checkbox is checked. A red box highlights the 'Associated Gene Name(s) [e.g. BRCA2]' search box, which contains a list of gene names: ENSGALG00000002638, ENSGALG000000026161, SOX10, NECAB2, ENSGALG00000014441, SS2, MAFA, DRD4, NXP2, EGR1, CALB1, ARHGEF9, ELFN1, WNT10A, HS6ST3, FADS6, SNTB1, ENSGALG00000019291. The 'Browse...' button is visible below the search box.

- Filter associated gene names in cuffdiff output through Biomart to obtain full gene names and GO terms

GO analysis using Ensembl Biomart

The screenshot displays the Ensembl Biomart interface. On the left, a table lists genes from the 'Gallus gallus genes (Galgal4)' dataset. A red box highlights the 'Associated Gene Name(s)' column, which contains gene names like NFKBIE, GARNL3, RHOBTB2, TMEM55A, etc. The right panel shows the 'Please select columns to be included in the output' dialog. Under the 'Ensembl' section, 'Description' is selected. Under the 'Phenotype' section, 'Phenotype description' and 'Source name' are selected. Under the 'EXTERNAL' section, 'GO Term Name' is selected. The 'Associated Gene Name' checkbox is also checked under the 'Ensembl' section.

Dataset
Gallus gallus genes (Galgal4)

Filters
Associated Gene Name(s)
[e.g. BRCA2]: [ID-list specified]

Attributes
Description
Associated Gene Name
GO Term Name

Dataset
[None Selected]

Please select columns to be included in the output and hit 'Results' when

Features ☐ **Variation (Germline)** ☐
☐ **Structures** ☐ **Variation (Somatic)** ☐
☐ **Homologs** ☐ **Sequences** ☐

GENE:

Ensembl

- ☐ Ensembl Gene ID
- ☐ Ensembl Transcript ID
- ☐ Ensembl Protein ID
- ☐ Ensembl Exon ID
- ☒ **Description**
- ☐ Chromosome Name
- ☐ Gene Start (bp)
- ☐ Gene End (bp)
- ☐ Strand
- ☐ Band
- ☐ Transcript Start (bp)
- ☐ Transcript End (bp)
- ☐ Transcription Start Site (TSS)
- ☐ Transcript length

Phenotype

- ☐ Phenotype description
- ☐ Source name

EXTERNAL:

GO

- ☐ GO Term Accession
- ☒ **GO Term Name**
- ☐ GO Term Definition

☒ **Associated Gene Name**

- ☐ Associated Gene Source
- ☐ Associated Transcript Name
- ☐ Associated Transcript Source
- ☐ Transcript count
- ☐ % GC content
- ☐ Gene type
- ☐ Transcript type
- ☐ Source (gene)
- ☐ Source (transcript)
- ☐ Status (gene)
- ☐ Status (transcript)
- ☐ Version (gene)
- ☐ Version (transcript)

☐ Study External Reference

- Filter associated gene names in cuffdiff output through Biomart to obtain full gene names and GO terms

GO analysis using Ensembl Biomart

	A	B	C	D
1	Description	Associated Gene Name	GO Term Name	
39	clusterin-like 1 (retinal) [Source:HGNC Symbol;Acc:HGNC:2096]	CLUL1	positive regulation of neuron death	
40	clusterin-like 1 (retinal) [Source:HGNC Symbol;Acc:HGNC:2096]	CLUL1	positive regulation of beta-amyloid formation	
41	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	transporter activity	
42	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	transport	
43	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	lipid binding	
44	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	extracellular vesicular exosome	
45	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	cholesterol binding	
46	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	fatty acid binding	
47	Gallus gallus peripheral myelin protein 2 (PMP2), mRNA. [Source:RefSeq]	PMP2	membrane organization	
48	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	methyltransferase activity	
49	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	O-methyltransferase activity	
50	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	melatonin biosynthetic process	
51	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	methylation	
52	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	identical protein binding	
53	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	protein homodimerization activity	
54	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	acetylserotonin O-methyltransferase activity	
55	Gallus gallus acetylserotonin O-methyltransferase (ASMT), mRNA. [Source:RefSeq]	ASMT	negative regulation of male gonad development	
56	Gallus gallus SOUL protein (SOUL), mRNA. [Source:RefSeq mRNA;Acc:RefSeq]	SOUL		
57	ethanolamine-phosphate phospho-lyase [Source:HGNC Symbol;Acc:HGNC:2096]	ETNPPL	catalytic activity	
58	ethanolamine-phosphate phospho-lyase [Source:HGNC Symbol;Acc:HGNC:2096]	ETNPPL	metabolic process	
59	ethanolamine-phosphate phospho-lyase [Source:HGNC Symbol;Acc:HGNC:2096]	ETNPPL	transaminase activity	
60	ethanolamine-phosphate phospho-lyase [Source:HGNC Symbol;Acc:HGNC:2096]	ETNPPL	pyridoxal phosphate binding	
61	dopamine receptor D5 [Source:HGNC Symbol;Acc:HGNC:3026]	DRD5	signal transducer activity	
62	dopamine receptor D5 [Source:HGNC Symbol;Acc:HGNC:3026]	DRD5	G-protein coupled receptor activity	
63	dopamine receptor D5 [Source:HGNC Symbol;Acc:HGNC:3026]	DRD5	dopamine neurotransmitter receptor activity	

- Search or sort spreadsheet for your favorite GO terms

GO analysis using Ensembl Biomart

1	Description	Associated Gene Name	GO Term Name
90	dopamine receptor D5 [Source:HGNC Symbol;Acc:HGNC:3026]	DRD5	regulation of female receptivity
91	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	protein binding
92	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	extracellular space
93	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	extracellular vesicular exosome
94	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	extracellular region
95	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	identical protein binding
96	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	extracellular matrix organization
97	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	basement membrane
98	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	behavioral response to pain
99	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	extracellular matrix
100	von Willebrand factor A domain containing 1 [Source:HGNC Symbol;Acc:HGNC:3026]	VWA1	interstitial matrix
101	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	signal transducer activity
102	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	G-protein coupled receptor activity
103	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	signal transduction
104	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	G-protein coupled receptor signaling pathway
105	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	phototransduction
106	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	integral component of membrane
107	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	visual perception
108	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	photoreceptor activity
109	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	membrane
110	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	protein-chromophore linkage
111	Gallus gallus rhodopsin (RHO), mRNA. [Source:RefSeq mRNA;Acc:NM_001001001]	RHO	response to stimulus

- Search or sort spreadsheet for your favorite GO terms (e.g. “phototransduction”)

GO analysis using Ensembl Biomart

	A	B	C	D	E	F
1	<u>group #</u>	<u>designed by</u>	<u>Gene Symbol</u>	<u>GO Term</u>	<u>Fold change</u>	<u>notes</u>
2	<i>example</i>	<i>Dr. Enke</i>	<i>GPR98</i>	<i>phototransduction</i>	6.02272	<i>This gene is awesome!</i>
3	Blue #1	CHD	NEUROD4	Notch signaling pathway	754.58	
4	Blue #1	JJD	RGR	phototransduction	43.27	
5	Blue #1	JPN	OPN4	phototransduction	301.73	
6	Red #2	Laura Morales	GUCA1A	Photoreceptor	110.48	
7	Red #2	Monique Waldman	PDE6C	phototransduction, visible light	509.76	
8	Red #2	Paul Roberson	GOT1	Notch Signaling Pathway	2.61	
9	Green #3	Se Chang	Sox9	Notch signaling pathway	4.56	
10	Green #3	Aiola Stoja	DLL4	Notch signaling pathway	19.93	
11	Green #3	Juan Rittgers	HES5	Notch Signalling Pathway	19.09	
12	Orange #4	Kristin McKenna	LFNG	Notch Signalling Pathway	15.55	
13	Orange #4	Kevin Decker	OPN3	phototransduction	14.2	
14	Orange #4	Zach Schuhmacher	ASIC2	phototransduction	8.5	
15	Pink #5	BWL	OPN4-5	phototransduction	28.4918	
16	Pink #5	AHL	cdk6	Notch Signalling Pathway	6.739	
17	Pink #5	TPL	OPN1MSW	phototransduction		
18	Yellow #6	CAH	PDE6C	phototransduction, visible light	484.52	
19	Yellow #6	BHY	NEURL	Notch Signalling Pathway	73.54	
20	Yellow #6	EAT	MIB1	notch signaling	5.49	

- Each student chooses 1 significant DEG in the “phototransduction” or “Notch” pathway for downstream analysis
- Populate GoogleSheet with relevant information from CuffDiff & GO analysis


```
graph TD; A[completed prior to course] --> B[RNA extraction For RNA-Seq]; B --> C[Illumina RNA-Seq]; C --> D[Bioinformatics analysis of raw sequence data];
```

completed prior to course

**RNA extraction
For RNA-Seq**

**Illumina
RNA-Seq**

**Bioinformatics
analysis of raw
sequence data**

UCSC Genome Browser on Chicken Nov. 2011 (ICGSC Gallus_gallus-4.0/galGal4)

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

chr12:19,501,119-19,506,117 4,999 bp.

chr12 12

Scale chr12: 19,501,500 | 19,502,000 | 19,502,500 | 19,503,000 | 19,503,500 | 19,504,000 | 19,504,500 | 19,505,000 | 1 kb galGal4

1323 _

1000 -

E18_Retina_mRNA

E18_Retina

0 _

RefSeq Genes

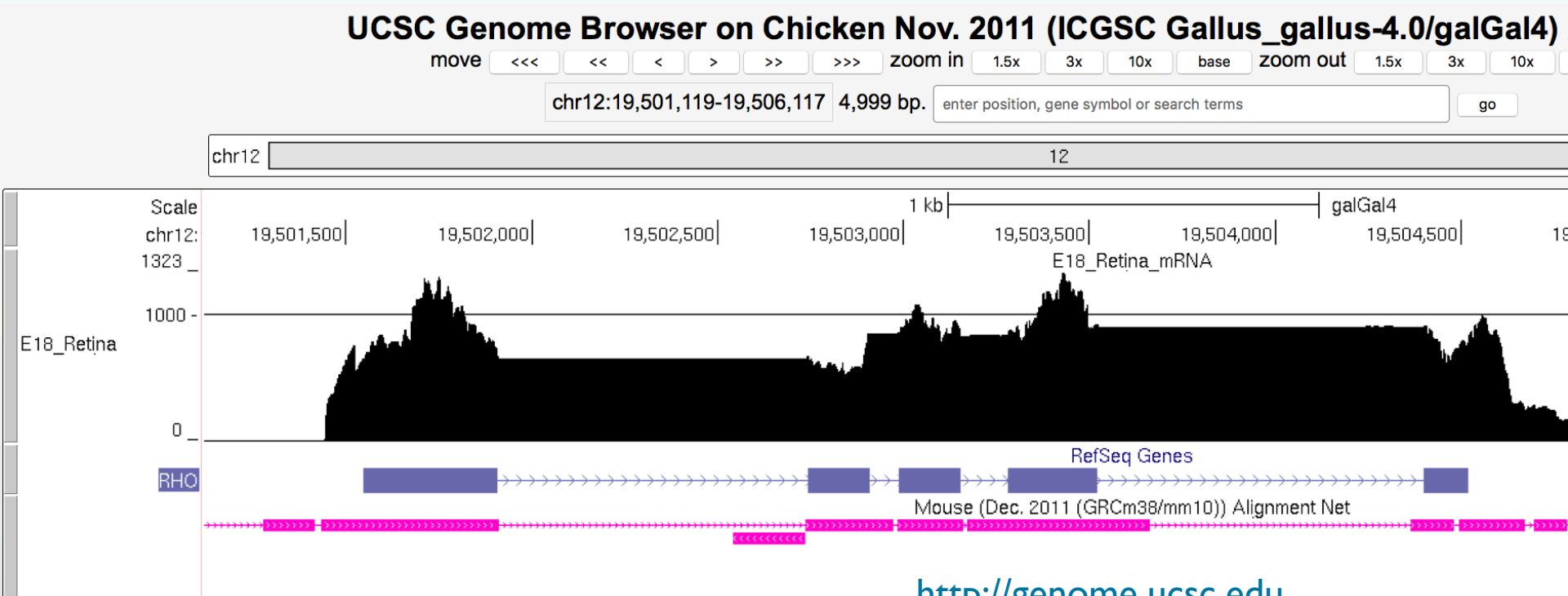
RHO

Mouse (Dec. 2011 (GRCm38/mm10)) Alignment Net

[illegible]

