
output: pdf_document

2016 General Notebook

Author: Andrew D. Nguyen

**Affiliation: Biology Department,
University of Vermont**

Contact: anbe642@gmail.com

Date started: 2016-05-13

Date end (last modified): 2017-01-01

Introduction:

I wish I started an online notebook earlier, but maybe it's not too late? Anyway, I'll use this doc to share my ideas and log the progress of my dissertation.

<a rel="license" href="<http://creativecommons.org/licenses/by/4.0/>">

This work is licensed under a <a rel="license"
href="<http://creativecommons.org/licenses/by/4.0/>">Creative
Commons Attribution 4.0 International License.

###General Lab protocols found [here](#) for heat shocks and RNA related experiments and [here](#) for protein related experiments.

###Table of Contents (Layout follows Page number: Date. Title of entry)

- [Page 1: 2016-05-13](#). Indirect genetic effects
- [Page 2: 2016-05-13](#). Comparing G matrices of different populations
- [Page 3: 2016-05-16](#). Complete ddRAD-seq samples: processing
- [Page 4: 2016-05-16](#). Aphaenogaster morphological IDs
- [Page 5: 2016-05-16](#). Sequencing qPCR amplicons; Curtis and ANBE experiments
- [Page 6: 2016-05-17](#). Phylogenetics results from 2016-05-16 (CIPRES RaxML analysis)
- [Page 7: 2016-05-17](#). ABI steponeplus machine maintenance.
- [Page 8: 2016-05-18](#). Phylogenetic results excluding pogo (CIPRES RAXML analysis)
- [Page 9: 2016-05-18](#). Agarose gel electrophoresis of qPCR

amplicons; Curtis and ANBE samples

- [Page 10: 2016-05-18.](#) RaxML ML pairwise distance matrix
- [Page 11: 2016-05-18.](#) ABI steponeplus machine maintenance update
- [Page 12: 2016-05-19.](#) Getting whole rad loci with pyRAD and/or population in Stacks
- [Page 13: 2016-05-20.](#) Evolution of proteome stability project
- [Page 14: 2016-05-24.](#) Evolution of proteome stability project: Polyacrylamide gels for colony level replicates (*A. rufis* vs *P. barbatus*)
- [Page 15: 2016-05-24.](#) Degenerate Hsp primer design from 2015-05-28
- [Page 16: 2016-05-24.](#) Sequencing analysis continued from [Page 5: 2016-05-16.](#)
- [Page 17: 2016-05-25.](#) Double check samples for SHC; JSG phytotron exp and MS.
- [Page 18: 2016-05-31.](#) Learning model selection and model averaging!

- [2016-06-01: SHC suggestion](#). Just adding all phylo components in a global model instead of screening them out initially
- [Page 19: 2016-06-01](#). Variance partitioning: thermal tolerance breadth example
 - [Table summary of variance partitioning for CTmax, CTmin, and tolerance breadth](#)
- [Page 20: 2016-06-02](#). Notes from climate cascade meeting (2016-06-01)
- [Page 21: 2016-06-02](#). Levine's test for raw residuals
- [Page 22: 2016-06-02](#). **Brute force fitting nls() functions in R!!**
- [Page 23: 2016-06-02](#). Literature reference for thermal niche paper to help write manuscript
- [Page 24: 2016-06-03](#). Proteome stability project: Organizational entry
- [Page 25: 2016-06-03](#). ggplot reference, updating a figure from [Page 20: 2016-06-02](#)
- [Page 26: 2016-06-03](#). What is a cell type?

- [Page 27: 2016-06-03](#). qPCR plate layout and using the loaner ABI steponeplus [Page 11: 2016-05-18](#)
- [Page 28: 2016-06-03](#). Papers showing differences between fast static vs slow dynamic temperature treatments.
- [Page 29: 2016-06-06](#). Isolating RNA: colony CJ8; showing Sylvia
- [Page 30: 2016-06-07](#). **Brute force fitting nls function in R revisited ** [Page 22: 2016-06-02](#)
- [Failed attempt with nls2()] (#id-section30.1).
- Success with **failwith()**
- [Page 31: 2016-06-08](#). Re-doing online notebook template
- [Page 32: 2016-06-08](#). qPCRs, 18s rRNA for Duke2, HF2, Kite 4, Kite8, 60 C annealing. Dilutions of future samples
- [Page 33: 2016-06-08](#). Climate cascade meeting
- [Page 34: 2016-06-09; 2016-06-10](#). qPCRs: Duke1, CJ2, SHC8, CJ5
- [Page 35: 2016-06-10](#). ABI steponeplus machine fix and sending back instrument.

- [2016-06-13 update](#)
- [Page 36: 2016-06-10.](#) Thoughts on Kingsolver & Woods 2016, AmNat
 - [2016-06-11.](#) Follow up model
 - [2016-06-13.](#) Predictions of thermostat model
- [Page 37: 2016-06-11.](#) Quantifying natural selection in natural populations
- [Page 38: 2016-06-13.](#) qPCR update for Duke1,CJ2,SHC8,CJ5. Randomizing samples treated at 25C(reference for basal expression) for qpcrs.
- [Page 39: 2016-06-13.](#) Post doc project idea: Assessing current impacts of climate change in natural populations.
- [Page 40: 2016-06-14.](#) qPCR's: Diluting samples for quantifying basal expression and repeats
- [Basal expression plate layout.](#)
- [Status update of samples.](#)
- [Page 41: 2016-06-15.](#) qPCRs to quantify basal expression
(Evolution of stress response project)
- [Page 42: 2016-06-15.](#) Evolution talks I want to attend.

- [Page 43: 2016-06-16](#). Figure for curve fitting: see [Success with failwithQ](#) and [Status update of samples](#).
- [Page 44: 2016-07-18](#). Summary statistics for modulation of Hsp paper.
- [Page 45: 2016-07-19](#). Meeting with [VGN proteomics facility](#)
- [Page 46: 2016-07-21](#). Reference samples for mapping index; Hsp modulation and thermal niche paper
- [Page 47: 2016-07-26](#). Learning mixed effects stat models
- [Page 48: 2016-07-27](#). Meeting with Steve Keller to discuss post doc idea (started here: [Page 37: 2016-06-11](#). Quantifying natural selection in natural populations)
- [Page 49: 2016-07-28](#). Quantitative genetics and the molecular basis of complex traits
- [Page 50: 2016-08-02](#). Picking a plant system for post doc idea
- [Page 51: 2016-08-02; 2016-08-03](#). Climate cascade meeting
- [Page 52: 2016-08-04](#). Following up stats, range limits project
- [Page 53: 2016-08-08](#). Post doc ideas part 2 ; see [Page 50: 2016-](#)

08-02. Picking a plant system for post doc idea

- [Page 54: 2016-08-10](#). Climate cascade meeting
- [Page 55: 2016-08-11](#). Overlaying raster files in a map in R
- [Page 56: 2016-08-16](#). range limits paper, data analysis of chill coma recovery time (CCRT) revisited
- [Page 57: 2016-08-25](#). Hsp modulation follow up stats
- [Page 58: 2016-08-29 and 30](#). Climate cascade meeting
- [Page 59: 2016-09-01](#). SHC lab meeting Fall 2016
- [Page 60: 2016-09-01](#). Paper notes: Paccard A, Van Buskirk J, Willi Y, Eckert CG, Bronstein JL. 2016. Quantitative Genetic Architecture at Latitudinal Range Boundaries: Reduced Variation but Higher Trait Independence. *The American Naturalist*.
- [Page 61: 2016-09-06](#). Playing with rpart with range limit data
- [Page 62: 2016-09-06](#). Climate cascade meeting
- [Page 63: 2016-09-07](#). PCA update for range limit data; see [Page 61: 2016-09-06](#). Playing with rpart with range limit data

- [Page 64: 2016-09-12](#). ref for time table, nsf post doc grant
- [Page 65: 2016-09-12](#). variable importance
- [Page 66: 2016-09-13](#). climate cascade meeting
- [Page 67: 2016-09-14](#). SICB meeting
- [Page 68: 2016-09-19; 2016-09-20](#). Climate cascade meeting
- [Page 69: 2016-09-21](#). qpcr redos for 18s rRNA
- [Page 70: 2016-09-26](#). selecting poplar clones
- [Page 71: 2016-09-26 and 2016-09-27](#). Climate cascade meeting
- [Page 72: 2016-09-27](#) . evolution of hsp gxp data analysis
- [Page 73: 2016-09-28](#). building ultrametric trees
- [Page 74: 2016-09-28](#). phylogenetic regressions (PGLS) and anovas
 - [2016-09-28](#). SHC suggestion: ancestral trait reconstrurction -> regressions/anovas; [following this tutorial](#)
 - [2016-09-29](#). PIC

- [Page 75: 2016-10-03 and 2016-10-04](#). Climate cascade meeting
- [Page 76: 2016-10-03 and 2016-10-04](#). Membrane stability
- [Page 77: 2016-10-04](#). Lab Safety Officer (LSO) meeting.
- [Page 78: 2016-10-05](#). Hsp gxp function valued trait fig
- [Page 79: 2016-10-06](#). SHC lab meeting: NSF post doc app
- [Page 80: 2016-10-07](#). Prepping climate cascade meeting
- [Page 81: 2016-10-11](#). ANCOVA models for testing interaction of hsp gxp parameter and habitat on CTmax
- [Page 82: 2016-10-11](#). variance partitioning in CTmax of aphaeno
- [Page 83: 2016-10-12](#). Testing effect of MAT on Hsp gxp and looking at correlations between phylogeny and climate.
 - [2016-11-01 adding full models with automated stepAIC](#).
- [Page 84: 2016-10-14](#). Updating climate cascade to do list.
- [Page 85: 2016-10-14](#). Paper note: Puentes, A., G. Granath, and J. Ågren. 2016. Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*. *Evolution* 70:2370-2386.

- [Page 86: 2016-10-14. Wiley House Style Guide](#)
- [Page 87: 2016-10-14. NSF post doc app meeting: Keller Lab](#)
- [Page 88: 2016-10-18. Climate cascade meeting](#)
- [Page 89: 2016-10-25. Climate cascade updated list](#)
- [Page 90: 2016-10-25. Meeting with M Pespeni on 2016-10-26 and Brent 2016-10-27](#)
- [Page 91: 2016-10-26. SICB meeting talk](#)
- [Page 92: 2016-10-27. Proteome stability project update](#)
- [Page 93: 2016-10-31. CTmax and Hsp reaction norm stats](#)
- [Page 94: 2016-10-31; 2016-11-01. Climate cascade meeting setup and notes](#)
- [Page 95: 2016-11-02. Ancestral trait reconstruction and CTmax PGLS ANBE common garden; corresponds with \[Page 74: 2016-09-28. phylogenetic regressions \\(PGLS\\) and anovas\]\(#\)](#)
- [Page 96: 2016-11-03. notes from skype meeting with KG, potential post doc opp](#)

- Page 97: 2016-11-04. ms in prep
 - Page 98: 2016-11-08. climate cascade meeting
 - Page 99: 2016-11-08. writing session with NJG and stats follow up
 - among colony variance
 - PCA of climate variables
 - CTmax response and regressions with PCA of climate variables, mat, Tmax, latitude
 - Hsp PCA
 - Ctmax variance partitioning into Hsp rxn norm, phylo, loca env
 - Page 100: 2016-11-14 & 2016-11-15. climate cascade meeting
 - Page 101: 2016-11-16. Hsp reaction norm stats; adding quadratic term
 - Page 102: 2016-11-22. climate cascade to do list
 - Page 103: 2016-12-06. climate cascade update
 - Page 104: 2016-12-19. climate cascade update
 - Page 105: 2016-12-20. Reading a few papers
-

<div id='id-section1'>

Page 1: 2016-05-13. Indirect genetic effects.

Q: How does the social environment impact traits of individuals? Or what is the contribution of indirect genetic effects on an individual?

In ant colonies, sisters are highly related if the queen mated once.

H1: Ant workers traits are more optimal when the rearing environment is of the same genotype compared to different genotype.

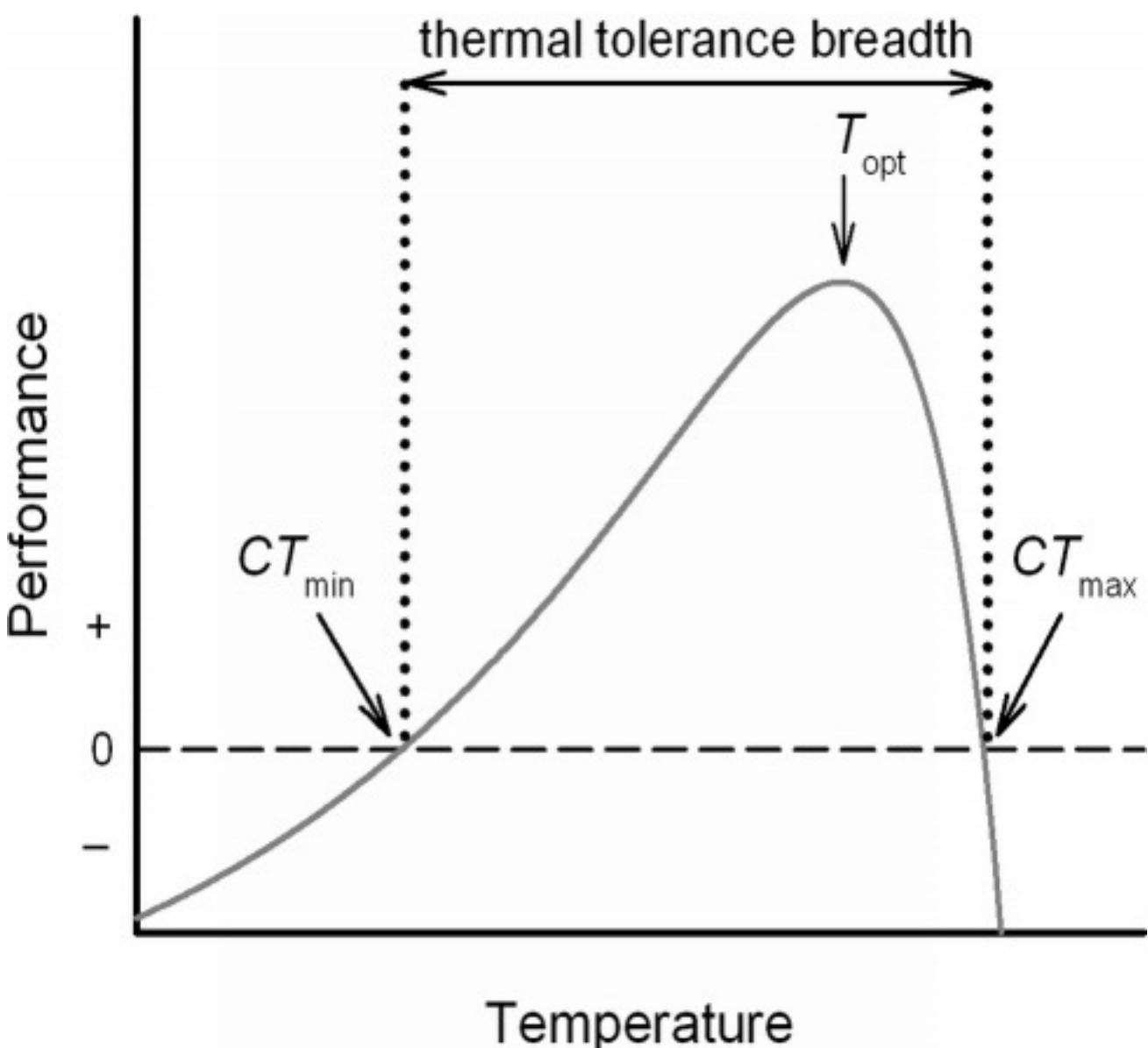
Experiment: Cross foster experiment. Each ant colony is a different genotype, take 20 ants and split them up so each colony rears each other's babys.

This isn't a new [idea: Linksvayer 2007](#). What would be interesting is to test the role of IGE in thermal ecology. Take a Northern(experiecnies cold) ant species and Southern (experiences warm) ant species and do a cross foster experiment. One outcome is that ants reared in the warm tolerant species will rear young in a way so that the baby has greater thermal tolerance than being reared by its own (cold tolerant genotype/species).

<div id='id-section2'>

Page 2: 2016-05-13. Comparing G matrices of different populations

Since I've been an RA since January 2015, I've been able to teach myself things. One of my emerging obsessions is understanding how multiple traits evolve or respond to selection. For example a thermal performance curve is multivariate and how can this curve change?



It can vary vertically, shift right to left (warmer-cooler variation), and/or exhibit generalist-specialist variation. [Kingsolver et al. 2105](#) has

a cool paper showing how you can construct a G matrix, decompose it with a PCA to look at the genetic correlations and it subsequently captures how G matrices can change or thermal performance curves can respond to selection together. So all positive loadings equals vertical shifts, positive relationships of loadings with temperature equals warmer-cooler variation, and a bell shaped curve equals the generalist-specialist variation.

Example table of loadings across each temperature:

Variation	15	20	25	30	35
Vertical	1	1	1	1	1
Warmer-cooler	-1	-.5	0	.5	1
Generalist-specialist	-1	.5	1	.5	-1

Whoa, what if you wanted to compare G matrices of different populations? One way is to do a PCA decomp with each G matrix constructed from each population. Then simply look at how the loadings change as a function of temperature between populations. Statistically, you can do an ANCOVA such as:

```
#Loadings is a continuous variable  
#Temperature can be a factor or continuous  
#Population is a factor  
aov(Loadings ~ Temperature * Population)
```

A cool paper by [Berger et al. 2013](#) has sort of done this (with out the ANCOVA). In table 3, they have gmax loadings (1st eigenvector of their G matrix) for each temperature for 3 populations: North, Central, South. So the Northern population exhibits warmer-cooler variation (high loadings low temps, negative loadings on high temps), whereas, Central and South exhibit vertical variation (all loadings are positive).

There is another cool paper to read about comparing G matrices by [Aguirre et al. 2014](#).

1. Random skewers method; simulate response to selection by calculating it with randomized betas
 2. Common subspace; no clue what this is
 3. Construct a tensor; sounds like a 3D G matrix
 4. Decompose G into eigenvectors; like Kingsolver, I believe
-

<div id='id-section3'>

Page 3: 2016-05-16. Complete ddRAD-seq samples: processing

ddrad-seq data are in! SHC processed short reads in STACKS and produced a fasta file.

From SHC:

Hi Andrew,

I have run all your samples against your index and through the STACKs pipeline - I used a minimum threshold of 5 reads to call a SNP, a maximum # of SNPs per tag of 6, and a minimum number of individuals that had to have a genotype call at a SNP of 10 individuals. The stats of genotype calls and heterozygosity across all your samples is in the excel spreadsheet - I highlighted those with <25% calls in yellow, and would not use those because they mess up the polarity inference for the SNPs and make the tree more ambiguous. The exception would probably be NOVCOC, since you will need an outgroup and none of the putative outgroup taxa meet the threshold. I've attached a NJ tree using all the >25% taxa plus NOVCOC, and it seems to resolve very nicely bootstrap-wise. I do not know what many of these samples are, so no clue if it is biologically reasonable.

You'll find your fasta file in my scratch space here:

Andrew_RADseq_051516/final_Andrew_sam_files/m5_output_refmap/
Sara

But, samples need to be redone:

Hi Andrew,

Just realized I did not adjust the barcode key for two samples in ddRAD10 that got moved during library prep - KITE5 and GF34-1. So their data are incorrect. Fixing now and should have a new version in a day or two.

Sara

So the following files should be disregarded but I'm keeping them just

to log them:

In the 2014_xanbe-common-garden_gxp_evolution/Data/Phylogenetics/20160516_complete_dataset_pl

- 20160516_SHC_Andrew_final_m5_filter6_ind10_NJtree.pdf
- 20160516_SHC_Andrew_het-summary_SNPs.xlsx
- 20160516_SHC_Andrew_SNP_sequences_m5_filter6_ind10.fas

But if you ignore KITE5 and GF34-1, here is the summary of results:

Sample	SNPs	Hets	Total	Prop.SNPs	Prop
FORMICA	47	1	173822	0.00	0
PB17-10_cat	203	0	173822	0.00	0
CAMPNSP	584	25	173822	0.00	0
PB17-14	1031	10	173822	0.01	0
PB07-23	1921	14	173822	0.01	0
09A	2587	32	173822	0.01	0
CREMATOGASTER_cat	3094	32	173822	0.02	0
Kite8r	3751	56	173822	0.02	0
TU64_cat	5217	45	173822	0.03	0
Sal13-14r	7905	78	173822	0.05	0
BK6-1	10743	182	173822	0.06	0
EXIT65	11612	120	173822	0.07	0
NOVCOC1	12013	34	173822	0.07	0

ALA4	18707	494	173822	0.11	0
KITE4_cat	36845	494	173822	0.21	0
AHF3r	39557	455	173822	0.23	0
Duke3r	53391	827	173822	0.31	0
FBR5r	61628	1072	173822	0.35	0
KITE5_cat	65777	1745	173822	0.38	0
KH1	69951	977	173822	0.40	0
KH2r	72601	1021	173822	0.42	0
BSK5r	73690	1573	173822	0.42	0
FBRAGG1	76194	830	173822	0.44	0
AHW7	76776	1298	173822	0.44	0
AHF1r	77515	1038	173822	0.45	0
KH3	78618	1099	173822	0.45	0
Avon19-1	78942	1001	173822	0.45	0
Avon19-3	80182	1137	173822	0.46	0
MA	80584	1546	173822	0.46	0
AHW2	80841	1405	173822	0.47	0
FBRAGG3	81143	1103	173822	0.47	0
AHF2	82047	1399	173822	0.47	0
CJ2r	82383	1026	173822	0.47	0
SHC2	84679	1541	173822	0.49	0

CJ4	84824	1375	173822	0.49	0
HW10	85989	1521	173822	0.49	0
SHC9r	87346	1526	173822	0.50	0
MIC2	88435	1198	173822	0.51	0
LPR4	90037	1529	173822	0.52	0
DUKE2	91310	1890	173822	0.53	0
Ala5r	91524	2161	173822	0.53	0
SHC10	91772	1614	173822	0.53	0
CJ6r	94419	1386	173822	0.54	0
CJ7	95005	2888	173822	0.55	0
LexSHC7r	96193	1810	173822	0.55	0
YATES1	96271	1921	173822	0.55	0
DUKE1	96675	1731	173822	0.56	0
SWSR45-1r	97057	652	173822	0.56	0
CJ8r	99904	1318	173822	0.57	0
LexSHC8r	102414	1934	173822	0.59	0
SHC5	102824	1916	173822	0.59	0
SHC3	102969	1891	173822	0.59	0
LEX9	103046	990	173822	0.59	0
CJ3r	103819	2001	173822	0.60	0
ALA1_cat	104644	2454	173822	0.60	0

DUKE7	104763	3081	173822	0.60	0
DUKE5	105184	2362	173822	0.61	0
LPR1	105777	1459	173822	0.61	0
LEX11	106302	1999	173822	0.61	0
DUKE6	106634	1284	173822	0.61	0
KH5	111245	1899	173822	0.64	0
Avon19-2	111264	1667	173822	0.64	0
Lex1r	112200	2215	173822	0.65	0
AHW4	112462	2571	173822	0.65	0
KH7	113614	1765	173822	0.65	0
NewSh20-2	114686	1843	173822	0.66	0
KH6	116788	1914	173822	0.67	0
Duke9r	117894	1385	173822	0.68	0
KH4	118160	1794	173822	0.68	0
ALA3_cat	118525	2965	173822	0.68	0
CJ1	119737	1712	173822	0.69	0
FBR4r	122054	1894	173822	0.70	0
Yates2r	122085	2440	173822	0.70	0
AHW1	122370	1423	173822	0.70	0
YATES3	124183	2700	173822	0.71	0
SHC6	124396	2577	173822	0.72	0

Mon22-2	124452	2148	173822	0.72	0
NP20-3	124543	2092	173822	0.72	0
CJ9	124795	2533	173822	0.72	0
Burn21-1	124846	2087	173822	0.72	0
KH8	125663	2139	173822	0.72	0
Can21-2	125727	2192	173822	0.72	0
KITE1	126422	3578	173822	0.73	0
GB33-1	127665	2376	173822	0.73	0
CJ5r	127719	2798	173822	0.73	0
Duke8r	128227	1555	173822	0.74	0
SHC4r	128586	2703	173822	0.74	0
Ted3r	129299	2332	173822	0.74	0
TED4_cat	131556	2828	173822	0.76	0
Unit22-1	134451	2447	173822	0.77	0
ALA2_cat	134714	3708	173822	0.78	0
Sap	135261	2478	173822	0.78	0
Pal21-3	135373	2400	173822	0.78	0
POP2	135796	3030	173822	0.78	0
Norr20-1	135922	2502	173822	0.78	0
FBRAGG2	136534	3678	173822	0.79	0
Duke4r	136812	3048	173822	0.79	0

Camb31-1	136979	2424	173822	0.79	0
KITE2	137173	2190	173822	0.79	0
Hamp23-1	137953	2639	173822	0.79	0
LEX5	139853	3058	173822	0.80	0
Pop1r	139912	3187	173822	0.80	0
GF34-1	140928	4088	173822	0.81	0
POP3	140937	3175	173822	0.81	0
LPR2	143401	2190	173822	0.82	0
SHC1	145375	2371	173822	0.84	0
AHW5	145662	2407	173822	0.84	0
Phil20-4	147770	2915	173822	0.85	0
AHW3	148236	3804	173822	0.85	0
MIC1	149191	2737	173822	0.86	0
LEX13	149260	3486	173822	0.86	0
TED6	154029	3347	173822	0.89	0
PMBE_cat	163739	3120	173822	0.94	0
KITE3	166928	6437	173822	0.96	0

Preliminary Tree; NJ:



SHC sent updated fasta file:

Your fasta file should be ready again - turns out that GF34-1 mapped very poorly and really should not be used. The new SNP yield/heterozygosity summary file is in the same directory for you.

Sara

Got rid of old fasta file, here is the updated file list:

- 20160516-Andrew_SNP_sequences_m5_filter6_ind10_het.tsv ; summary
- 20160516_SHC_Andrew_SNP_sequences_m5_filter6_ind10.fas ; unmodified names
- 20160516_Andrew_SNP_sequences.fas; relabeled to match my sampling sheet; got rid of “_trimmed90_filtered”

Summary table of updated fasta file:

Sample	SNPs	Hets	Total	Proportion_loci_w
FORMICA	43	1	174008	
PB17-10_cat	203	0	174008	
CAMPNSP	590	23	174008	
PB17-14	1034	9	174008	
PB07-23	1924	14	174008	
09A	2608	34	174008	
CREMATOGASTER_cat	3087	32	174008	

Kite8r	3688	53	174008	
GF34-1	4035	28	174008	
TU64_cat	5180	45	174008	
Sal13-14r	7892	78	174008	
BK6-1	10723	174	174008	
EXIT65	11573	122	174008	
NOVCOC1	12003	34	174008	
ALA4	18742	500	174008	
KITE4_cat	36913	498	174008	
AHF3r	39632	458	174008	
Duke3r	53458	813	174008	
FBR5r	61790	1073	174008	
KH1	70047	977	174008	
KH2r	72760	1024	174008	
BSK5r	73728	1575	174008	
FBRAGG1	76234	832	174008	
AHW7	76850	1278	174008	
AHF1r	77526	1043	174008	
KH3	78767	1102	174008	
Avon19-1	79026	995	174008	
Avon19-3	80160	1125	174008	

MA	80715	1536	174008	
AHW2	80937	1402	174008	
FBRAGG3	81176	1122	174008	
AHF2	82223	1396	174008	
CJ2r	82528	1023	174008	
SHC2	84811	1527	174008	
CJ4	85003	1371	174008	
HW10	85935	1512	174008	
SHC9r	87518	1514	174008	
MIC2	88542	1199	174008	
LPR4	90158	1530	174008	
DUKE2	91423	1896	174008	
Ala5r	91632	2171	174008	
SHC10	91826	1595	174008	
CJ6r	94504	1388	174008	
CJ7	95178	2898	174008	
LexSHC7r	96265	1803	174008	
YATES1	96479	1934	174008	
DUKE1	96531	1570	174008	
SWSR45-1r	97061	654	174008	
CJ8r	100052	1315	174008	

LexSHC8r	102556	1914	174008	
SHC5	102976	1895	174008	
LEX9	103074	994	174008	
SHC3	103077	1882	174008	
CJ3r	103816	1963	174008	
ALA1_cat	104771	2433	174008	
DUKE7	104940	3087	174008	
DUKE5	105313	2376	174008	
LPR1	105841	1459	174008	
LEX11	106390	1984	174008	
DUKE6	106792	1291	174008	
Avon19-2	111266	1661	174008	
KH5	111410	1902	174008	
Lex1r	112257	2203	174008	
AHW4	112475	2552	174008	
KH7	113763	1762	174008	
NewSh20-2	114753	1863	174008	
KH6	116912	1917	174008	
Duke9r	117978	1390	174008	
KH4	118263	1797	174008	
ALA3_cat	118653	3003	174008	

CJ1	119837	1716	174008	
FBR4r	122154	1887	174008	
Yates2r	122241	2424	174008	
AHW1	122370	1435	174008	
YATES3	124252	2669	174008	
SHC6	124556	2553	174008	
Mon22-2	124561	2157	174008	
NP20-3	124747	2105	174008	
CJ9	124875	2508	174008	
Burn21-1	124936	2101	174008	
Can21-2	125784	2198	174008	
KH8	125792	2150	174008	
KITE1	126638	3576	174008	
GB33-1	127656	2385	174008	
CJ5r	127851	2772	174008	
Duke8r	128355	1556	174008	
SHC4r	128604	2669	174008	
Ted3r	129289	2348	174008	
TED4_cat	131758	2863	174008	
Unit22-1	134508	2472	174008	
ALA2_cat	134818	3729	174008	

Pal21-3	135398	2411	174008	
Sap	135413	2487	174008	
POP2	135928	3004	174008	
Norr20-1	136013	2506	174008	
FBRAGG2	136626	3680	174008	
Duke4r	136895	3035	174008	
Camb31-1	137074	2448	174008	
KITE2	137322	2185	174008	
Hamp23-1	138088	2646	174008	
Pop1r	139982	3140	174008	
LEX5	139987	3014	174008	
POP3	141037	3140	174008	
LPR2	143432	2185	174008	
SHC1	145541	2382	174008	
AHW5	145766	2409	174008	
Phil20-4	147887	2925	174008	
AHW3	148314	3796	174008	
MIC1	149322	2762	174008	
LEX13	149401	3461	174008	
TED6	154109	3362	174008	
KITE5_cat	157748	5246	174008	

PMBE_cat	163881	3111	174008	
KITE3	167083	6441	174008	

Parsed 20160516_Andrew_SNP_sequences.fas:

- got rid of samples with low number of SNPs
 - FORMICA
 - PB17-10
 - CAMPNSP
 - PB17-14
 - PB07-23
 - 09A
 - CREMATOGASTER
 - Kite8r
 - GF34-1

Grabbing number of samples from command line:

```
grep '^>' 20160516_Andrew_SNP_sequences.fas | wc -l
107
```

107 samples!

Next step is to reconstruct relationships of SNP Matrix

1. Use [CIPRES](#)
2. Use RAxML-HPC BlackBox (8.2.8) to reconstruct ML tree
3. I also need to estimate the ML distance matrix with computer in ant room.

For ML distance matrix with raxml, you need a fasta file and tree.

Piece of code I've tried before:

```
*##for anbe tree, calculate pairwise ml distance matrix  
nohup nice -n 19 ./raxmlHPC -fx -p 12345 -s  
~/Desktop/2015_ANBE_common_garden/20150818_Andrew_SNP_sequ  
-m GTRGAMMA -t  
~/Desktop/2015_ANBE_common_garden/RAxML_bestTree.20150819_c  
-n 20150828_commongarden_pairwise_ML_distance &
```

<div id='id-section4'>

Page 4: 2016-05-13. Aphaenogaster morphological IDs

For JSG phytotron project (and also partly Lchick's thermal niche paper).