Computational module

Wk#3 (computational): Sequence Retrieval & Annotation



UCSC Genome Browser:

- Highly intuitive genome browser and sequence repository for many genomes (no plant genomes unfortunately; try Ensembl)

Wk#3 (computational): Sequence Retrieval & Annotation

Rho mRNA

>NM_001030606.1

atgaacgggacggaaggccaagacttctacgtgccatgtccaacaagac
cggggtggtgcggagccccttcgagtacccccagtactacctggtgagc
cctggaagttctcggcgctggctgcctacatgttcacgtgatcctgctc
ggcttccccgtcaacttcctcacgctgtacgtcacatccagcacaagaa
actccggacgcctctaaactacatcctcctgaacctgggtggtcgccgacc
tctttatggtctttggaggttcacgaccaccatgtacacctcgatgaac
gggtactttgtctttggagtaaacaggggtgtacatcgagggcttctttgc
tacgctgggcggtgaaatcgctctctgggtcactcgtcgtcctggcgtgg
aacgatatgtggtggtctgtaagcccatgagcaacttcgcttcggggag
aaccacgccatcatgggggtcgcttctcctggatcatggccatggcctg
cgcagccccccgctgttcgggtggtcacggtacatccccgagggcatgc
agtgtcgtgtgggacgactactacacgctgaagcgggagatcaacaac
gaatcctttgtcatctacatgttcgtgggtccacttcatgatccactggc
cgctcatcttctctgctatgggaacctggctgcactgtgaaggaggctg
ccgcccagcagcaggagctgtccaccacccagaaggcagagaaggaagt
acccgcattggtgatcatcatggtcatcgcttcctcatctgctgggttcc
ctacgcccagcgtcgctttctacatcttccaccaaccagggtcagactttg
gacctatcttcatgacctcccgccattctttgccaagagctctgccatc
tacaatcctgtgatctacatcgtaataaacaacagttccgtaactgcat
gatcacaacctctgctgcggaagaacctggtggcgatgaggacacgt
ctgctggaaagacagagacctcctccgtctccaccagccaggtgtccct
gcatag

Rho genomic DNA

>galGal4_refGene_NM_001030606 range=chr12:19501548-

ATGAACGGGACGGAAGGCCAAGACTTCTACGTGCCATGTCCAACAAGAC
CGGGGTGGTGCGGAGCCCCCTTCGAGTACCCCCAGTACTACCTGGCTGAGC
CTTGAAGTTCTCGGCGTGGCTGCCATACATGTTTCATGCTGATCCTGCTC
GGCTTCCCCGTCAACTTCTCAGCTGTACGTACCATCCAGCACAGAA
ACTCCGGACGCCTCTAACTACATCCTCCTGAACCTGGCGGTGCGCGACC
TCTTTATGGTCTTCGGAGGCTTCACGACCACCATGTACACCTCGATGAAC
GGGTACTTTGTCTTTGGAGTAACAGGGTGTACATCGAGGGCTTCTTTC
TACGCTGGGCGgtgagttctccaagtttccaaaaatctgacagaagtttc
cccagattttgggtggggggggtgcccagctgctgctgaaagccgggtg
ctgctgtgccaggccctgctgctgctgctgctgctgctgctgctgctgctg
ctgggtgggaagtgaagaaagtggcaaatggcacttcttttttaagcctt
gttttggtagccacctgttccatctcctgcaaccacagcgtaacctga
aaagtaaccaacagcagattgtgagctgctggatttgcacatccctttg
gcttgccttctgacgggagaggtgacagtttcttctggcattggggggt
gagagttactgcagcgcagcgcagccagcagagccgcttactgtgcc
gccttactgtgcccgtccagccggtgatggacagtgtaaataagggccgt
gtgtctgagaacggcctcctccagccagctccgctccagccgactgca
ggcagctgttggcagggccggggcccccaggcacgggtgacccgagcagaa
cctgggggcagtgctgcagctcagagtgggtgcagcgtgtgtgtgggag
gagggcaggatccagtggtgggtctcaaagcagagcagcagcagcgtcac
cgctggggctctggaggcatcttctgactccttcccaccacactggtggg
gtactcgtcctctgctccgttctgtttctgagcacagggtggaggttca
tgctgatgggggtcttccccatgccagcaatggcaggcctttgtcctccc
atcactcgaggtccccgcacatctctacccctctctccccgggtgcagGTGAA
ATTGCTCTCTGGTCACTCGTCGTCCTGGCCGTGGAACGATACGTGGTGGT
CTGTAAGCCCATGAGCAACTTCCGCTTCGGGGAGAACCACGCCATCATGG
GGTTCGCGTTCTCCTGGATCATGGCCATGGCCTGCGCAGCCCCCCCCCTG
TTCGGCTGGTACAGgtagggaggggggtgggcatggagcgcagggtgtgg
gcagctctgcccggcgctgaccaccccgcgtgctcgcagGTACATCCCCG
AGGGCATGCAGTGCTCGTGTGGGATCGACTACTACAGCTGAAGCCGGAG
ATCAACAACGAATCCTTTGTCTATCTACATGTTTCGTGGTCCACTTCATGAT
CCCACTGGCCGTATCTTCTTCTGCTATGGGAACCTGGTCTGCACTGTGA
AGGAGgtgggtgctgccccgggggtgggggtgctgccccgggtccacgcgg

UCSC Genome Browser:

- Students retrieve DNA & mRNA sequence for their candidate genes

Wk#3 (computational): Sequence Retrieval & Annotation

Rho mRNA

>NM_001030606.1
atgaacgggacggaaggccaagacttctacgtgccatgtccaacaagac
cggggtggtgcgagagccccttcgagtacccccagttactacctggtgagc
cctggaagttctcggcgctggctgcctacatgttcatgctgatcctgctc
ggcttccccgtcaacttcctcacgctgtacgtcaccatccagcacaagaa
actccggacgcctctaaactacatcctcctgaacctgggtggtcgccgacc
tctttatggtctttggaggttcacgaccaccatgtacacctcgatgaac
gggtactttgtctttggagtaaacagggtgtacatcgagggcttctttgc
tacgctgggcggtgaaatcgctctctgggtcactcgtcgtcctggccgtg
aacgatattggtggtctgtgaagcccatgagcaacttcgcttcggggag
aaccacgccatcatgggggtcgcgttctcctggatcatggccatggcctg
cgcagccccccgctgttcgggtggtcacggtacatccccgagggcatgc
agtgtcgtgtgggacgactactacacgctgaagcgggagatcaacaac
gaatcctttgtcatctacatgttcgtgggtccacttcatgatccactggc
cgctcatcttctctgctatgggaacctggtctgcactgtgaaggaggctg
ccgcccagcagcaggaggtctgccaccaccacagaaggcagagaaggaagt
acccgcatggtgatcatcatggtcatcgcttctcctcatctgctgggttcc
ctacgccagcgtcgctttctacatcttccaccaccagggtcagactttg
gacctatcttcatgacctcccgccattctttgccaagagctctgccatc
tacaatcctgtgatctacatcgtaataaacaacagttccgtaactgcat
gatcacaacctctgctgcgggaagaacctggtggcgatgaggacacgt
ctgctggaaagacagagacctcctccgtctccaccagccaggtgtccct
gcatag

Exonic: UPPER CASE

Intronic: lower case

Rho genomic DNA

>galGal4 refGene NM_001030606 range=chr12:19501548-

ATGAACGGGACGGAAGGCCAAGACTTCTACGTGCCATGTCCAACAAGAC
CGGGGTGGTGCGGAGCCCCCTTCAGTACCCCCAGTACTACCTGGCTGAGC
CTTGAAGTTCTCGGCGTGGCTGCCATACATGTTTCATGCTGATCCTGCTC
GGCTTCCCCGTCAACTTCTCAGCTGTACGTACCATCCAGCACAGAA
ACTCCGGACGCCTCTAACTACATCCTCCTGAACCTGGCGGTGCGCGACC
TCTTTATGGTCTTCGGAGGCTTCACGACCACCATGTACACCTCGATGAAC
GGGTACTTTGTCTTTGGAGTAACAGGGTGTACATCGAGGGCTTCTTTGC

TACGCTGGGCGgtgagttctccaagtttccaaaaatctgacagaagtttc
cccagatttttgggtgggggggggatgccagctgctgctgaaagccgggtg
ctgctgtgccaggccctgctgctgctgctgctgctgctgctgctgctgctg
ctgggtgggaagtgaagaagtggcaaatggcacttcttttttaagcctt
gttttggtagccacctgttccatctcctgcaaccacagcgtaaccatga
aaagtaaccaacagcagattgtgagctgctggatttgcacatccctttg
gcttgccttctgacgggagaggtgacagtttcttctggcattgggggct
gagagttactgcagcgcagcgcagccagcagagccgcttactgtgcc
gccttactgtgcccgtccagccggtgatggacagtgtaaataagggccgt
gtgtctgagaacggcctcctccagccagctccgctccagccgactgca
ggcagctgttggcagggccggggcccccaggcacgggtgacccgagcagaa
cctgggggcagtgctgcagctcagagtgggtgcagcgtgtgtgtgggag
gagggcaggatccagtggtgggtctcaaagcagagcagcagcagcgtcac
cgctggggctctggaggcatcttctgactccttccaccacactggtggg
gtactcgtcctctgctccgttctgtttctgagcacagggtggaggttca
tgctgatgggggtcttccccatgccagcaatggcaggcctttgtcctccc
atcactcgaggtccccgcatctctaccctctctccccgggtgcagGTGAA

ATTGCTCTCTGGTCACTCGTCGTCCTGGCCGTGGAACGATACGTGGTGGT
CTGTAAGCCCATGAGCAACTTCCGCTTCGGGGAGAACCACGCCATCATGG
GGTTCGCGTTCTCCTGGATCATGGCCATGGCCTGCGCAGCCCCCCCCCTG
TTCGGGTGCTGACGgtgggaggggggtgggagggagggagggagggaggg
gcagctctggcgggcgctgaccaccccgctgctcgagGTACATCCCCG

AGGGCATGCAGTGTCTGTGGGATCGACTACTACAGCTGAAGCCGGAG
ATCAACAACGAATCCTTTGTCTATCTACATGTTCTGTTGGTCCACTTCATGAT
CCCACTGGCCGTCATCTTCTTCTGCTATGGGAACCTGGTCTGCACTGTGA
ACGAGgtgggtgctgccccgggggtgggggtgtggcggggtccacggg

UCSC Genome Browser:

- Students retrieve DNA & mRNA sequence for their candidate genes

Rho mRNA

atgaacggggacggaaggccaagacttctacgtgcccatgtccaacaagac
cgggggtgggtcgcgagcccttctcagtagccccagtagtacctgtcgtagc
cctggaagtctctcggcgtggtcgtcctacatgttctatgctgacctgctc
gcttccccgtcaactcttctcagctgtcagtcaccatccagcacaaaga
actccggagcgctctaaactacatcctcctgaacctgggtggtcgccgacc
tctttatggctctttggagggttcacgaccaccatgtacacctcgatgaac
gggtactttgtctttggagtaacagggtgctacatcgagggtctctttg
tacctgggcgggtgaaatcgctctgtgctactcgtgctctggccgtgg
aacgatattgtgtgtgtctgtaagcccatgagcaacttcgcttcggggag
aaccacgccatcatgggggtcgcggtctctcctggatcatggccatggcctg
cgcagcccccccgctgttcggtctggtcacggtacatccccgaggcgatgc
agtctcgtgtgggactgactacacgctgaagccggagatcaacaac
gaatcctttgtcatctacatgttctggtgctcactcatgatcccaactggc
cgtcactcttcttctgctatgggaacctgggtctgcactgtgaaggagggtg
ccgccacgacgacaggagtctgccaccaccacagaaaggcagagaagggaagt
accgcagctgtgatcatcatggtcattcgctctcctcatctgctgggttcc
taccgcagcgtcgcttctacatcttaccacaaacagggtcagacttgc
gacctatcttcatgacctccccggcattctttgccaaagactctgccatc
tacaatcctgtgatctacatcgtaatgaacaaacagttccgtaactgcat
gatcacaaacctctgctgcggaagaacctcgctgggcgatgaggacacgt
ctgctggaagacagagacctcctccgtctccaccagccaggtgtccctt
gcatag

Intronic: lower case

```
>galGal4_refGene_NM_001030606 range=chr12:19501548-
```