ID	<u>Colony.ID</u>	Species	Vouchers	<u>Bernice.morp</u>
ApGXL-01-A	MagSpr3	carolinensis		
ApGXL-01-B	MagSpr4	rudis		
ApGXL-01-C	MagSpr7	carolinensis		
ApGXL-	HW1	rudis		rudis

02-A				
ApGXL-02-B	HW5	rudis		
ApGXL-02-C	HW7	rudis	voucherNCSU	rudis
ApGXL-03-A	FMU4	.		
ApGXL-04-A	UNF8	rudis		rudis
ApGXL-04-B	UNF9	rudis		rudis
ApGXL-04-C	UNF1	carolinensis		carolinensis
ApGXL-05-B	GSMNP4	picea		picea
ApGXL-05-D	GSMNP5	picea		picea
ApGXL-06-A	DW2	rudis		rudis
ApGXL-06-B	DW1	rudis		rudis
ApGXL-07-A	BRP2	picea	voucherNCSU	picea
ApGXL-07-B	BRP9	picea	voucherNCSU	
ApGXL-08-A	Ijams6	rudis		rudis

ApGXL-08-D	IJams1	rudis		rudis
ApGXL-09-A	RC12	rudis		rudis
ApGXL-10-A	LVA9	rudis		rudis
ApGXL-10-B	LVA12	rudis		rudis
ApGXL-10-C	LVA11	fulva		fulva
ApGXL-10-F	LVA9	rudis		rudis
ApGXL-11-A	WP9	rudis	voucherNCSU	rudis
ApGXL-11-B	WP11	rudis	voucherNCSU	rudis?
ApGXL-11-C	WP3	fulva	voucherNCSU	fulva

ApGXL-11-D	WP6	rudis		rudis
ApGXL-12-A	NOCK6	picea		rudis
ApGXL-12-D	NOCK8	rudis		rudis
ApGXL-13-A	HSP6	picea		picea
ApGXL-13-B	HSP7	picea		picea
ApGXL-13-C	HSP9	picea	voucherNCSU	picea

ApGXL-13-D	HSP12	picea		picea
ApGXL-15-A	DSF4	picea	voucherNCSU	picea
ApGXL-15-B	DSF11	picea	voucherNCSU	picea
ApGXL-15-C	DSF8	picea		picea
ApGXL-15-D	DSF12	picea	voucherNCSU	picea
APGXL-16-A	BRM4	picea		picea
APGXL-16-B	BRM/BRF8	picea		picea
ApGXL-17-A	Bard10	picea	voucherNCSU	picea
ApGXL-17-B	Bard9	picea	voucherNCSU	picea
ApGXL-17-C	Bard3	picea		picea
ApGXL-18-A	Notch1	fulva	voucherNCSU	picea

ApGXL-18-C	Notch4	rudis		picea
ApGXL-18-D	Notch2	fulva	voucherNCSU	picea
ApGXL-19-A	HF001	picea		picea
ApGXL-20-A	APB10	picea	voucherNCSU	picea
ApGXL-20-B	APB3a	picea		picea
ApGXL-20-C	APB3b	picea		picea
ApGXL-20-D	APB8	picea		picea
ApGXL-21-A	Bear6	picea		picea
ApGXL-21-B	Bear5	picea		picea
ApGXL-21-C	Bear3	picea		picea
ApGXL-22-A	SEB1	.		picea
ApGXL-22-B	SEB8	picea		picea
ApGXL-22-C	SEB9	picea		picea

ApGXL-23-A	MM1	picea	voucherNCSU	picea
ApGXL-23-B	MM2	picea		picea
ApGXL-23-C	MM4	picea	voucherNCSU	picea
ApGXL-24-A	EW09	picea		picea
ApGXL-24-B	EW4	.		picea
ApGXL-25-A	RW3	picea	voucherNCSU	picea
ApGXL-25-C	RW1	.		
ApGXL-25-D	RW5	picea		picea
ApGXL-26-A	MB1	picea	voucherNCSU	picea
ApGXL-26-B	MB3	picea	voucherNCSU	
ApGXL-26-C	MB4	picea	voucherNCSU	picea
ApGXL-26-D	MB2	picea	voucherNCSU	picea
ApGXL-26-E	MB6	picea	voucherNCSU	picea

ApGXL-27-A	KBH4b	picea	voucherNCSU	picea
ApGXL-27-B	KBH1	picea	voucherNCSU	picea
ApGXL-28-A	Brad1	picea		picea
ApGXL-28-B	Brad6	picea	voucherNCSU	picea
Aphaen 15	Aphaen 15			
Aphaen A2	Aphaen A2			
Aphaen12	Aphaen12			
Aphaen17	Aphaen17			
Aphaen18	Aphaen18	rudis		
AphaenA	AphaenA	rudis	voucherNCSU	
AphaenB	AphaenB			
BARD11	BARD11			
BARD2	BARD2	picea		
BARD5	BARD5	fulva		
Blank	Blank	rudis	voucherNCSU	
Brad2	Brad2	picea	voucherNCSU	
Brad3	Brad3			

BRP-2B	BRP-2B	picea		
BRP08	BRP08			
BRP1	BRP1	picea	voucherNCSU	
BRP10	BRP10			
BRP11	BRP11	picea	voucherNCSU	
BRP3	BRP3	picea	voucherNCSU	
BRP5	BRP5	picea	voucherNCSU	
BRP6	BRP6			
BRP7	BRP7	picea	voucherNCSU	
DF-3A	DF-3A	rudis	voucherNCSU	
DF1-A	DF1-A	rudis	voucherNCSU	
FMU6	FMU6	rudis	voucherNCSU	
HSP1	HSP1	picea		
HSP4	HSP4			
HSP5	HSP5	picea		
HW8	HW8			
HW9	HW9			
KBH6	KBH6			
KBH8	KBH8	picea	voucherNCSU	
LVA1	LVA1	fulva	voucherNCSU	
LVA13	LVA13	rudis	voucherNCSU	

LVA2	LVA2	rudis	voucherNCSU
LVA3	LVA3	rudis	voucherNCSU
MAGSPR6	MAGSPR6	rudis	voucherNCSU
NSP2	NSP2	picea	voucherNCSU
NSP3	NSP3	rudis	
NSP7	NSP7	fulva	
OLDRC1	OLDRC1	fulva	
OldRC3	OldRC3	fulva	
OldRC4	OldRC4	rudis	
OldRC6	OldRC6	rudis	
OLDRC7	OLDRC7	rudis	
RC02	RC02	fulva	voucherNCSU
RC04	RC04	rudis	
RC06	RC06	rudis	voucherNCSU
RC09	RC09	rudis	voucherNCSU
RC10	RC10	rudis	voucherNCSU
RC11	RC11	rudis	
RC13	RC13	rudis	voucherNCSU
RC14	RC14	rudis	
RC15	RC15	rudis	
RC16	RC16	rudis	

Seb 2A	Seb 2A			
SEB3A	SEB3A			
UNF4A	UNF4A	rudis		
UNF7A	UNF7A	miamiana		
YM01	YM01	rudis		
YM02	YM02	rudis		
YM03	YM03	rudis		

<div id='id-section5'>

Page 5: 2016-05-13. Sequencing qPCR amplicons; Curtis and ANBE experiments

Sample list and plate layout for sanger sequencing. Amplicons ~ 100bps and were Qiagen PCR purified following manufacturer's instructions. Added ~3 ng template, with 2 uM primer in 11.6 uL volume. Curtis' chamber samples are on [here](#) and my own ANBE gene expression experiment. Submitting to [vermont cancer center](#).

If interested in protocols , see [here](#).

Well	<u>Template.Name</u>	<u>Primer.Name</u>
A1	HF 5-1	18s_F328

B1	HF 5-1	18s_R427
C1	HF 7-1	18s_F328
D1	HF 7-1	18s_R427
E1	DF 13-A	18s_F328
F1	DF 13-A	18s_R427
G1	DF 14-A	18s_F328
H1	DF 14-A	18s_R427
A2	DF 8-B	hsp83_F1583
B2	DF 8-B	hsp83_R1682
C2	DF 5C-4	hsp83_F1583
D2	DF 5C-4	hsp83_R1682
E2	HF 8-1	hsp83_F1583
F2	HF 8-1	hsp83_R1682
G2	HF 2-2	hsp83_F1583
H2	HF 2-2	hsp83_R1682
A3	DF 1-D	hsp70_F1468
B3	DF 1-D	hsp70_R1592
C3	DF 10-3	hsp70_F1468
D3	DF 10-3	hsp70_R1592
E3	HF2 8-2	hsp70_F1468
F3	HF2 8-2	hsp70_R1592

G3	HF2 4-1	hsp70_F1468
H3	HF2 4-1	hsp70_R1592
A4	HF2 7-2	hsp40_F541
B4	HF2 7-2	hsp40_R641
C4	HF2 5-2	hsp40_F541
D4	HF2 5-2	hsp40_R641
E4	DF A1-B	hsp40_F541
F4	DF A1-B	hsp40_R641
G4	DF A8-B	hsp40_F541
H4	DF A8-B	hsp40_R641
A5	HF2 5-3	actin_F984
B5	HF2 5-3	actin_R1095
C5	HF2 8-1	actin_F984
D5	HF2 8-1	actin_R1095
E5	DF 3-A	actin_F984
F5	DF 3-A	actin_R1095
G5	DF 7-A	actin_F984
H5	DF 7-A	actin_R1095
A6	Exit65	70_1468
B6	BK	70_1468
C6	Ted6	70_1468

D6	DUKE6	70_1468
E6	ALA1	70_1468
F6	KH2	70_1468
G6	FB2	70_1468
H6	Exit65	70_1592
A7	BK	70_1592
B7	Ted6	70_1592
C7	DUKE6	70_1592
D7	ALA1	70_1592
E7	KH2	70_1592
F7	FB2	70_1592
G7	Exit65	83_1583
H7	BK	83_1583
A8	TED3	83_1583
B8	DUKE6	83_1583
C8	ALA1	83_1583
D8	KH2	83_1583
E8	FB2	83_1583
F8	Exit65	83_1682
G8	BK	83_1682
H8	TED3	83_1682

A9	DUKE6	83_1682
B9	ALA1	83_1682
C9	KH2	83_1682
D9	FB2	83_1682
E9	PB1710	83_279
F9	POP2	83_279
G9	SHC2	83_279
H9	cremato	83_279
A10	ex	83_279
B10	bk	83_279
C10	TED6	83_279
D10	PB1710	83_300
E10	POP2	83_300
F10	SHC2	83_300
G10	cremato	83_300
H10	ex	83_300
A11	bk	83_300
B11	TED6	83_300
C11	DUKE6	hsp40_541
D11	ALA1	hsp40_541
E11	KH2	hsp40_541

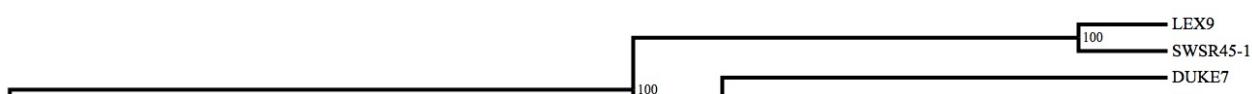
F11	FB2	hsp40_541
G11	EX	hsp40_541
H11	BK	hsp40_541
A12	Ted6	hsp40_541
B12	DUKE6	hsp40_641
C12	ALA1	hsp40_641
D12	KH2	hsp40_641
E12	FB2	hsp40_641
F12	EX	hsp40_641
G12	BK	hsp40_641
H12	Ted6	hsp40_641

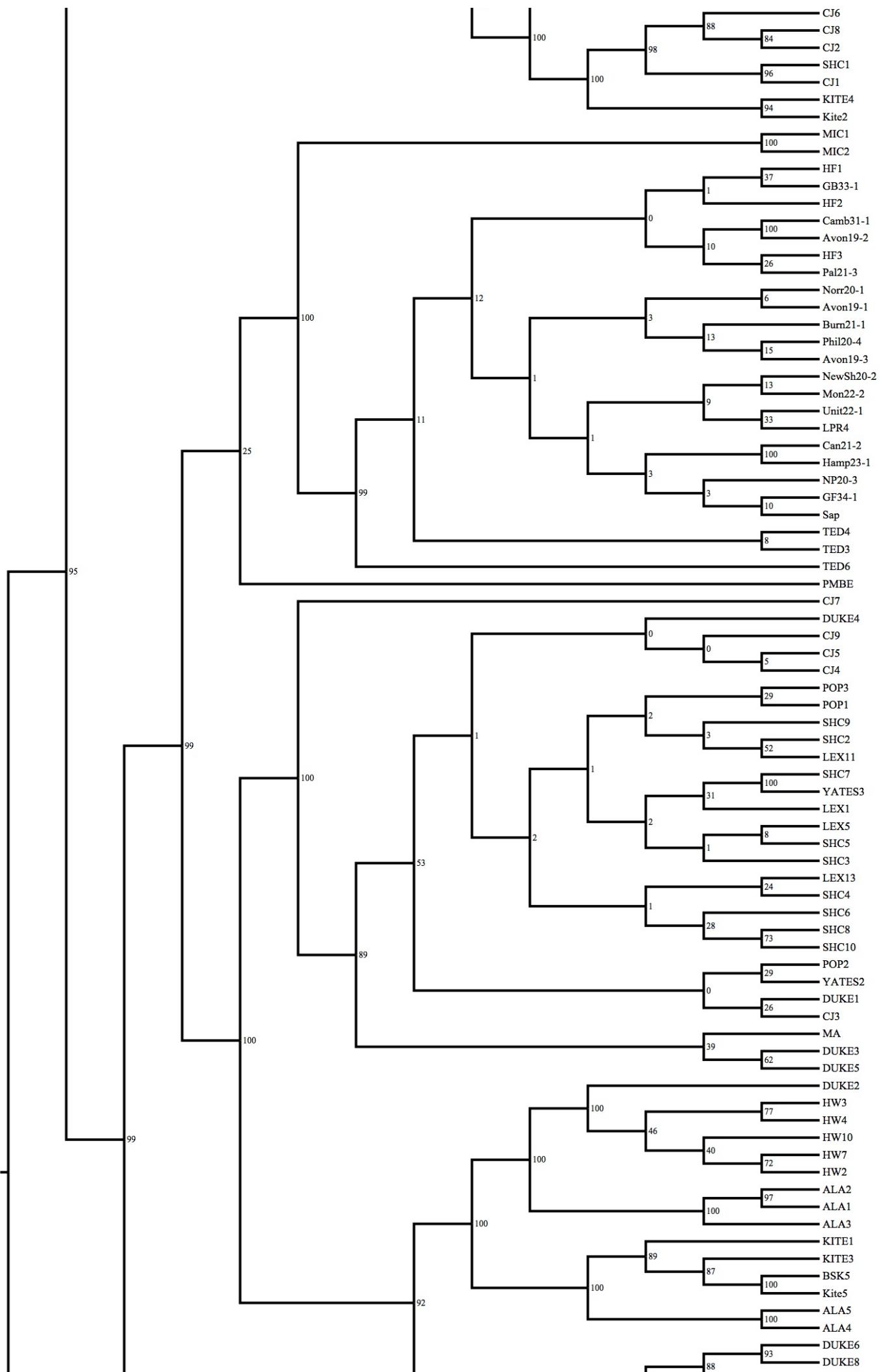
<div id='id-section6'>

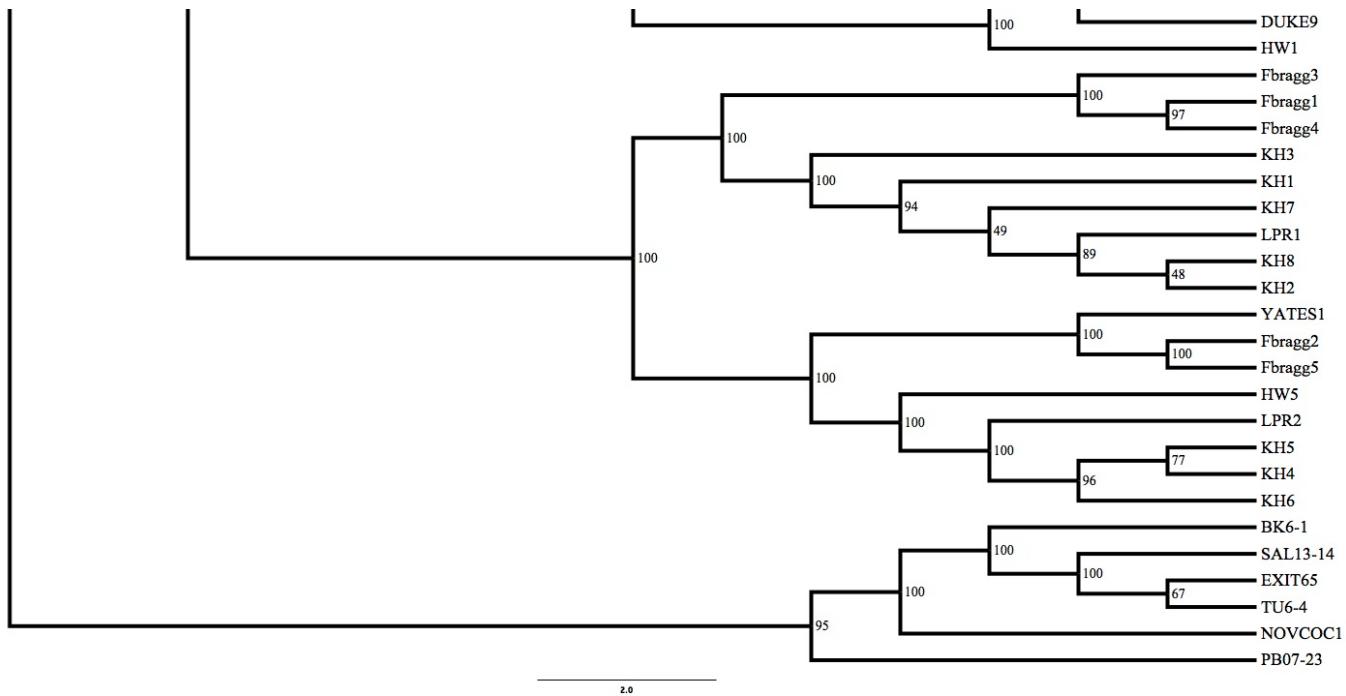
Page 6: 2016-05-17 Phylogenetics results from 2016-05-16 (CIPRES RaxML analysis)

Results from [2016-05-16](#) ML tree using RaxML black box on CIPRES.

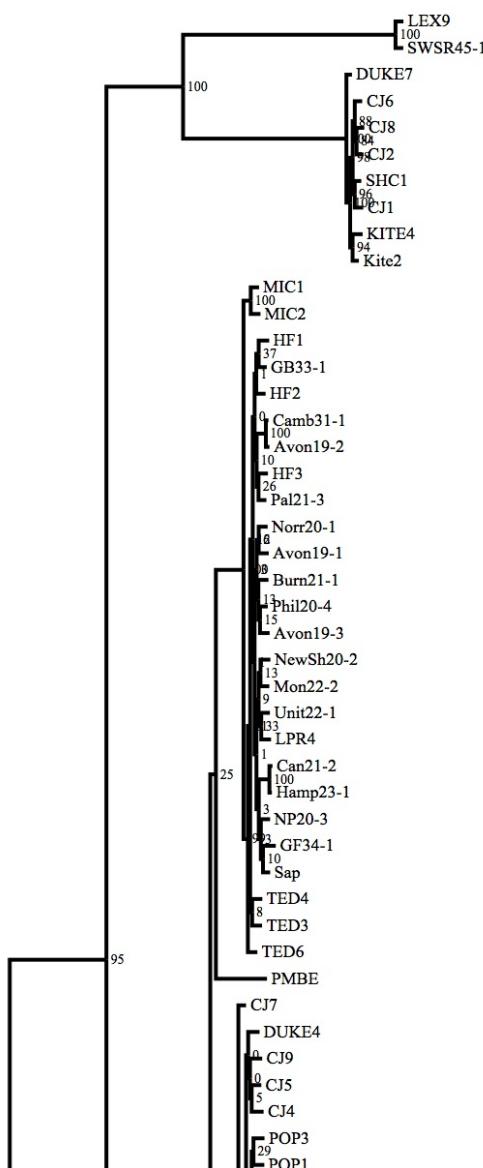
Transformed branch lengths

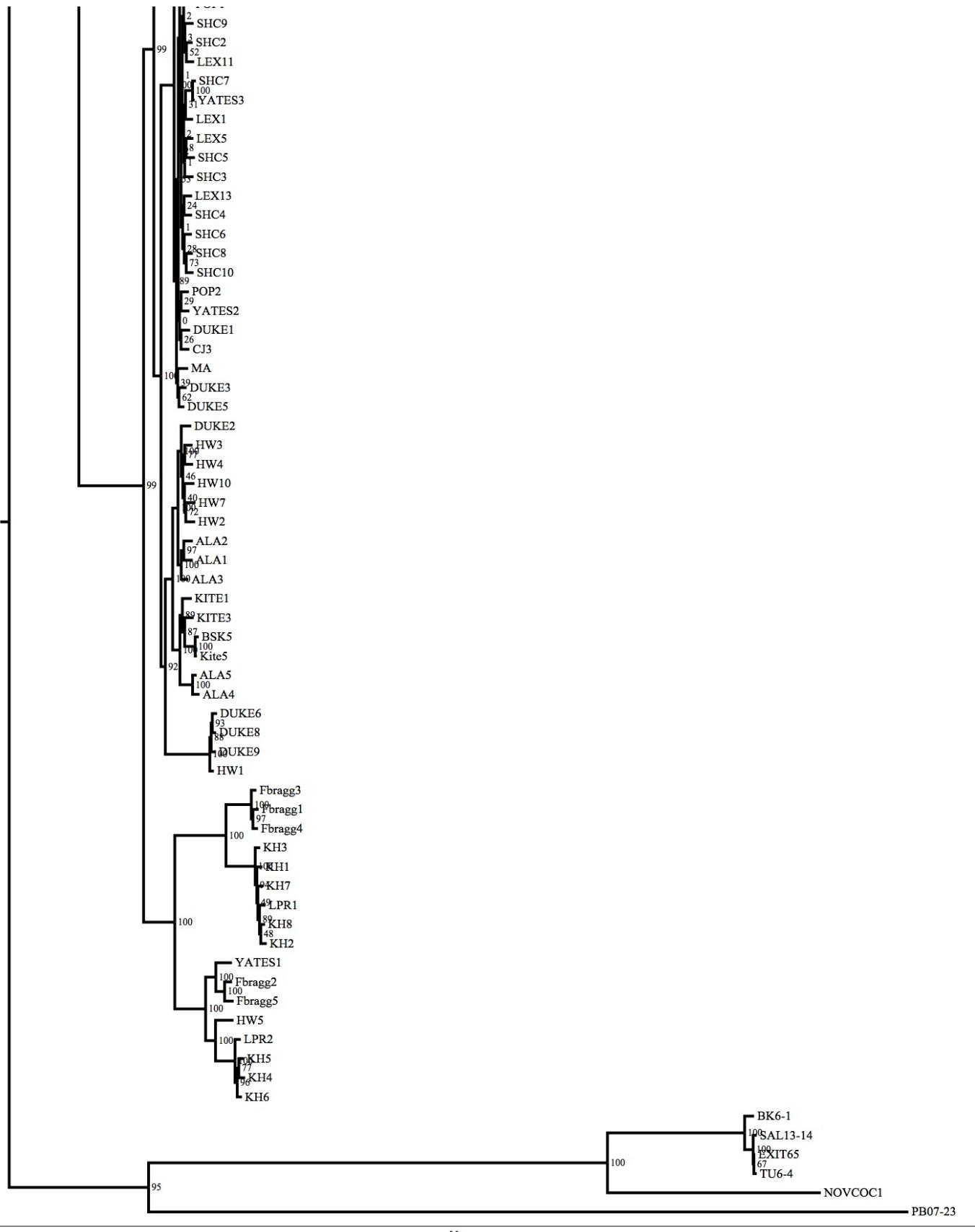






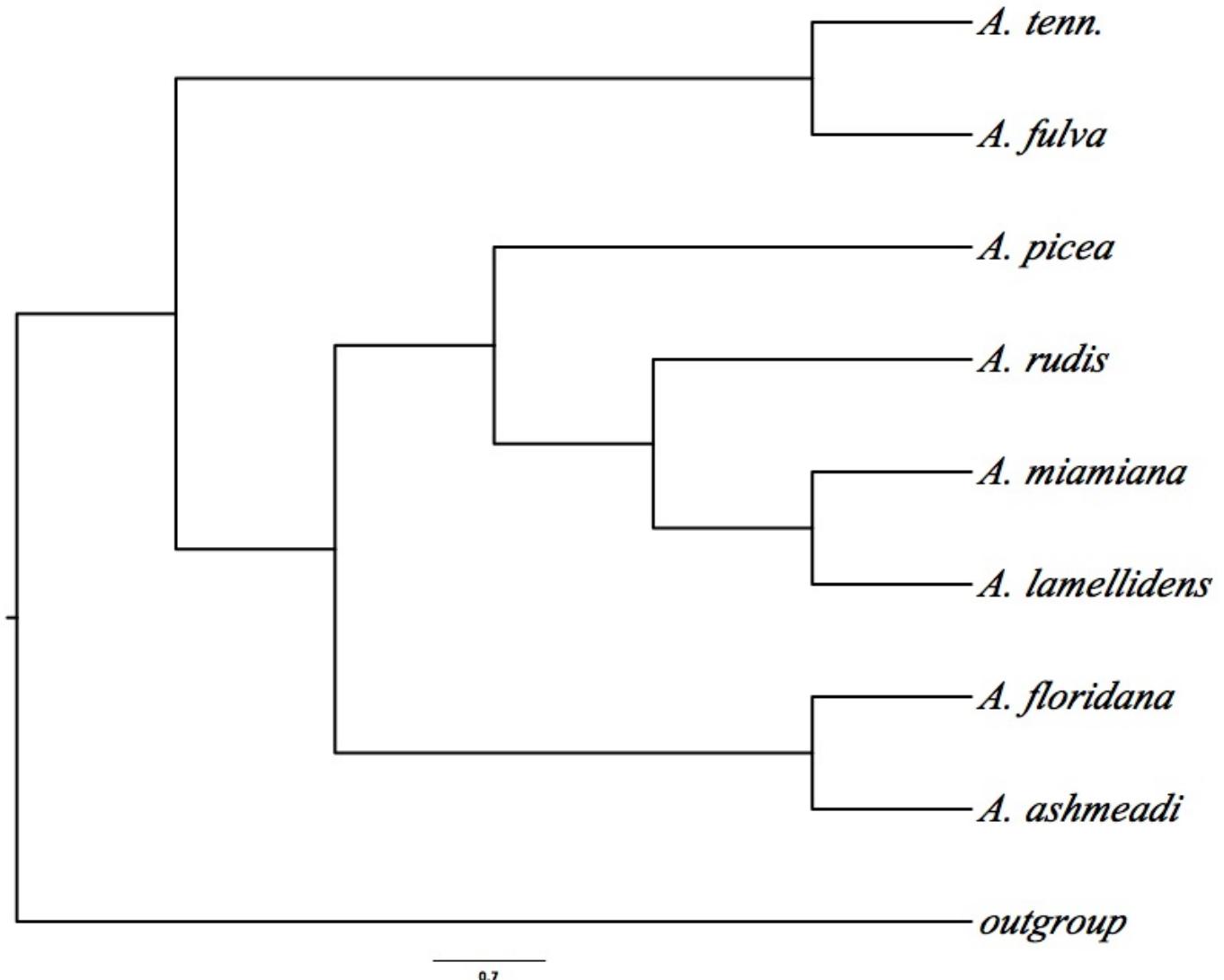
Untransformed branch lengths





Notes: I left a pogo sample in there. LPR4 and HW5 look switched.

###Summary of tree by species:



When comparing with the NJ tree, the placement of *A. picea* is different.

- ML tree: *A. picea* is sister to *A. rufa*, *A. miamiana*, *A. lamellidens*
- NJ tree: *A. picea* is sister to *A. rufa*, *A. miamiana*, *A. lamellidens*, *A. ashmeadi*, *A. floridana*

##Rerunning analysis without PB07-23 to double check this sample doesn't skew ingroup relationships.

<div id='id-section7'>

Page 7: 2016-05-17. ABI steponeplus machine maintenance.

Machine Problem: It freezes mid run without giving an error, even while operating stand alone. Sometimes when it freezes, the door wont release plate. And it also has trouble connecting to laptop even after restart.

Machine Info: [ABI steponeplus](#)

1. serial #: 272007769
2. ref: 4376592
3. University #: A92219

Under contract, no cost.

Contact info:

- Jeremy, 1-800-955-6288 option 3, then option 1
- issue#: 405638599

They need to send to Indonesia for repair. 1 month eta.

20160519 update: tracking number for box (for us to put machine in and send to them)- 6506 8693 8148

Also:

**Hi Andrew,*

You should receive a Loaner within 2-3 business days.

Thanks,

Foi Taua

Didn't know we were getting a loaner. He didn't mention cost.

20160520 update: Machine sent out

<div id='id-section8'>

Page 8: 2016-05-18.Phylogenetic results without pogo sample

The results of phylogenetic analysis of SNP matrix from [Page 3: 2016-05-16](#). Complete ddRAD-seq samples: processing. I excluded pogos, and it still needs further parsing.

1. Get rid of LPR4, BSK. LPR4 is not in the right place. Also there was a labeling problem with this sample. BSK, have no clue what this sample actually is. It also had a labeling problem. BSK does not match any sites, but had Kite on the side. It is in the right place, but still have no clue which kite colony.
2. Parse out bootstraps below 100
3. Need to relabel kite samples so that they're lower case.
4. Add in samples:

- HW6-rudis
- LPR4-ashmeadi
- 09A and 10A-rudis

###Getting rid of LPR4 and BSK5

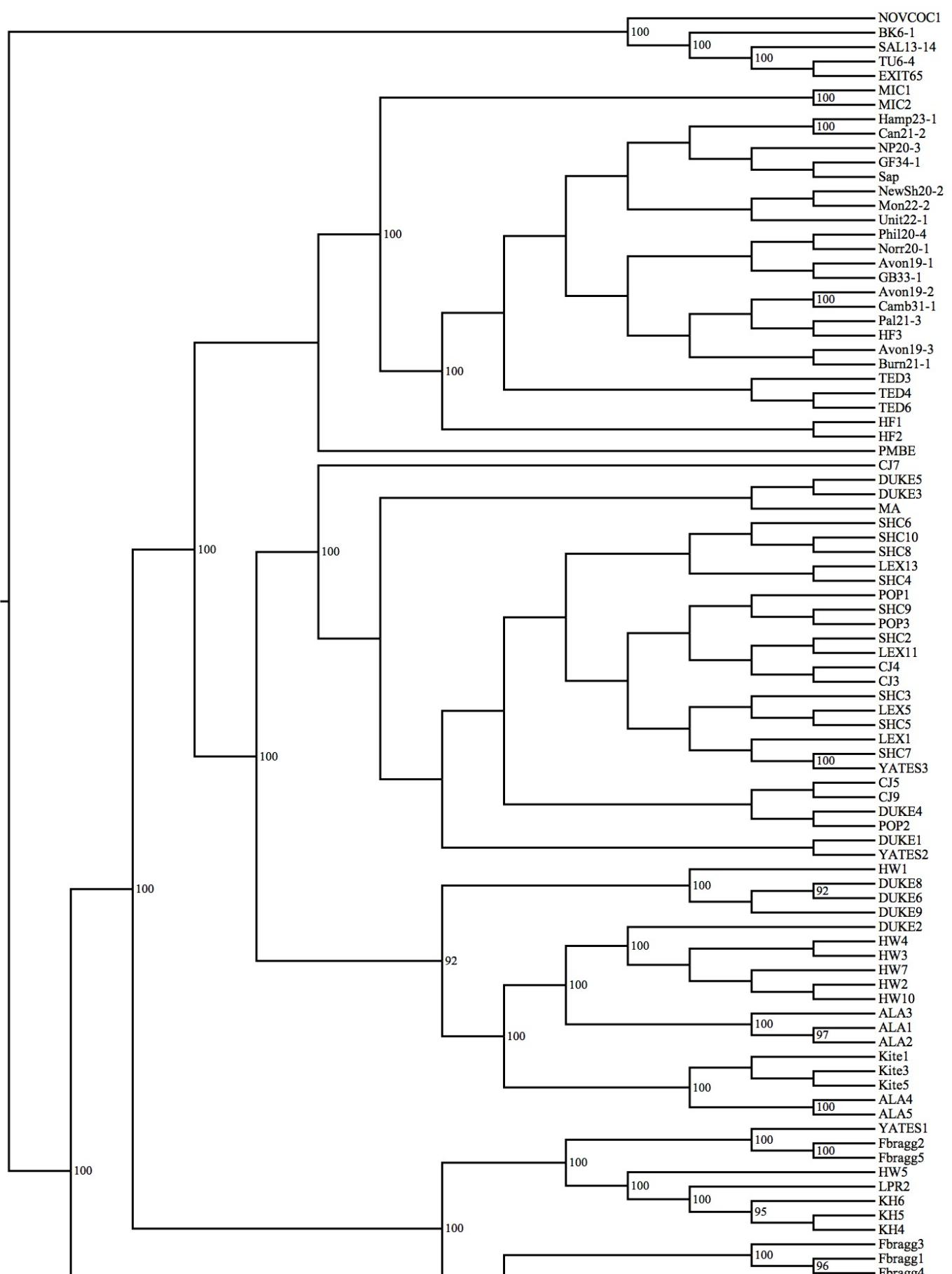
```
library(ape)
x<-read.tree("20160518_ML_tree_BL_BS_RAxML.newick")
plot(x)
length(x$tip.label)
x2<-drop.tip(x,c("LPR4","BSK5"))
length(x2$tip.label)# checking length
plot(x2) # plot to see
write.tree(x2,"20160518_ML_tree_BL_BS_RAxML_parsed.newick"
) # new file name
```

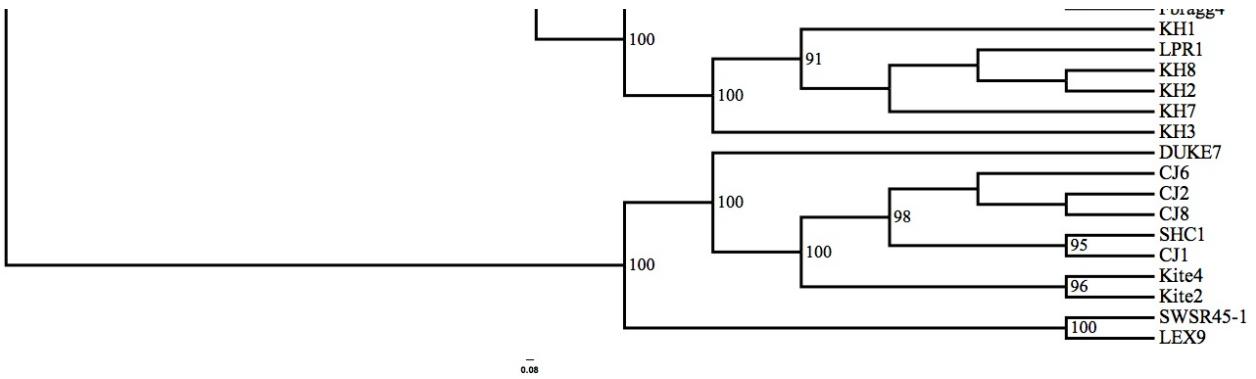
Parsing out bootstraps below 100

```
x2$node.label<-as.numeric(as.character(x2$node.label))
x2$node.label<-ifelse(x2$node.label>90,x2$node.label,"")
x2$node.label[1]<-""
x2$node.label
```

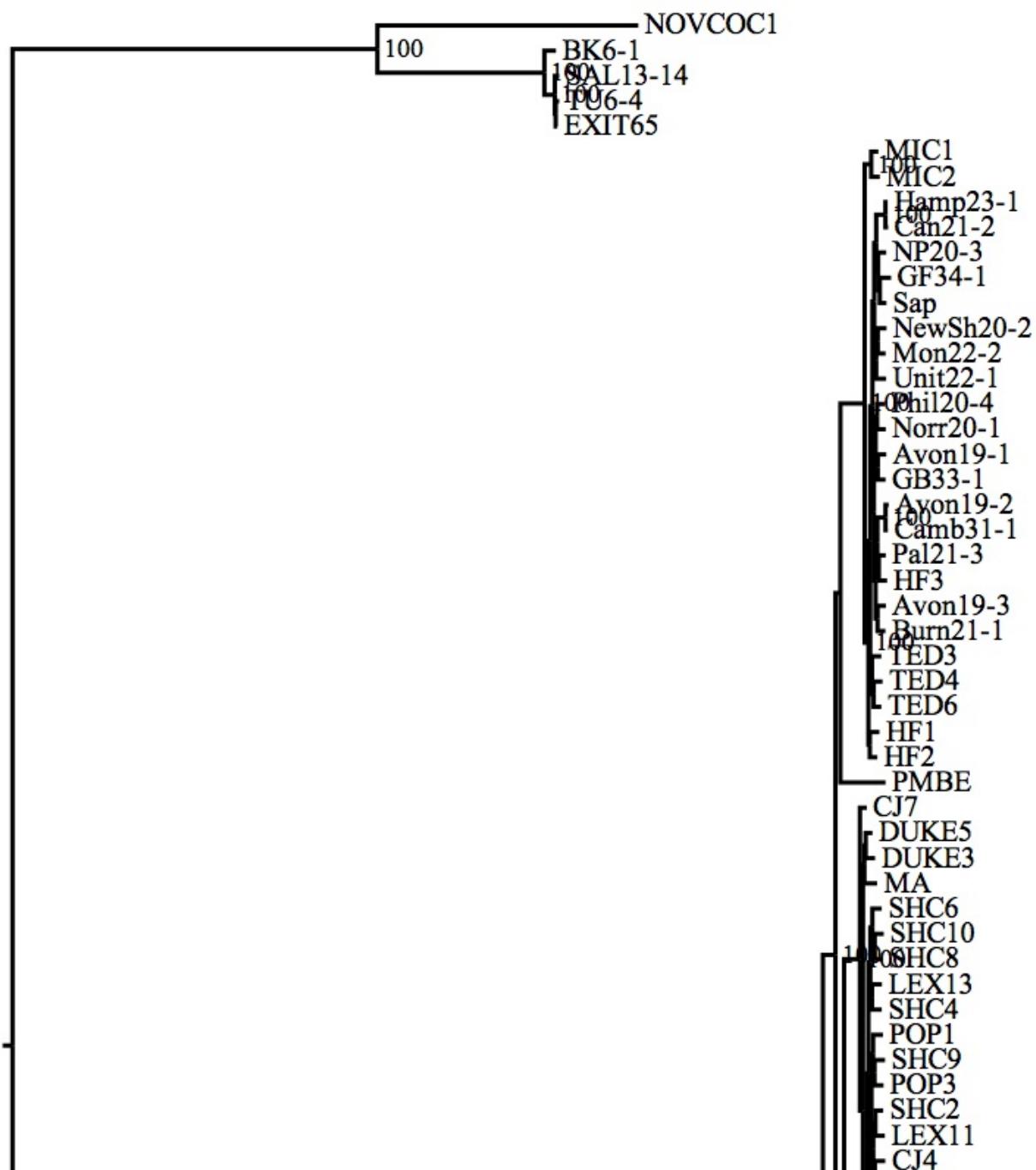
I'll hold off on adding samples to a phylogeny.