ATGAACGGGACGGGAAGGCCAAGACCTTCTACGTGCCCATGTCCAACAAGAC
CGGGGTGGTGGCGGAGGCCCTTCGAGATACCCCCAGTACTACCTGGCTGAGC
CCTGGAAGTTCTCGGCGCTGGCTGCCATGTTTACTGTGCTACCTGCTC
GGCTTCCCCGTCAACTTCTCAGCTGTACGTACCATTACAGCAAGAAB
ACTCCGGACGCCTCTAAACTACATCCTCCTGAACCTGGCGGTGCGCGACC
TCTTTATGGTCTTTCGAGGCTTCACGACCACCATGTACACCTCGATGAAC
GGGTACTTTGTCTTTGGAGTAACAGGGTGCTACATCGAGGGCTCTTTTC
TACGCTGGGCGgtgagtcttctccaagttctccaaaaatctgcagaagttt
cccagatttttgggtgggggggagtgccagagctgctgtcgaagcgggtg
ctgctgcccagggcctgctgctgctgcggtgctgctgcagcactt
ctgggtgggaagtgaagaaagtggcaaatggcacttctttttaagcctt
gttttggtagccactgttccatctcctgcgaaccacagcgtaacctga
aaagtcaaccaacagagattgtgcagctgtggatttgccattcccttg
gcttgcttctgcagggaaggtgacagttgttcttgctgacttgggggtg
gagagttactgcagcgcagcgcagcccagcagagccgccttcaactgtgcc
gccttcaactgtgccgtccagccggtgatggacagtgtaataagggccgt
gtgtctgagaacggcctcctccagccagctccgctccagccggtactgca
ggcagctgttggaaggccggggcccccaggcagcgggtgacccgagcaga
cctggggcgagtgctgcagctcagagtgggtgcagcgtgtgtgtggag
gagggcaggatccagtggtgggtctcaaagcagagcagcacagcggtcac
cgctggggctctggaggcatcttctgactcctcccaccacactgggtgg
gtactcgctctctgtccgcttctgttcttgatgacacaggggtgaggttcc
tgctgatgggggtctccccatcccgcaatggcaggccttgtctctccc
atcactcgaggtcccccgcattcttaccctctctccccgggtgcagGTGAA
ATTGCTCTCTGGTCACTCGTCTCTGGCCGTGGAAACCATACGTGGTGGT
CTGTAAGCCCATGAGCAACTCCGCCTTCGGGGAGAACCACGCCATCATGG
GGGTGCGCTCTCTGGATCATGGCCATGGCTTGGCAGCCCCCCCCCGT
TTCGGCTGGTACAGTtagggaggggctgggcatggagcagggtgtgtg
gcagctctgcccgggcgctgaccaccccgcgtgctcgcagGTACATCCCCG
AGGGCATCGAGTCTCTGTGGGATCGACTACTACAGCTGAAGCCGGAG
ATCAACAACCAATCCTTTGTGCATCTACATGTTCGTGGTCCACTTCATGAT
CCCATGCCCGATCTCTTCTTCTGCTATGGAAACCTGGTCTGCATGTGA
AGGAGgtgggtgctgcccgggggggtggggctgtgcccgggctccacgcgg

- **Students retrieve DNA & mRNA sequence for their candidates genes**

Wk#3 (computational): Sequence Retrieval & Annotation

Rho mRNA

Rho genomic DNA

GgRho mRNA.apex

Sequence Start Length End ORF Tm %GC Linear

938 1056 936<2> 2<2> 937<0> <0> 50% ☒ Dam/Dcm

* 10 * 20 * 30 * 40 * 50 * 60 *

1 atgaacgggacggaaggccaagacttctacgtgccatgtccaacaagaccggggtggtgcggag
66 ccccttcgagtagcccccagtagactacgtggctgagccctggaagtctcggcgctggctgcctaca
131 tgttcacgtgatcctgctcggcttcccgcgtcaacttctcgcgtgtacgtcaccatccagcac
196 aagaaactccggacgctctaaactacatcctcctgaacctgggtgctcgcgaacctctttatggt
261 ctttggagggtttacgaccaccatgtacacctcgatgaacgggtactcttcttggagtaacag
326 ggtgctacatcgaggggtcttcttgcacgtgggcgggtaaatcgctctctggtcactcgtcgtc
391 ctggccgtggaacgatatgtggtggtctgtaagcccatgagcaacttcgcttcggggagaacca
456 cgccatcatgggggtcgcgttctcctggatcatggccatggcctgcgcagccccccgctgttcg
521 gctggtcacggacatccccgagggcatgcagtgtcgtgtgggtagcactactacacgtgaag
586 ccggagatcaacaacgaatccttctgcatctacatgttcgtggccacttcatgatcccactggc
651 cgtcatcttcttctgcatgggaacctggtctgactgtgaaggaggctgcccagcagcagcag
716 agtctgccaccaccagaaaggcagagaaggaaagtgacccgcatgggtgatcatatggtcatcgcc
781 ttctcatctgctgggttccctacgcccagcgtcgtcttctacatcttcaccaaccagggtcaga
846 ctttggaccatcttcatgaccatcccgcatcttcttgcgaagagctctgcatctacaatcctg
911 tgatctacatcgtaataaacaacaggttcgtaactgcatgacacacccctctgctgcggcaag
976 aacccgctggcgatgaggacacgtctgctggaagacagagacctcctcgtctccaccagcca
1041 ggtgtccccctgcatag

*annotated exon junctions in mRNA

5 exons

Gg Rho DNA primer examples.apex

Sequence Insert@ Linear

958 2972 958<0> ☒ Dam/Dcm

Feature	Direction	Type	Location ↓
F1 primer	>>>	misc_feature	109..129
R1 primer	<<<	misc_feature	227..248
F2 primer	>>>	misc_feature	330..349

* 10 * 20 * 30 * 40 * 50 * 60 * 70 *

1 ATGAACGGGACGGAAGCCAAGACTTCTACGTGCCATGTCCAACAAGACCGGGGTGGTGCGGAGCCCTTCC
74 AGTACCCCCAGTACTACCTGGCTGAGCCCTGGAAGTCTCGGGCGCTGGCTGCCTACATGTTTCATGCTGATCCT
147 GCTCGGCTTCCCCGTCACCTTCTCAGCTGTACGTACCATCCAGCACAAGAACTCCGACGCCCTCTAAAC
220 TACATCCCTCGTAACCTGGCGGTGCGGACCTCTTTATGGTCTTCGAGGCTTCACGACCACCATGTACACCT
293 CGATGAACGGGTACTTGTCTTTGGAGTAACAGGGTGTCTACATCGAGGGCTTCTTGTCTACGCTGGGCGgtga
366 gttctccaaagtttccaaaaatgtacagaaagtttcccgagatttgggtgggggggggatgcccagctgctgct
439 gaaagccgggtgctgctgctgccagggccctgctgctgctgcgggctgcggtgtctgcagcaacttctgggtgggaa
512 gtgaagaaagtggcaaatggcacttctcttttaagccttgttttgtagccacctgttccatctcctgcgaac
585 cacagcgtaaacatgaaaagtaaccaacagcagattgtgcagctgctggatttgcacatcccttggcttgc
658 ttctgacgggagaggtgacagtttgttctggcattggggggctgagagttactgcagcgcagcgcagccagc
731 agagccgcttctactgtgcgccttcaactgtgcctccagccggtgatggacagtgtaaataaggccctgtg
804 tctgagaacggcctctccagccagctccgctccagccgactgcagggcagctgttggcagggccgggccc
877 caggcacgggtgacccgagcagaacctgggggagctgctgcagctcagagtggtgcagcgtgtgtgtggga
950 ggagggcagggatccagtggtgggtctcaaaagcagagcagcagcagcggtcacgctggggctctggaggca
1023 tctgactccttcccaccacactgggtgggtactcgtcctctgctccgttctgttctgagcacagggtggagg
1096 ttcaactgctgatgggggtcttcccagccagcaatggcagggccttcttctccatcactcaggttcccgc
1169 atctctaccctctctcccgggtgcagGTGAATTGCTCTCTGGTCACTCGCTCGCTCGGCGTGAACAGATA
1242 CTGGTGGTCTGTAAGCCCATGAGCAACTTCGCTTCGGGGAGAACACGCCATCATGGGGGTGCGGTTCTCC

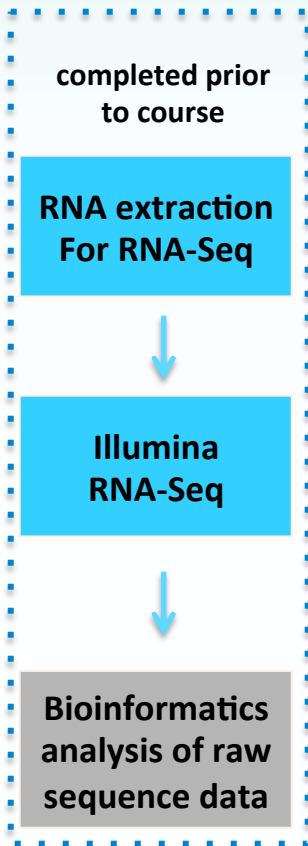
exon/intron junctions: 362, 1196, 1365, 1440, 1606, 1736, 1976, 2853

ApE Sequence Editor:

- Students annotate exon junctions in mRNA in preparation for qPCR primer design (primers usually span exon junctions)

Wk#3 (computational): qPCR Primer Design

using IDT PrimerQuest Web Browser



week #3

cDNA synthesis

Gene ontology analysis

Sequence annotation

qPCR primer design

Popular free software for primer design:

- Primer3plus

(<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>)

- NCBI Primer-BLAST

(<http://www.ncbi.nlm.nih.gov/tools/primer-blast>)

- IDT PrimerQuest

(<http://www.idtdna.com/primerquest/home/index>)

- Students annotate exons junctions in mRNA in preparation for qPCR primer design (primers usually span exon junctions)

Key:

Wet lab module

Computational module

Wk#3 (computational): qPCR Primer Design using IDT PrimerQuest Web Browser

PrimerQuest

Assay Design / Results

IDT PrimerQuest <http://www.idtdna.com/primerquest/home/index>

Sequence Entry

▼ Enter sequence(s) manually

```
ACCAACCAGGGCTCAGACTTTGGACCCATCTTCATGACCATCCCGGCATTCTTTGCCAAGAGCTCTGCCATCT  
ACAATCCTGTGATCTACATCGTAATGAACAAACAGGTAACAGCTGATTGCCTCTGTAGTGTTCCTTTGT  
AGAAGCTGCGGGGAGCTCCACACCTCCAGCAGCGGTGAGCCGTGAGGGGATCCAGGGAGCTGCGTGCT  
GATGTGCAGGAGAGTGAAGTGCAGCTCACTCCTTCACACTGTGCAGCTCTCATTGATGCAGACACCCACCTG  
AGCAATGTGCAGCTGGCCAGCGGAGCAGCTGGGAGTGCCTCCCCAGCCCTGAGCACTGGGAGGGAGA  
CTTCACCAGCAGTGGCCTGTTGCAGACACGCTGGAGGGTTCCAAGAGTTCCCACTACAGCATGTGGGAT  
CTGACTCGCTTGTCTCCTATGGTAGTTCCTCCAATCTTATGAGGAGCAGCTGAGGGAGTTGGGAT
```

Hints

- Paste your sequence into the Textbox
- Add up to 50 sequences in **FASTA** format
- Sequence length must be greater than 80 bases
- PrimerQuest accepts only nucleic acid bases

Sequence Name

Gg Rho

Clear Sequence Entry

- Navigate to IDT PrimerQuest (<http://www.idtdna.com/Primerquest/Home/Index>)
- Paste the mRNA sequence from your annotated ApE file into sequence entry box
- Select "Show custom design parameters"
- Set Design Parameters for "qPCR Intercalating Dyes (primers only)"
- Change the following Parameters:
 - Primer Tm (C): min=59; opt=60; max=61
 - Primer CG%: min=40; opt=50; max=60
 - Primer size (nt): min=17; opt=22; max=30
 - 3' GC clamp (nt) = 1