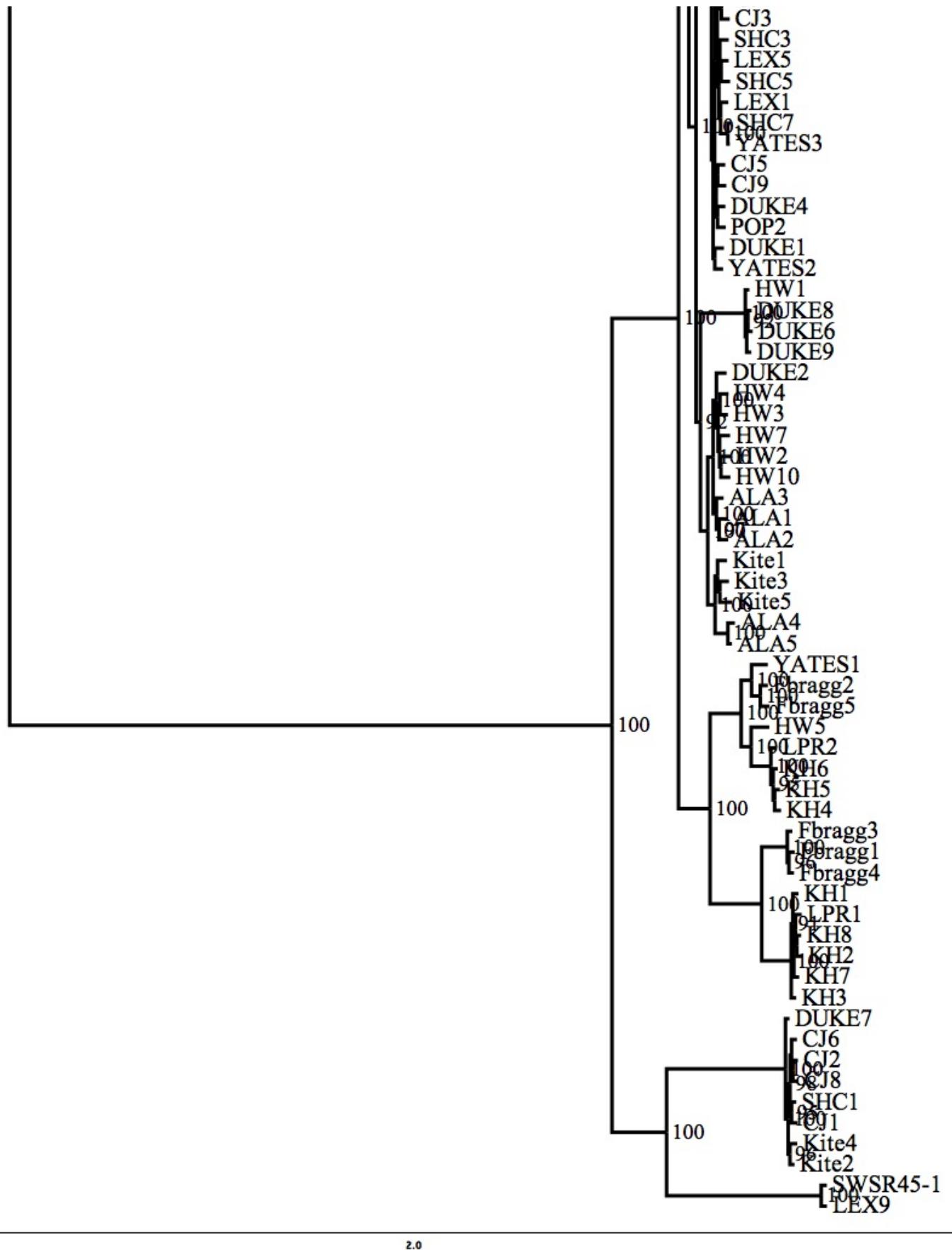
Transformed BL tree with 90 BS cutoff





Untransformed BL tree with 90 BS cutoff





2.0

Summary: Same topology without pogo sample.

<div id='id-section9'>

Page 9: 2016-05-18. Agarose gel electrophoresis of qPCR amplicons; Curtis and ANBE samples

We wanted to check for specificity on a gel. Although, agarose gels don't completely pick up primer dimers. Even so, we acquire fluorescence at a higher temperature where those primer dimers disappear.

Sample list

Lane	Section	Sample	Gene	Primer_pair
1	Top	Ladder		
2	Top	Exit65	hsc70-4 h2	1468+1592
3	Top	BK	hsc70-4 h2	1468+1592
4	Top	Ted6	hsc70-4 h2	1468+1592
5	Top	DUKE6	hsc70-4 h2	1468+1592
6	Top	ALA1	hsc70-4 h2	1468+1592
7	Top	KH2	hsc70-4 h2	1468+1592
8	Top	FB2	hsc70-4 h2	1468+1592
9	Top	Exit65	hsp83	1592+1682

10	Top	BK	hsp83	1592+1682
11	Top	TED3	hsp83	1592+1682
12	Top	DUKE6	hsp83	1592+1682
13	Top	ALA1	hsp83	1592+1682
14	Top	KH2	hsp83	1592+1682
15	Top	FB2	hsp83	1592+1682
16	Top	PB1710	hsp83	279
17	Top	POP2	hsp83	279
18	Top	SHC2	hsp83	279
19	Top	cremato	hsp83	279
20	Top	Ladder		
1	Bottom	Ladder		
2	Bottom	ex	hsp83	279
3	Bottom	bk	hsp83	279
4	Bottom	TED6	hsp83	279
5	Bottom	DUKE6	hsp40	541+641
6	Bottom	ALA1	hsp40	541+641
7	Bottom	KH2	hsp40	541+641
8	Bottom	FB2	hsp40	541+641
9	Bottom	EX	hsp40	541+641
10	Bottom	BK	hsp40	541+641

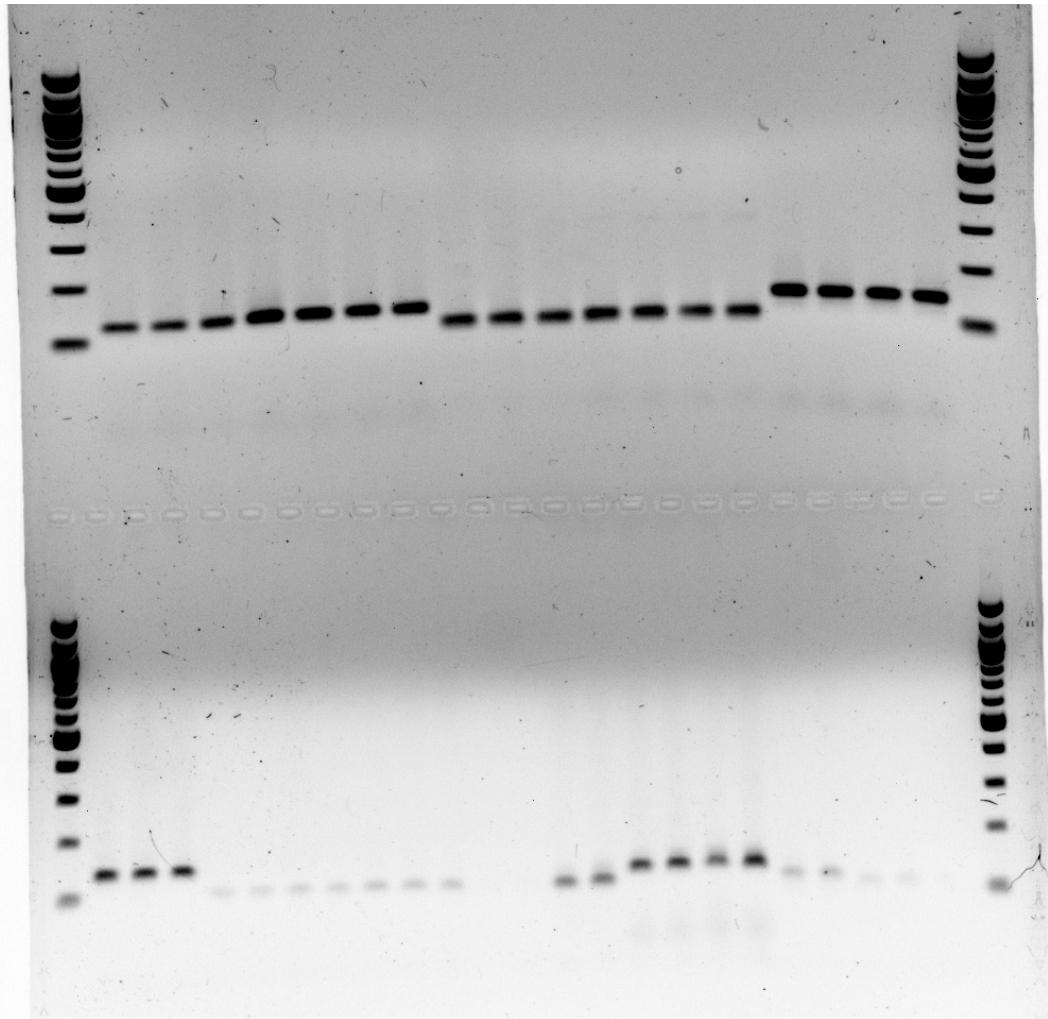
11	Bottom	Ted6	hsp40	541+641
12	Bottom	HF	hsp83	1592+1682
13	Bottom	HF	hsp83	1592+1682
14	Bottom	DF	hsp83	1592+1682
15	Bottom	DF	hsp83	1592+1682
16	Bottom	HF	hsc70-4 h2	1468+1592
17	Bottom	HF	hsc70-4 h2	1468+1592
18	Bottom	DF	hsc70-4 h2	1468+1592
19	Bottom	DF	hsc70-4 h2	1468+1592
20	Bottom	DF	actin	
21	Bottom	DF	actin	
22	Bottom	HF	hsp40	541+641
23	Bottom	HF	hsp40	541+641
24	Bottom	DF	hsp40	541+641
25	Bottom	Ladder		

Protocol:

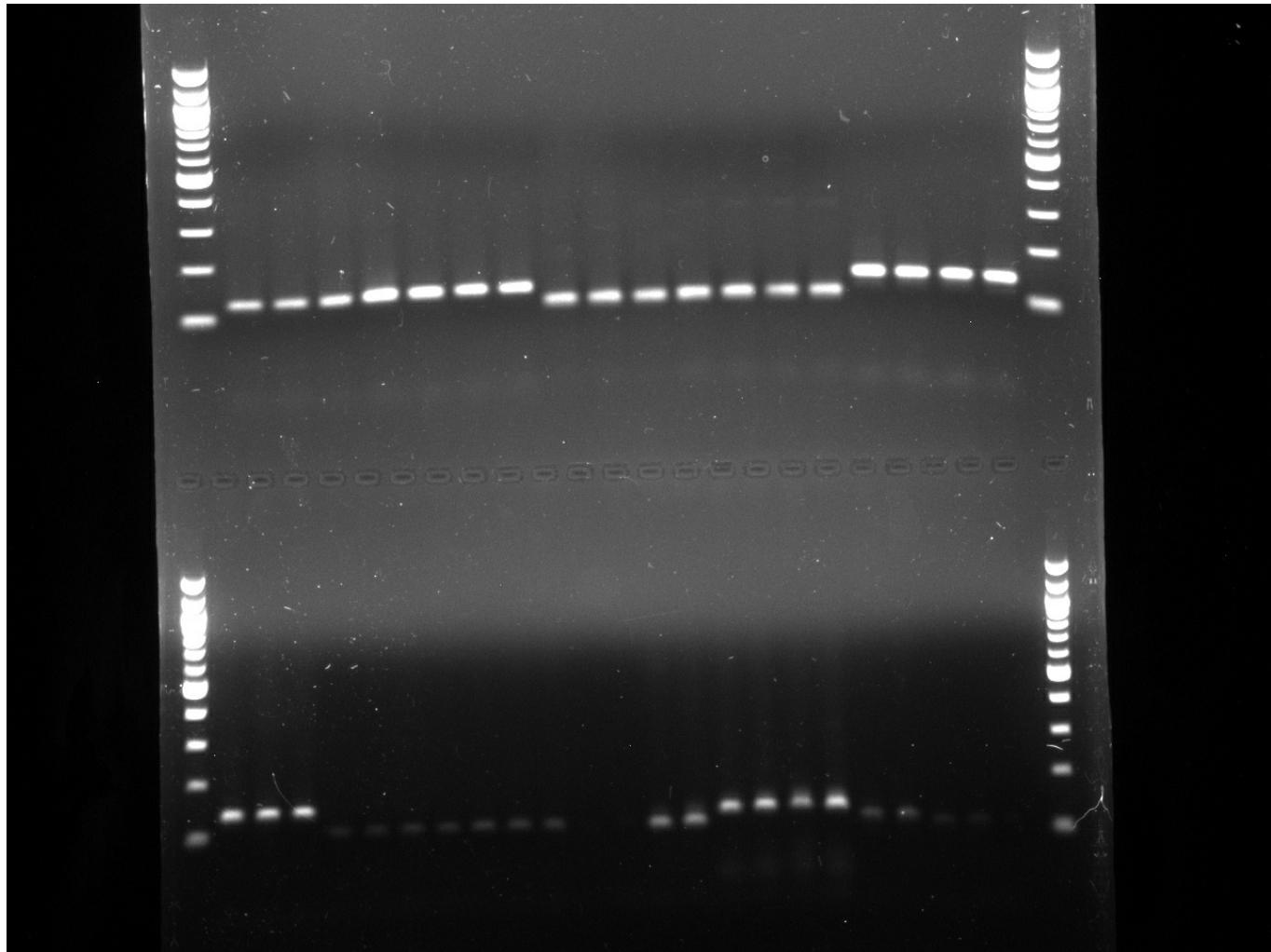
1. Mixed ladder: 6.5 dye (6x) + 8 uL 100bp ladder+ 25.5 ul h20 to make 40 uL total—makes 4 lanes worth at 10 uL each lane
2. For ANBE add 10 uL qpcr amplicon with 2 uL 6 x dye.
3. For Curtis, add 5 uL qpcr amplicon with 1 uL 6 x dye.
4. Electrophoresed on 1.5 % agarose gel , 125 Volts for 45

minutes.

###Grayscaled whole:



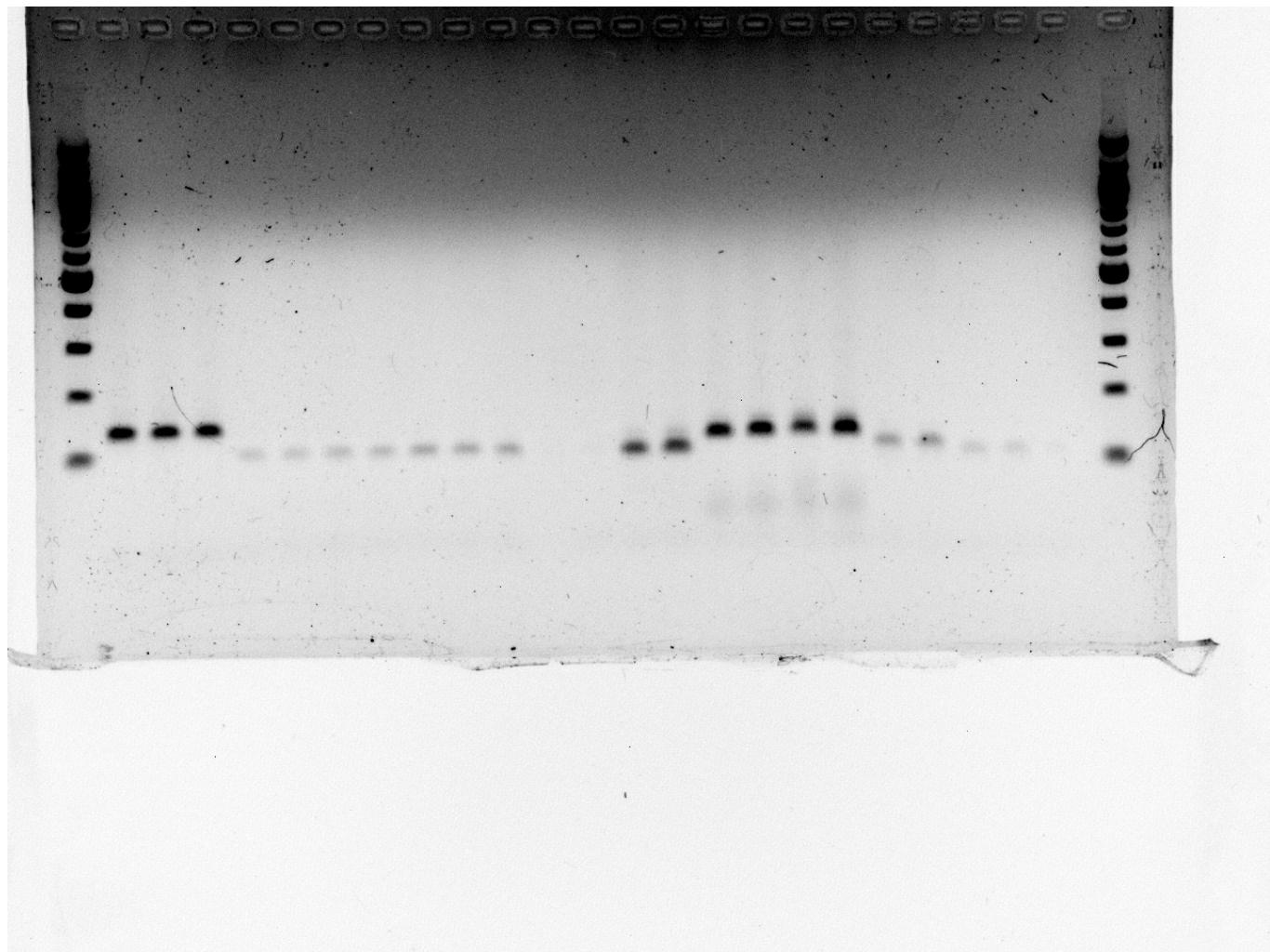
Black whole:



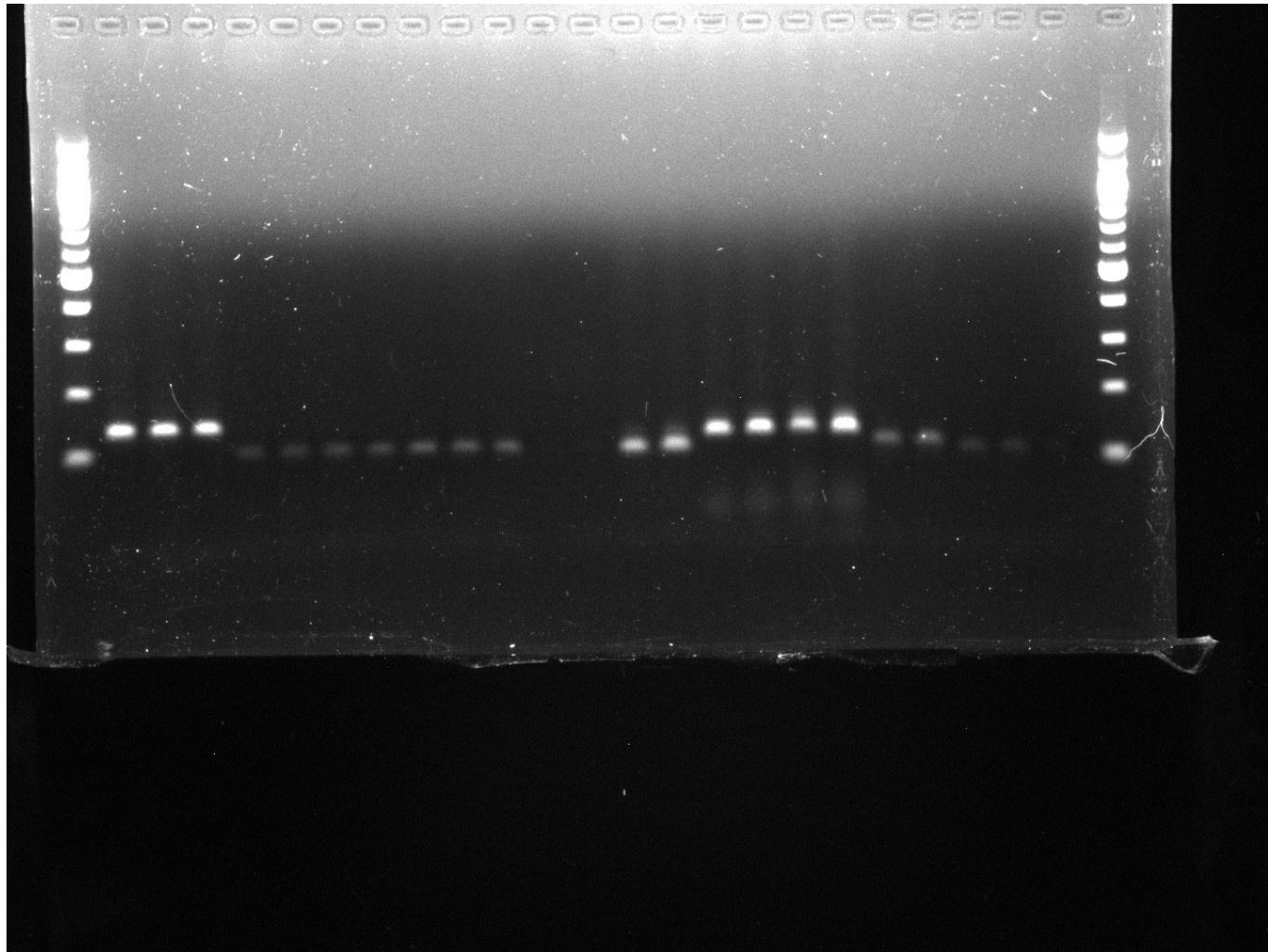
##The bottom is hard to see:

Showing pictures that focus on bottom part

###Grayscaled bottom:



Black bottom:



###Summary: Amplicons are specific. NO double bands.

<div id='id-section10'>

Page 10: 2016-05-18. RaxML ML pairwise distance matrix

- Used SNP matrix found here: [Page 3: 2016-05-16](#)
- Used input tree(pre-parsed) found here: [Page 8: 2016-05-18](#)

Code for RaxML

```

./raxmlHPC -f x -p 12345 -s
~/Desktop/2015_ANBE_common_garden/20160516_Andrew_SNP_sequ
-m GTRGAMMA -t
~/Desktop/2015_ANBE_common_garden/20160518_ML_tree_unparsed.
-n 20150618_ML_pairwise_distance_ANBEsamples

```

Snippet of output:

V1	V2	V3
HW5	ALA1	0.094440
HW5	BK6-1	0.512869
HW5	POP3	0.096510
HW5	MA	0.092071
HW5	CJ1	0.277364
HW5	Camb31-1	0.096856
HW5	DUKE9	0.113134
HW5	ALA2	0.098850
HW5	KH4	0.032412
HW5	Unit22-1	0.097533

<div id='id-section11'>

machine maintenance update

Update from, [Page 7: 2016-05-17](#). ABI steponeplus machine maintenance.

*Hi Andrew,

Thank you for your recent request to have your StepOne Plus serial number 272007769 sent in to our Global Repair Center. Attached you will find the necessary paperwork to ensure that your unit is received correctly and promptly.

1. Your RMA is 405638599

2. Please review and complete the attached decontamination form, and print out 2 copies.

For 9700/9800's, Please put both the TOP and BASE serial numbers on the decontamination certificate.

3. Please DO NOT include your power cord with your instrument (remove from unit and keep it).

4. Please DO NOT include any consumables (trays, tubes, etc.).

5. Place a copy of the completed decontamination form INSIDE and OUTSIDE of the box.

6. Print out the FedEx label, (link will arrive via separate email).

Service of your instrument cannot begin without the completed decontamination form.

Best Regards,

Foi Taua

Remote Service Center

T 800 955 6288 option 3, 1 to reach Remote Service Center

F 760 930 2300

5791 Van Allen Way • Carlsbad • CA • 92008 • United States

instrumentservices@lifetech.com

[www.lifetechnologies.com*](http://www.lifetechnologies.com)



2016-05-26 update: we received loaner.

<div id='id-section12'>

Page 12: 2016-05-19. Getting whole rad loci with [pyRAD tutorial](#) and/or stacks

Previous analyses concatenate SNPs, but many studies use whole rad loci.

Computer cluster:

Reference for [mason cluster](#)

path of raw ddrad data

/N/dc2/scratch/scahan/Andrew_RADseq_051516/Data/

SHC email:

If you want to explore/analyze the RADseq data yourself:

/N/dc2/scratch/scahan/Aphaenogaster_RADfiles_051516/

You should find in each lane directory the raw .fq file from the sequencer, a barcode key file, the demultiplexed sample .fq files, and the trimming, filtering and mapping files from the pipeline.

The earliest lanes (1&2) might have fewer files because the process was not yet regularized back then. The STACKs portion of the pipeline is specific to each project, so they all have their own directories in the main scratch space (e.g.,

Andrew_RADseq_051516, Bernice_051516,

Phytotron_analyses_051516, etc.). All directories at this level have their date suffix modified every two weeks, so job scripts that point to a particular path have to get edited to the current date suffix. Some of the ddRAD lane directories also have a date suffix because they were secondarily moved from the main level into the Aphaenogaster directory.

Trying pyRAD tutorial. Looks “easy”.

No access to dependencies:

1. scipy
2. vsearch
3. muscle

20160520 update, working on Mason compute cluster:

Hi Andrew,

*First, I'd suggest you add "module load python" to your
~/.modules file, which*

*will load the python 2.7.3 module each time you login. It's not
terribly*

*current, but it is the version under which we install python
packages on Mason.*

You'll find that numpy and scipy are both available there.

*As for muscle and vsearch, I'll let you know when we get those
packages*

installed.

Matt

I could use the population function/module in stacks.

<div id='id-section13'>

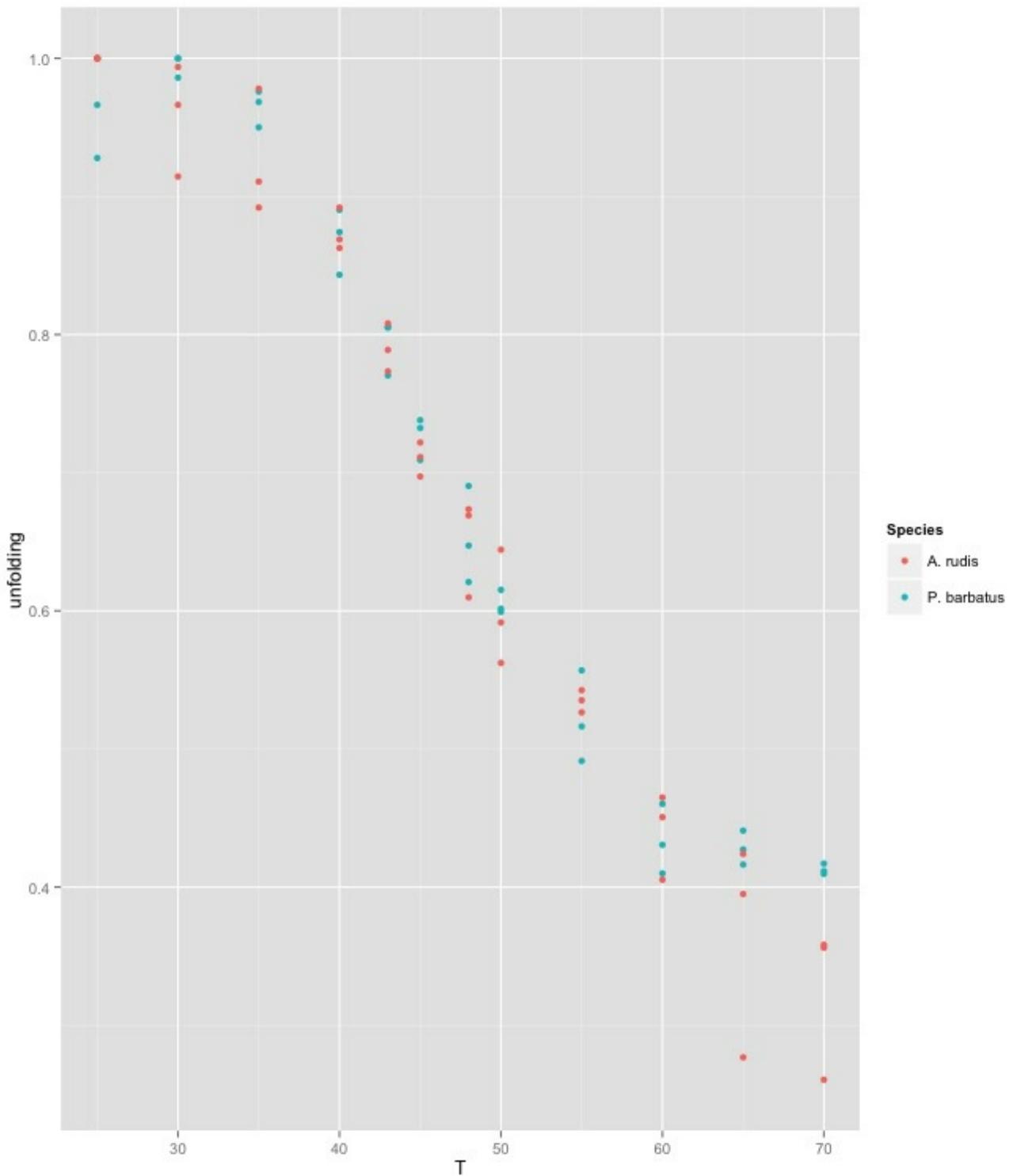
Page 13: 2016-05-20. Evolution of proteome stability project

We are interested in the adaptive variation in how proteins unfold between 2 different ant species. [Github repo](#)

We isolated native proteins, subjected them to temperature treatments for 10 min. Then ultracentrifuged to pull down aggregates, then

quantified. [protocols here](#)

###Figure:



###Function I am fitting to these points:

$$\text{min} + \frac{1-\text{min}}{(1+e^{(-\text{slope}(\text{Tm}-\text{Temp}))})}$$

$$min + \frac{1-min}{(1+e^{(-slope(Tm-Temp))})}$$

Code for curve fitting, also loading libraries

```

library(plyr)
library(ggplot2)
library(tidyr)
library(minpack.lm)

nls.fit<-function(data=data){
  y<-nlsLM(unfolding ~ min+ (1-min)/(1+exp((-slope*(Tm-T)
))),data=data,
            start=list(slope=.5,Tm=45,min=.3),
            trace=TRUE,control=nls.control(warnOnly = TRUE
, tol = 1e-05, maxiter=1000))
  #return(y)
  return(summary(y)$coefficients)
}

```

function to visualize curves by simply putting in parameters

```

fud<-function(T=seq(25,50,1),Tm=40,slope=.5,max=1,min=0){
  y<-min+ (max-min)/(1+exp((-slope*(Tm-T))))
}

```

```
    return(y)
```

```
}
```

How I implemented th code:

```
mod1<-ddply(x.par,.(Species,Colony),nls.fit)
mod1$parameter<- rep(c("slope","Tm","min"),length(mod1$Species)/3)
knitr:::kable(mod1)
```

Table summary of results from fitting curves.