- Each student designs 2 F/R primer sets for their target gene of interest

Wk#3 (computational): qPCR Primer Design

using IDT PrimerQuest web browser

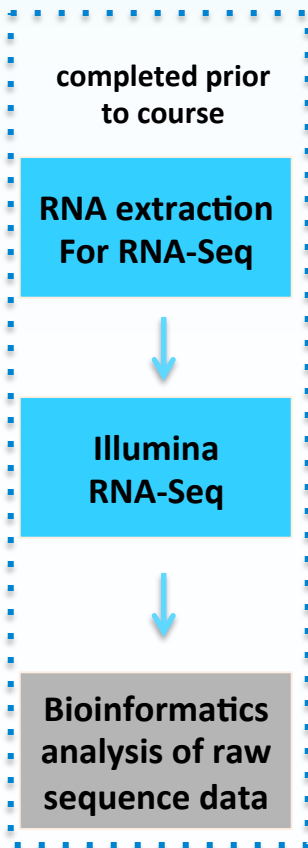
1	group #	designed by	Oligo Name	Gene Symbol	Oligo Sequence 5'-3'	amplicon bp	up/down reg	GO Term	Fold change
2	example	Dr. Enke	Gg-GPR98-RT-F1	GPR98	GCTGCAGAAATCCCTCTG	82 bp	Up	phototransduction	6.02272
3	Blue #1	CHD	Gg-NEUROD4-RT-F1	NEUROD4	TGCTCTGCAAAGGGTTATC	89bp	Up	Notch signaling pathway	754.58
4	Blue #1	CHD	Gg-NEUROD4-RT-R1	NEUROD4	GTGCTTCTCCAGGAATAAAGG	89bp	Up	Notch signaling pathway	754.58
5	Blue #1	CHD	Gg-NEUROD4-RT-F2	NEUROD4	AACCTGAGGAGGGTGATG	95bp	Up	Notch signaling pathway	754.58
6	Blue #1	CHD	Gg-NEUROD4-RT-R2	NEUROD4	GACAGAGCCCAGATGTAATTC	95bp	Up	Notch signaling pathway	754.58
7	Blue #1	JJD	final G protein-coupled r	RGR	TGAAGACGCTGGTCATTTG	19	Up	phototransduction	43.27
8	Blue #1	JJD	final G protein-coupled r	RGR	TAATGGCAGGAATCATTCCGG	20	Up	phototransduction	43.27
9	Blue #1	JJD	final G protein-coupled r	RGR	GATTCCAGGGCTTCTTGAC	19	Up	phototransduction	43.27
10	Blue #1	JJD	final G protein-coupled r	RGR	CTGCAGCTTGCTCCTTG	17	Up	phototransduction	43.27
11	Blue #1	JPN	Gg-OPN4-RT-F1	OPN4	TGTGGAGGAAATTCAGAGAAG	85	Up	phototransduction	301.73
12	Blue #1	JPN	Gg-OPN4-RT-R1	OPN4	CTGAGATGCGGATGTTCTAC	85	Up	phototransduction	301.73
13	Blue #1	JPN	Gg-OPN4-RT-F2	OPN4	CCAGACGAACCAACATAAC	100	Up	phototransduction	301.73
14	Blue #1	JPN	Gg-OPN4-RT-R2	OPN4	GTACAGCCCGATTCACTATC	100	Up	phototransduction	301.73
15	Red #2	Laura Morales	Gg-GUCA1A-RT-F1	GUCA1A	AAGCTGCGGTGGTATTTC	92	Up	Photoreceptor	110.48
16	Red #2	Laura Morales	Gg-GUCA1A-RT-R1	GUCA1A	GCTCGGATGGCTTTGATG	92	Up	Photoreceptor	110.48
17	Red #2	Laura Morales	Gg-GUCA1A-RT-F2	GUCA1A	CTCTGCACAGGATAATGTC	84	Up	Photoreceptor	110.48
18	Red #2	Laura Morales	Gg-GUCA1A-RT-R2	GUCA1A	AGGTACCCAAGACATCTCAG	84	Up	Photoreceptor	110.48
19	Red #2	Monique Waldman	Gg-PDE6C-RT-F1	PDE6C	GGAGGTGCAAAGCAAGG	105	Up	phototransduction, visible light	509.76
20	Red #2	Monique Waldman	Gg-PDE6C-RT-R1	PDE6C	CTGTCCATCATGGGAATTGG	105	Up	phototransduction, visible light	509.76
21	Red #2	Monique Waldman	Gg-PDE6C-RT-F2	PDE6C	AATGAAATCAGCTGCTCCAC	97	Up	phototransduction, visible light	509.76
22	Red #2	Monique Waldman	Gg-PDE6C-RT-R2	PDE6C	ACTCTCATCTTGCAAGTAGGG	97	Up	phototransduction, visible light	509.76
23	Red #2	Paul Roberson	Gg-GOT1-RT-F1	GOT1	CTGAGTGGAAGGACAATGTG	117	Up	Notch Signaling Pathway	2.61
24	Red #2	Paul Roberson	Gg-GOT1-RT-R1	GOT1	GATCTGTGATGTGGTTCCAG	117	Up	Notch Signaling Pathway	2.61
25	Red #2	Paul Roberson	Gg-GOT1-RT-F2	GOT1	CTGGATGACATGGAGAAAGC	98	Up	Notch Signaling Pathway	2.61
26	Red #2	Paul Roberson	Gg-GOT1-RT-R2	GOT1	TTCCAATCGTCTGGAGTAG	98	Up	Notch Signaling Pathway	2.61
27	Green #3	Se Chang	Gg-Sox9-RT-F1	Sox9	GTGGAGGCTGCTGAATG	68	Down	Notch signaling pathway	4.56
28	Green #3	Se Chang	Gg-Sox9-RT-R1	Sox9	CGTGGTTGGTACTTGTAGTC	68	Down	Notch signaling pathway	4.56
29	Green #3	Se Chang	Gg-Sox9-RT-F2	Sox9	AACGCCCTTCATGGTGTG	81	Down	Notch signaling pathway	4.56
30	Green #3	Se Chang	Gg-Sox9-RT-R2	Sox9	CATTGAGCAGCCTCCAC	81	Down	Notch signaling pathway	4.56

pricing: *18 students * 4 primers each = ~\$300

- Each student designs 2 F/R primer sets for their target gene of interest
- Populate GoogleSheet with 5'-3' primer info for ordering
- hold them accountable for both correct primer design as well as correctly inputting into the sheet!

Wk#4 (computational): Basic RNA-Seq Analysis in R Studio

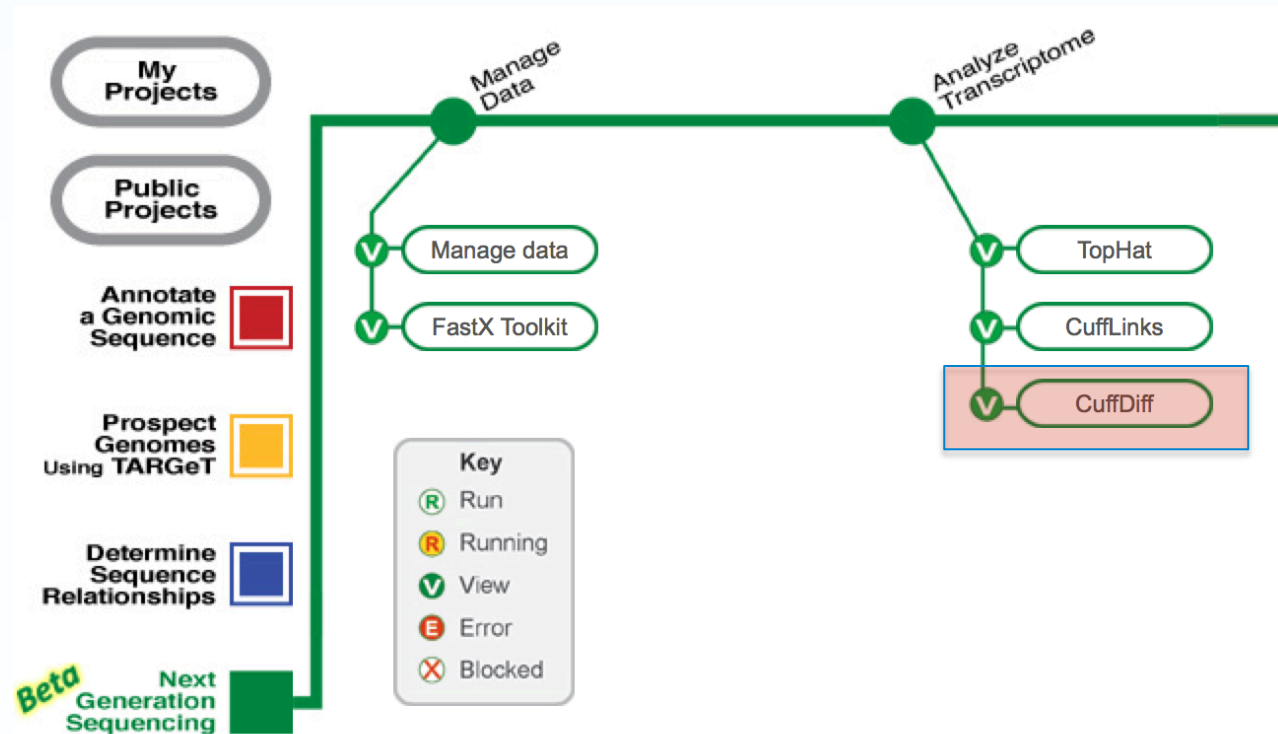
- analysis of CuffDiff output beyond the list of DEGs
-ie what's the quality of the data set?



week #4

qPCR primer standard curves

RNA-Seq Data Analysis in R Studio



Key:

Wet lab module

Computational module

Wk#4 (computational): Basic RNA-Seq Analysis in R Studio

Cuffdiff

< Back

View:

Genes

E8_retina

E18_retina

Transcripts sorted by Q-Value



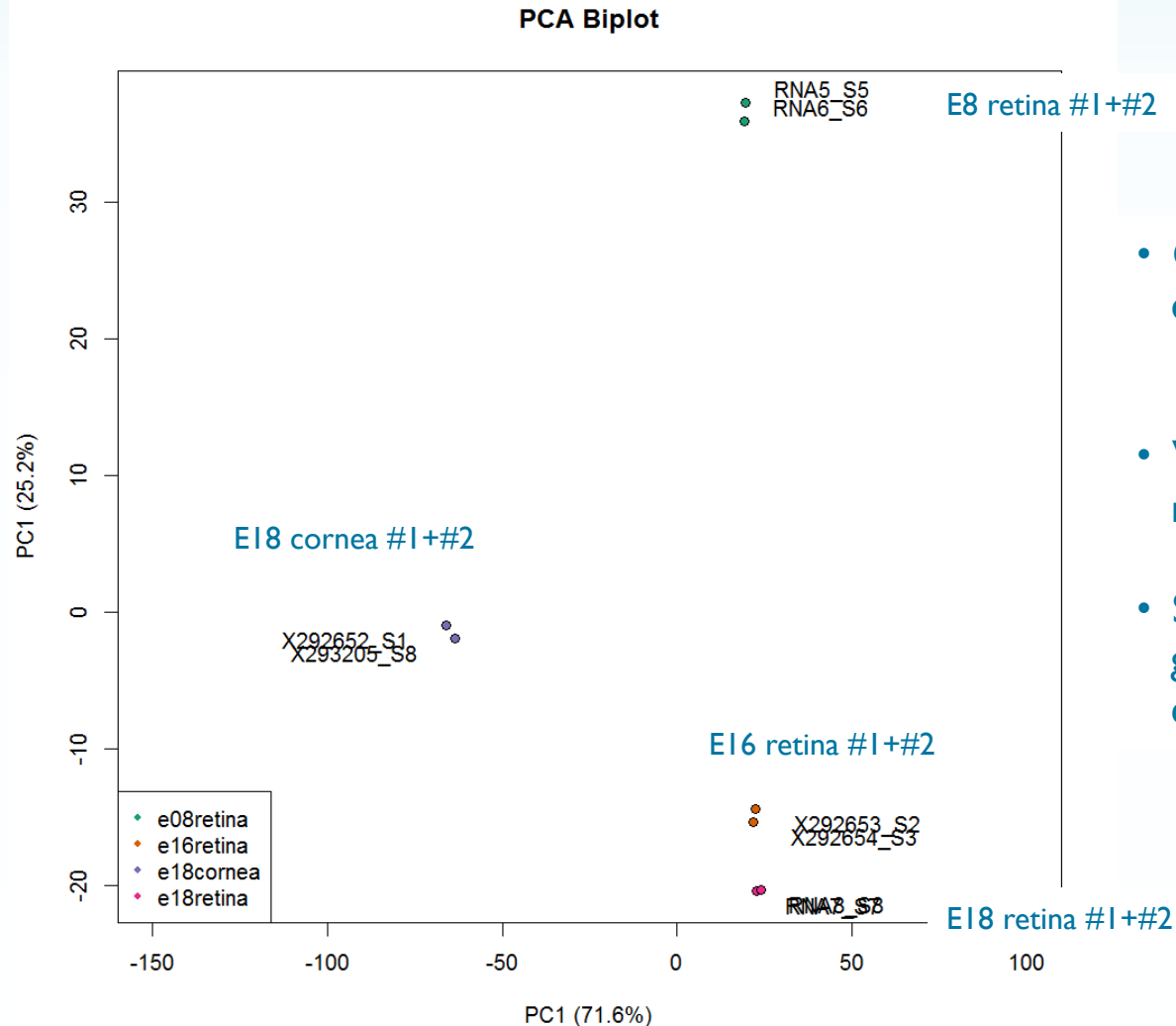
Export data to spreadsheet

Search:

Transcript ⬆	Nearest Ref Id ⬆	Gene ⬆	Alias ⬆	Fold Change ⬆	Direction ⬆	Sample 1 FPKM ⬆	Sample 2 FPKM ⬆	Q-Value ⬆	Description ⬆	
ENSGALT00000000003	-	ENSGALG00000000003			PANX2	2.16	UP	9.25776	19.9978	0.0001
ENSGALT00000000026	-	ENSGALG00000000019			AAK1	3.89	UP	18.9927	73.8016	0.0001
ENSGALT00000000064	-	ENSGALG00000000049			CSDC2	3.10	UP	44.2807	137.052	0.0001
ENSGALT00000000070	-	ENSGALG00000000055			LAMTOR3	1.74	UP	80.5424	140.282	0.0001
ENSGALT00000000071	-	ENSGALG00000000056			DAPP1	8.58	DOWN	10.8211	1.2606	0.0001
ENSGALT00000000074	-	ENSGALG00000000059			TUBB3	3.10	DOWN	869.907	280.966	0.0001
ENSGALT00000000104	-	ENSGALG00000000076			NRBP2	3.23	UP	2.94817	9.53326	0.0001
ENSGALT00000000141	-	ENSGALG00000000102			UBE2T	9.17	DOWN	65.529	7.14856	0.0001
ENSGALT00000000161	-	ENSGALG00000000112			PLP1	25.25	UP	3.97947	100.472	0.0001
ENSGALT00000000171	-	ENSGALG00000000122			GNB2L1	2.30	DOWN	1302.8	566.928	0.0001
ENSGALT00000000190	-	ENSGALG00000000137			SNRPE	4.04	DOWN	229.053	56.72	0.0001
ENSGALT00000000206	-	ENSGALG00000000150			RPL9	2.44	DOWN	2565.02	1052.75	0.0001
ENSGALT00000000210	-	ENSGALG00000000154			DENND2D	9.75	UP	3.29344	32.0971	0.0001
ENSGALT00000000221	-	ENSGALG00000000165			CDS2	2.50	UP	32.2086	80.6627	0.0001
ENSGALT00000000224	-	ENSGALG00000000168			ADORA1	0.10	UP	1.07010	10.0007	0.0001

- 12,939 differentially regulated transcripts between E8 retina & E18 retina
- Are the same genes differentially expressed in sample replicates?

Principal Component Analysis: PCA Plot



- Globally compares Cuffdiff data for each sample replicate
- Visualization of “do my sample replicates look the same”
- Students are given the code to and generate these plots essentially by copy/pasting (not ideal)

Thanks Taylor B & Alanna H!

Wk#4 (computational): Basic RNA-Seq Analysis in R Studio

Cuffdiff

< Back

View:

Genes

E8_retina

E18_retina

Transcripts sorted by Q-Value



Export data to spreadsheet

Search:

Transcript ⬆	Nearest Ref Id ⬆	Gene ⬆	Alias ⬆	Fold Change ⬆	Direction ⬆	Sample 1 FPKM ⬆	Sample 2 FPKM ⬆	Q-Value ⬆	Description ⬆	
ENSGALT00000000003	-	ENSGALG00000000003			PANX2	2.16	UP	9.25776	19.9978	0.0001
ENSGALT00000000026	-	ENSGALG00000000019			AAK1	3.89	UP	18.9927	73.8016	0.0001
ENSGALT00000000064	-	ENSGALG00000000049			CSDC2	3.10	UP	44.2807	137.052	0.0001
ENSGALT00000000070	-	ENSGALG00000000055			LAMTOR3	1.74	UP	80.5424	140.282	0.0001
ENSGALT00000000071	-	ENSGALG00000000056			DAPP1	8.58	DOWN	10.8211	1.2606	0.0001
ENSGALT00000000074	-	ENSGALG00000000059			TUBB3	3.10	DOWN	869.907	280.966	0.0001
ENSGALT00000000104	-	ENSGALG00000000076			NRBP2	3.23	UP	2.94817	9.53326	0.0001
ENSGALT00000000141	-	ENSGALG00000000102			UBE2T	9.17	DOWN	65.529	7.14856	0.0001
ENSGALT00000000161	-	ENSGALG00000000112			PLP1	25.25	UP	3.97947	100.472	0.0001
ENSGALT00000000171	-	ENSGALG00000000122			GNB2L1	2.30	DOWN	1302.8	566.928	0.0001
ENSGALT00000000190	-	ENSGALG00000000137			SNRPE	4.04	DOWN	229.053	56.72	0.0001
ENSGALT00000000206	-	ENSGALG00000000150			RPL9	2.44	DOWN	2565.02	1052.75	0.0001
ENSGALT00000000210	-	ENSGALG00000000154			DENND2D	9.75	UP	3.29344	32.0971	0.0001
ENSGALT00000000221	-	ENSGALG00000000165			CDS2	2.50	UP	32.2086	80.6627	0.0001
ENSGALT00000000224	-	ENSGALG00000000168			ADORA1	2.16	UP	1.27216	12.2227	0.0001

- 12,939 differentially regulated transcripts between E8 retina & E18 retina
- Are they all statistically & biologically significant?

Volcano Plot: statistical & biological significance

Volcano Plot E8 retina vs E18 retina



- Scatter plot of each DEG
- Stat significance on y axis (Q-value)
- Bio significance on x axis (log fold change)
- Students are given code to and copy/pasting to generate plots (not ideal)

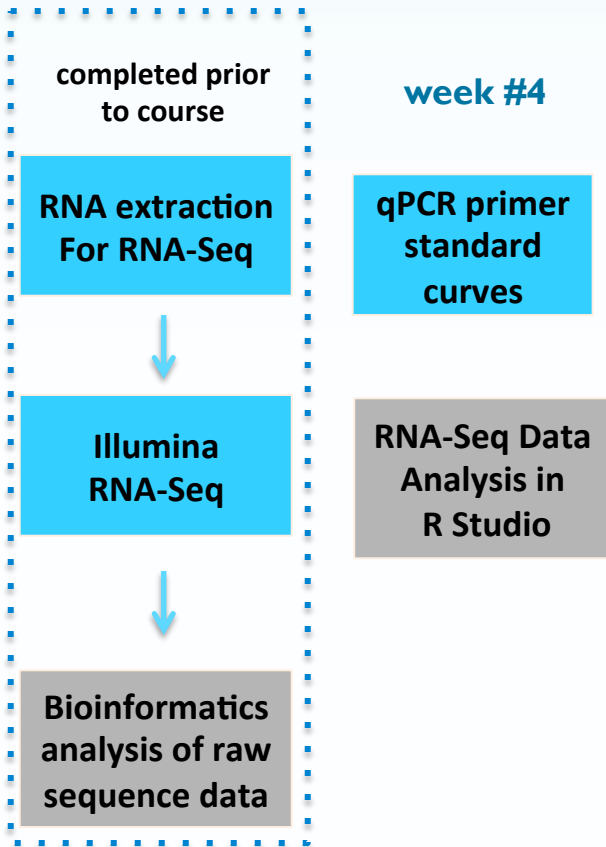
Thanks Josh D!

Wk#4-6 (wet lab): qPCR Analysis

- Bio-Rad CFX96 Real Time Cycler: ~\$27K



- iTaq cDNA synthesis kit (100 rctns): \$328
- iTaq SYBR Green PCR mix (1000 rctns): \$353
- qPCR plates and seals: ~\$300
- Multichannel pipettes recommended: ~\$600-\$800 each



Key:

Wet lab
module

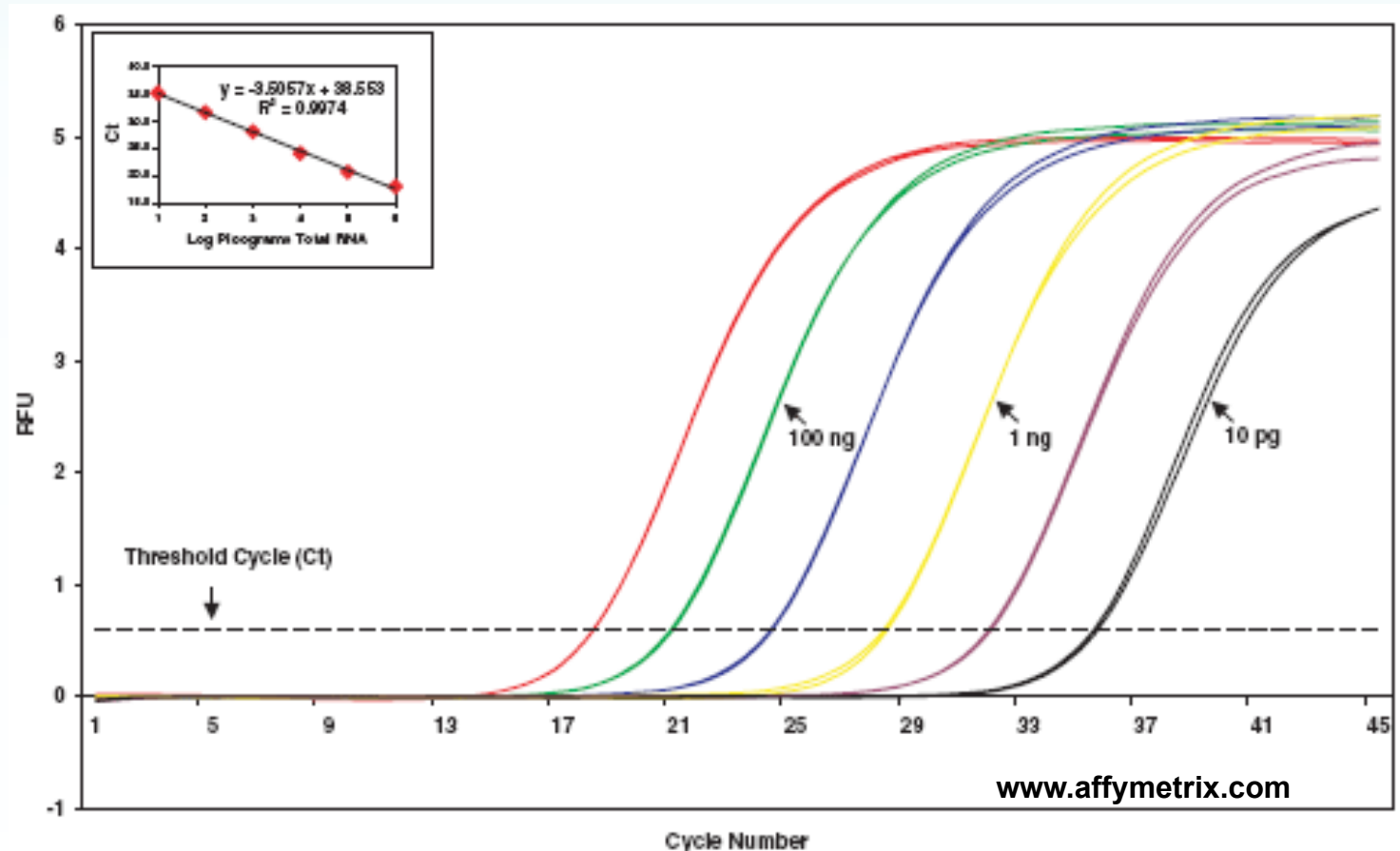
Computational
module

PCR primers are individually validated for “amplification efficiency” using a standard curve assay

week #4

qPCR primer
standard
curves

RNA-Seq Data
Analysis in
R Studio



- Each student amplifies a cDNA dilution series with 2 primer sets to ID efficiency

Wk#4 (wet lab): qPCR analysis of primer efficiency

Plate #1 (groups 1-2)

cDNA: 1:10 1:100 1:1K 1:10K 1:10 1:100 1:1K 1:10K

12		Red #2, student #3 F2/R2 primer					
11		Red #2, student #3 F1/R1 primer					
10		Red #2, student #2 F2/R2 primer					
9		Red #2, student #2 F1/R1 primer					
8		Red #2, student #1 F2/R2 primer					
7		Red #2, student #1 F1/R1 primer					
6		Blue #1, student #3 F2/R2 primer					
5		Blue #1, student #3 F1/R1 primer					
4		Blue #1, student #2 F2/R2 primer					
3		Blue #1, student #2 F1/R1 primer					
2		Blue #1, student #1 F2/R2 primer					
1		Blue #1, student #1 F1/R1 primer					
A	B	C	D	E	F	G	H

Plate #2 (groups 3-4)

DNA: 1:10 1:100 1:1K 1:10K 1:10 1:100 1:1K 1:10K

12		Green#3, student #3 F2/R2 primer						
11		Green#3, student #3 F1/R1 primer						
10		Green#3, student #2 F2/R2 primer						
9		Green#3, student #2 F1/R1 primer						
8		Green#3, student #1 F2/R2 primer						
7		Green#3, student #1 F1/R1 primer						
6		Orange #4, student #3 F2/R2 primer						
5		Orange #4, student #3 F1/R1 primer						
4		Orange #4, student #2 F2/R2 primer						
3		Orange #4, student #2 F1/R1 primer						
2		Orange #4, student #1 F2/R2 primer						
1		Orange #4, student #1 F1/R1 primer						
	A	B	C	D	E	F	G	H

Plate #3 (groups 5-6)

cDNA: 1:10 1:100 1:1K 1:10K 1:10 1:100 1:1K 1:10K

12																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															</
----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

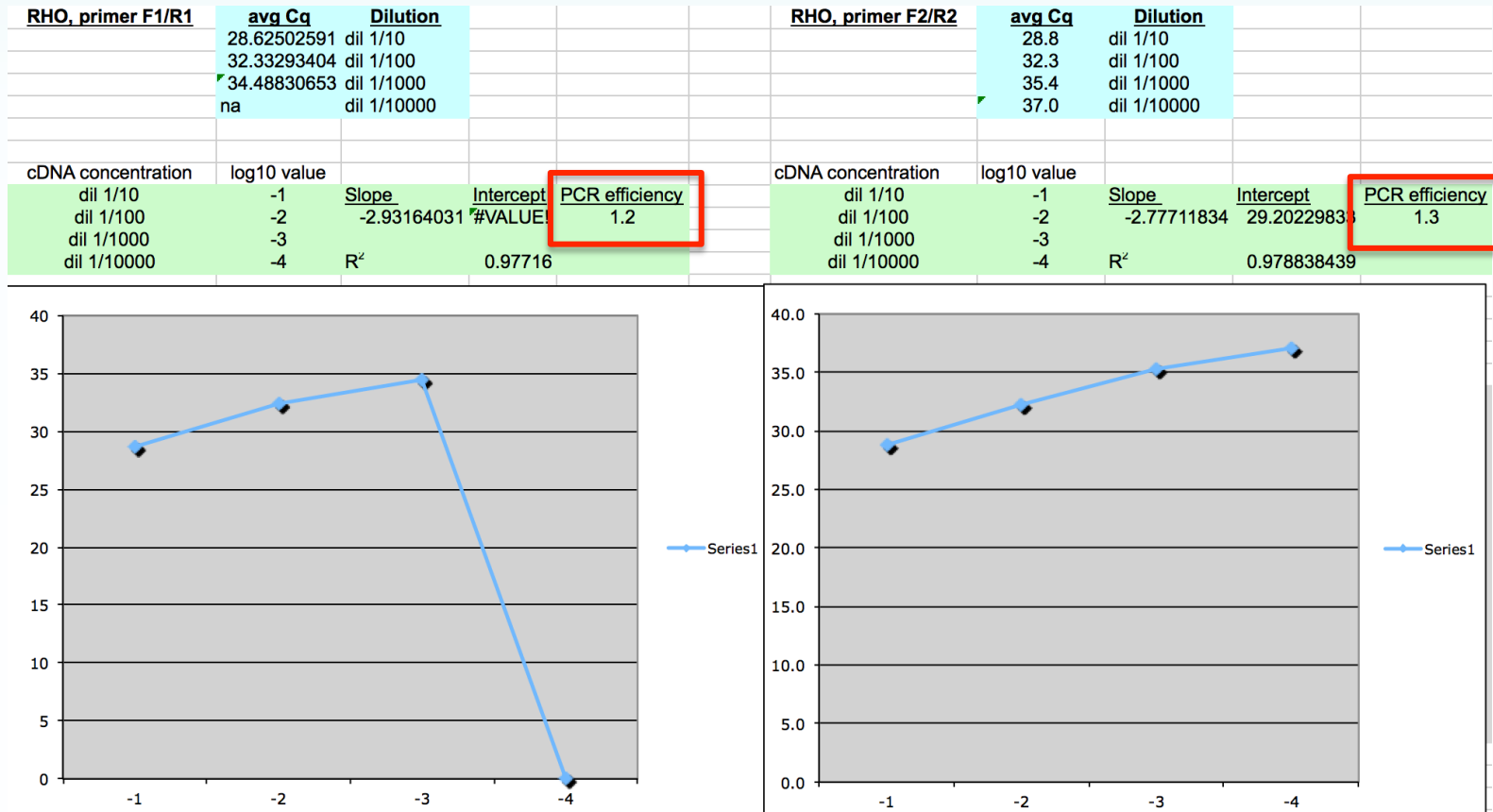
- Each student amplifies a cDNA dilution series with 2 primer sets to ID efficiency
- 18 students X 2 primers sets X 4 dilutions in duplicate = 288 well (3X 96 well plates)

Wk#4 (wet lab): qPCR analysis of primer efficiency

week #4

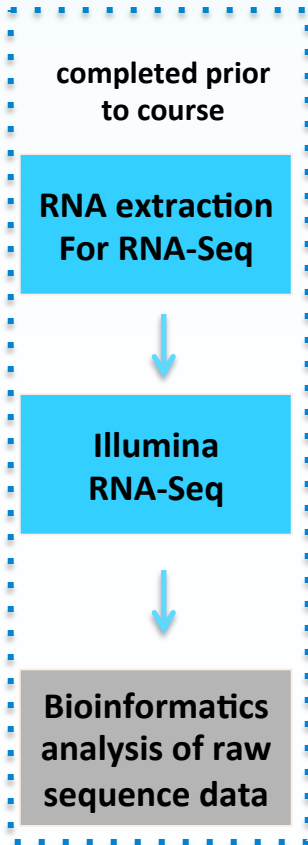
qPCR primer
standard
curves

RNA-Seq Data
Analysis in
R Studio



- Each student amplifies a cDNA dilution series with 2 primer sets to ID efficiency
- Ideally each group IDs 2 “good” primers pairs from the 6 sets screened

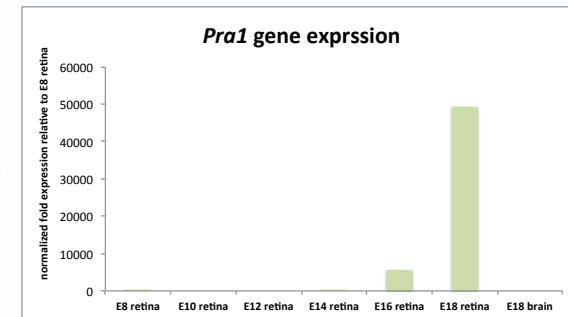
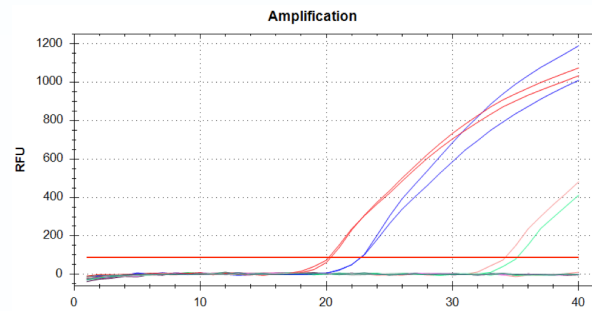
Wk#5-6: qPCR Analysis of DEGs (set up & analysis)



week #5-6

qPCR setup

qPCR data analysis



Required equipment/associated costs:

- PCR primers ~\$300
- SYBR Green PCR mix- \$436/1K rxtns
- Real Time PCR cyclor

Key:

Wet lab
module

Computational
module

Example of in class qPCR experiment

Plate #1 (groups 1+2)											
cDNA:											
no	E18	E18	E16	E14	E12	E10	E8				
RT	brain										
H	G	F	E	D	C	B	A				
		Nicole, GRP98 rep #1						1			
		Nicole, GRP98 rep #2						2			
		Matt, DTX1 rep #1						3			
		Matt, DTX1 rep #2						4			
		Tracy, SLITRK6 rep #1						5			
		Tracy, SLITRK6 rep #2						6			
		Colin, ZIC2 rep #1						7			
		Colin, ZIC2 rep #2						8			
		Aman, SNAI2 rep #1						9			
		Aman SNAI2 rep #2						10			
		Austin, OPN4-1 rep #1						11			
		Austin, OPN4-1 rep #2						12			

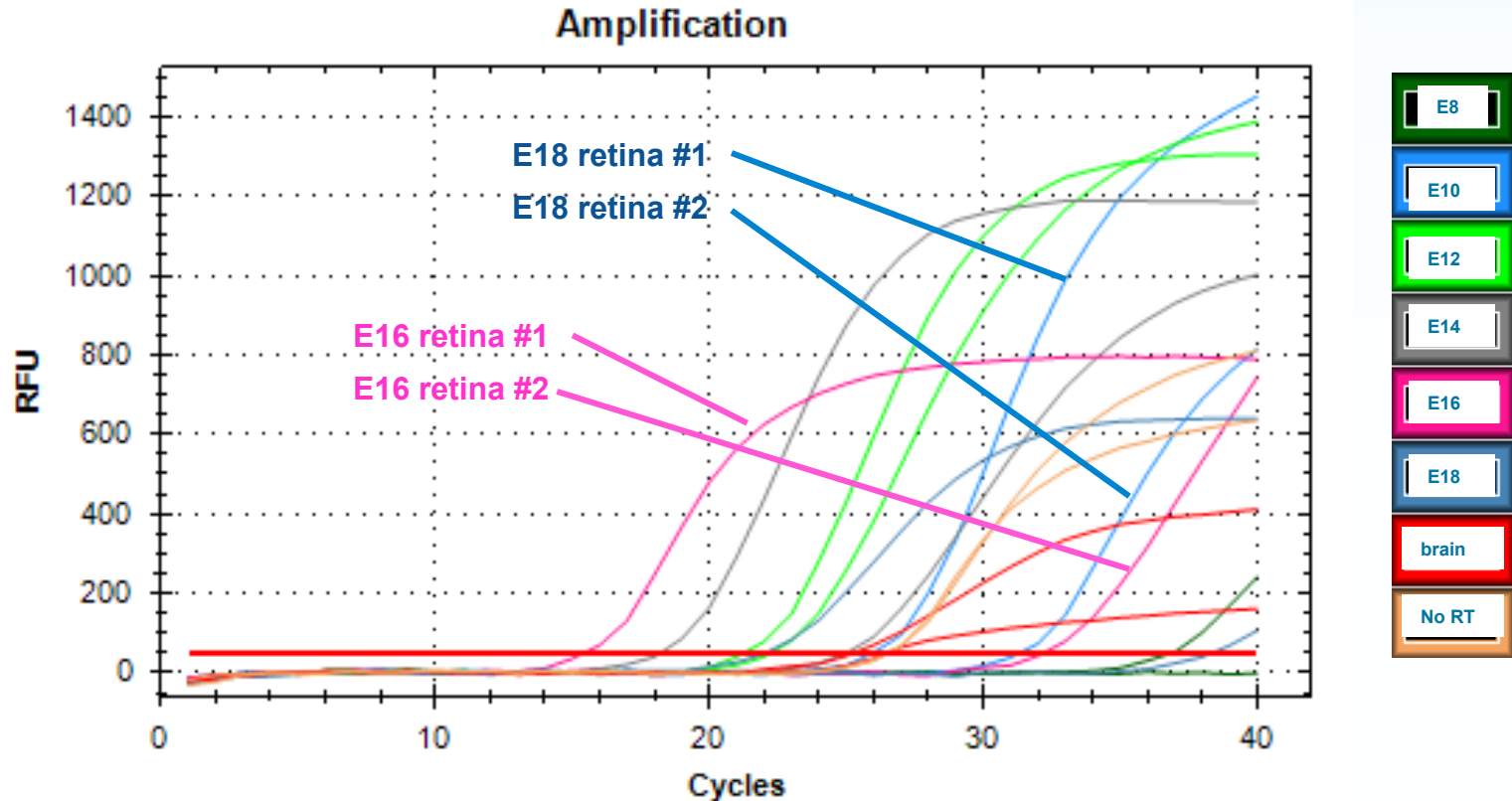
Plate #2 (groups 3+4)											
cDNA:											
no	E18	E18	E16	E14	E12	E10	E8				
RT	brain										
H	G	F	E	D	C	B	A				
		Osscar, FSCN2 rep #1						1			
		Osscar, FSCN2 rep #2						2			
		Erin, REEP6 rep #1						3			
		Erin, REEP6 rep #2						4			
		Katie, PTPRK rep #1						5			
		Katie, PTPRK rep #2						6			
		Jon, NR2E1 rep #1						7			
		Jon, NR2E1 rep #2						8			
		Alexandra, DLL1 rep #1						9			
		Alexandra, DLL1 rep #2						10			
		Gunnar, Sox2 rep #1						11			
		Gunnar, Sox2 rep #2						12			