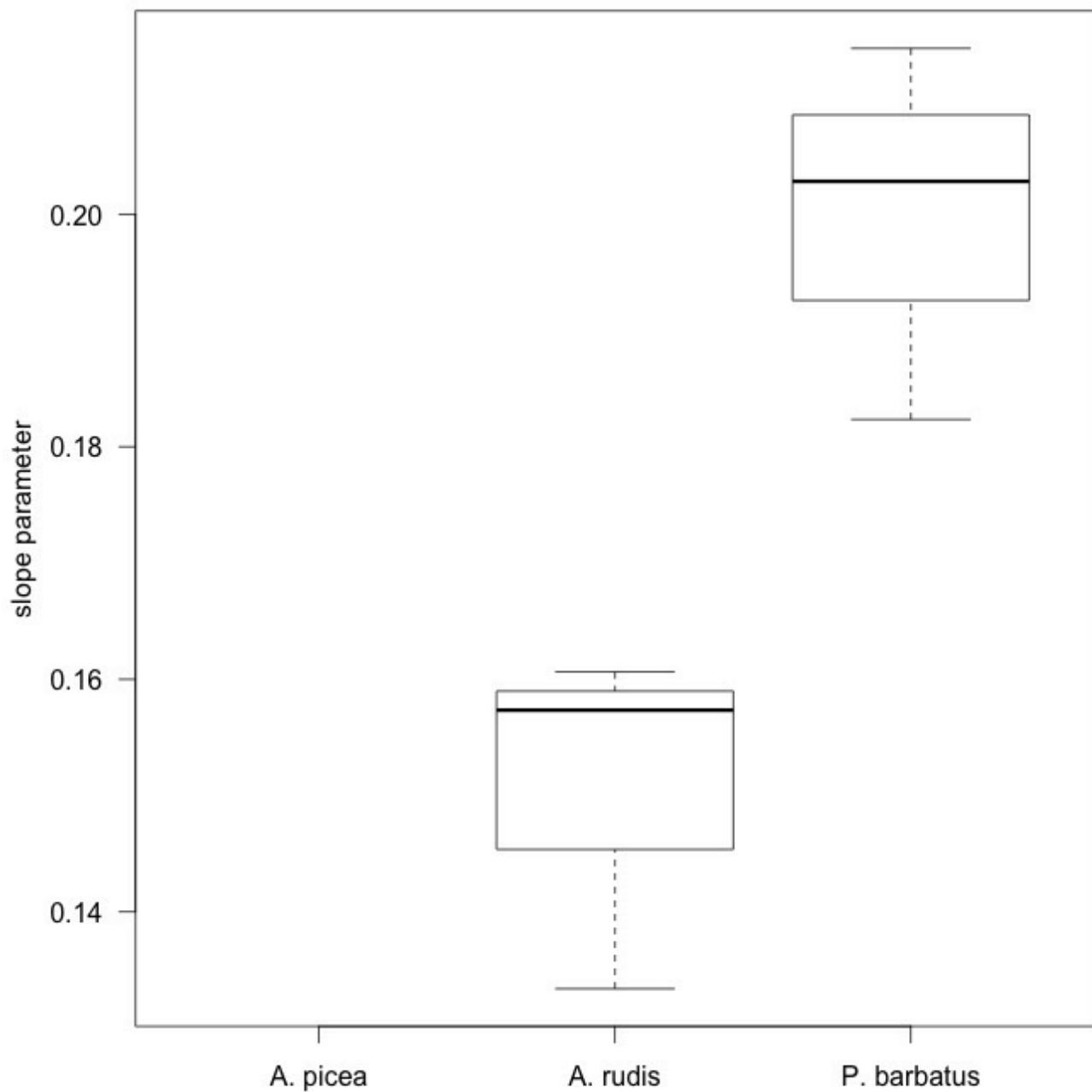
Species	Colony	Estimate	Std. Error	t value	Pr(> t
A. rufus	Duke 1	0.1606280	0.0206403	7.782238	0.000027
A. rufus	Duke 1	47.2920297	0.9451544	50.036301	0.000000
A. rufus	Duke 1	0.3637620	0.0293990	12.373285	0.000000
A. rufus	Lex 13	0.1333902	0.0159832	8.345673	0.000015
A. rufus	Lex 13	49.7593929	1.2760137	38.995972	0.000000
A. rufus	Lex 13	0.2161279	0.0451703	4.784737	0.000994
A. rufus	Yates 2	0.1573466	0.0220329	7.141430	0.000054
A. rufus	Yates 2	47.9849648	1.0899761	44.023870	0.000000
A. rufus	Yates 2	0.3637813	0.0336777	10.801853	0.000001
P. barbatus	WWRQ-45	0.2142567	0.0165774	12.924625	0.000000

P. barbatus	WWRQ- 45	45.9987927	0.3837543	119.865208	0.000000
P. barbatus	WWRQ- 45	0.4032438	0.0126671	31.834069	0.000000
P. barbatus	WWRQ- 53	0.1823480	0.0173963	10.482009	0.000002
P. barbatus	WWRQ- 53	47.2858982	0.5958843	79.354167	0.000000
P. barbatus	WWRQ- 53	0.4013122	0.0184886	21.705927	0.000000
P. barbatus	WWRQ- 8	0.2028211	0.0245990	8.245113	0.000017
P. barbatus	WWRQ- 8	45.5664742	0.6340253	71.868543	0.000000
P. barbatus	WWRQ- 8	0.4280916	0.0194756	21.980921	0.000000

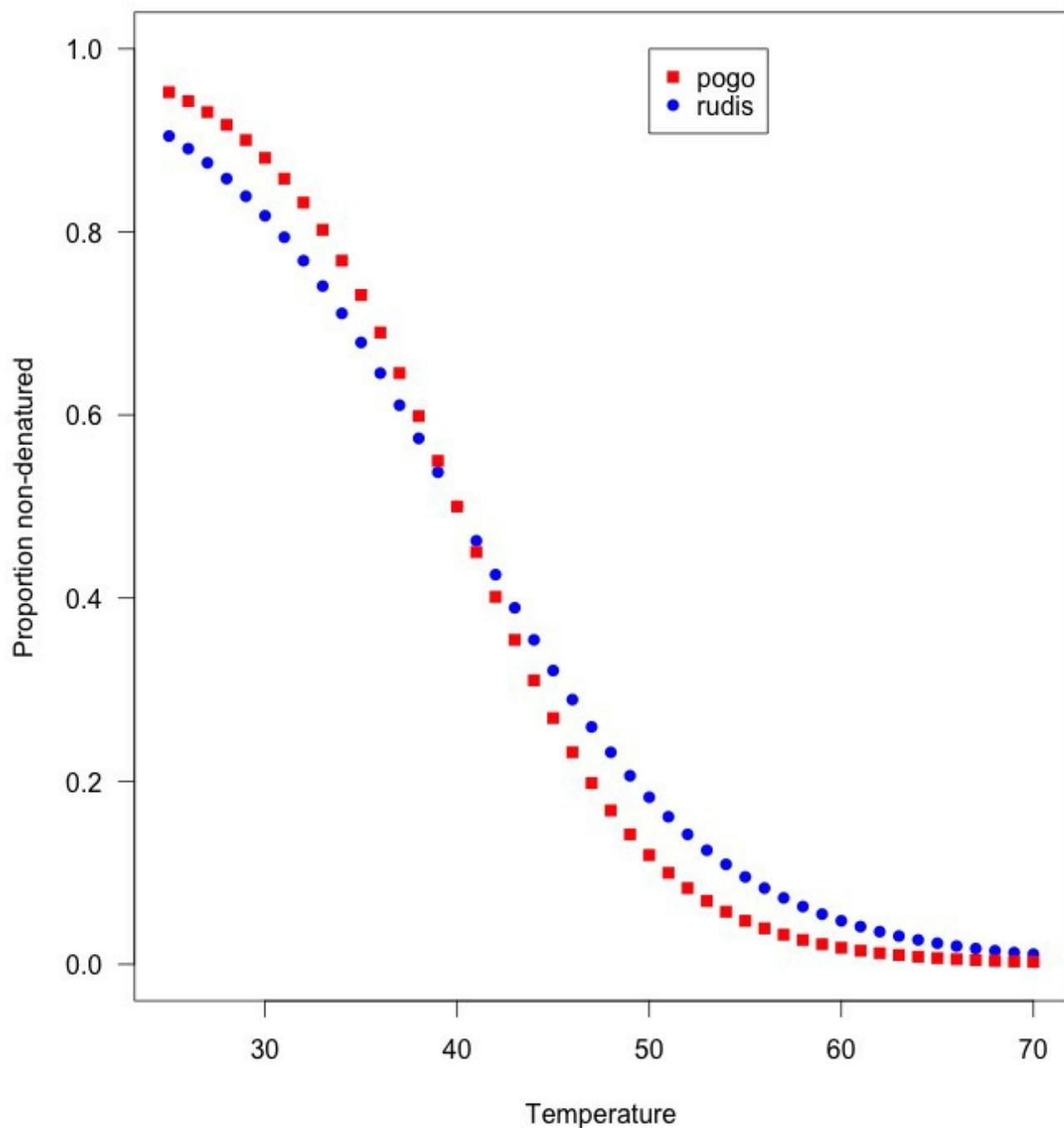
Only slope was significant

```
summary(aov(Estimate~Species,data=slope))

      Df  Sum Sq Mean Sq F value Pr(>F)
Species   1 0.003654 0.003654   15.15 0.0177 *
Residuals 4 0.000965 0.000241
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
   ' ' 1
```



**Figure of unfolding if only changing
slope (eye balled mean slope, so pogo
= .2,rufis=.15)**



<div id='id-section14'/>

Page 14: 2016-05-24. Evolution of proteome stability project:

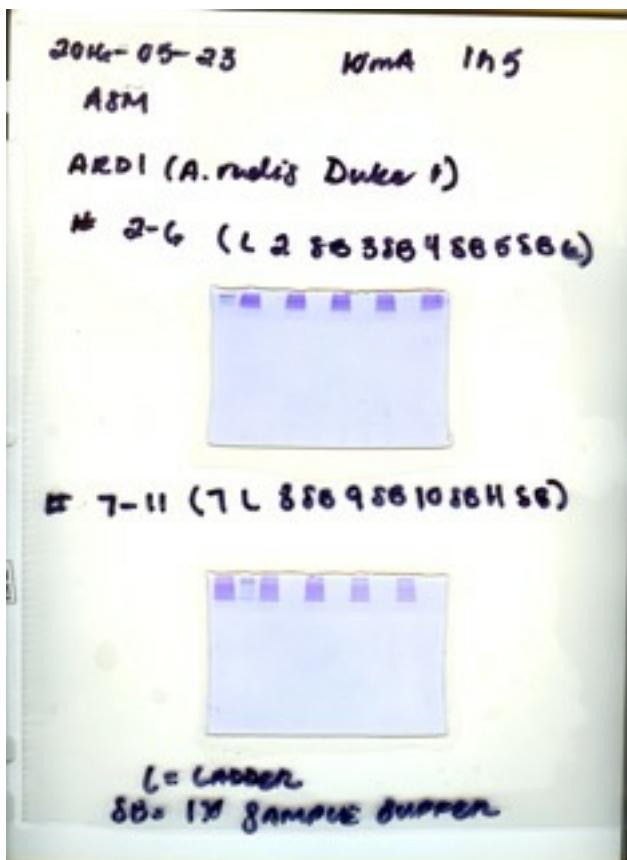
Polyacrylamide gels for colony level replicates (*A. rufis* vs *P. barbatus*)

Amanda Meyer is working on this project:

1. Samples stored at -80C, we took 30 uL and speed vacuumed (took 1 hr) and then resuspended in 60 uL of 1X sample buffer
2. We took 25 uL of sample and added in 5 ul of gapdh (20ng/uL in sample buffer)
3. Loaded on polyacryladmide and electrophoresed.

Polyacrylamide Gels:

1. Duke1 (*A. rufis*)



2. Yates2 (*A. rufis*)

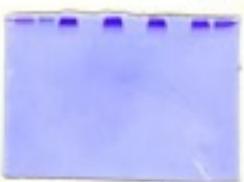
2016-05-23 10mA 0h45
A8M

ARY2 (*A. mollis* Yates 2)

2-6 (L 2 883884885,6)



7-11 (L 2 88388488108811)



L = LADDER
S0 = 1% SAMPLE BUFFER

Note: For Yates2, the gels are reversed. (Bottom gel starts at 30C)

3. WWRQ45 (*P. barbatus*)

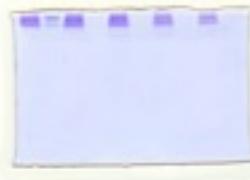
2016-05-23 10mA 1h5
A8M

WWRQ-45 Pogo

2-6 (L 2 883884885,6)

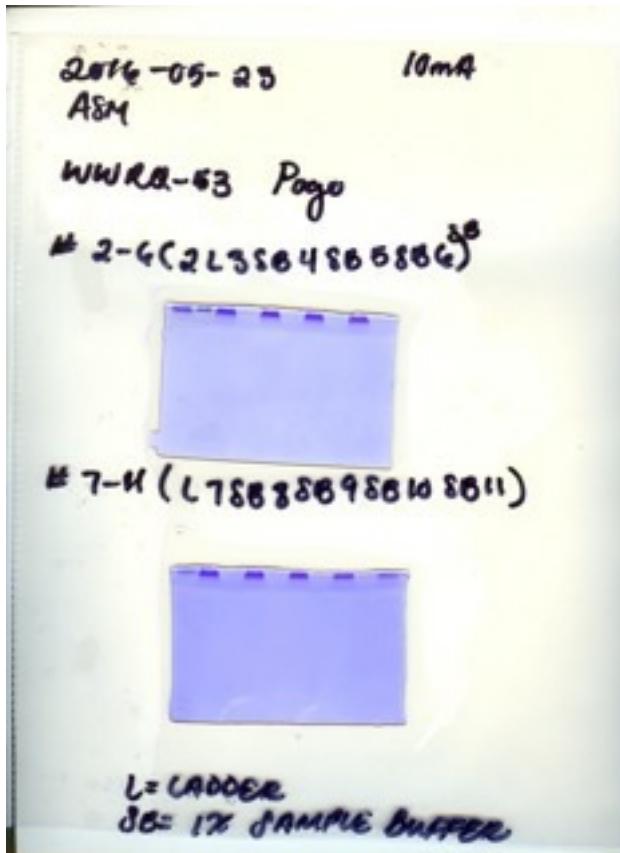


7-11 (L 2 88388488108811)



L = LADDER
S0 = 1% SAMPLE BUFFER

4. WWRQ53 (*P. barbatus*)



Next steps:

Need to destain and trypsin digest.

<div id='id-section15'>

Page 15: 2016-05-24. Degenerate Hsp primer design from 2015-05-28

Primer design from 2015-05-28 referenced here

n	name	sequence
1	hsc70-4h2_1175F	TTCTGYTGGAYGTDACTCC

2	hsc70-h2_1345R	TCGCTCTCTCHCCYTCRTARAC
3	hsc70-4h2_1468F	GCGATYGARAAATCTACVGGC
4	hsc7004h2_1592R	TGYTCRTCYTCCGATCGGT
5	hsc70-4h1_1291F	ACYTAYGCCGACAATCARCC
6	hsc70-4h2_1390R	CGCTTRAGCTCGAAYTTDCCC
7	hsc70-4h1_1506F	CACYATYACCAAYGACAARG
8	hsc70-4h1_1605R	YTCCTTCTGCTTCTCRTCCCTC
9	hsp40_118F	GCCTTRCGATATCATCCTGA
10	hsp40_248	CCYTCCTCGCCRAATTATC
11	hsp40_541F	AAAGATCGYGCYCARGATCC
12	hsp40_641R	GCYCGTCTRCATATYTTCATC
13	hsp40_869F	TRTGCGGTACTRTYGTCAAG
14	hsp40_968R	TGGAACCTYTTGACNGTRTC
15	hsp83_278F	ACDATYCTTGATTCTGGYATTGG
16	hsp83_392R	CCAAACTGTCCAATCATGGA
17	hsp83754F	GATGTYGGHGAGGATGA
18	hsp83_880R	GATTCTYGTCCARATCGG
19	hsp83_1583F	AATTCGAYGGAAARCAGYTGG
20	hsp83_1682R	AAYTTGGCYTTGTCYTCCCTC
21	hsp83_1807F	ATGGAGAGRATCATGAAGGC
22	hsp83_1917R	CARRTTCTCCATGATRGGATGATC

23	nedd_510F	TAATCATTCCAGTCAGCGG
24	ned_614R	TCAGATACGTCTCCGTTGTC
25	nedd_556F	TATCATGCATAACATTCCGAC
26	nedd_683R	ATCGTAATATCTGCACTTGYTC
27	nedd_956F	ATGGTGAAGTTCTACGCGAG
28	nedd_1088R	TAAGGTAGCCACGTTGATCG
29	nedd_1222F	CAAGTAGCACCTAATGGTAGA
30	nedd_1316R	GGTATAGARCTTGGTCTTCC
31	nedd_1351F	GATTTAGATCAATTAGGACCDCTTC
32	nedd_1460R	GGATCTTCCCATTGTGTTGT
33	nedd_2375F	GGAGAGTCGTTTGTCAATTCAAG
34	nedd_2459R	CCATTCAATTGGAACACGTGATG

I don't use all of these anymore. But here are the ones that I've tested for specificity (from agarose gel electrophoresis , sequencing, and melt curve analysis following qPCR) and efficiency (titrate amplicon across a dynamic range to compare slope equals -3.2).

1. hsc70-4 h2; 1468F + 1592R
 2. hsp83; 278F+392R and 1583F + 1682 R
 3. hsp40 541F+ 641R
 4. NEDD; 956F+ 1088R (This is off the top of my head, so I need to double check this!)
-

<div id='id-section16'>

Page 16: 2016-05-24](#id-section16).

Sequencing analysis continued from

Page 5: 2016-05-16.

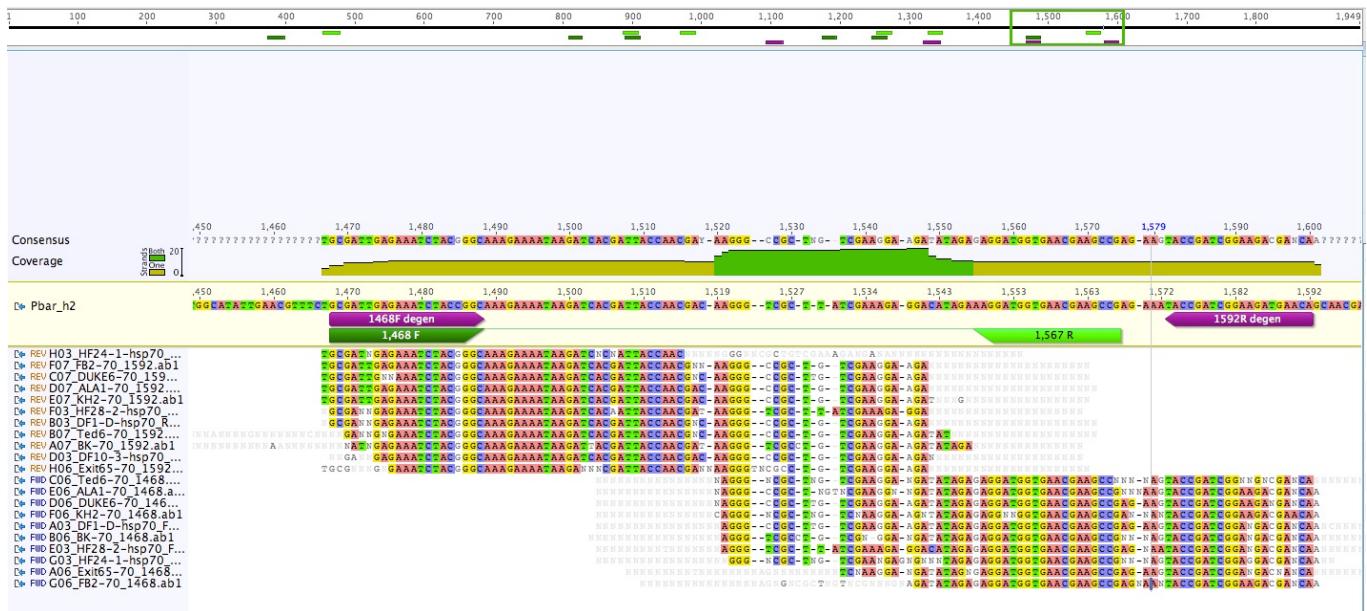
Sharing screenshots of sanger sequenced samples mapped to reference transcript (*P. barbatus*)

- I used the software [Geneious v6](#) to analyze sequence data.
- Sample structure on figure: Well_colony.id_gene_primer#
- The pics and raw sequence data can be found: [here](#)

◦ Path:

/Dissertation_temperature_adaptation_ants/Dissertation_P:
common-
garden_gxp_evolution/Data/sequencing/Sanger/

1. hsc70-4 h2 1468F + 1592R



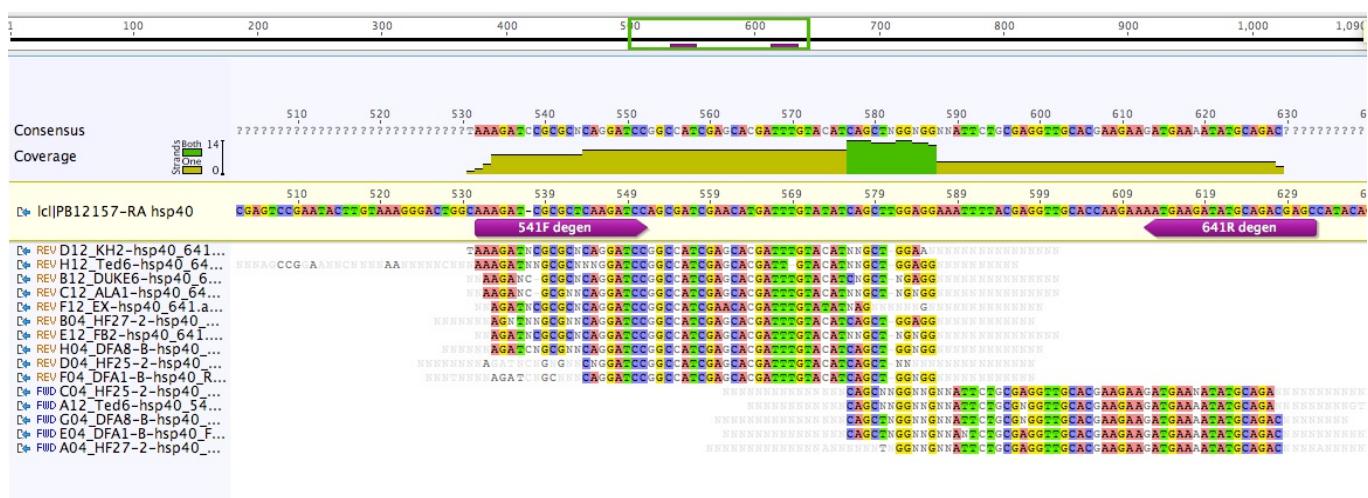
2. hsp83 278F+392R



3. hsp83 1583F + 1682 R



4. hsp40 541F+ 641R



Summary of results:

Most of samples mapped really well! Generally, the sequencing with the forward primer recovers the reverse primer, and vice versa.

```
<div id='id-section17'>
```

```
###Page 17: 2016-05-25. Double check samples for SHC; JSG  
phytotron exp and MS.
```

```
###email sent 2016-05-18:
```

Ok - your list is missing 20-B (AP2), which is on your tree. The two samples with no morphological ID are your RW2 (25-C) and your BP2 (07-B), which will have to get omitted. The only remaining samples whose placements are problematic are your RW1, which was ID'd picea but comes out in that odd basal clade with the intermediate NK samples, and LA4, which was ID'd as rufa but falls out in the middle of picea. Looking at the latter one, however, this is the mysterious LVA9, which was written down as the source for two different experimental colonies (not possible, since they were supposed to be queenright) and there is no way to know if the sample Bernice looked at is the same as the RADseq sample or the assayed colony. So there is good reason to throw that one out as well.

Need to double check these samples.

2016-05-26 UPdate- Excluding:

1. 25-C / RW2
 2. 07-B / BP2
 3. 10-F / LA4
-

<div id='id-section18'>

Page 18: 2016-05-31. Learning model selection and model averaging!

###I'm learning model averaging!

Basically, there is uncertainty in parameter estimates of a stat model (linear regression) and we should explore how many stats model compare to each other, usually by AIC.

From [Burnham and Anderson 2002](#)

If data analysis relies on model selection, then inferences should acknowledge model selection uncertainty. If the goal is to get the best estimates of a set of parameters in common to all models (this includes prediction), model averaging is recommended. If the models have definite, and differing, interpretations as regards understanding relationships among variables, and it is such understanding that is sought, then one wants to identify the best model and make inferences based on that model. Hence, reported parameter estimates should then be from the selected model (not model averaged values).

However, even when selecting a best model, also note the competing

models, as ranked by their Akaike weights. Restricting detailed comparisons to the models in a 90% confidence set on models should often suffice. If a single model is not strongly supported, $w_{min} \geq 0.9$, and competing models give alternative inferences, this should be reported. It may occur that the basic inference(s) will be the same from all good models. However, this is not always the case, and then inference based on a single best model may not be sound if support for even the best model is weak (in all-subsets selection when $R > 1,000$, w_{min} can be very small, e.g., < 0.01).

General Steps:

1. Construct global model. Pick predictors you think are most important.
2. I used MuMin package in R with the dredge() function to construct subsets of global model.
3. Pick out top model set: subset based on... top 2/6/10 AIC or delta 4 AIC.
4. Average models from top set.

Picking predictors I think are important

Decomposing phylogeny with PCOA, looking at eigenvalues:

Eigenvalues	Relative_eig	Rel_corr_eig	Broken_stick	Cum_cc
0.362	0.563	0.407	0.114	

0.086	0.134	0.102	0.087
0.052	0.081	0.065	0.073
0.020	0.032	0.030	0.064
0.016	0.025	0.025	0.057
0.014	0.022	0.023	0.052
0.011	0.017	0.020	0.047
0.010	0.015	0.018	0.043
0.008	0.013	0.017	0.040
0.007	0.011	0.016	0.037
0.005	0.008	0.013	0.034
0.005	0.008	0.013	0.032
0.004	0.007	0.013	0.030
0.004	0.007	0.012	0.028
0.004	0.006	0.012	0.026
0.004	0.006	0.012	0.024
0.003	0.005	0.011	0.022
0.003	0.005	0.011	0.021
0.003	0.005	0.011	0.019
0.003	0.005	0.011	0.018

We have ~40 samples, so use 10:1 rule (sample: predictor). Regress

first 4 Axes (60% of variation) against Ctmax.

```
Ctmax.sel<-lm(Ctmax~Axis.1+Axis.2+Axis.3+Axis.4,data=merg  
)
```

```
summary(Ctmax.sel)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	42.43692	0.06339	669.409	<2e-16 ***
Axis.1	11.87677	0.65817	18.045	<2e-16 ***
Axis.2	2.87094	1.35038	2.126	0.0408 *
Axis.3	3.72343	1.73540	2.146	0.0391 *
Axis.4	-2.25911	2.76538	-0.817	0.4197

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Residual standard error: 0.3959 on 34 degrees of freedom

Multiple R-squared: 0.908, Adjusted R-squared: 0.8971

F-statistic: 83.85 on 4 and 34 DF, p-value: < 2.2e-16

Looks like first 3 axes are significant: choose these in regression models with Tmax

Check correlation between bioclim variables and phylogenetic

components

	Axis.1	Axis.2	Axis.3	Axis.4	bio5	bio6	bio7
Axis.1	1.000	0.000	0.000	0.000	0.882	0.745	-0.454
Axis.2	0.000	1.000	0.000	0.000	0.159	0.139	-0.089
Axis.3	0.000	0.000	1.000	0.000	0.151	0.301	-0.327
Axis.4	0.000	0.000	0.000	1.000	-0.044	-0.090	0.099
bio5	0.882	0.159	0.151	-0.044	1.000	0.772	-0.411
bio6	0.745	0.139	0.301	-0.090	0.772	1.000	-0.897
bio7	-0.454	-0.089	-0.327	0.099	-0.411	-0.897	1.000
merg\$nb	-0.258	0.023	-0.321	0.072	-0.412	-0.728	0.757

Model subsets

Construct full model to test interaction between Tma and each eigenvector(part of phylogeny)

```
#Ctmax = upper thermal limit  
# Axis1 - picea rudis split  
# Axis2 - N-S rudis clade split  
# Axis 3 - Pica split  
# Rearing temp: 20(23?) and 26  
#Bio 5 = Tmax
```

```
lm(Ctmax~bio5*Axis.1+bio5*Axis.2+bio5*Axis.3+Rearing.temp
```

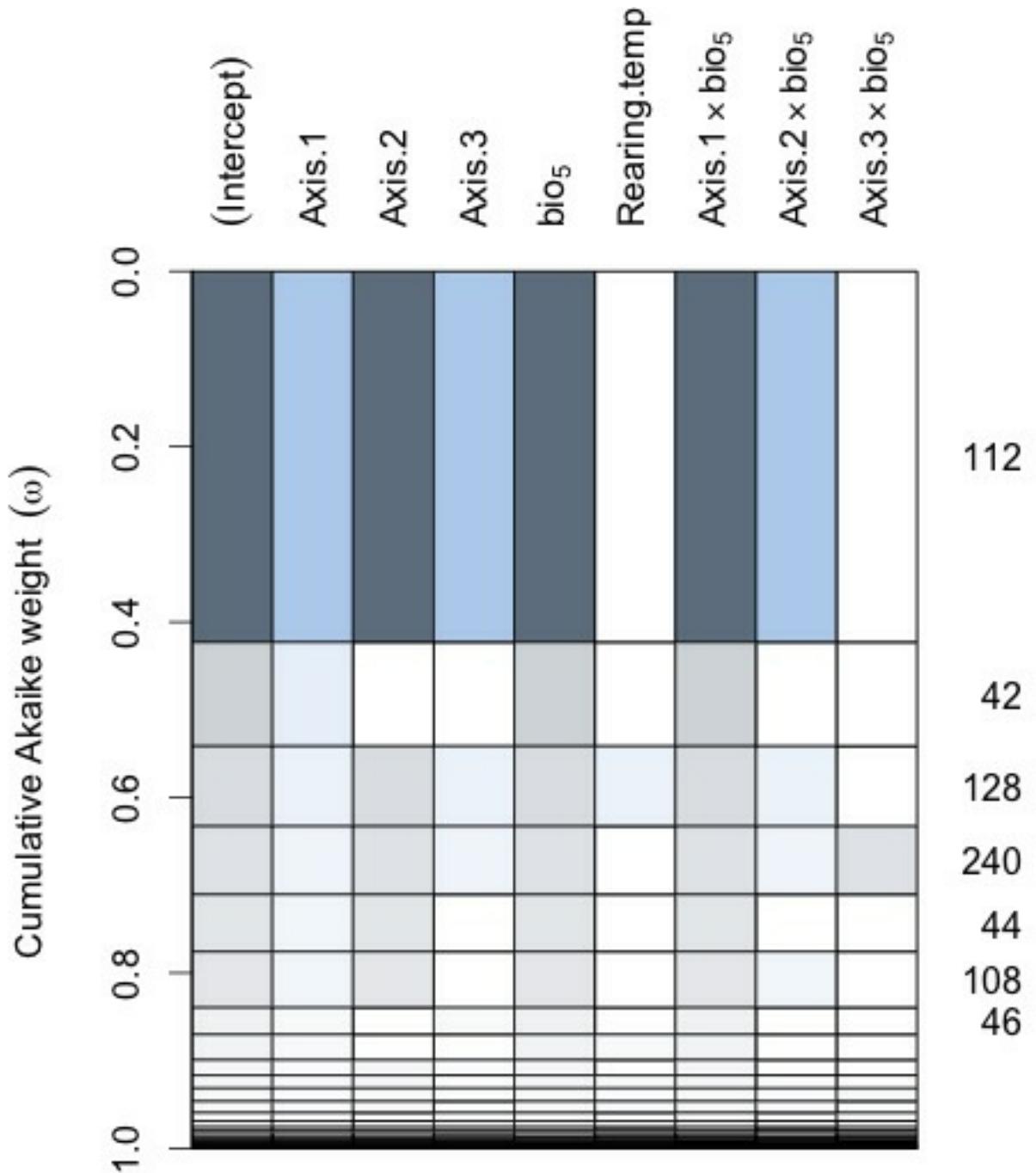
```
, data=merg)
```

Showing table of model subsets generated from dredge() function

	(Intercept)	Axis.1	Axis.2	Axis.3	bio
112	38.89633	-64.54042	163.379439	-7.1344145	0.114202
42	40.27150	-32.80201	NA	NA	0.068810
128	38.47028	-66.64409	164.835542	-7.4660612	0.121369
240	38.99929	-63.44634	165.175057	-17.1348640	0.110586
44	39.58611	-45.63541	-2.113326	NA	0.090878
108	40.67312	-34.07965	68.733204	NA	0.054957
46	40.41404	-31.25480	NA	0.5303359	0.063974
58	40.27478	-32.80931	NA	NA	0.068790
48	39.02228	-53.61002	-2.839872	-1.2211435	0.109601
60	39.47928	-45.71692	-2.160369	NA	0.091941
256	38.58207	-65.43836	166.850595	-18.6312806	0.117385
124	40.68974	-34.04638	68.875613	NA	0.054743
174	40.17971	-36.89376	NA	-39.0161170	0.070702
62	40.43434	-31.27748	NA	0.5367575	0.063799
176	38.62972	-58.09940	-4.104072	36.3065961	0.123395
64	38.74224	-54.92127	-3.032774	-1.3987999	0.114295

76	42.07767	13.63416	119.882620		NA	0.015140
8	42.43692	11.87677	2.870941	3.7234291		NA
190	40.09493	-37.03074		NA	-40.5928212	0.071616
6	42.43692	11.87677		NA	3.7234291	NA

###Cumulative AIC weights



2016-06-01 continued : Actually model averaging

top 2 AIC

```
>summary(model.avg[a.max[1:2]])
```

Full model-averaged coefficients (with shrinkage):

	Estimate	Std. Error	Adjusted SE	z value	Pr(> z)
(Intercept)	39.19741	1.81718	1.88065	20.842	< 2e-16 ***
Axis.1	-57.59157	22.32729	22.90080	2.515	0.0
Axis.2	127.60890	83.60797	84.76171	1.506	0.1
Axis.3	-5.57240	3.87984	3.94483	1.413	0.1
bio5	0.10426	0.06385	0.06611	1.577	0.1
Axis.1:bio5	2.29329	0.74450	0.76259	3.007	0.0
Axis.2:bio5	-4.10371	2.67227	2.70829	1.515	0.1

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '
	1				

Relative variable importance:

	Axis.1	bio5	Axis.1:bio5	Axis.2	Axis.3
3 Axis.2:bio5					
Importance:	1.00	1.00	1.00	0.78	0.78
	0.78				
N containing models:	2	2	2	1	1
	1				

###top 6 AIC