Plate #3 (groups 5+6)											
cDNA:											
no	E18	E18	E16	E14	E12	E10	E8				
RT	brain										
H	G	F	E	D	C	B	A				
		Emily, Hes5 rep #1						1			
		Emily, Hes5 rep #2						2			
		Lina, CLN6 rep #1						3			
		Lina, CLN6 rep #2						4			
		Lemar, BBS2 rep #1						5			
		Lemar, BBS2 rep #2						6			
		Quinlin, OPN4 rep #1						7			
		Quinlin, OPN4 rep #2						8			
		Camden, NRARP rep #1						9			
		Camden, NRARP rep #2						10			
		Ashton, DFNB31 rep #1						11			
		Ashton, DFNB31 rep #2						12			

- Each student designs 1 qPCR primer set (18 genes/class analyzed)
- Students analyze 7 cDNA samples + no RT control run in duplicate
- B-actin housekeeping gene is run in tandem

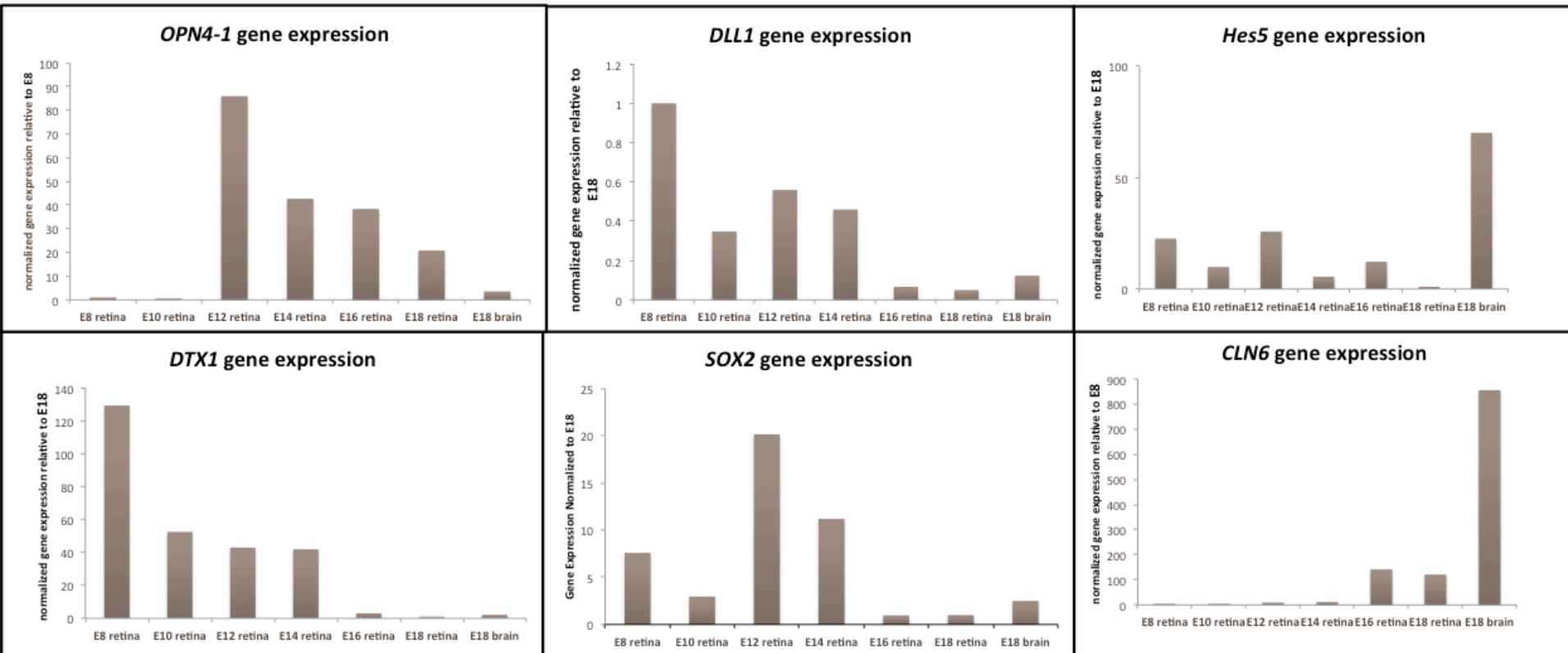
Example of class gene expression data

extra time built into wk#6 to repeat



- Too much variability to interpret!
- We will repeat this week...VERY CAREFULLY!

Example of class gene expression data

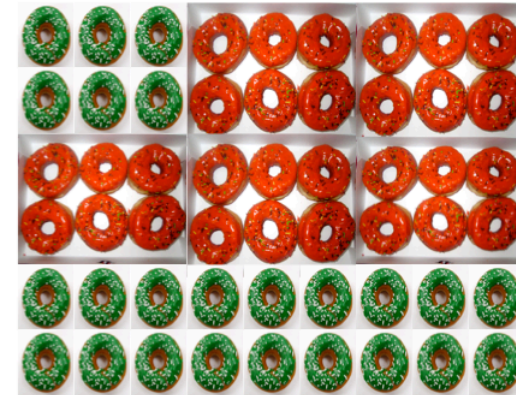
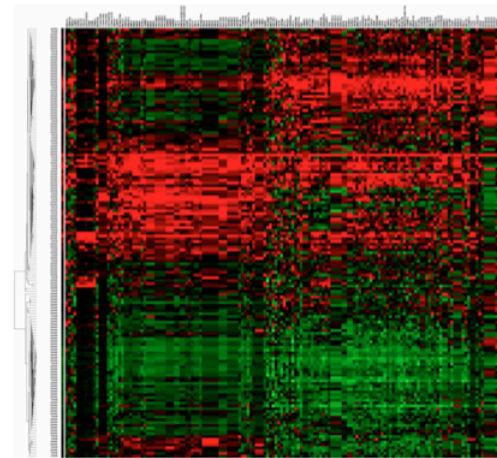
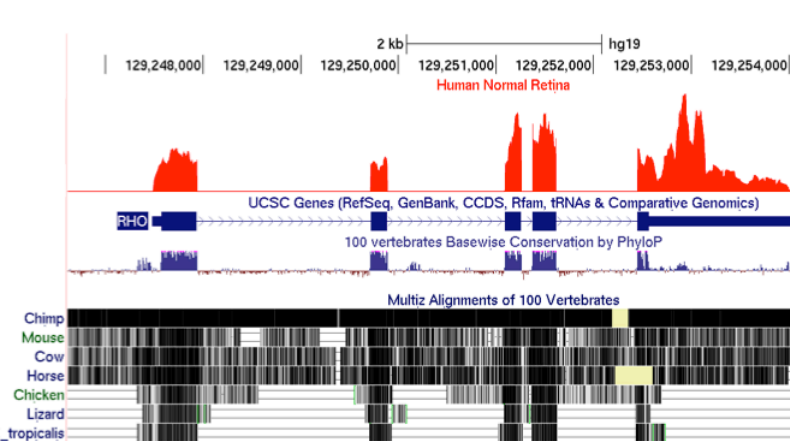


- Spring 15 class validated 6 out of 18 genes; spring 16 class 0/18 genes
- Technical difficulties included poor/no amplification and replicate variability

Wk#7: Student Poster Session

The JMU Advanced Molecular Biology Class
invites you to

The Donuts & Transcriptomics Poster Session



Wed April 27th

8:30 – 10 am

Biosciences #2009 & 2nd floor atrium

Breakfast provided

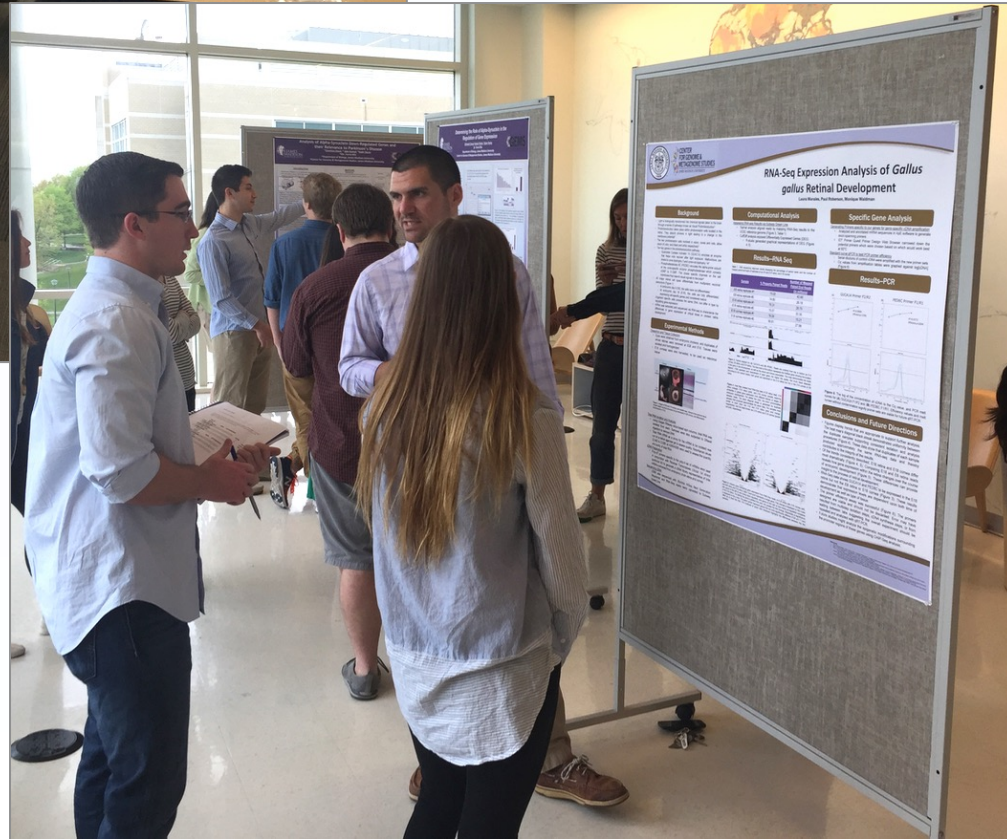
contact Dr. Enke or Dr. Rife for more details

In partnership with:

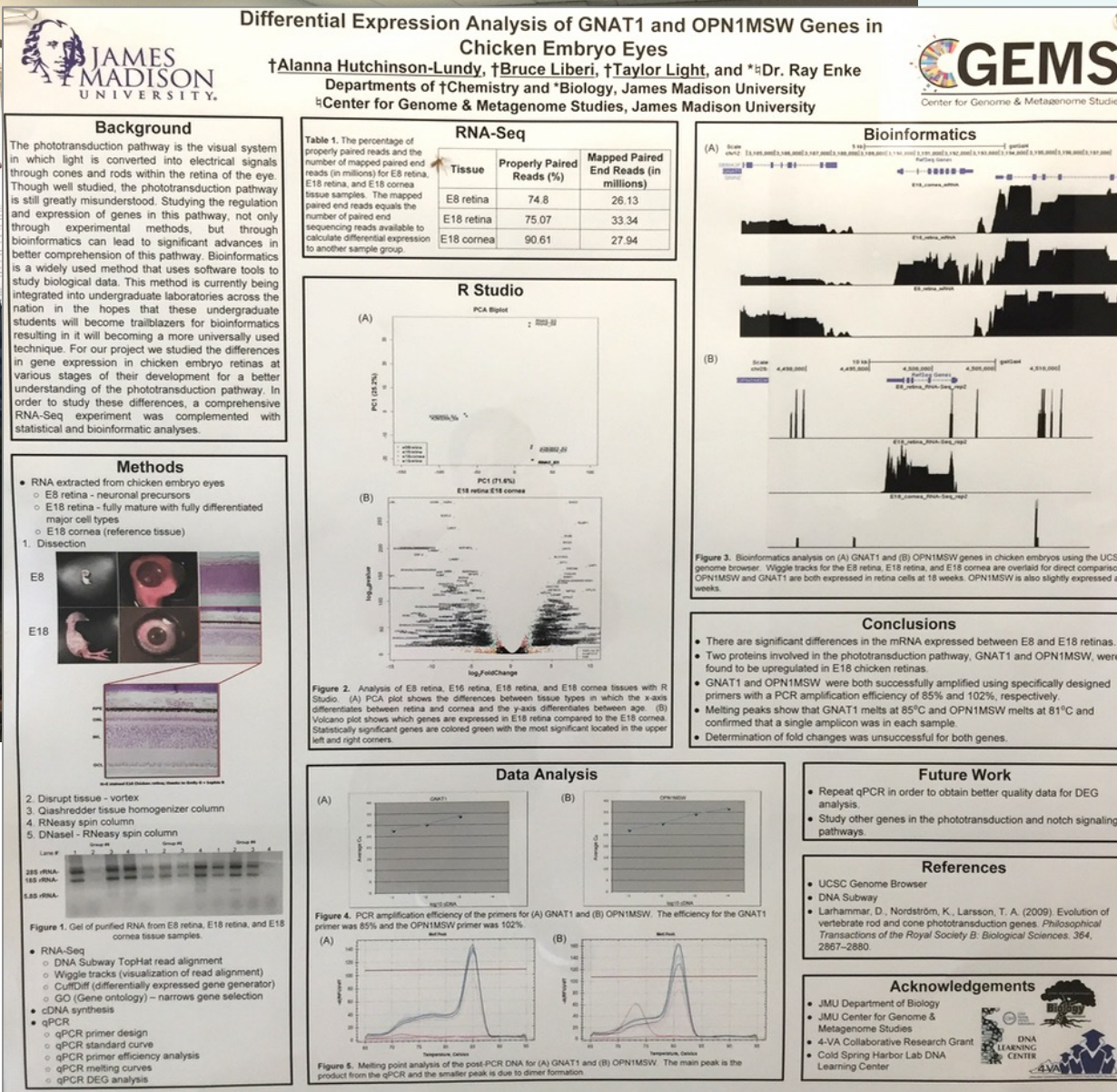


RNA-Seq for the **NextGeneration**

Wk#7: Student Poster Session



Wk#7: Student Poster Session



Reflections & Future Directions I

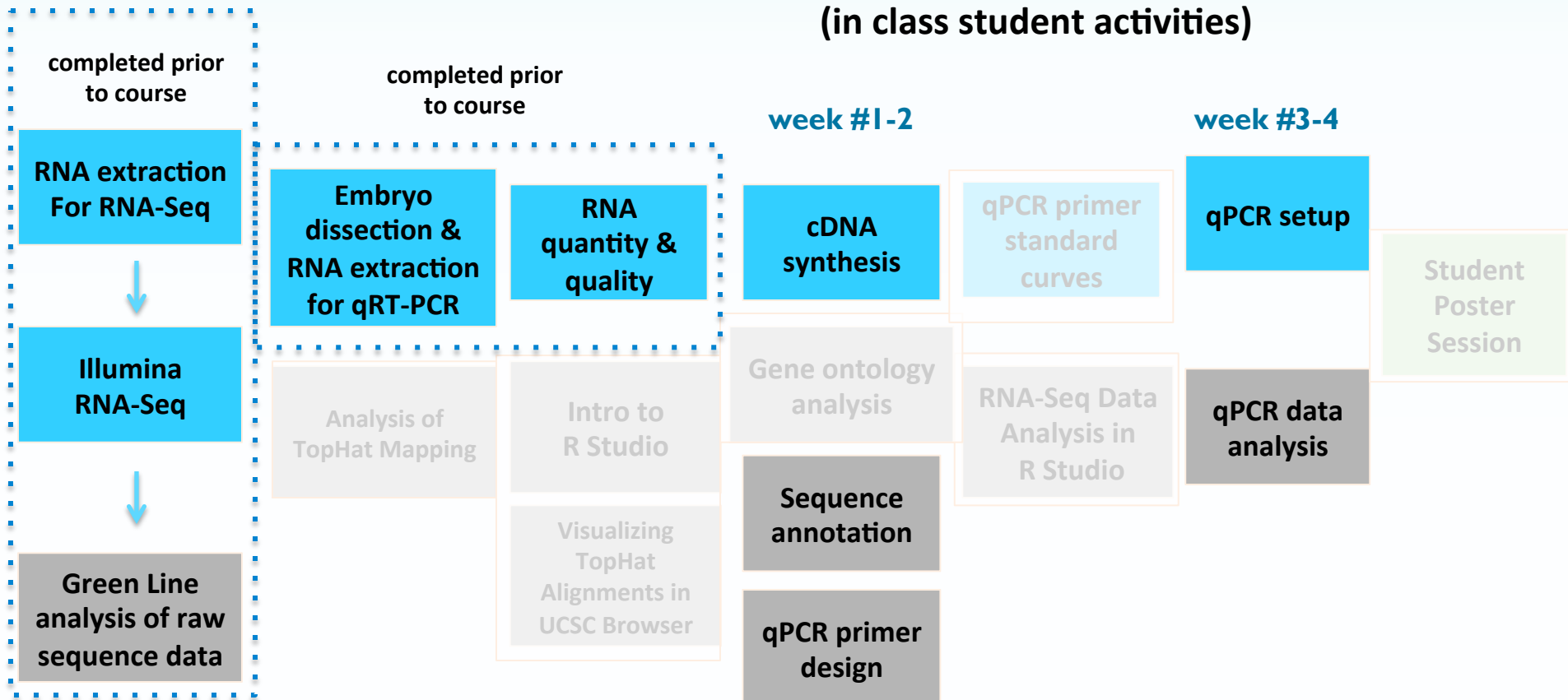
- Hope to include more student Green Line analysis in future
 - i.e. FastQC analysis of sequencing quality
 - unsure about feasibility of having students actually run Green Line
- Hope to expand R Studio analysis modules
 - requires me becoming more comfortable with R Studio
- Hope to make qPCR analysis less variable and more universally successful
 - possibly adding several more weeks to repeat experiments

Reflections & Future Directions II

- **(If possible) mixing wet bench & computational activities is key**
 - no budget for wet lab? Find collaborators!
- **Make your in class modules dovetail with your research program**
 - maintains research focus during the academic year
 - course budgets can be used to propel your research
 - *planning & troubleshooting efforts for class will benefit your lab research
- **Start small, build up over several iterations/semesters**
 - “Rome wasn’t built in a day”* (even Secaucus NJ took a while)
 - start with activities you feel most comfortable with & build out
 - post your protocols to the Wiki & Website!

Evolution of activities over 3 semesters

(in class student activities)



* 1st iteration of RNA-Seq module

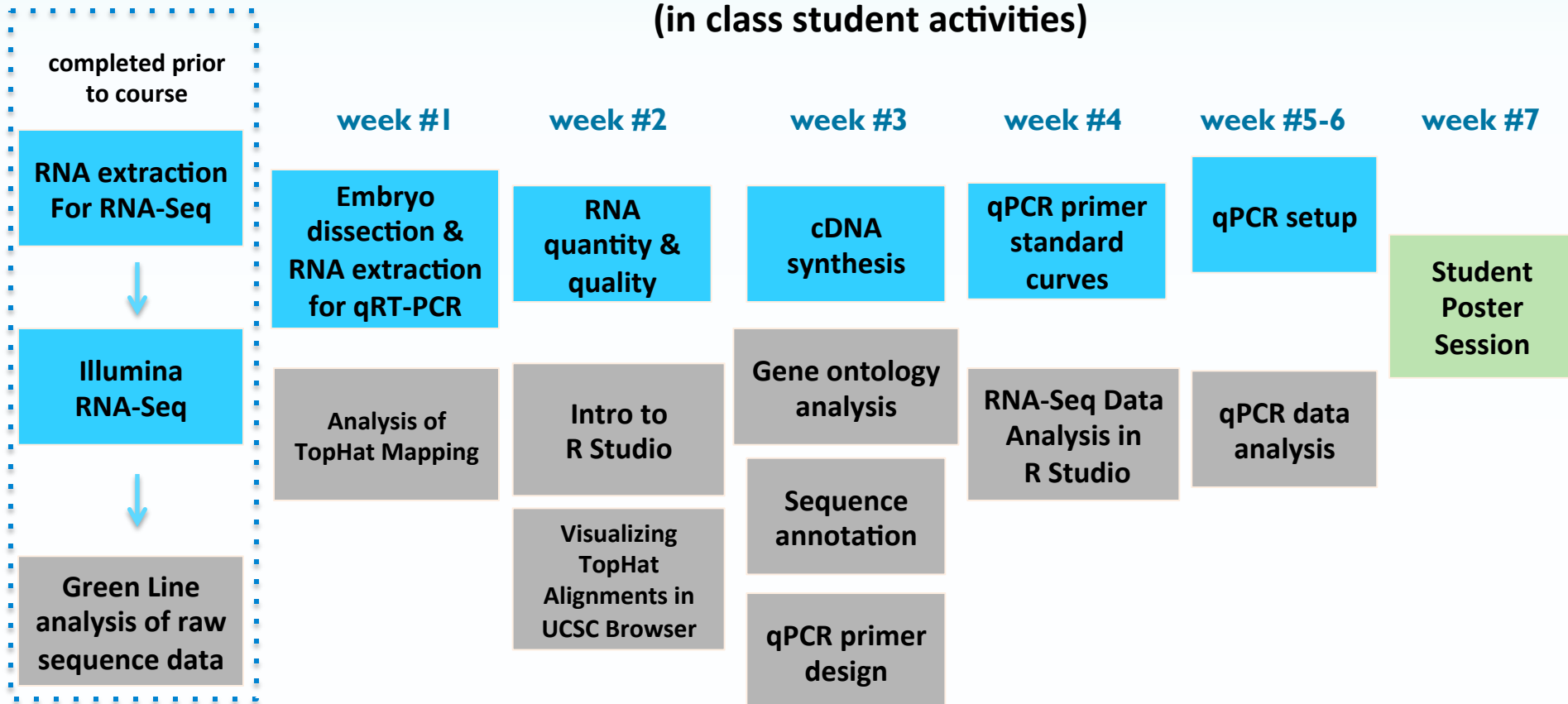
Key:

Wet lab
module

Computational
module

Evolution of activities over 3 semesters

(in class student activities)



*most recent iteration of RNA-seq module

Key:

Wet lab
module

Computational
module

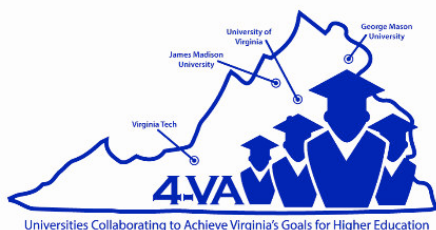
Thanks!

2014 RNA-Seq Working Group



- CHSL DNA Learning Center
- JMU Biology & GCEMS
- Christophe Langouet
- Annamarie Meinsen
- Stephen Turner (UVA)
- JMU/UVA 4-VA Funding

contact: enkerajmu.edu
Twitter: @Enke_Lab



Twitter: @JMU_CGEMS



RNA-Seq for the Next Generation