```
>summary(model.avg[a.max[1:6]])
```

Full model-averaged coefficients (with shrinkage):

	Estimate	Std. Error	Adjusted SE	z value	Pr <
(> z)					
(Intercept)	39.242396	1.895875	1.960765	20.014	< 2e-16 ***
Axis.1	-56.401968	22.642086	23.231252	2.428	0.01519 *
Axis.2	120.584643	84.389355	85.517189	1.410	0.15852
Axis.3	-5.992746	25.128163	26.111410	0.230	0.81848
bio5	0.101921	0.065747	0.068039	1.498	0.13414 **
Axis.1:bio5	2.251255	0.751716	0.770321	2.922	0.00347 **

```
Axis.2:bio5 -3.881627 2.688644 2.723787 1.425 0  
.15413  
Rearing.temp 0.001041 0.006764 0.006986 0.149 0  
.88151  
Axis.3:bio5 0.034305 0.915567 0.952226 0.036 0  
.97126  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' ' 1
```

Relative variable importance:

```
Axis.1 bio5 Axis.1:bio5 Axis.2 Axis.  
2:bio5 Axis.3
```

```
Importance: 1.00 1.00 1.00 0.86 0.78  
0.71
```

```
N containing models: 6 6 6 5 4  
3
```

```
Rearing.temp Axis.3:bio5
```

```
Importance: 0.11 0.09
```

```
N containing models: 1 1
```

###top 10 AIC

```
>summary(model.avg(a.max[1:10]))
```

Full model-averaged coefficients (with shrinkage):

	Estimate	Std. Error	Adjusted SE	z value	Pr <
(> z)					
(Intercept)	3.931e+01	1.901e+00	1.965e+00	20.004	<2e-16 ***
Axis.1	-5.462e+01	2.281e+01	2.337e+01	2.337	0
	.01945 *				
Axis.2	1.086e+02	8.793e+01	8.891e+01	1.221	0
	.22190				
Axis.3	-5.407e+00	2.393e+01	2.486e+01	0.218	0
	.82782				
bio5	9.962e-02	6.590e-02	6.817e-02	1.461	0
	.14391				
Axis.1:bio5	2.187e+00	7.598e-01	7.775e-01	2.813	0
	.00491 **				
Axis.2:bio5	-3.499e+00	2.803e+00	2.833e+00	1.235	0
	.21689				
Rearing.temp	9.893e-04	7.724e-03	7.987e-03	0.124	0
	.90143				
Axis.3:bio5	3.092e-02	8.693e-01	9.041e-01	0.034	0
	.97272				

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Relative variable importance:

	Axis.1	bio5	Axis.1:bio5	Axis.2	Axis.2:bio5	Axis.3
Importance:	1.00	1.00	1.00		0.81	0.70

```

0.69

N containing models: 10      10      10      7      4
5

Rearing.temp Axis.3:bio5

Importance: 0.15      0.08

N containing models: 3      1

```

top 4 delta AIC

```
>summary(model.avg(a.max, subset = delta < 4))
```

Full model-averaged coefficients (with shrinkage):

	Estimate	Std. Error	Adjusted SE	z value	Pr(> z)
(Intercept)	39.242396	1.895875	1.960765	20.014	< 2e-16 ***
Axis.1	-56.401968	22.642086	23.231252	2.428	0.01519 *
Axis.2	120.584643	84.389355	85.517189	1.410	0.15852
Axis.3	-5.992746	25.128163	26.111410	0.230	0.81848
bio5	0.101921	0.065747	0.068039	1.498	0.13414
Axis.1:bio5	2.251255	0.751716	0.770321	2.922	0.00347 **
Axis.2:bio5	-3.881627	2.688644	2.723787	1.425	0.99999

.15413

Rearing.temp	0.001041	0.006764	0.006986	0.149	0
	.88151				
Axis.3:bio5	0.034305	0.915567	0.952226	0.036	0
	.97126				

Signif. codes:	0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Relative variable importance:

Axis.1	bio5	Axis.1:bio5	Axis.2	Axis.	
2:bio5	Axis.3				
Importance:	1.00	1.00	1.00	0.86	0.78
	0.71				

N containing models:	6	6	6	5	4
	3				

Rearing.temp	Axis.3:bio5	
Importance:	0.11	0.09
N containing models:	1	1

###Comparing output to stepwise AIC both directions

```
> summary(stepAIC(full.max,direction="both"))

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 38.89633   1.77085  21.965 < 2e-16 ***
```

```
bio5          0.11420    0.06223   1.835  0.075805 .
Axis.1       -64.54042   19.60370  -3.292  0.002429 ** 
Axis.2        163.37944   55.72789   2.932  0.006179 ** 
Axis.3       -7.13441    2.85109  -2.502  0.017640 *  
bio5:Axis.1   2.53663    0.63426   3.999  0.000351 *** 
bio5:Axis.2   -5.25404   1.76036  -2.985  0.005402 ** 

---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1
```

```
Residual standard error: 0.3216 on 32 degrees of freedom
```

```
Multiple R-squared:  0.9428,    Adjusted R-squared:  0.93
```

```
21
```

```
F-statistic: 87.96 on 6 and 32 DF,  p-value: < 2.2e-16
```

```
<div id='id-section18.5'/>
```

```
#SHC suggestion: Just include all phylo axes in all analyses
```

```
full.max<-lm(Ctmax~bio5*Axis.1+bio5*Axis.2+bio5*Axis.3+bi
o5*Axis.4+Rearing.temp,data=merg)
```

```
##Showing top 2 AIC
```

```
summary(model.avg(a.max[1:2]))
```

```
Full model-averaged coefficients (with shrinkage):
```

	Estimate	Std. Error	Adjusted SE	z value	Pr(> z)
(Intercept)	39.19741	1.81718	1.88065	20.842	< 2e-16 ***
Axis.1	-57.59157	22.32729	22.90080	2.515	0.01191 *
Axis.2	127.60890	83.60797	84.76171	1.506	0.13220
Axis.3	-5.57240	3.87984	3.94483	1.413	0.15778
bio5	0.10426	0.06385	0.06611	1.577	0.1477
Axis.1:bio5	2.29329	0.74450	0.76259	3.007	0.0264 **
Axis.2:bio5	-4.10371	2.67227	2.70829	1.515	0.12971

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1
	' '	1			

Relative variable importance:

	Axis.1	bio5	Axis.1:bio5	Axis.2	Axis.3	Axis.2:bio5
Importance:	1.00	1.00	1.00	0.78	0.78	0.78
N containing models:	2	2	2	1	1	1
	1					

```
##Showing stepwise variable selection
```

```
> summary(stepAIC(full.max,direction="both"))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	38.89633	1.77085	21.965	< 2e-16	***
bio5	0.11420	0.06223	1.835	0.075805	.
Axis.1	-64.54042	19.60370	-3.292	0.002429	**
Axis.2	163.37944	55.72789	2.932	0.006179	**
Axis.3	-7.13441	2.85109	-2.502	0.017640	*
bio5:Axis.1	2.53663	0.63426	3.999	0.000351	***
bio5:Axis.2	-5.25404	1.76036	-2.985	0.005402	**

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1
. 1

Residual standard error: 0.3216 on 32 degrees of freedom

Multiple R-squared: 0.9428, Adjusted R-squared: 0.93
21

F-statistic: 87.96 on 6 and 32 DF, p-value: < 2.2e-16

#2016-06-02 update:

full mod construction for all traits

```
#Ctmax
```

```
full.max<-lm(Ctmax~bio5*Axis.1+bio5*Axis.2+bio5*Axis.3+bi
```

```
o5*Axis.4+Rearing.temp,data=merg)
```

#Ctmin

```
full.min<-lm(Ctmin~bio6*Axis.1+bio6*Axis.2+bio6*Axis.3+bi  
o6*Axis.4+Rearing.temp,data=merg)
```

#thermal tolerance breadth

```
TNB.full<-lm(nb~Axis.1*bio7+Axis.2*bio7+Axis.3*bio7+Axis.4*bio7+Rearing.temp,data=merg)
```

Probably a poor way to show output, but you can see the consistency with model averaging at different criteria for selecting top model (Top 2/6/10 AIC, Δ 4, 95 conf int):

Ctmax	Ctmin					TNB											
top 2 AICc	top 2 AICc					top 2 AICc											
Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)			
(Intercept)	39.1976	1.817176	1.880051	2.04246	0	(Intercept)	6.708118	0.227695	0.235429	26.49319	0	(Intercept)	27.05454	1.95524	1.968461	13.7385	0
AxIs.1	-59.50197	29.37729	22.8008	1.514828	0.011909	AxIs.2	-2.139192	2.377636	2.414152	0.908274	0.363734	b7o6	0.438122	0.021645	0.022378	19.57783	0
AxIs.2	127.6088	83.60797	84.76171	1.505502	1.212915	AxIs.3	-5.572395	3.879844	3.944825	1.412584	0.157778	AxIs.4	0.104264	0.063855	0.061111	1.577106	0.147761
b105	2.293286	0.744503	0.762593	3.00722	0.0262636	AxIs.5:bio5	-4.103714	2.672266	2.708291	1.515241	0.129711	AxIs.5:bio5	-4.103714	2.672266	2.708291	1.515241	0.129711
top 4 AICc	top 6 AIC					top 6 AIC					top 6 AIC						
Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)			
(Intercept)	39.19771	1.817176	1.880051	2.03039	0	(Intercept)	6.708188	0.227695	0.235429	26.49315	0	(Intercept)	27.05454	1.95524	1.968461	13.7385	0
AxIs.1	-59.50197	29.37729	22.8008	1.514828	0.011909	AxIs.2	-2.139192	2.377636	2.414152	0.908274	0.363734	b7o6	0.438122	0.021645	0.022378	19.57783	0
AxIs.2	128.5516	84.09567	85.20416	1.507484	0.131764	AxIs.3	-6.557056	24.85762	25.84512	0.253706	0.799723	AxIs.4	0.106761	0.064593	0.066955	1.594518	0.11082
b105	2.351012	0.7372743	0.752659	3.102352	0.001952	AxIs.5:bio5	-4.139189	2.678603	2.715903	1.524056	0.127495	AxIs.5:bio5	-4.139189	2.678603	2.715903	1.524056	0.127495
Rearing.term	0.001023	0.006706	0.006926	0.147723	0.882561	Rearing.term	-0.000455	0.010265	0.010626	0.042838	0.965831	AxIs.4	0.040415	0.730963	0.759753	0.053195	0.957577
AxIs.3:bio5	0.037078	0.907576	0.943915	0.035711	0.971513	AxIs.3:bio5	-0.00116	0.007016	0.007252	0.118736	0.905485	AxIs.4:bio5	-0.00116	0.007016	0.007252	0.118736	0.905485
top 10 AICc	top 10 AIC					top 10 AIC					top 10 AIC						
Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)			
(Intercept)	3.015-01	0.099779	0.106311	20.01693	0	(Intercept)	6.708105	0.227695	0.235429	27.12047	0	(Intercept)	26.29557	1.971391	2.030906	13.34027	0
AxIs.1	-5.49E-01	22.98736	23.54748	2.330393	0.971945	AxIs.2	-2.306237	2.491911	2.461931	0.93759	0.348882	b7o6	0.438149	0.020648	0.026666	16.17379	0
AxIs.2	1.13E-02	87.18011	88.20794	1.280664	0.200312	AxIs.3	-5.52E-02	22.98846	23.88232	0.231331	0.817058	AxIs.4	0.134918	0.913886	0.938079	1.438284	0.885564
b105	9.98E-02	0.065951	0.068222	1.462655	0.143562	AxIs.5:bio5	2.204204	0.766279	0.789393	2.803832	0.00505	AxIs.5:bio5	2.204204	0.766279	0.789393	2.803832	0.00505
Rearing.term	8.61E-06	0.007016	0.007252	0.118736	0.905485	Rearing.term	-0.00116	0.007016	0.007252	0.118736	0.905485	AxIs.4	-5.58E-02	0.843673	0.874007	0.006338	0.994095
AxIs.3:bio5	2.85E-02	0.834372	0.867772	0.032824	0.973815	AxIs.3:bio5	-0.00116	0.007016	0.007252	0.118736	0.905485	AxIs.4:bio5	-0.00116	0.007016	0.007252	0.118736	0.905485
delta 4	delta 4					delta 4					delta 4						
Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)			
(Intercept)	3.92E-01	1.839977	1.959535	2.008481	0	(Intercept)	6.70814	0.368598	0.380575	17.6202	0	(Intercept)	26.99363	1.983194	2.048548	13.17696	0
AxIs.1	-5.72E-01	22.60477	23.20901	2.463546	0.013761	AxIs.2	-2.097715	2.400133	2.438602	0.885162	0.380803	b7o6	0.434141	0.051851	0.053574	6.372708	0
AxIs.2	1.24E-02	83.3521	84.5345	1.471524	0.141149	AxIs.3	-0.10200	24.04585	24.9866	0.244186	0.807087	AxIs.4	0.084009	1.175858	1.215284	0.069127	0.944889
b105	6.10E-02	0.065748	0.068063	1.515669	0.129603	AxIs.5:bio5	2.167097	0.774457	0.799384	2.803832	0.00505	AxIs.5:bio5	2.167097	0.774457	0.799384	2.803832	0.00505
Rearing.term	2.28E-03	0.749622	0.768723	2.96354	0.003041	Rearing.term	-0.001052	0.011709	0.012089	0.087037	0.930642	AxIs.4	-0.018674	0.95272	0.958921	0.194847	0.984453
AxIs.3:bio5	-4.00E-06	0.652554	0.692134	1.487272	0.136043	AxIs.3:bio5	-0.00116	0.018487	0.191488	0.060101	0.95132	AxIs.4:bio7	0.847238	4.894996	4.961944	0.170747	0.864423
AxIs.4	3.76E-02	0.705181	0.732951	0.051308	0.95908	AxIs.4:bio5	-0.023578	0.126854	0.128488	0.183501	0.854405	AxIs.4:bio7	-0.004577	0.51494	0.526154	0.122734	0.902318
95 conf int	95 conf int					95 conf int					95 conf int						
Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)	Estimate	st SE	Adjusted SE	z value	Pr(> z)			
(Intercept)	39.34503	1.925652	1.990504	19.76637	0	(Intercept)	6.700288	0.445664	0.460353	14.55468	0	(Intercept)	26.99134	1.983194	2.048548	13.17696	0
AxIs.1	-54.06306	23.24734	23.8115	2.2704	0.02318	AxIs.2	-2.082293	2.257452	2.621191	0.794407	0.426959	b7o6	0.434177	0.056532	0.058389	5.853388	0
AxIs.2	105.8731	88.45648	89.42203	1.183971	0.236425	AxIs.3	-0.226675	1.259187	1.393486	0.172151	0.863319	AxIs.4	0.171162	2.726291	2.351.91	0.072861	0.941936
b105	-5.43-038	26.52153	27.5459	0.197212	0.843662	AxIs.5:bio5	-2.167097	0.774457	0.799384	0.060333	0.95054	AxIs.5:bio5	-2.167097	0.774457	0.799384	0.060333	0.95054
Rearing.term	0.001016	0.008617	0.008481	0.120188	0.904334	Rearing.term	-0.001052	0.007016	0.007252	0.093333	0.95054	AxIs.5:bio6	-0.244608	3.82073	3.966705	0.0753	0.959365
AxIs.4	2.167887	41.79217	43.1077	0.050453	0.956969	AxIs.6:bio5	-0.366728	3.789057	3.750744	0.041045	0.940745	AxIs.6:bio6	-0.323789	6.259504	6.360073	0.200984	0.834383
AxIs.3:bio5	0.037162	0.968321	1.006521	0.036921	0.970548	AxIs.7:bio5	-0.210874	1.176293	1.246462	0.038477	0.969308	AxIs.7:bio6	-0.100084	7.755699	7.968682	0.1256	0.900404
AxIs.4:bio5	-0.066848	1.259095	1.298786	0.051469	0.958951	AxIs.8:bio5	-0.1320819	6.259504	6.360073	0.200984	0.834383	AxIs.8:bio6	-0.033818	0.346496	0.355516	0.095127	0.924214
AxIs.5:bio5	-0.023578	0.126854	0.128488	0.183501	0.854405	AxIs.9:bio5	-0.086117	0.981984	1.005278	0.085664	0.931733	AxIs.9:bio6	-0.045477	0.51494	0.526154	0.122734	0.902318

Ctmax	X	X.1	X.2	X.3
	Estimate	st SE	Adjusted SE	z value
top 2 AICc				
(Intercept)	39.1974113	1.81717555	1.88065146	20.84246
Axis.1	-57.5915669	22.32728931	22.90079918	2.514828
Axis.2	127.6089015	83.60796871	84.76171296	1.505502
Axis.3	-5.5723952	3.87984439	3.94482507	1.412584
bio5	0.1042642	0.06385464	0.06611108	1.577106
Axis.1:bio5	2.2932857	0.74450301	0.76259317	3.00722
Axis.2:bio5	-4.1037142	2.67226625	2.70829115	1.515241
top 6 AICc				
(Intercept)	39.09770833	1.858677254	1.925749929	20.30258
Axis.1	-58.89921078	22.04271462	22.67258208	2.597816
Axis.2	128.5516463	84.09957359	85.29260627	1.507183
Axis.3	-6.557056311	24.86761943	25.84512303	0.253705
bio5	0.106760957	0.064593101	0.066955014	1.594517
Axis.1:bio5	2.335012045	0.732742988	0.752658592	3.102352
Axis.2:bio5	-4.139188942	2.678608124	2.715902917	1.524056

Rearing.temp	0.001023202	0.006706439	0.006926472	0.147723
Axis.4	0.040415014	0.730962899	0.759753398	0.053194
Axis.3:bio5	0.033707897	0.907576414	0.943914621	0.035710
top10 AICc				
	Estimate	st SE	Adjusted SE	z value
(Intercept)	3.93E+01	1.899791736	1.963911039	20.01693
Axis.1	-5.49E+01	22.98798186	23.54781004	2.330392
Axis.2	1.13E+02	87.18011426	88.20794399	1.280663
Axis.3	-5.52E+00	22.98845903	23.88232484	0.231330
bio5	9.98E-02	0.06595076	0.068221677	1.462654
Axis.1:bio5	2.20E+00	0.766278694	0.783913018	2.803832
Axis.2:bio5	-3.64E+00	2.782129711	2.81413218	1.292064
Rearing.temp	8.61E-04	0.007016252	0.007252224	0.118735
Axis.4	-5.58E-03	0.843672767	0.874007493	0.006385
Axis.3:bio5	2.85E-02	0.834371842	0.867772028	0.032823
delta 4				
	Estimate	st SE	Adjusted SE	z value
(Intercept)	3.92E+01	1.893976934	1.959535391	20.00841

Axis.1	-5.72E+01	22.60477016	23.20900583	2.463456
Axis.2	1.24E+02	83.35210021	84.53449839	1.471524
Axis.3	-6.10E+00	24.04585344	24.98659536	0.244185
bio5	1.03E-01	0.065747663	0.068062898	1.515669
Axis.1:bio5	2.28E+00	0.749621892	0.768722868	2.963540
Axis.2:bio5	-4.00E+00	2.655253834	2.692134467	1.487272
Rearing.temp	9.52E-04	0.006474441	0.006686524	0.142389
Axis.4	3.76E-02	0.705181274	0.732950521	0.051308
Axis.3:bio5	3.14E-02	0.875514464	0.910565657	0.034446

95 conf int

	Estimate	st SE	Adjusted SE	z value
(Intercept)	39.34503249	1.92565228	1.99050365	19.76637
Axis.1	-54.06305882	23.2473386	23.81149971	2.270460
Axis.2	105.8730629	88.45648472	89.42203144	1.183970
Axis.3	-5.432382412	26.52153131	27.54590024	0.197212
bio5	0.098505527	0.066665352	0.068957648	1.428493
Axis.1:bio5	2.167761235	0.774452089	0.79221653	2.736324
Axis.2:bio5	-3.410786773	2.820916563	2.850966	1.196361
Rearing.temp	0.001015674	0.008166926	0.008450716	0.120187
Axis.4	2.178787485	41.79217286	43.10770462	0.050542

Axis.3:bio5	0.037161754	0.968321351	1.006521472	0.036920
Axis.4:bio5	-0.06684784	1.259095332	1.298786073	0.051469

<div id='id-section19'>

Page 19: 2016-06-01 Variance partitioning: thermal tolerance breadth example

Partitioning variation into phylogenetic (Axes 1-4), ecological (Tmax or Tmin or TAR), and phylogenetic + ecological components using the varpart() function in the vegan R package:

- a+b= phylo
 - b= shared
 - c+b= ecological
 - a = phylo independent of ecology
 - c = ecology independent of phylo
- code for model construction:

```
#Ctmax
#varpar
full<-varpart(merg$Ctmax,~Axis.1+Axis.2+Axis.3+Axis.4,~bi
o5,data=merg)
full
```

output

Partition table:

	Df	R.squared	Adj.R.squared	Testable
[a+b] = X1	4	0.90796	0.89713	TRUE
[b+c] = X2	1	0.75637	0.74979	TRUE
[a+b+c] = X1+X2	5	0.90814	0.89422	TRUE

Individual fractions

[a] = X1 X2	4	0.14444	TRUE
[b]	0	0.75270	FALSE
[c] = X2 X1	1	-0.00291	TRUE
[d] = Residuals		0.10578	FALSE

Use **function 'rda'** to test significance of fractions of interest

Looking at plots

```
plot(full)
```



```
#global model: a+b+c  
anova(rda(merg$Ctmax~Axis.1+Axis.2+Axis.3+Axis.4+bio5, dat  
a=merg))  
  
#fraction a+b  
ab<- rda(merg$Ctmax~Axis.1+Axis.2+Axis.3+Axis.4, data=merg)
```

```

anova(ab)

#frac b+c

bc<- rda(merg$Ctmax~bio5,data=merg)

anova(bc)

#fraction a (phylo)

a<- rda(merg$Ctmax~Axis.1+Axis.2+Axis.3+Axis.4+Condition(b
io5),data=merg)

anova(a)

#fraction c (eco)

c<- rda(merg$Ctmax~Condition(Axis.1+Axis.2+Axis.3+Axis.4)+
bio5,data=merg)

anova(c)

```

Only showing code for CTmax I also applied variance partitioning for Ctmin and thermal tolerance breadth

```

<div id='id-section19.5'>

###Summary of results: Proportion of variance assigned to each
component

```

Please scroll right to see the whole table, this table is wide

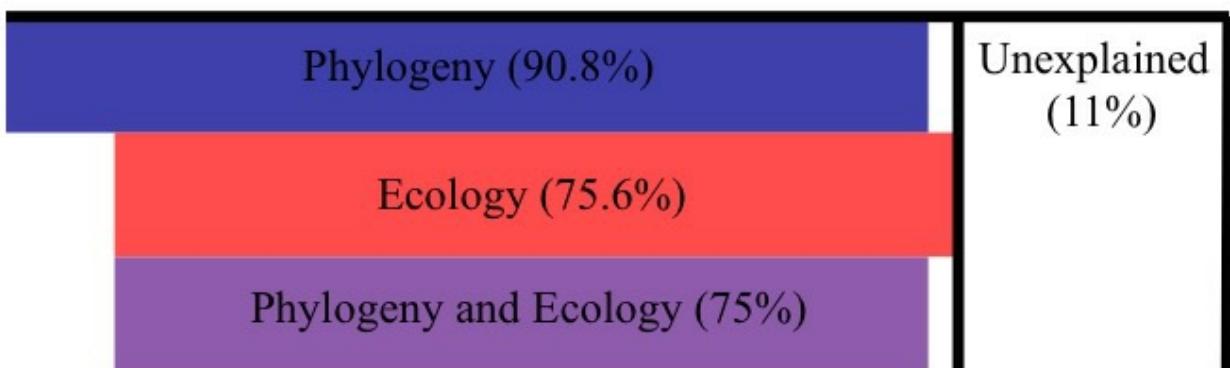
Trait	Independent.Phylogeny	Independent.Ecology	Phylo
Ctmax	0.14	0	0.90
Ctmin	0	0.31	0.64
Tolerance Breadth	0	0.45	0.17

Note-Bolded values represents significant variance component. The combined phylogeny and ecology variance component can not be tested for significance, only indirectly measured. **The ecological component is represented by Tmax for Ctmax, Tmin for CTmin, and TAR for tolerance breadth.**

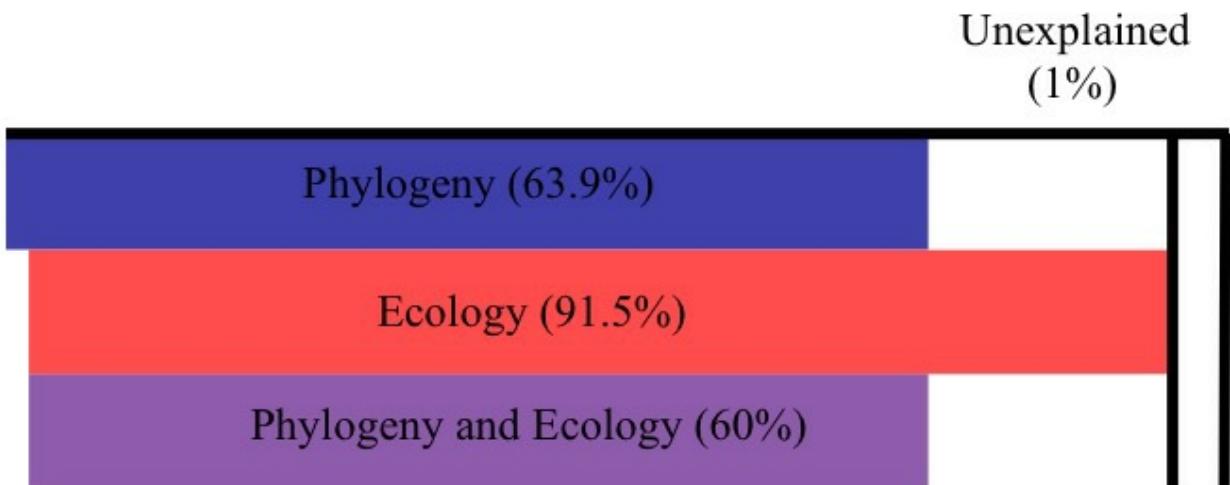
Different way to represent proportion of variance explained by each component

CTmax

Variance partitioning CTmax



Variance partitioning CTmin



```
<div id='id-section20'>
```

```
###Notes from climate cascade meeting, 2016-06-01
```

I have meetings with SHC and NJG every week, so I'll start logging our discussions here

We talked about the analysis from the thermal niche paper:

1. NJG and SHC don't have strong feelings about model averaging.

2. FOr the 4 panel field vs phytotron CTmax and Ctmin figure, keep separate lines for each species
3. For thermal tolerance breadth, make 1 line
4. Include variance partitioning analysis: Estimate amount of variance that go into phylogenetic components, ecological component, and their shared component.
5. For CTmax , perform a Levine's test on the raw residuals from the regression line for picea (field vs phytotron).
6. NJG: What does the literature say? Do people compare field vs common garden often? Do people assay thermal tolerance in the field alone?

Writing this up

SHC suggestion for results: Talk about field, then phyto, then present thermal tolerance breadth for phytotron.

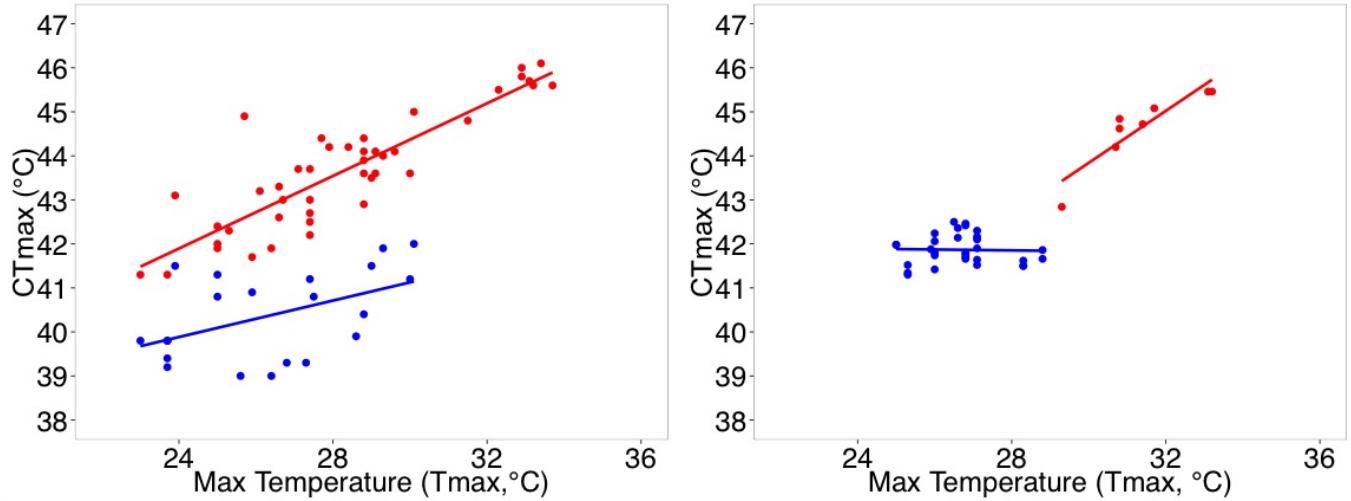
For the phytotron gxp paper:

1. Remake boxplot to include Axis 2
-

<div id='id-section21'>

Page 21: 2016-06-02. Levine's test for raw residuals

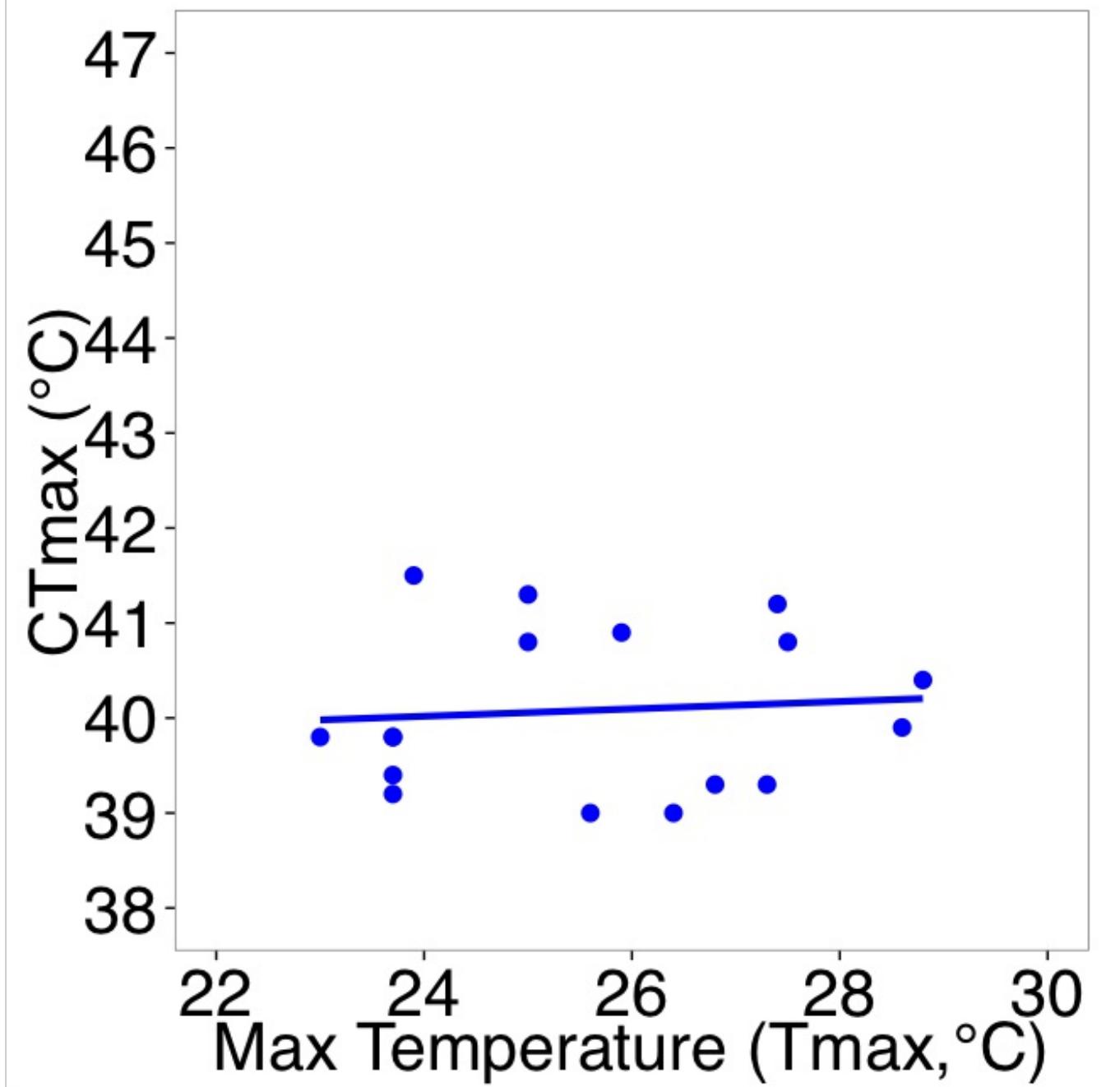
We(SHC) suspect that the variance in field samples (for CTmax) is larger than the ones in the phytotron for A. picea.



In this fig, just look at the blue line. Left is field, right panel is phyto.

There is a cline in the field samples, but the cline goes away when comparing similar Tmax range as phyto:

This is a re-analysis for the field samples



To test differences in variances, we'll perform a Levine's test on the raw residuals.

About the test: [some background](#)

**Using the [car package](#) in R

```
###Raw residuals in long format
```

Raw Residuals	Field_V_phyto

1.2426873	field
1.0498107	field
0.1956326	field
-0.8463195	field
1.4852558	field
-0.2966277	field
0.7426873	field
0.8078586	field
0.6459408	field
-0.8070045	field
-0.2070045	field
-0.6070045	field
-0.2070045	field
-0.1799154	field
-1.1114907	field
-0.8269702	field
-1.0805318	field
-0.1320043	phyto
-0.4520043	phyto
0.0980829	phyto
0.0980829	phyto

-0.1842485	phyto
0.0157515	phyto
-0.2211002	phyto
-0.3411002	phyto
-0.2292049	phyto
-0.3492049	phyto
-0.3492049	phyto
-0.1640741	phyto
-0.1040741	phyto
-0.2040741	phyto
-0.0640741	phyto
-0.0720043	phyto
0.1879957	phyto
0.4388998	phyto
0.0388998	phyto
0.2988998	phyto
0.2388998	phyto
0.2739434	phyto
0.4939434	phyto
0.5959259	phyto
0.5559259	phyto

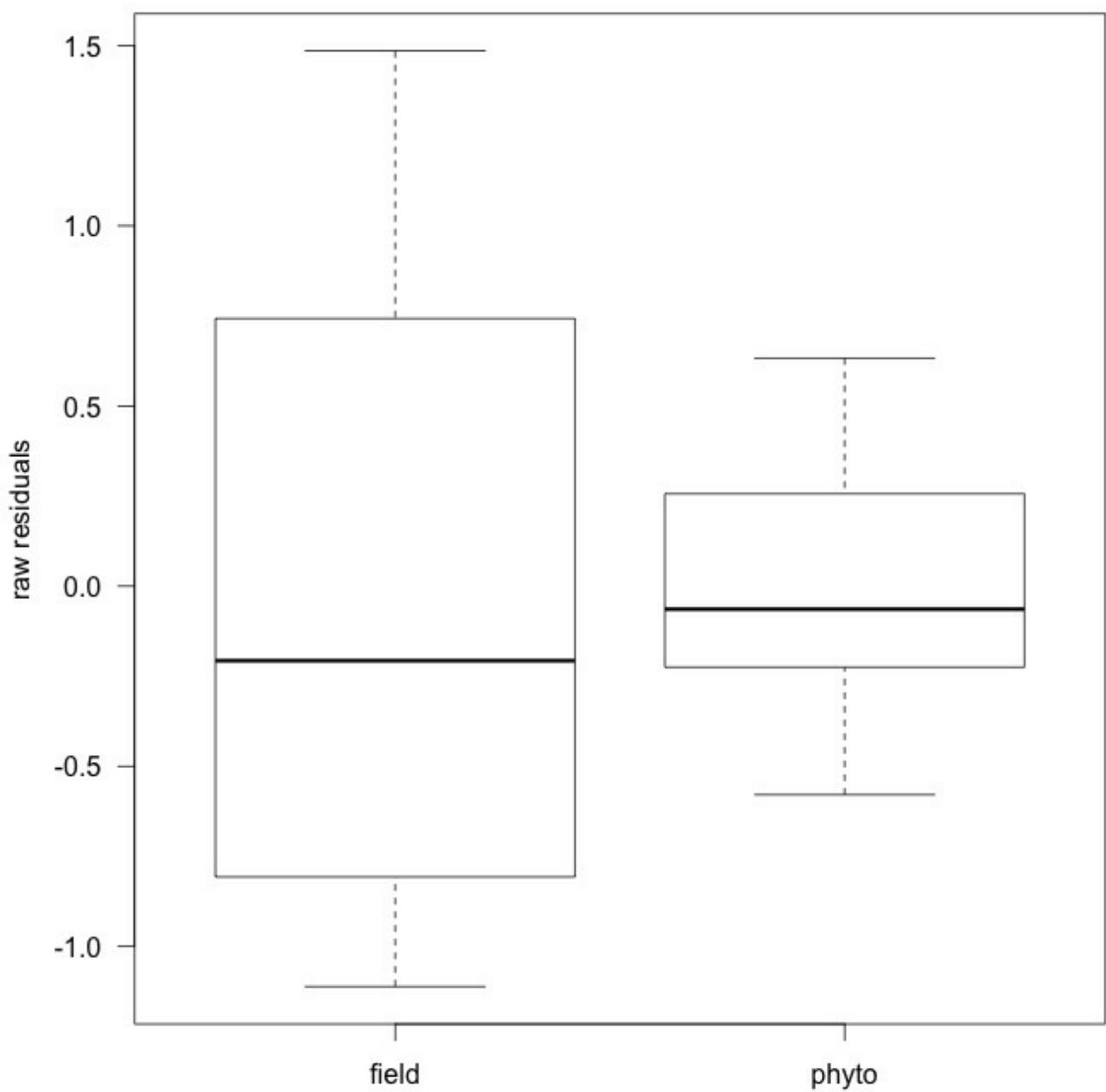
0.3679957	phyto
0.6329521	phyto
0.0070044	phyto
-0.5389433	phyto
-0.5789433	phyto
-0.3589433	phyto

Code:

```
library(car)
#levene's test
leveneTest(lt[,1],lt[,2])

Levene's Test for Homogeneity of Variance (center = median)
Df F value    Pr(>F)
group  1 16.299 0.0002028 ***
               46
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
               ' ' 1

#visualizing
boxplot(lt[,1]~lt[,2],ylab="raw residuals",las=1)
```



##Summary: Yes, sig diff in variance between field and phyto.

<div id='id-section22'/>

fitting nls() functions in R!!

I googled *how to fit nls even when failing to converge in R* and found this [gem](#).

Basically, use [nls2\(\)](#) to brute force fit curves. I have not tried it, but putting it here as a ref.

```
<div id='id-section23'>  
###Page 23: 2016-06-02. Literature reference for thermal niche  
paper to help write manuscript
```

Probably not comprehensive, but here it is:

Thermal breadth = 1 if they analyze it, 0 if they don't.

###Table:

Type	Author	Year	Journal	Taxa	Rearing
Meta-analysis	Addo-Bediako et al.	2000	Proceedings of the royal society b	Insects	
Lab acclimation	Deere & Chown	2006	American Naturalist	Mites	1; 5; 10
Field	Compton et al.	2007	Experimental marine biology and ecology	Bivalve	
Lab acclimation	Calosi et al.	2008	Biology letters	Beetles	14.5; 20

Lab acclimation	Calosi et al.	2008	Journal of biogeography	Beetles	14.5; 2
Field	Sinervo et al.	2010	Science	Lizards	
Lab acclimation	Calosi et al.	2010	Journal of Animal Ecology	Beetles	14.5; 2
Lab acclimation	Anert et al.	2011	Integrative and Comparative Biology	Plants	20-24
Meta-analysis	Sunday et al.	2011	Proceedings of the royal society b	Terrestrial and Marine	
Common garden	Overgaard et al.	2011	American Naturalist	Fruit Fly	25;29
Common garden	Krenek et al.	2012	Plosone	Paramecium	22
Meta-analysis	Grigg & Buckley	2012	Biology letters	Lizards	
Short acclimation	Sheldon & Tewksbury	2014	Ecology	Beetles	20
Common garden	Sheth & Angert	2014	Evolution	Plants	20-25
Meta-analysis	Khaliq et al.	2014	Proceedings of the royal society b	Birds and Mammal	

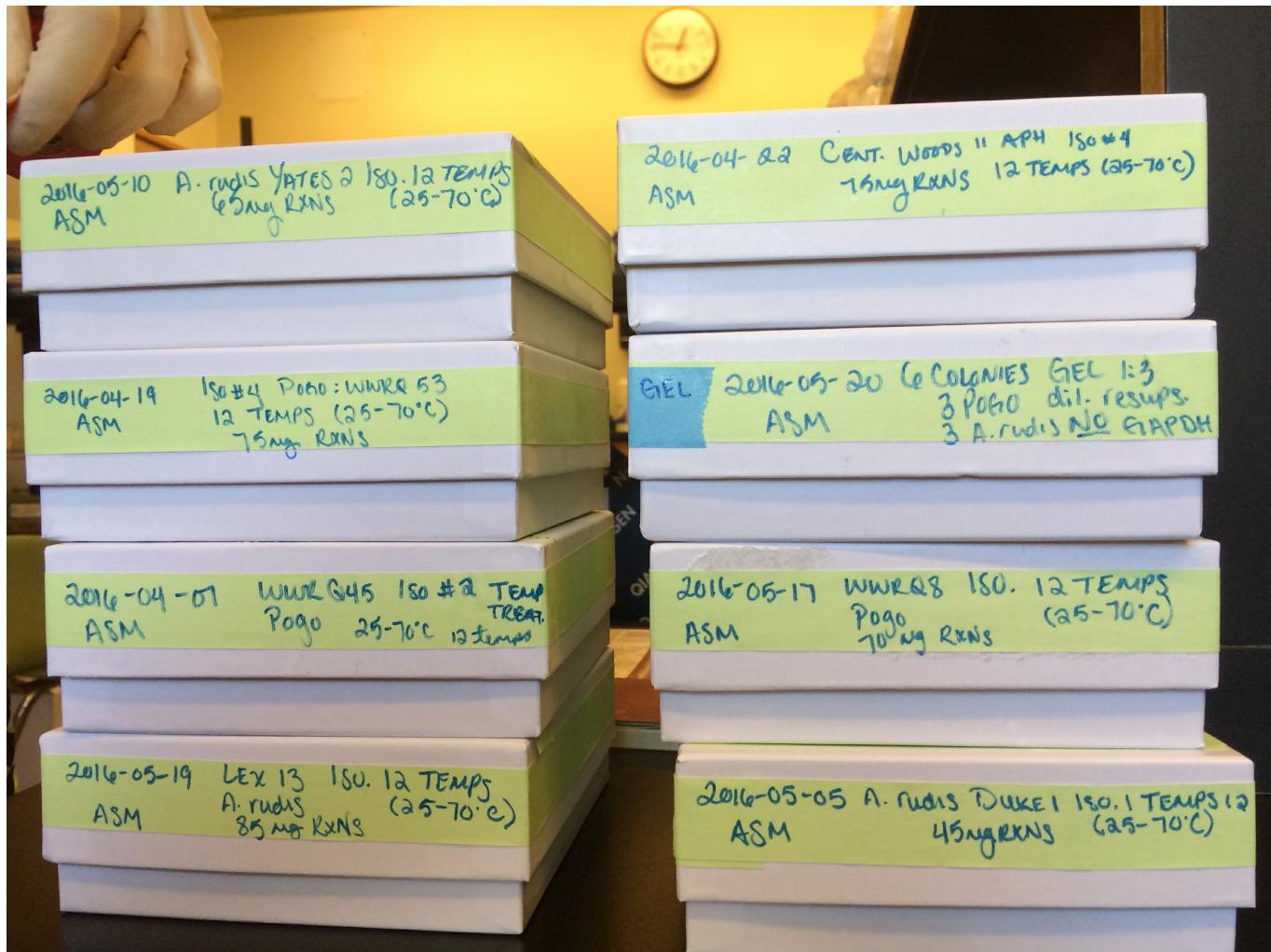
Short acclimation	Sheldon et al.	2015	Global Ecology and Biogeography	Lizards	29
Lab acclimation	Bonino et al.	2015	Zoology	Lizards	20-40
	Velasco et al.	2016	Journal of biogeography		
Meta-analysis	Lancaster	2016	Nature Climate Change	Insects	
Lab acclimation	Gutierrez-Pesquera et al.	2016	Journal of biogeography	Frogs (tadpoles)	20

<div id='id-section24'>

Page 24: 2016-06-03. Proteome stability project: Organizational entry

Today is Amanda's last day, so sad. She was working on the proteome stability project. Here I'll log all the organizational info that I'll need in the future:

1. Where are the samples stored: Amanda and I both transferred the gel pieces, native and total protein boxes per colony to the -80C downstairs.



2. For the TMT labeling, what order will they be labeled? See table below
3. What else needs to be done? Wai and Bethany will resuspend our tryptic peptides, take some out and run some of the samples on LTQ to see if we have peptides. If we have peptides, then Wai will do the labeling for us! Wow!

Organization table: 3 pogo and 3 rudis colonies treated across 10 temperatures, that will be TMT labelled. LTQ run means that a

subsample will be taken out to run on mass spec to check for peptides. ug of sample indicates how much protein we have.

Species	Replicate	Colony	Temperature	Sample...	Sample
P. barbatus	1	WWR45	30.1	2	P45-2
P. barbatus	1	WWR45	36.0	3	P45-3
P. barbatus	1	WWR45	41.2	4	P45-4
P. barbatus	1	WWR45	43.9	5	P45-5
P. barbatus	1	WWR45	46.3	6	P45-6
P. barbatus	1	WWR45	48.2	7	P45-7
P. barbatus	1	WWR45	50.3	8	P45-8
P. barbatus	1	WWR45	55.1	9	P45-9
P. barbatus	1	WWR45	61.2	10	P45-10
P.	1	WWR45	65.2	11	P45-11

barbatus						
A. rудis	1	Duke 1	30.1	2	ARD1-2	
A. rудis	1	Duke 1	36.0	3	ARD1-3	
A. rудis	1	Duke 1	41.2	4	ARD1-4	
A. rудis	1	Duke 1	43.9	5	ARD1-5	
A. rудis	1	Duke 1	46.3	6	ARD1-6	
A. rудis	1	Duke 1	48.2	7	ARD1-7	
A. rудis	1	Duke 1	50.3	8	ARD1-8	
A. rудis	1	Duke 1	55.1	9	ARD1-9	
A. rудis	1	Duke 1	61.2	10	ARD1-10	
A. rудis	1	Duke 1	65.2	11	ARD1-11	
P. barbatus	2	WWRQ53	30.1	2	P53-2	
P. barbatus	2	WWRQ53	36.0	3	P53-3	
P. barbatus	2	WWRQ53	41.2	4	P53-4	
P. barbatus	2	WWRQ53	43.9	5	P53-5	
P. barbatus	2	WWRQ53	46.3	6	P53-6	
P. barbatus	2	WWRQ53	48.2	7	P53-7	
P.	2	WWRQ53	50.3	8	P53-8	

barbatus						
P. barbatus	2	WWRQ53	55.1	9	P53-9	
P. barbatus	2	WWRQ53	61.2	10	P53-10	
P. barbatus	2	WWRQ53	65.2	11	P53-11	
A. rудis	2	Yates 2	30.1	2	ARY2-2	
A. rудis	2	Yates 2	36.0	3	ARY2-3	
A. rудis	2	Yates 2	41.2	4	ARY2-4	
A. rудis	2	Yates 2	43.9	5	ARY2-5	
A. rудis	2	Yates 2	46.3	6	ARY2-6	
A. rудis	2	Yates 2	48.2	7	ARY2-7	
A. rудis	2	Yates 2	50.3	8	ARY2-8	
A. rудis	2	Yates 2	55.1	9	ARY2-9	
A. rудis	2	Yates 2	61.2	10	ARY2-1	
A. rудis	2	Yates 2	65.2	11	ARY2-1	
P. barbatus	3	WWRQ8	30.1	2	P8-2	
P. barbatus	3	WWRQ8	36.0	3	P8-3	
P. barbatus	3	WWRQ8	41.2	4	P8-4	
P.	3	WWRQ8	43.9	5	P8-5	

barbatus						
P. barbatus	3	WWRQ8	46.3	6	P8-6	
P. barbatus	3	WWRQ8	48.2	7	P8-7	
P. barbatus	3	WWRQ8	50.3	8	P8-8	
P. barbatus	3	WWRQ8	55.1	9	P8-9	
P. barbatus	3	WWRQ8	61.2	10	P8-10	
P. barbatus	3	WWRQ8	65.2	11	P8-11	
A. rудis	3	Lex 13	30.1	2	ARL13-	
A. rудis	3	Lex 13	36.0	3	ARL13-	
A. rудis	3	Lex 13	41.2	4	ARL13-	
A. rудis	3	Lex 13	43.9	5	ARL13-	
A. rудis	3	Lex 13	46.3	6	ARL13-	
A. rудis	3	Lex 13	48.2	7	ARL13-	
A. rудis	3	Lex 13	50.3	8	ARL13-	
A. rудis	3	Lex 13	55.1	9	ARL13-	
A. rудis	3	Lex 13	61.2	10	ARL13-	
A. rудis	3	Lex 13	65.2	11	ARL13-	

###A note: showing actual temperature treatments from thermal cycler

Thermocylcer.Actual.Temp	Temperature
25.0	25
30.1	30
36.0	35
41.2	40
43.9	43
46.3	45
48.2	48
50.3	50
55.1	55
61.2	60
65.2	65
70.1	70

<div id='id-section25'>

**Page 25: 2016-06-03. ggplot
reference, updating a figure from [Page](#)**

20: 2016-06-02

For JSG gxp ms that SHC is writing. Adding axis 2 into boxplot for hsp40 basal xp.

###code for manipulating data so that I convert different axes into factors! There is probably a better way of doing this, but...

```
mergy<-subset(merg,merg$Axis.2> -0.1) # excluding axis 2 samples  
sub<-subset(merg,merg$Axis.2< -0.1)# taking out samples separating axis 2  
sub$axis3_desig<-rep("zAxis 2 A. picea",3) #naming factors based on axis2  
mergy$axis3_desig<-ifelse(mergy$Axis.3<= -0.044,"North",ifelse(mergy$Axis.3>0.05,"South","A. picea")) # axis 3 designations!  
mergy<-rbind(mergy,sub) # combine them!  
mergy<-mergy[-54,] # 54th row has an NA
```

##ggplot settings I like:

```
T<-theme_bw() + theme(text=element_text(size=30),axis.text=element_text(size=30),  
legend.text=element_text(size=28),panel.grid.major=element_blank(),  
legend.position="none",panel.grid.minor.x = element_blank)
```

```
(),
panel.grid = element_blank(),legend.key = element_blank()
)
```

###Code to make fig

```
meds <- c(by(mergy$B_40, mergy$axis3_desig, median))

Axis3_b40_v3<-ggplot(data=mergy,aes(x=factor(axis3_desig)
,y=B_40,fill=factor(axis3_desig)))+
geom_boxplot() +T+
ylab(expression(paste("Hsp40 basal expression (",2^past
e(Delta,Delta,"CT"),")")))
+scale_x_discrete(expression(paste(italic(
A. rудis)," clade")),labels=expression(paste(ita
lic("A.
picea ")," North "," South ", " A
xis 2")))+
scale_y_continuous(limits=c(-1,11),breaks=seq(0,11,1))+
scale_fill_manual(name = "", values = c("gray","deepsky
blue4", "firebrick","purple"))+guides(fill=FALSE)
Axis3_b40_v3
```

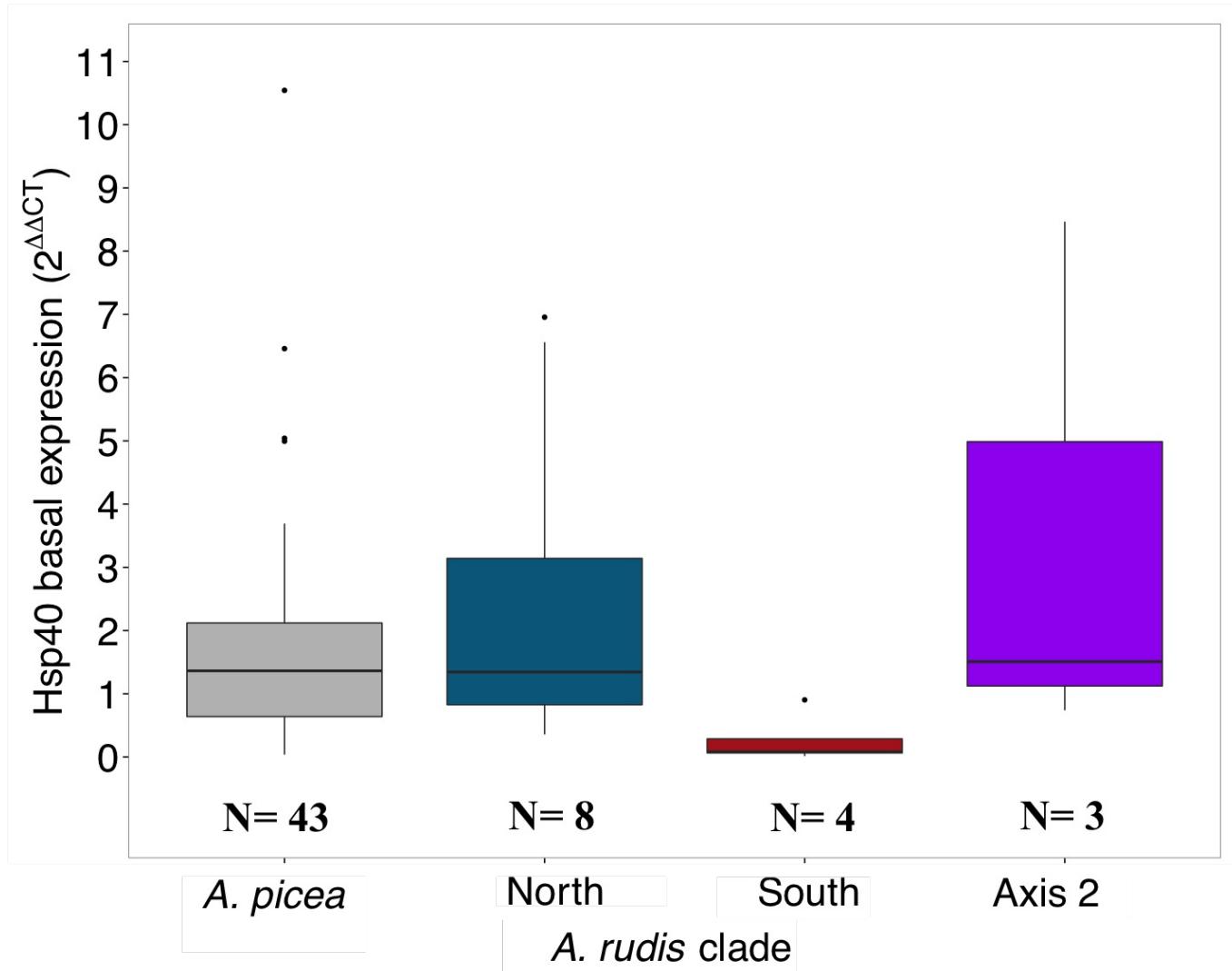
**highlighting part of code
where I can incorporate math**

symbols into the y axis:

```
ylab(expression(paste("Hsp40 basal expression (", 2^paste(Delta, Delta, "CT"), ")")))
```

###Final fig

I did play around with the fig in ppt first



<div id='id-section26'>

I was having lunch with Federico and thought: When I go to seminars and cell biologists use markers to indicate cell types, how do they know? What exactly is a cell type?

I've seen scientists using 1 marker to say, this is a this type of cell.

And in our graduate seminar series, there is a group that studies the physiology of taste receptors. They could not ID specific cell types and that part of the biology was unknown (not the typ 1/2/3's, but the VNO?).

So I thought: Why not try to do single cell transcriptomics for 10 cells per group (what you think is a group). Then I'd explicitly test for differences using a discriminant analysis, or classification analysis. This approach could lead to a more quantitative justification for designating cell types.

Then we imagined that a totipotent cell that differentiates into diverse cell types can also follow or resemble a phylogenetic tree.

Feder suggests to read:

[Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis](#)

[Gene expression divergence recapitulates the developmental hourglass model](#)

```
<div id='id-section27'>
```

```
###Page 27: 2016-06-03. qPCR plate layout and using the loaner ABI  
steponeplus Page 11: 2016-05-18
```

I started up the aBI steponeplus loaner today.



My usual 96 well plate layout is in my physical notebook, but I'll share it here:

X	X1	X2	X3	X4	X5
A	Colony1:T1	Colony1:T2	Colony1:T3	Colony1:T4	Colony1:T5
B	Colony1:T1	Colony1:T2	Colony1:T3	Colony1:T4	Colony1:T5

C	Colony2:T1	Colony2:T2	Colony2:T3	Colony2:T4	Colony2:T5
D	Colony2:T1	Colony2:T2	Colony2:T3	Colony2:T4	Colony2:T5
E	Colony3:T1	Colony3:T2	Colony3:T3	Colony3:T4	Colony3:T5
F	Colony3:T1	Colony3:T2	Colony3:T3	Colony3:T4	Colony3:T5
G	Colony4:T1	Colony4:T2	Colony4:T3	Colony4:T4	Colony4:T5
H	Colony4:T1	Colony4:T2	Colony4:T3	Colony4:T4	Colony4:T5

For each plate, I can run the 12 points of a performance curves for 4 colonies in duplicates. Each colony takes up 24 wells: 12 (T1-T12) temperature treatments and then ran in duplicates. Conditions for qPCR found [here](#).

Usual temperatures:

1. T1 - 25 C
2. T2 - 28 C
3. T3 - 30 C
4. T4 - 31.5 C
5. T5 - 33 C
6. T6 - 35 C
7. T7 - 36.5 C
8. T8 - 38.5 C
9. T9 - 40 C
10. T10 - 41 C
11. T11 - 25 C (middle of run)
12. T12 - 25 C (end of run)

T11 and T12 are in there to serve as a time control (When I do the delta delta CT calculation, I'll include those to wash out the effect of time.

###Actual samples I ran today for hsp40 541-641 primer pair, 55 C annealing.

Colonies:

1. Duke2
2. HF2
3. Kite4
4. Kite8

##Summary of results:

All colonies had double peaks. So they're not usable. For these colonies, only hsp83 279-392 primers worked. Next, do 18s rRNA for housekeeping gene.

Silvia asked me to show her how to isolate RNA next Monday (2016-06-03), so I can isolate CJ8(a colony I thought I isolated RNA from, but I didn't). **It is in box 54**

<div id='id-section28'>

###Page 28: 2016-06-03. Papers showing differences between fast static vs slow dynamic temperature treatments.

There is a large argument in the literature about how to best temperature treat ectotherms. One thing to point out, many fruit fly studies plot their heat tolerance traits against latitude, why not against local temperatures (Tmax, MAT)?

Here are a list of papers that find no clinal variation for slow ramp, but do for fast static experiments.

1. [Castaneda et al. 2015; Evolution](#)
- 2.

[Sgro et al. 2010; Journal of Evolutionary Biology](#) shows complex patterns between slow, hardening, and fast heat shocks across latitude.

Our group has argued that different temperature treatments represent different aspects of their thermal biology. Meaning:

1. Fast heat shocks, whether dynamic or static = basal heat tolerance
 2. Slow heat shocks, whether dynamic or static = phenotypic plastic response in heat tolerance or acclimation or partial hardening response.
-

```
<div id='id-section29'>  
###Page 29: 2016-06-06. Isolating RNA: colony CJ8; showing Sylvia
```

Isolated RNA and converted to cDNA. [link to protocols](#)

Colony: CJ8

Samples:

1. CJ8 25
2. CJ8 28
3. CJ8 30
4. CJ8 31.5
5. CJ8 33
6. CJ8 35
7. CJ8 36.5
8. CJ8 38.5
9. CJ8 40
10. CJ8 41
11. CJ8 mid
12. CJ8 last
13. CJ8 25 -2
14. CJ8 41 -2

Results of RNA isolation: we have RNA, now convert 50ng to cDNA

N	Date	species	colony	box.condition	temp	Qubit_q
799	20160606	fulva	CJ8	box54	25	
800	20160606	fulva	CJ8	box54	28	
801	20160606	fulva	CJ8	box54	30	
802	20160606	fulva	CJ8	box54	31.5	

803	20160606	fulva	CJ8	box54	33	
804	20160606	fulva	CJ8	box54	35	
805	20160606	fulva	CJ8	box54	36.5	
806	20160606	fulva	CJ8	box54	38.5	
807	20160606	fulva	CJ8	box54	40	
808	20160606	fulva	CJ8	box54	41	
809	20160606	fulva	CJ8	box54	mid	
810	20160606	fulva	CJ8	box54	last	
811	20160606	fulva	CJ8	box54	25_2	
812	20160606	fulva	CJ8	box54	41_2	

Master mix for cDNA conversion

cDNA.synthesis	X1.rxn	X17.rxns
10xBuffer	2.0	34.0
dNTP	0.8	13.6
multiscribe RT	1.0	17.0
Rnase	1.0	17.0
Primer	2.0	34.0
H2O	3.2	54.4
total rxn	10.0	170.0

Steps:

1. Put pcr strip tubes on ice.
 2. Add h20 specified above table
 3. Add RNA specified above table. Final volume should be 10uL
 4. Aliquot 10 uL of master mix to all tubes.
 5. PCR, see protocol link at beginning of post.
-

```
<div id='id-section30'>  
###Page 30: 2016-06-07. **Brute force fitting nls function in R  
revisited ** Page 22: 2016-06-02
```

I tried this on my desktop to play with the data quick and dirty, but it should go in my dissertation repo:

So the main problem I had in the past was that nls would stop if it the fit was poor, nls2() will brute force fit curves.

Here is my mock dataset:

```
knitr::kable(m)
```

Colony	temp	FC_hsc70_1468	FC_Hsp83_279	FC_Hsp83_158
SHC6	25.0	0.8180765	1.2727190	1.374114

SHC6	28.0	0.8999074	1.3778736	2.307771
SHC6	30.0	0.7922560	0.9294879	1.139005
SHC6	31.5	0.8561583	1.1546421	0.867914
SHC6	33.0	3.3855425	1.9787656	1.854011
SHC6	35.0	7.1917199	2.5450325	3.500944
SHC6	36.5	19.4708137	3.4314556	4.363093
SHC6	38.5	30.8610304	4.2174121	6.758014
SHC6	40.0	32.5603639	4.6504188	7.540167
SHC6	41.0	26.0984907	2.8898872	NA
Avon	25.0	1.1732547	1.2784472	1.139045
Avon	28.0	1.4387152	1.5022087	1.333696
Avon	30.0	0.8752047	1.1583008	1.290241
Avon	31.5	1.1998622	1.0781117	1.213210
Avon	33.0	2.0881946	2.1492356	2.648233
Avon	35.0	6.7926522	NA	4.821989
Avon	36.5	10.7125651	4.0352515	5.500039
Avon	38.5	22.8261858	7.0736972	9.003823
Avon	40.0	NA	NA	NA
Avon	41.0	32.0884860	10.1245880	12.781207
KH7	25.0	0.8116304	0.7712080	0.930432
KH7	28.0	1.0896696	1.1911849	1.121952

KH7	30.0	1.1757139	1.2275952	1.340302
KH7	31.5	1.3429711	2.1066143	1.744253
KH7	33.0	3.7095882	3.2454970	2.812335
KH7	35.0	7.6945833	3.1906332	3.051582
KH7	36.5	19.6792961	7.5792950	5.624946
KH7	38.5	25.7475125	6.0603869	5.538655
KH7	40.0	47.1850131	12.1240032	9.799137
KH7	41.0	44.5367758	11.4567417	9.655169
test	25.0	1.0000000	10.0000000	5.000000
test	28.0	1.0000000	9.0000000	5.000000
test	30.0	2.0000000	8.0000000	5.000000
test	31.5	3.0000000	7.0000000	5.000000
test	33.0	4.0000000	6.0000000	5.000000
test	35.0	5.0000000	5.0000000	5.000000
test	36.5	6.0000000	4.0000000	5.000000
test	38.5	7.0000000	3.0000000	5.000000
test	40.0	8.0000000	2.0000000	5.000000
test	41.0	9.0000000	1.0000000	5.000000

1. Now I have to fit curves (boltzmann function) to for each colony and each gene (FC_70, FC_83, and FC_40). You can see I have a test colony with made up numbers, these should be poor fits.

I'm using `nls2()` and this curve estimates the critical temperature T_m , slope (a), and max expression

```
Boltz<-function(data=x){  
  B<-nls2(gxp ~ (1+(max-1)/(1+exp((Tm-T)/a))), data=data,  
  start=list(max=80, Tm=35, a=1.05), trace=TRUE, control=nls.c  
  ontrol(warnOnly = TRUE, tol = 1e-05, maxiter=1000))  
  #summary(B)  
  return(summary(B)$parameters)  
}
```

2. I'll need to convert it long format, it is in wide right now.

```
names(m)  
[1] "Colony"          "temp"           "FC_hsc70_1468" "FC_H  
sp83_279"  
[5] "FC_Hsp83_1583"   "FC_hsp40_424"   "T"  
>  
mlong<-gather(m,gene,gxp,FC_hsc70_1468:FC_hsp40_424)
```

3. fit for each colony and gene with ddply + Boltz functions

```
fits<-ddply(mlong,.Colony,gene),Boltz)  
fits<-cbind(fits,rep(c("max","Tm","slope"),length(fits$Co  
lony))) # adding parameter column  
names(fits)[7]<- "parameter"># renaming column
```

```
knitr::kable(fits)
```

```
<div id='id-section30.1'>  
###Won't fit with test colony
```

Trying fits by removing test colony

```
mlong<-subset(mlong,mlong$Colony!="test")  
fits<-ddply(mlong,.(Colony,gene),Boltz)
```

```
###Output table!
```

Colony	gene	Estimate	Std. Error	t value	p value
Avon	FC_hsc70_1468	35.8189402	1.3830780	25.897990	0.0000000
Avon	FC_hsc70_1468	37.7704625	0.1824726	206.992555	0.0000000
Avon	FC_hsc70_1468	1.5075619	0.1117296	13.492950	0.0000000
Avon	FC_Hsp83_279	13.0621490	1.7746986	7.360207	0.0000000
Avon	FC_Hsp83_279	38.5802879	0.7637267	50.515830	0.0000000
Avon	FC_Hsp83_279	2.1031077	0.3554831	5.916195	0.0000000
Avon	FC_Hsp83_1583	16.8751069	2.4307114	6.942456	0.0000000
Avon	FC_Hsp83_1583	38.4508017	0.9001894	42.714125	0.0000000
Avon	FC_Hsp83_1583	2.4352914	0.3611821	6.742558	0.0000000
Avon	FC_hsp40_424	21.9643380	12.1034762	1.814713	0.0000000

Avon	FC_hsp40_424	40.8933831	2.9441107	13.889893	(
Avon	FC_hsp40_424	2.6054162	0.6918408	3.765919	(
KH7	FC_hsc70_1468	57.0478157	12.0292674	4.742418	(
KH7	FC_hsc70_1468	38.2671391	1.0235944	37.385060	(
KH7	FC_hsc70_1468	1.7874874	0.5009719	3.568039	(
KH7	FC_Hsp83_279	18.8164697	14.2023236	1.324887	(
KH7	FC_Hsp83_279	39.5972751	5.0039209	7.913250	(
KH7	FC_Hsp83_279	2.9760831	1.4783205	2.013152	(
KH7	FC_Hsp83_1583	16.7337144	10.3102857	1.623012	(
KH7	FC_Hsp83_1583	40.1004665	4.0390733	9.928135	(
KH7	FC_Hsp83_1583	3.0388325	1.0845309	2.801979	(
KH7	FC_hsp40_424	19.9496194	14.9270787	1.336472	(
KH7	FC_hsp40_424	41.3533804	3.8900811	10.630467	(
KH7	FC_hsp40_424	2.6777066	0.8223794	3.256048	(
SHC6	FC_hsc70_1468	30.1357724	1.3518947	22.291509	(
SHC6	FC_hsc70_1468	36.0181917	0.2145002	167.916817	(
SHC6	FC_hsc70_1468	0.7601739	0.1966529	3.865562	(
SHC6	FC_Hsp83_279	3.9378751	0.3837209	10.262341	(
SHC6	FC_Hsp83_279	34.4580183	0.8580317	40.159376	(
SHC6	FC_Hsp83_279	1.2755059	0.6850160	1.862009	(
SHC6	FC_Hsp83_1583	8.6530046	1.6923497	5.113012	(

SHC6	FC_Hsp83_1583	36.6782852	1.1214736	32.705437	(
SHC6	FC_Hsp83_1583	1.8095631	0.6243422	2.898352	(
SHC6	FC_hsp40_424	8.3707957	1.0694746	7.827017	(
SHC6	FC_hsp40_424	35.6669753	0.9166608	38.909679	(
SHC6	FC_hsp40_424	1.8169999	0.6708063	2.708680	(

looks like it works when there is no poor fit.

<div id='id-section30.2'>

#Ok, I figured out how to suppress errors and let the function loop with [failwith\(\) function](#).

```
m<-read.csv("20160607_gxp_test.csv")
m$T<-m$temp
str(m)

#change to long format
mlong<-gather(m,gene,gxp,FC_hsc70_1468:FC_hsp40_424)
str(mlong)

#mlong<-subset(mlong,mlong$Colony!="test")
fits<-ddply(mlong,.Colony,gene),failwith(f=Boltz)) ## the magical code here
```

##Table of outputs

Colony	gene	Estimate	Std. Error	t value

Avon	FC_hsc70_1468	35.8189402	1.3830779	25.897991	(
Avon	FC_hsc70_1468	37.7704625	0.1824726	206.992559	(
Avon	FC_hsc70_1468	1.5075619	0.1117296	13.492950	(
Avon	FC_Hsp83_279	13.0621489	1.7746986	7.360207	(
Avon	FC_Hsp83_279	38.5802879	0.7637267	50.515830	(
Avon	FC_Hsp83_279	2.1031077	0.3554832	5.916195	(
Avon	FC_Hsp83_1583	16.8751071	2.4307113	6.942456	(
Avon	FC_Hsp83_1583	38.4508017	0.9001893	42.714127	(
Avon	FC_Hsp83_1583	2.4352914	0.3611821	6.742558	(
Avon	FC_hsp40_424	21.9649309	12.1044659	1.814614	(
Avon	FC_hsp40_424	40.8935313	2.9442708	13.889188	(
Avon	FC_hsp40_424	2.6054554	0.6918546	3.765900	(
KH7	FC_hsc70_1468	57.0473854	12.0288922	4.742530	(
KH7	FC_hsc70_1468	38.2671031	1.0235676	37.386005	(
KH7	FC_hsc70_1468	1.7874685	0.5009659	3.568045	(
KH7	FC_Hsp83_279	18.8160754	14.2013489	1.324950	(
KH7	FC_Hsp83_279	39.5971341	5.0036704	7.913618	(
KH7	FC_Hsp83_279	2.9760359	1.4782900	2.013161	(
KH7	FC_Hsp83_1583	16.7333374	10.3095588	1.623090	(
KH7	FC_Hsp83_1583	40.1003166	4.0388773	9.928580	(
KH7	FC_Hsp83_1583	3.0387896	1.0845105	2.801992	(

KH7	FC_hsp40_424	19.9504446	14.9288152	1.336372	(
KH7	FC_hsp40_424	41.3536013	3.8903675	10.629742	(
KH7	FC_hsp40_424	2.6777587	0.8223999	3.256030	(
Phil	FC_hsc70_1468	14.4816051	0.6238735	23.212404	(
Phil	FC_hsc70_1468	34.8148669	0.2209902	157.540295	(
Phil	FC_hsc70_1468	0.8480438	0.2387966	3.551322	(
Phil	FC_Hsp83_279	4.6238796	0.4489827	10.298570	(
Phil	FC_Hsp83_279	33.7411733	0.7422000	45.461025	(
Phil	FC_Hsp83_279	1.2133128	0.5981040	2.028598	(
Phil	FC_hsp40_424	4.3629872	0.2614315	16.688838	(
Phil	FC_hsp40_424	34.6387089	0.3401929	101.820776	(
Phil	FC_hsp40_424	0.7043699	0.3427897	2.054816	(
SHC6	FC_hsc70_1468	30.1357991	1.3519005	22.291433	(
SHC6	FC_hsc70_1468	36.0181969	0.2145014	167.915909	(
SHC6	FC_hsc70_1468	0.7601800	0.1966547	3.865558	(
SHC6	FC_Hsp83_279	3.9379010	0.3837369	10.261982	(
SHC6	FC_Hsp83_279	34.4580679	0.8580653	40.157863	(
SHC6	FC_Hsp83_279	1.2755764	0.6850461	1.862030	(
SHC6	FC_Hsp83_1583	8.6530046	1.6923498	5.113012	(
SHC6	FC_Hsp83_1583	36.6782851	1.1214737	32.705435	(
SHC6	FC_Hsp83_1583	1.8095631	0.6243422	2.898351	(

SHC6	FC_hsp40_424	8.3707958	1.0694747	7.827016	(
SHC6	FC_hsp40_424	35.6669753	0.9166608	38.909677	(
SHC6	FC_hsp40_424	1.8169999	0.6708063	2.708680	(
test	FC_hsc70_1468	9.8719349	0.9800918	10.072460	(
test	FC_hsc70_1468	35.6649510	0.8966939	39.773830	(
test	FC_hsc70_1468	2.9884380	0.4909301	6.087299	(
test	FC_hsp40_424	8.0828867	0.1090835	74.098170	(
test	FC_hsp40_424	30.3192228	0.1219349	248.650901	(
test	FC_hsp40_424	1.1145318	0.1136478	9.806893	(

###Notice:

That not all genes have fitted parameters! nice! ie. test hsp83's!

###Now we need to:

1. Predict new sets of values for each gene/colony
2. Visualize actual vs predicted values!

###Code to predict new values

- first, the plotting function

```
fud<-function(T=seq(25,70,.1),Tm=40,slope=1.8,max=50){
  y<-1+ (max-1)/(1+exp(((Tm-T)/slope)))
  return(y)
}
```

```
plot(fud())
```

- OK, now the data manipulation

```
#grab fitted lines from estimates
#change to wide format
library(reshape2)
feeder<-dcast(fits2,Colony+gene~parameter,value.var="Estimate")

list_predictions<-sapply(split(feeder,list(feeder$Colony,
feeder$gene)),function(x) {fud(T=seq(25,45,.1),Tm=x$Tm,slope=x$slope,max=x$max)})

predi<-as.data.frame(do.call("rbind", list_predictions),stringAsFactors=FALSE)
predi$Sample<-row.names(predi)

nom<-as.data.frame(matrix(unlist(strsplit(predi$Sample,"[. ]")),ncol=2,byrow=TRUE)) #messing with the names
names(nom)<-c("Colony","gene")
predictions<-cbind(predi,nom)
##gotta change to long format
conv<-gather(predictions,Colony,gxp,V1:V201)[,-4]
#need to sort
conv<-conv[order(conv$Sample),] #dont forget to order!!!
```

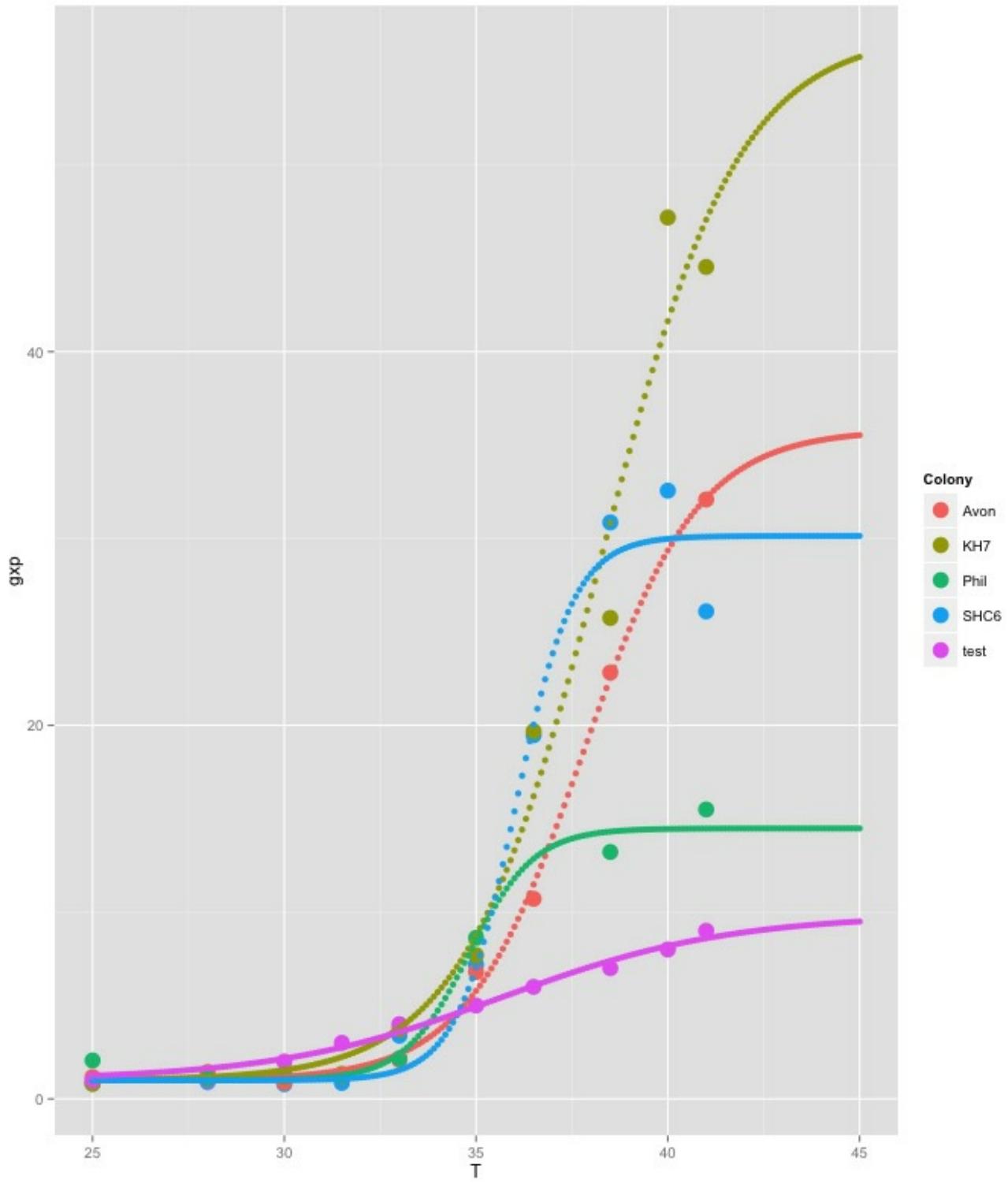
```
plong<-cbind(conv, rep(seq(25,45,.1), nrow(predi)))
names(plong)[5]<- "T"
head(plong)
```

##Plotting with ggplot

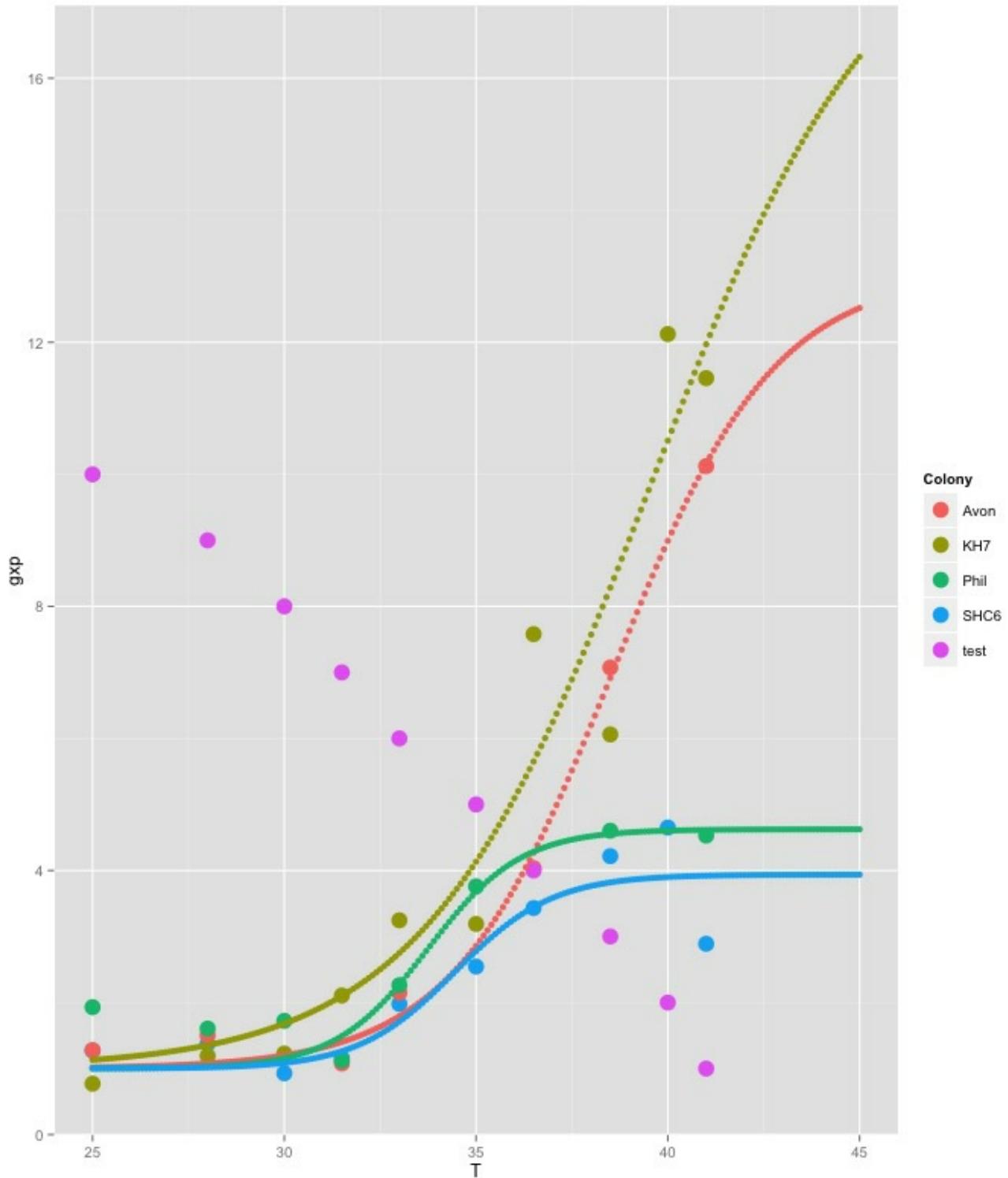
- **for hsc70-4 h2**

- lines = predicted fit from function
- points = empirical

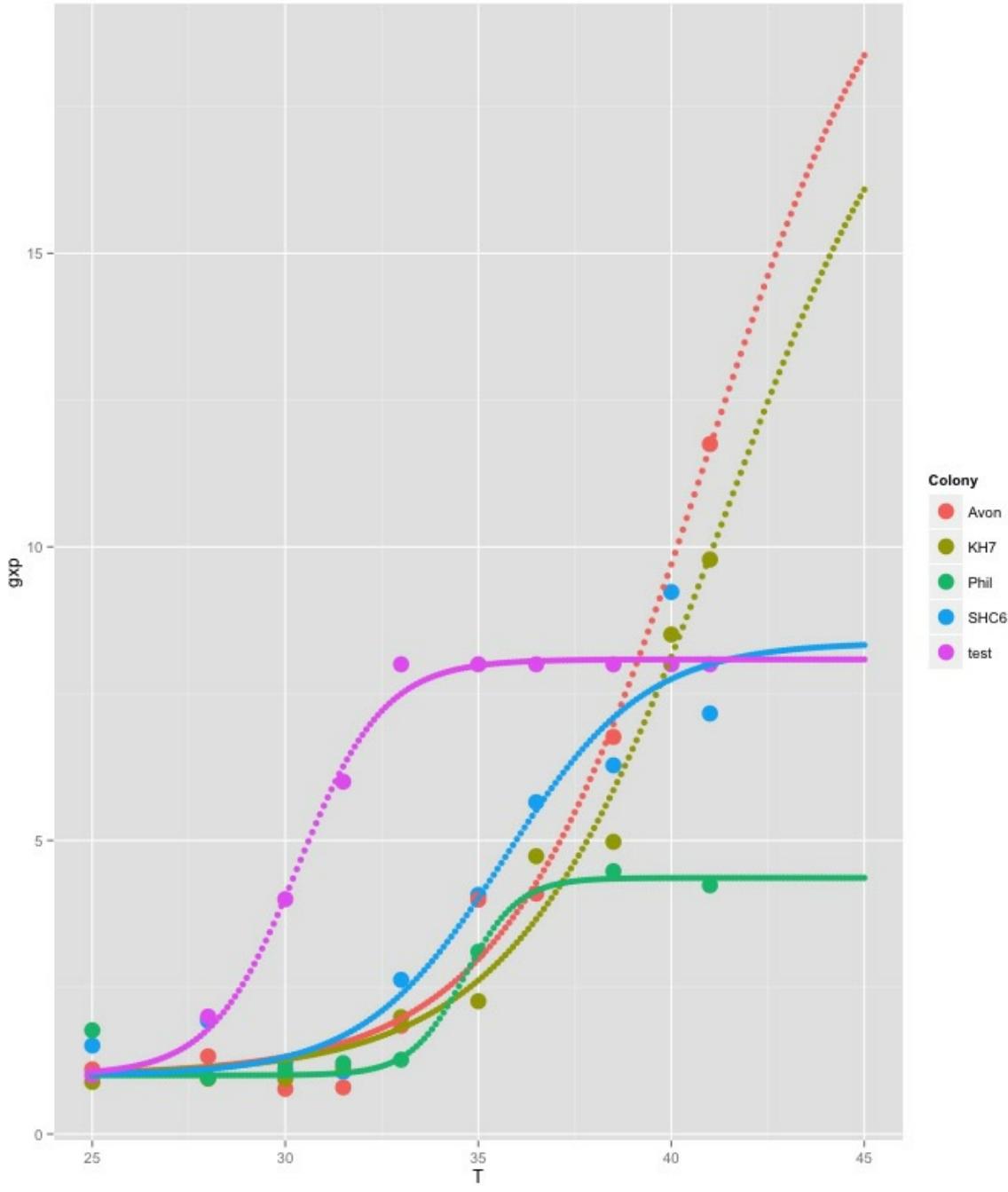
```
b<-subset(plong, plong$gene=="FC_hsc70_1468")
qplot(x=T, y=gxp, data=subset(mlong, mlong$gene=="FC_hsc70_1
468"), colour=Colony)+geom_point(size=5)+xlim(25,45)+geom_
point(aes(y=gxp, x=T, colour=Colony), data=b)
```



• hsp83 279



- **hsp40 541 **



<div id='id-section31'>

Page 31: 2016-06-08. Redoing online notebook template

I updated my online notebook template. I probably should have done this from the start. But there is a table of contents with 200 entries

with automatic links to those entries.

Code for automatically generating table of contents:

```
* [Page 1: Date](#id-section1). Title  
* [Page 2: Date](#id-section2). Title
```

For table of contents, you want this syntax:

1. I used R with a series of paste functions to get the right syntax
2. Exported to csv and just pasted it into the markdown

```
#constructing table of contents  
one<- rep("* [Page",200)  
two<- seq(1:200)  
  
three<- paste(one,two)  
four<- paste(three,":","]",sep="")  
  
five<- paste(four,"(#id-section",two,").",sep="")  
six<- data.frame(five)  
write.csv(six,"ffff.csv")
```

Code for automatically generating

entries with titles that correspond with table of contents

For this you want this syntax:

```
-----  
<div id='id-section1'>
```

1. R manipulations

```
b<-rep("-----",200)  
c<-rep("<div id='id-section",200)  
d<-seq(1:200)  
e<-paste(d,"'/>",sep="")  
  
m<-paste(c,e,sep="")  
m  
  
i<-rep("### Page",200)  
i2<-paste(i,rep(1:200))  
i3<-paste(i2,":",sep="") # can even add year here  
  
m1<-paste(b,m,i3,sep="  
")  
write.csv(m1,"testy.csv")
```

2. Export to csv
 3. You do need to get rid of header and first column manually, save and close (in excel)
 4. Open in textwrangler and you'll see that the line breaks appear. Then get rid of quotes.
-

<div id='id-section32'>

Page 32: 2016-06-08. qPCRs, 18s rRNA for Duke2, HF2, Kite 4, Kite8, 60 C annealing

Ran qpcr plate (96 well) on loaner ABI steponeplus. Samples were already 1/10 diluted, and for 18s, I dilute 1/10 again to have a 1/100 dilution.

Colonies:

1. Duke2
2. HF2
3. Kite4
4. Kite8

5. Made master mix: added 550 uL sybr green, 21 uL F+R primer, and 84 uL h20
6. Dispensed 6 uL into plate
7. Added 4 uL of cDNA (1/100 dilution) into plate
8. qPCR, 60 C annealing

Summary:

Single peaks from melt curve analysis indicating single amplicon. The threshold was set to 0.5.

###Updated summary of whole project so far:

Progress	X18s	hsc70.4_1468_1592_degen	hsp83_279_392_degen
works	59		51
double peaks	2		11
total	61		62

#Dilutions of future samples

Dilute 1/10: 5 uL of sample + 45 uL of h20 in 12 strip pcr tubes.

Sample colonies:

1. CJ2
 2. CJ5
 3. Duke1
 4. SHC8
-

<div id='id-section33'>

Page 33: 2016-06-08. Climate cascade meeting.

SHC can't make it. KM going to process samples. ANBE + NJG meet

1. Evolution poster: Go over figures and conclusions
2. Update gxp curve fitting
- 3.

NJG suggestions:

- For figure4, gray out points and put pretreatment temps beside each line.
 - Figure 3, plot hardening ability vs basal cold tolerance.
 -
-

<div id='id-section34'>

Page 34: 2016-06-09; 2016-06-10.

qPCRs: Duke1, CJ2, SHC8, CJ5

1. hsc70-4 h2 1468, 60C annealing results: only Duke1 worked
 2. hsc70-4 h2 1468, 55C annealing results: none worked
 3. hsp83 279 prim, 55C annealing results: all worked
 4. hsp40 541 prim, 55C annealing results: all worked , although some replicates excluded due to non-specificity
 5. 18s rRNA, 60 C annealing results: Samples were diluted 1/10.
-

<div id='id-section35'>

Page 35: 2016-06-10. ABI steponeplus machine fix and sending back instrument.

machine repaired

Dear Andrew,

The repair of your instrument on service reference notification 405638599 has been completed and is now on its way back to you. For your record the reference tracking number is 650686939762 I will be sending you a separate email with the decontamination forms and FedEx labels to return the loaner you received during the repair of your instrument. Please send this loaner back in a

timely fashion as we do have other customers in need of this loaner.

Thank you,

Leticia C.

Instrument Services

Life Sciences Solutions

Sending back loaner

Dear Andrew,

Attached you will find the necessary paperwork to ensure that the loaner unit is returned correctly and promptly.

1. Your RMA is 14635-69
2. Please review and complete the attached decontamination form and print out 2 copies.
3. Please remember to place the instrument in the “Ship Prep” position prior to packing the instrument.
4. Please DO NOT include your power cord with your instrument (remove from unit and keep it).
5. Please DO NOT include any consumables (trays, tubes, etc.).
6. Place a copy of the completed decontamination form INSIDE and OUTSIDE of the box.
7. Print out the FedEx label, (link will arrive via separate email).

The return transaction cannot be processed until the completed decontamination form and the instrument are received.

Thank you,

Leticia C.
Instrument Services
Life Sciences Solutions

<div id='id-section35'>

2016-06-13 update

We received the repaired machine back.

Here is the decomtamination form for the loaner.

<div id='id-section36'>

Page 36: 2016-06-10. Thoughts on Kingsolver & Woods 2016, AmNat. ref here

reference:

- Kingsolver JG, Woods HA. 2016. Beyond Thermal Performance Curves: Modeling Time-Dependent Effects of Thermal Stress on Ectotherm Growth Rates. *The American Naturalist* 187:283-294.

This paper models growth rate under heat stress over time. The authors use Hsp gene and protein expression as a measure of cost and

ingestion rate as a trait that inputs energy into an animal.

Fig 1:

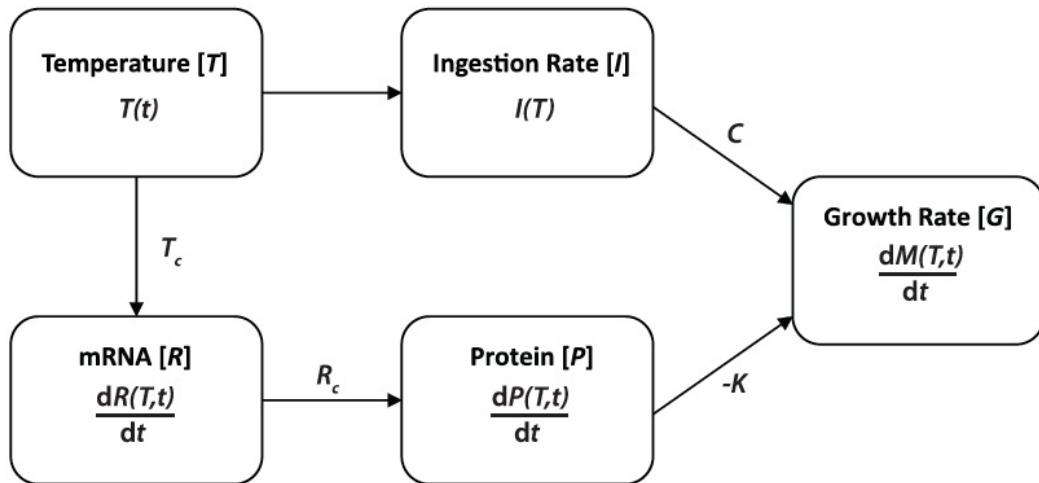
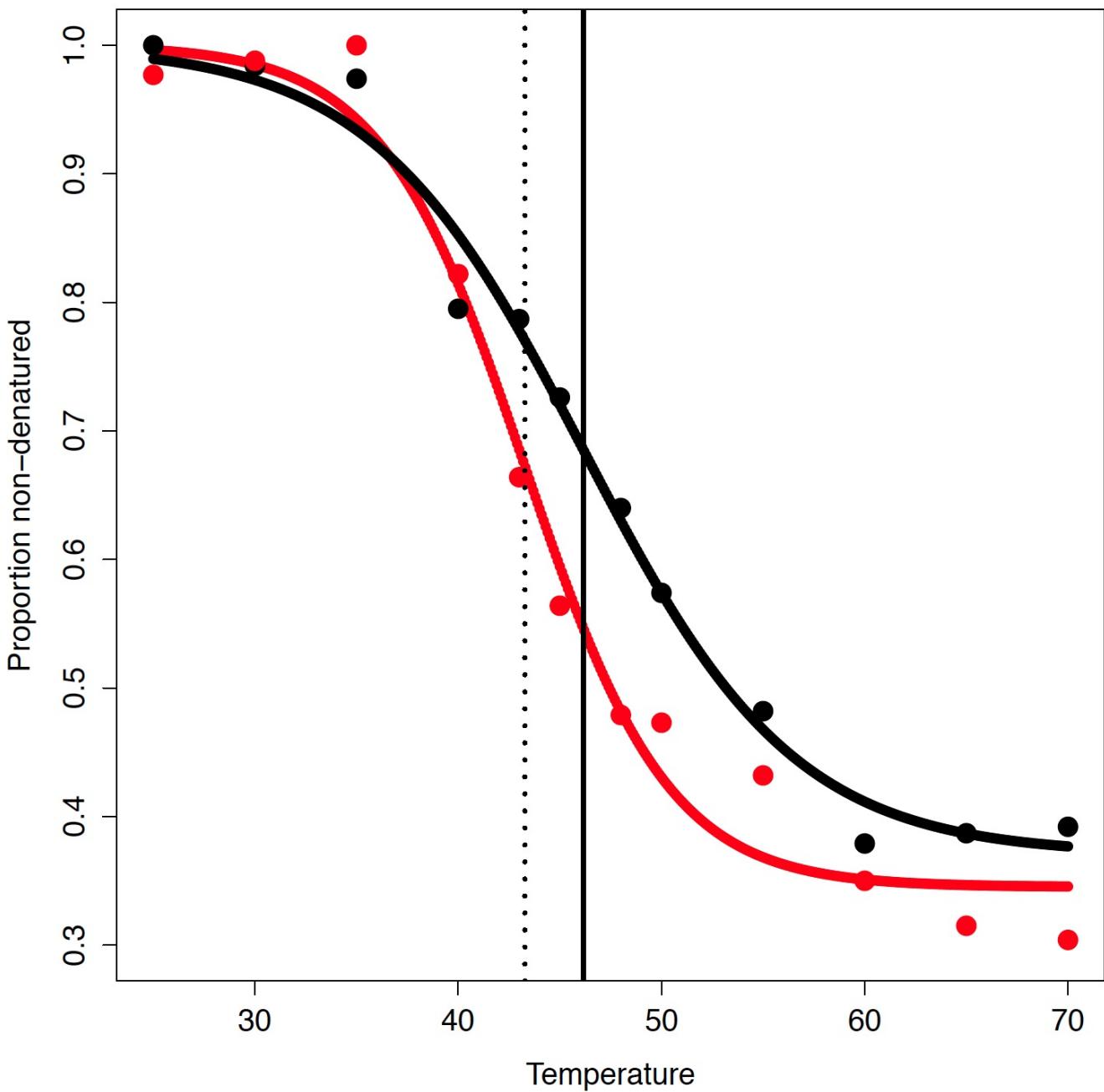


Figure 1: Diagram of the main components of the model. See text and equations (1)–(4).

The physiology is more complicated than this. First, increasing Hsp gene expression is costly in itself, so there should be a separate cost term. While the actual Hsp protein expression is costly to invest into too, there is a cost for using them and also having unstable proteomes. Also, organisms can get rid of unstable proteins through degradation and halting translation which would offset the costs of Hsp (gene or protein) expression and using it. Basically, I'm saying the actual cost incurred come in the form of macromolecular damage (proteome stability) and the response to macromolecular damage (Hsp expression). Not sure if proteome stability cost needs to included

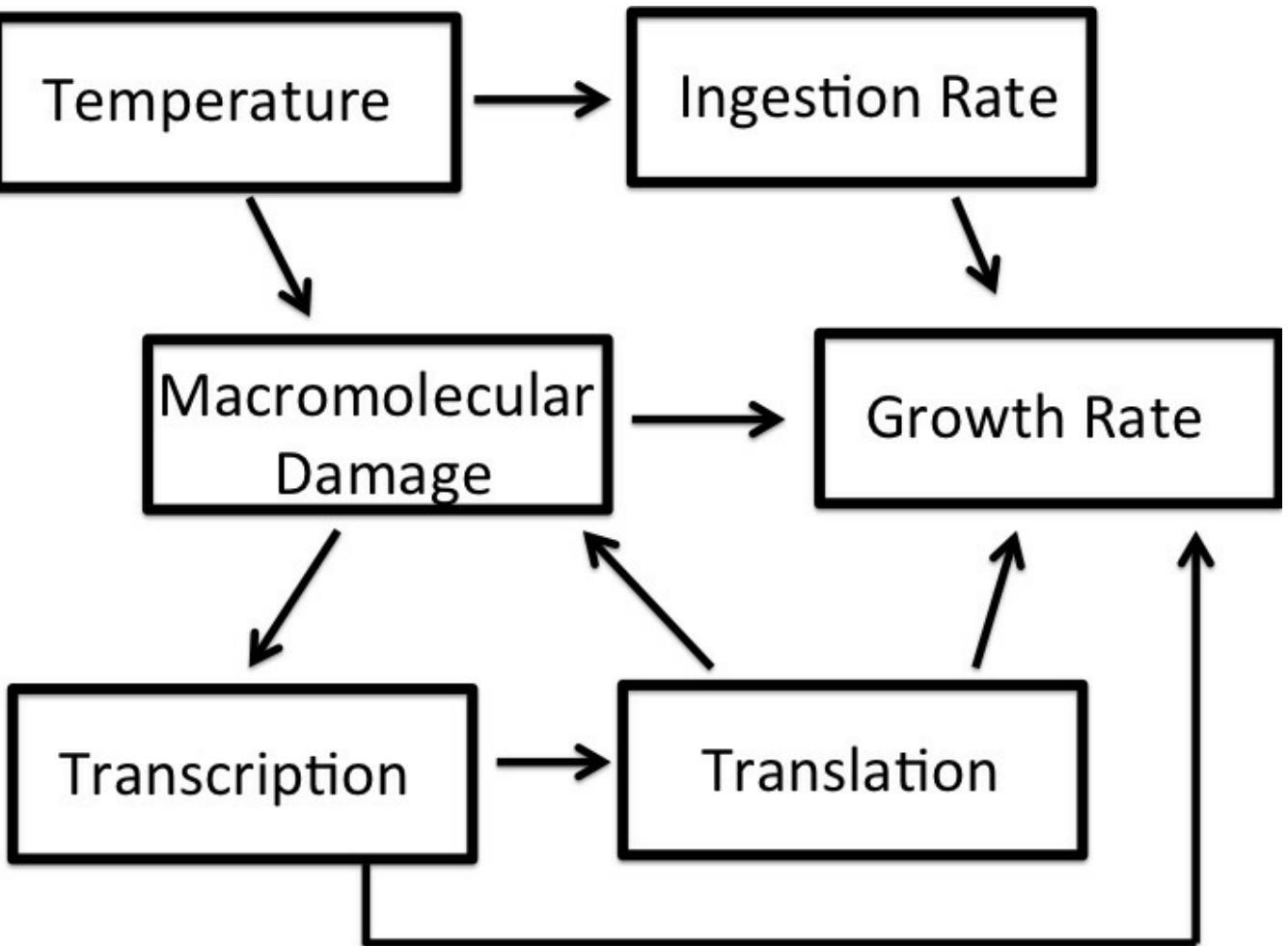
But here is a fig for proteome stability (prop non-denatured) as a function of Temperature:



The black line is 10 min incubations, the red line is 20 min. I fit a non-linear logistic curve to it [link](#). This captures the incurring costs associated with temperature AND time without an acclimation response. It'd be interesting to develop a model from this...

<div id='id-section36.5'>

2016-06-11. Follow up model



I've included potentially important physiological components.

Macromolecular damage includes unstable proteins and damage to membranes. For simplicity, it can just represent unstable proteins. On second thought, it should be macromolecular stability, assuming there is an optimal stability of membranes and proteins for growth. So temperature directly affects macromolecular stability and given a certain amount of damage(instability), it elicits a physiological response (transcription + translation) . Transcription includes all the transcripts that turn on and turn off. If the net effect is using more energy to turn on/off over higher temperatures, this incurs a cost. Same with translation, but there is also a cost of “using” the proteins. For example, Hsp mediated folding uses ATP. However, the

combination of altering translation rates and using the proteins offsets the costs of macromolecular damage which directly affects growth.

Anyway, I'd call this the "thermostat" model.

- Craig EA, Gross CA. 1991. Is hsp70 the cellular thermometer? Trends in Biochemical Sciences 16:135-140.

<div id='id-section36.6'>

2016-06-13. Predictions of thermostat model

1. There is some temperature where the costs associated with macromolecular damage exceeds any type of physiological response (transcription, translation), resulting in inhibited growth.
2. Under sublethal temperature stress, the negative long term outcome of inducing a physiological response may be tempered by increasing ingestion rate.
3. Although a physiological response is costly, at sublethal levels, the combination of gene/protein expression(downregulating unstable proteins) and upreg of Hsps may have a net positive effect on proteome stability, which is related to growth.

Note: There is a cool paper by Hoekstra & Montooth that shows how Hsp70 expression covaries with metabolic rate.

- Hoekstra LA, Montooth KL. 2013. Inducing extra copies of the Hsp70 gene in *Drosophila melanogaster* increases energetic demand. *BMC Evolutionary Biology* 13:68.

Other thoughts:

1. One cool thing about the model is that you can add transcriptome, proteome data as parameters into the model. How?
- Count the costs of each transcript (# of basepairs) and subtract response from baseline to get relative response. One could argue that overall, Hsp expression is not costly because other transcripts can be downregulated at the same time. I don't think anybody has tried to explore this in transcriptome datasets.
2. In aquatic systems, oxygen limitation seems to be the mechanism for upper thermal limits. Is there a way to make one global model so that we can make predictions for any ectotherm?

<div id='id-section37'>

**Page 37: 2016-06-11. Quantifying
natural selection in natural
populations**

I've been reading more Kingsolver (specifically [Kingsolver et al. 2001](#); [Kingsolver & Diamond 2011](#)) which led me to think about quantifying modes of selection in nature. Basically all you need to do is regress traits against relative fitness (fitness of individual / mean fitness of population). The slope is the magnitude of direction selection. Also, if you want to detect disruptive or stabilizing, then you can add a quadratic term. It'd be interesting to apply this technique to assess the fitness consequences of climate change. So take a species and measure fitness and traits along a transect to pick up the warm-edge, core, and cool-edge populations.

Refs for me to read:

- [Lande & Arnold 1983](#)
 - [Arnold & Ward 1984](#) for measuring natural and sexual selection
 - [Mitchell-Olds & Shaw 1987](#) for pitfalls in these regressions analyses.
 - [Heisler & Damuth 1987](#) for multilevel selection and it also introduces contextual analyses.
-

<div id='id-section38'>

Page 38: 2016-06-13. qPCR update for Duke1,CJ2,SHC8,CJ5. Randomizing samples treated at 25C(reference for

basal expression) for qpcrs.

Running qpcr for Duke1/CJ2/SHC8/CJ5; hsc70-4 h2 50C annealing.

Randomizing procedure

- Load data set as csv in R
- Code for sampling randomly:

```
write.csv(sample(d$colonies), "ra.csv")
```

- Changed csv so that I have rows and columns
- Here is the layout:

Row	Column	Colony
A	1	Ala1
A	2	KITE8
A	3	Yates2
A	4	FBRAGG3
A	5	CJ4
A	6	BK
A	7	HW7
A	8	KH3
A	9	DUKE9

A	10	SHC8
A	11	CJ2
A	12	HF2
B	1	shc7
B	2	MA
B	3	PB07-23
B	4	CJ8
B	5	Lex9
B	6	ApGxL10A
B	7	Phillips
B	8	hf3
B	9	PB17-10
B	10	CJ6
B	11	Ala4
B	12	CJ5
C	1	PB17-14
C	2	DUKE8
C	3	KH1
C	4	Greenfield
C	5	fbragg1
C	6	Avon19.1

C	7	CampNSP
C	8	KH6
C	9	KH5
C	10	DUKE2
C	11	SHC9
C	12	LPR2
D	1	KITE4
D	2	FBRAGG4
D	3	KH7
D	4	DUKE1
D	5	PMBE
D	6	DUKE6
D	7	CJ7
D	8	fbragg5
D	9	CJ1
D	10	LPR4
D	11	YATES3
D	12	POP1
E	1	kh2
E	2	Bingham
E	3	SHC3

E	4	ApGxL09A
E	5	Ted6
E	6	DUKE7
E	7	SHC6
E	8	DUKE4
E	9	DUKE5
E	10	Ted4
E	11	EXIT65
E	12	sidewalk (formica)
F	1	POP2
F	2	fbragg2
F	3	SHC2
F	4	LEX13
F	5	SHC5
F	6	cremat
F	7	SHC10
F	8	pop3
F	9	SR45
F	10	AS4

I'll arrange these samples in rows of 12 in pcr strip tubes, dilute 1/10 and then I can multichannel the samples into a 96 well qpcr plate.

<div id='id-section39'>

Page 39: 2016-06-13. Post doc project idea: Assessing current impacts of climate change in natural populations.

Alternate title: Quantifying the intensity of selection associated with climate change.

Question: Are populations experiencing selection associated with climate change out in nature?

Hypothesis: The magnitude and direction of selection acts on different parts of their range depending on their thermal environment.

Predictions:

1. Individuals at the warm edge of their range experience positive directional selection for a thermal trait.
2. Individuals at the core experience stabilizing selection for a thermal trait.
3. Individuals at the cool edge experience negative directional selection for a thermal trait.

Approach: Measure phenotypic selection on physiological, behavioral traits across a cline for a given species. A good system to measure phenotypic selection are ants because alates are direct measurement

of fitness. So the product of # of alates by their weights will give a measurement of fitness. Then, regress different traits on relative fitness to obtain a selection gradient. I can detect disruptive and stabilizing selection by adding a quadratic term in the regression model. I don't want to automatically assign individuals to warm-edge, cool edge, core. I'd sample along a cline (10-20 sites?). Also, there may be differences in the phenology for alates to develop, so I'd probably need to sample 3-4 times a year?

Some key traits:

1. Colony size (# of workers, # of larvae, # of pupae, Colony biomass really)
2. Thermal tolerance (CTmax, Ctmin, KO-time, hardening ability)
3. Morphology (leg length, average worker weight)

Some things to think about:

1. I read somewhere (find it) that what one really wants is the life time reproductive success (LRS). But this is almost impossible to measure. In this sense, it is more accurate to say I'm measuring episodic selection (Angiletta 2009)?
2. Also, one should be comparing within a generation. There may be different age classes of colonies, but it may be reasonable to assume that if the colony has alates, then they belong to a similar age class.
3. I'd need to do some pop gen to determine the population level structure so that I can empirically assign individuals to

populations.

Another thought: Phenotypic selection seems like a good way to associate higher and lower phenotypic levels.

1. For example, I have CTmax data and the underlying stress response measured. CTmax is a component of fitness, so if I regress the stress response onto the relative fitness of $\text{CTmax}(\text{CTmax of individual/ population CTmax mean})$, then I can determine a selection gradient.
 2. I can also measure phenotypic selection for allele frequencies!
(Dr. Goodnight's suggestion)
-

<div id='id-section40'>

Page 40: 2016-06-14. qPCR's: Diluting samples for quantifying basal expression and repeats

Diluting samples for basal expression:

I diluted 1x cDNA samples 1:10, so I added 5 uL cDNA with 45 uL water. I added 25C-mid samples (because of technical mistake in diluting) for some colonies to replace 25C samples that were started at the beginning of heat shock.

1. F10: Duke8 41 (switched with AS4)

2. F11: SHC10 mid
3. F12: AS4 25C
4. G1: yates3 mid
5. G2: shc2 mid
6. G3: exit65 mid
7. G4: greenfield mid

I also diluted the 1:10 cDNA samples again at 1:10 to run 18s rRNA.
So I added 2 uL cDNA into 18 uL water.

All in all, it took ~ 3 hours from organization to completion.

<div id='id-section40.1'>

Updated plate layout:

Row	Column	Colony
A	1	Ala1
A	2	KITE8
A	3	Yates2
A	4	FBRAGG3
A	5	CJ4
A	6	BK
A	7	HW7
A	8	KH3

A	9	DUKE9
A	10	SHC8
A	11	CJ2
A	12	HF2
B	1	shc7
B	2	MA
B	3	PB07-23
B	4	CJ8
B	5	Lex9
B	6	ApGxL10A
B	7	Phillips
B	8	hf3
B	9	PB17-10
B	10	CJ6
B	11	Ala4
B	12	CJ5
C	1	PB17-14
C	2	DUKE8
C	3	KH1
C	4	Greenfield
C	5	fbragg1

C	6	Avon19.1
C	7	CampNSP
C	8	KH6
C	9	KH5
C	10	DUKE2
C	11	SHC9
C	12	LPR2
D	1	KITE4
D	2	FBRAGG4
D	3	KH7
D	4	DUKE1
D	5	PMBE
D	6	DUKE6
D	7	CJ7
D	8	fbragg5
D	9	CJ1
D	10	LPR4
D	11	YATES3
D	12	POP1
E	1	kh2
E	2	Bingham

E	3	SHC3
E	4	ApGxL09A
E	5	Ted6
E	6	DUKE7
E	7	SHC6
E	8	DUKE4
E	9	DUKE5
E	10	Ted4
E	11	EXIT65
E	12	sidewalk (formica)
F	1	POP2
F	2	fbragg2
F	3	SHC2
F	4	LEX13
F	5	SHC5
F	6	cremat
F	7	SHC10
F	8	pop3
F	9	SR45
F	10	Duke 8 41
F	11	SHC10 mid

F	12	AS4
G	1	yates3 mid
G	2	shc2 mid
G	3	exit65 mid
G	4	gf mid

Repeats ran alongside CJ8

Ran hsp83 279 55 C annealing for following coloines:

1. Fbragg1
2. CJ1
3. CJ8
4. KH1; 1 row
5. FB4; 1 row

results: Fb4 not work

Ran hsp40 541 prim 55C annealing for the same colonies as above.

results: CJ8 and KH1 worked

Ran 18s rRNA for following colonies:

1. CJ1
2. CJ8
3. KH1

results: all worked

<div id='id-section40.5'>

#Update of samples:

Status	X18s	hsc70.4_1468_1592_degen	hsp83_279_392_dege
works	67	58	6
double peaks	0	9	
total	67	67	6

<div id='id-section41'>

Page 41: 2016-06-15. qPCRs to quantify basal expression. (Evolution of stress response project)

I probably should have mentioned this earlier, but since all the samples are on 1 plate, I'll be quantifying 4 genes in a replicated randomized block design.

So for each gene, run 2 plates. Samples on the plate were already previously randomized.

1. **Ran** 18s rRNA plate 1, 55 C annealing temp.

2. **Ran** hsc70-4 h2 1468 plate 1, 55 C annealing temp.
 3. **Ran** hsp83 279 plate 1, 55 C annealing temp.
 4. **Ran** hsp40 541 plate 1, 55 C annealing temp.
-

<div id='id-section42'>

Page 42: 2016-06-15. Evolution talks I want to attend.

Not a comprehensive list, but a start.

Day	Speaker	Room	Time	Title
Monday, June 20	Tangwancheroen, Sumaetee	MR10C	1:30PM	Adaptation via divergence in gene regulation along a temperature cline: cis and trans effects on HSP expression the copepod <i>Tigriopus californicus</i>
Monday, June 20	Lyons,Marta	BallroomC	2:00PM	Predicting range contractions in niche conserved plethodontid salamanders comparing correlative and biophysical

				niche models
Saturday, June 18	Gilbert, Kimberly	MR6B	1:30PM	Local maladaptation interacts with expansion load during species range expansions
Saturday, June 18	Kingsolver,Joel	BallroomC	9:15AM	Elevational clines in plastic and evolutionary responses of montane butterflies to climate change
Sunday, June 19	Nunney,Leonard	MR9AB	2:45PM	Adapting to a changing environment: modeling the interaction of directional evolution and plasticity
Sunday, June 19	Muir,Chris	BallroomA	8:30AM	What is evolutionary physiology?
Sunday, June 19	Garcia,Matteo	MR7	9:00AM	Performance determines division of labor in leafcutting ants

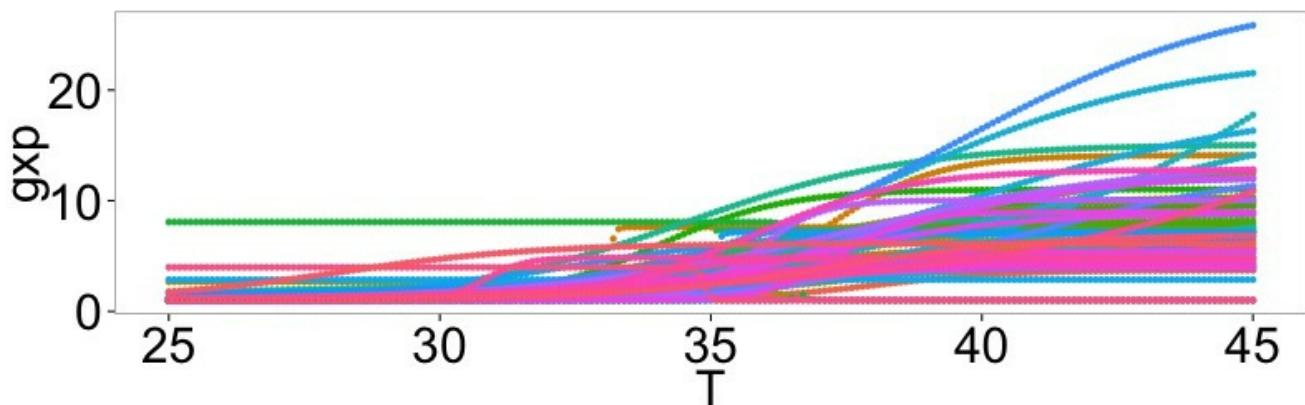
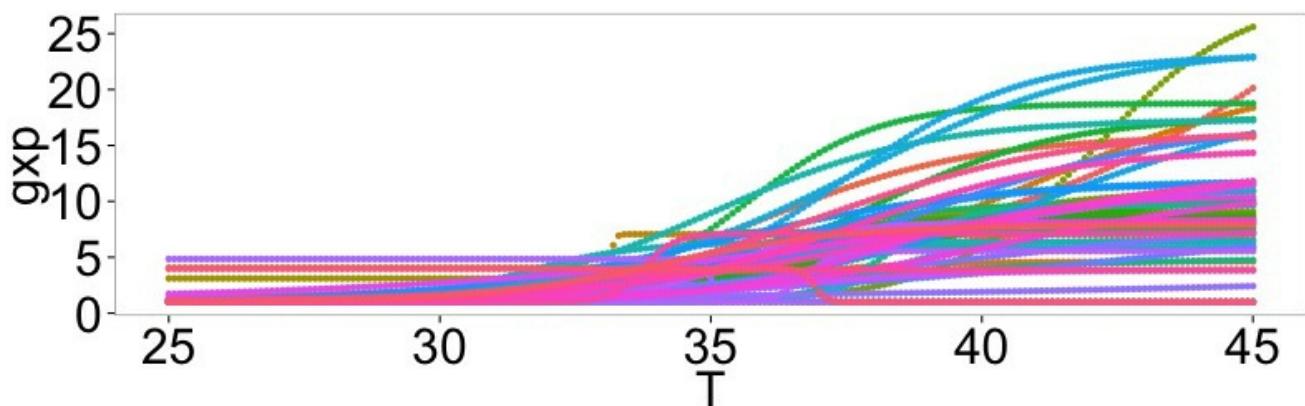
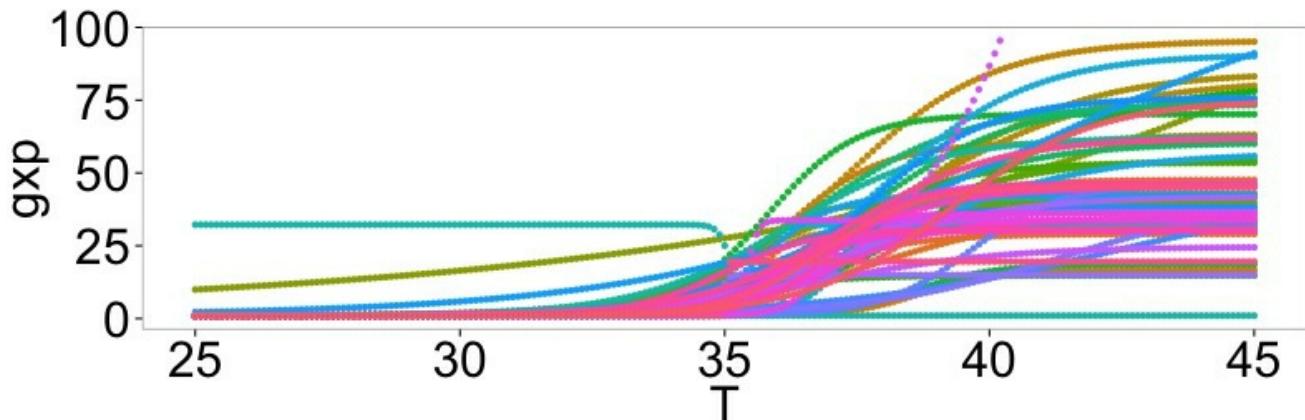
Sunday, June 19	Campbell Staton, Shane	MR9C	9:15AM	Polar Vortex cold wave elicits rapid physiological, regulatory and genetic shifts in populations of the green anole <i>Anolis</i> <i>carolinensis</i>
Sunday, June 19	Fumagalli, Sarah	MR7	9:30AM	The evolution of cooperation between unrelated individuals
Sunday, June 19	Catullo,Renee	BallroomC	10:15AM	Extending spatial modelling of climate change responses beyond the realized niche: estimating, and accommodating physiological limits and adaptive evolution
Sunday, June 19	Powell,Scott	MR9AB	10:15AM	Diversification of complex social phenotypes: insights from the turtle ants
Sunday,	Sexton, Jason	MR6A	10:45AM	Does species

June 19				niche breadth predict plant performance in novel environments? An experimental test in Australian Alps plants
Sunday, June 19	Rosauer,Dan	BallroomC	10:45AM	Distribution models below species level
Sunday, June 19	Chau,Linh	MR7	10:45AM	Gene Duplication in the Evolution of Sex- and Caste-biased Gene Expression in Social Insects
Sunday, June 19	Gunderson,Alex	BallroomA	11:00AM	The physiology of adaptive radiation
Sunday, June 19	Angert,Amy	BallroomA	11:15AM	Linking physiology to biogeography in monkeyflowers
Sunday, June 19	Parker,Joseph	MR9AB	11:15AM	An inordinate fondness for rove beetles: evolution and diversification of ant social parasites

```
<div id='id-section43'>
```

```
###Page 43: 2016-06-16. Figure for curve fitting: see Success with failwithO and Status update of samples.
```

Hsp70, 40, 83 from top to bottom



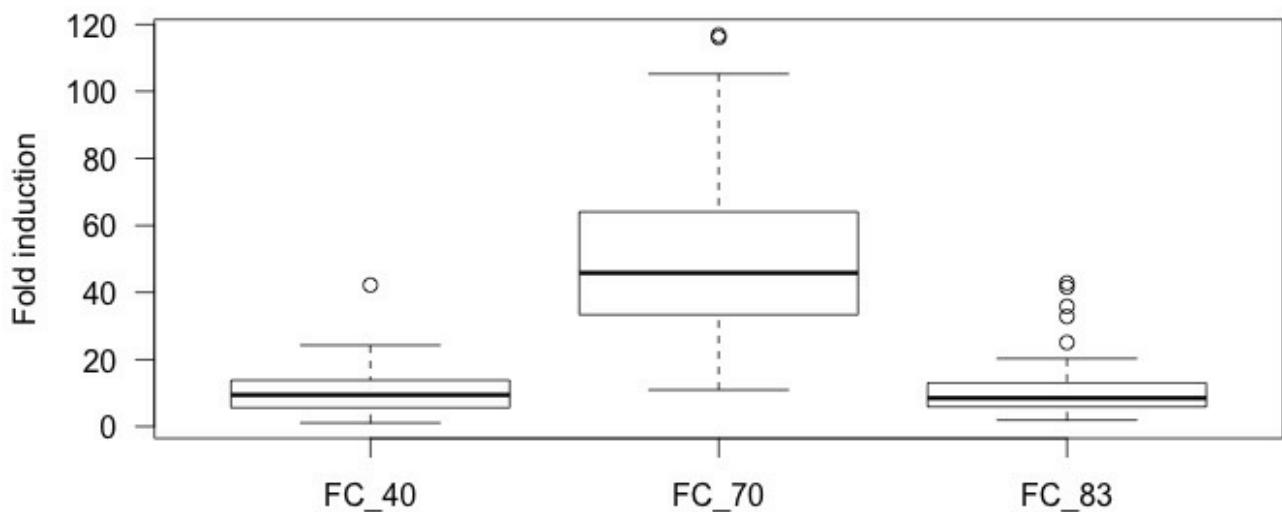
```
<div id='id-section44'>
```

Page 44: 2016-07-18. Summary statistics for modulation of Hsp paper.

Overall means

	mean xp
FC_83	11.218868
FC_70	50.227915
FC_40	10.535062
B_83	1.735492
B_70	1.446917
B_40	1.935067

Comparison among genes



medians

Rearing_Temp	Induction83	Basal83	Induction70	Basal70
20	7.046216	0.9032384	48.88187	0.477379
26	10.441149	1.5197949	39.13139	2.231829

means

Rearing_Temp	Induction83	Basal83	Induction70	Basal70
20	9.352522	1.262254	55.45230	0.640272
26	14.320365	2.334319	42.62233	2.565149

<div id='id-section45'>

Page 45: 2016-07-19. Meeting with VGN proteomics facility

Meeting with Wai and Bethany to finish up the comparative proteomics project (Amanda was working on this).

I went over our experimental protocol. Wai suggested to do searches with MASCOT and SEQUEST to ID more proteins.

Timeline:

- Next week for TMT labelling
 - First week of August for sending me a dataset
-

<div id='id-section46'>

Page 46: 2016-07-21. Reference samples for mapping index; Hsp modulation and thermal niche paper.

From SHC:

- FMU4 (ApGxL-03A)
- WP9 (ApGxL-11A)
- BRF4 (ApGxL-16A)
- SEB9 (ApGxL-22C)

- MB6 (ApGxL-26E)
-

<div id='id-section47'>

###Page 47: 2016-07-26. Learning mixed effects stat models

Mixed effects stat models let you include random or fixed variables, implemented in (lme4 package) (<http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf>). The difference? Summarized [here](#) in dynamic ecology blog.

As I understand it:

(Using sites as an example...)

Fixed effect...

- variable you're interested in
- continuous or categorical
- estimates values at each site, so if you have a lot of sites, it'll use more degrees of freedom
- syntax: $y \sim x + s$

Random effect...

- variable you want to control (blocking)
- categorical/discrete (**Can not have continuous variable as a random effect**)
- estimates variance among all sites, conserves degrees of freedom (also cant calculate p values)

- syntax: $(y_{x,\text{random}=1|s})$
- rule of thumb: sites should have roughly >5 levels (5 sites)
- comment in blog post says you can think of RE as groups having different slopes and or intercepts

Typing this out seems to make more sense. Now to go over some of the syntax...

- [see this](#)
- [and this](#)

This [tutorial](#) gives a good explanation.

It's hard to get p-values from mixed effects models, so one strategy is to make a full and null model with and without the variable of interest and running an anova. **Don't use REML** when doing these comparisons.

More syntax...

```
politeness.model = lmer(frequency ~ attitude + gender + 
  1|subject) + (1|scenario), data=politeness)
```

This syntax (1|variable) specifies subject and scenario as random effects. **It is a random intercept model.**

This specifies a random slope model:

```
politeness.model = lmer(frequency ~ attitude + gender + (1+attitude|subject) + (1+attitude|scenario),  
data=politeness,REML=FALSE)
```

This allows subjects and items to have difference slopes and intercepts.

Only thing changed is the random effect

Best practice to fit random slopes and intercepts! (Grueber et al. 2011, Journal of Evolutionary Biology; and the tutorial advocates for this because it reduces type I and II errors)

Notes, assumptions similar to fixed effects models

1. Check for collinearity and influential data points
2. check residuals, Q-Qplots
3. One of the main shifts from linear models to mixed effect models was to account for non-independence (measuring outcome of same individual)

random effects note

So, a random effect is generally something that can be expected to have a nonsystematic, idiosyncratic, unpredictable, or “random” influence on your data. In experiments, that’s often “subject” and “item”, and you generally want to

generalize over the idiosyncrasies of individual subjects and items.

fixed effects note

Fixed effects on the other hand are expected to have a systematic and predictable influence on your data.

Writing this up in a methods section

We used R (R Core Team, 2012) and lme4 (Bates, Maechler & Bolker,

2012) to perform a linear mixed effects analysis of the relationship

between pitch and politeness. As fixed effects, we entered politeness and

gender (without interaction term) into the model. As random effects, we

had intercepts for subjects and items, as well as by-subject and by-item

random slopes for the effect of politeness. Visual inspection of residual

plots did not reveal any obvious deviations from homoscedasticity or

normality. P-values were obtained by likelihood ratio tests of the full

model with the effect in question against the model without the effect in question.

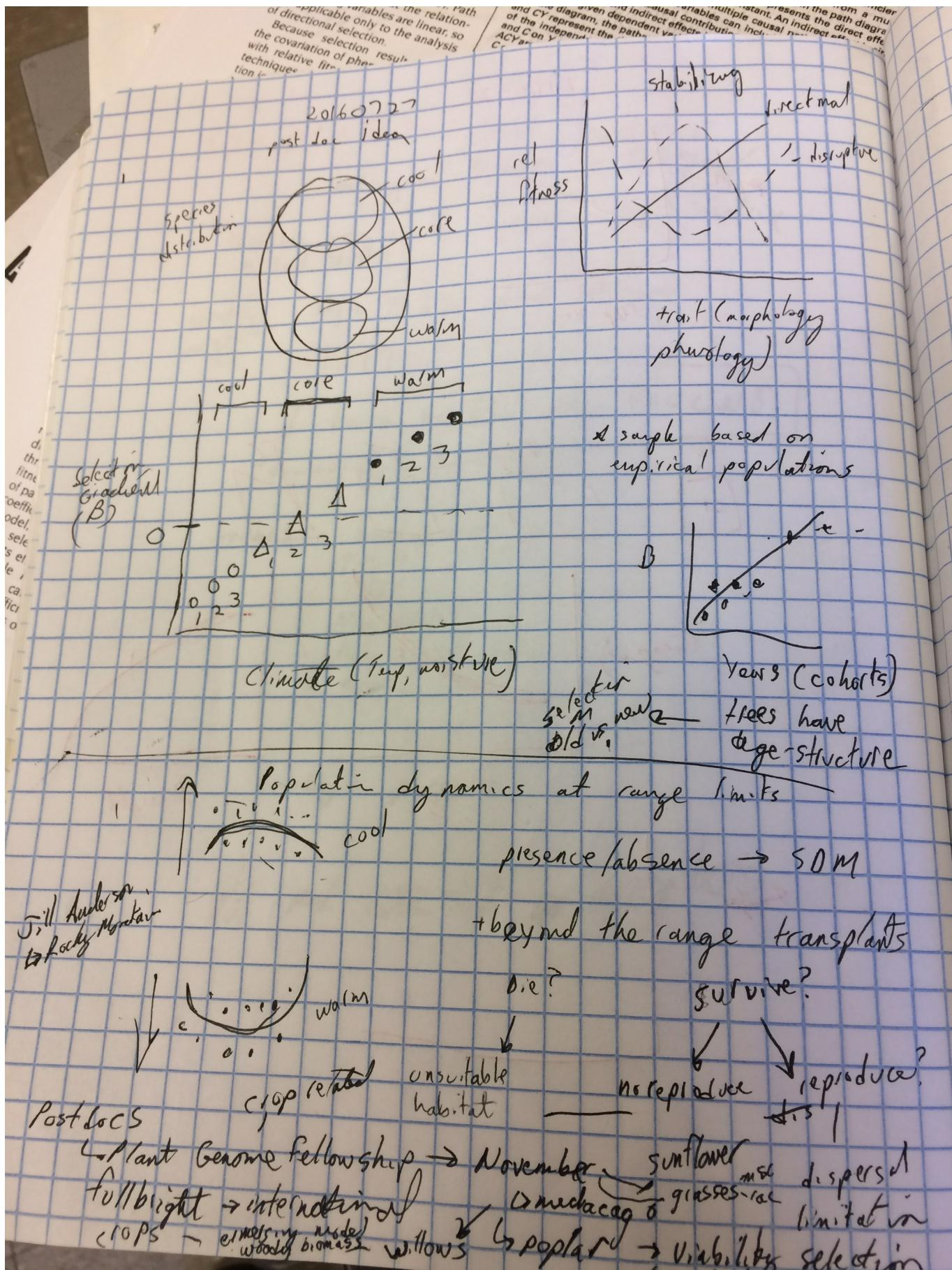
<div id='id-section48'>

Page 48: 2016-07-27. Meeting with Steve Keller to discuss post doc idea (started here: [Page 37: 2016-06-11.](#)

Quantifying natural selection in natural populations)

Raw notes from notebook:

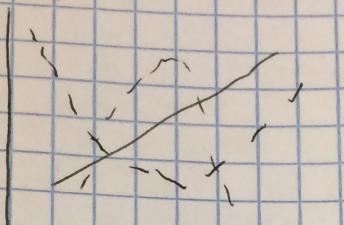
Page 1



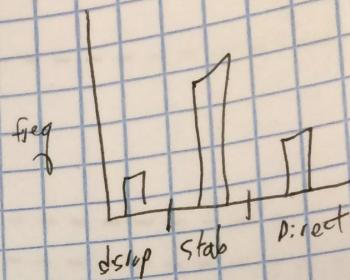
Variables between measured variables may be obtained for each dependent variable in the diagram below represents the variables hold constant. An indirect effect can include multiple causal paths. Total contributions to effects is the sum of all paths.

Identifying adaptive loci or alleles

direct
disruptive



Allele frequency



Julie Ellerson

QG of climate & phenotypic selection

do manipulate

manipulation
phenotypic
env

start cohorts
at diff times

control development

pollinator

floral display

Induce herbivory

plant stress responses

challenges → climate change

are diff across

photo period ~~are the same~~ are diff

satisfied

manipulate temp, keep photoperiod constant

Duant genetics in natural pop w/o breeding design

(Falconer McKay) h^2 , can be estimated based on trait matrix

estimated

Thoughts+ retyping notes:

1. One challenge Steve brought up was that photoperiod is different across latitudes and is not changing with climate. So when scientists do reciprocal transplants between north and south populations, photoperiod is a confounding effect with temperature/climate.
2. Selection gradients may not be increasing with climate if there is insufficient genetic variation to respond to selection. It could decrease. I need to think more carefully about how to connect selection gradients with population level dynamics. (I still need to read Ruth Shaw's aster modeling papers).
3. Right now, as I've pitched it, I have no manipulations which is something I need to determine whether temperature is actually increasing selection gradients.
 - start cohorts at different times to control development
 - for biotic interactions, manipulate floral display for pollinators
 - induce herbivory- plant stress responses
4. I could estimate kinship matrix in natural populations with many markers(thousands) and apply quantitative genetics techniques to identify constraints between different traits.

Post doc grants:

1. Plant genome fellowship due in November (focused on crop or crop related plants)
 - systems: sunflower, grasses, medicago, poplar(viability selection, high early life stage mortality), willows
2. Fullbright for international opportunities

```
<div id='id-section49'>
```

```
###Page 49:2016-07-28. Quantitative genetics and the molecular  
basis of complex traits
```

Molecular biologists and quantitative genetics are interested in, at some level, the molecular basis of complex traits. However, each field uses different approaches to this problem. Traditionally, the molecular biologist will manipulate a gene within a single genotype to observe its effect on a phenotype. On the other hand, a quantitative geneticist will take many different genotypes, shuffle genes around by mating individuals with each other, and then statistically assign the effect of genotypes in general on a phenotype.

It'd be interesting to merge both approaches: Knock out or in a gene for many genotypes within a mating design. This way, you can observe the effect on a particular gene within many different genotypes. Just a thought!

Paaby lab is doing a bit of this. She gave a talk earlier this year?

Anyway, she picked a well known developmental pathway in worms (*C. elegans*) and used RNAi for many different species (I think) for a panel of genes.

```
<div id='id-section50'>
```

```
###Page 50: 2016-08-02. Picking a plant system for post doc idea
```

I plan on applying for Plant Genome Research Program (PGRP).

Previous [awards](#). I need a plant with a sequenced genome which is a crop or crop-related. [List of sequenced genomes](#), list of genomes with “good” [annotations](#)

***Mimulus guttatus* (Monkey Flower)**

[Cool paper](#) showing that there are annuals and perennials which vary in morphology under a common garden. So, I could compare selection gradients for annuals vs perennials.

- Read this lab’s [papers](#) because they are interested in similar things.

Leavenworthia alabamica

Annual endemic to [Alabama](#) and it has low population size.

Populations/individuals vary in their reproductive mode: self compatible, self incompatible. So it’d be interesting to see how selection acts on these different reproductive forms.

Papers:

- [Herman & Schoen 2016](#) and this [one](#)
- [Secondary loss in self incompat](#)
- [Compares selection gradients of self comp and self incompt plants] (<http://www.amjbot.org/content/99/3/488.full>)

***Panicum virgatum* (switch grass)**

Perennial with wide distribution from Canada to Mexico. We could look at episodic selection under a common garden across latitude. If you have performance on the y axis and x axis is lat(climate), and we have a mating design, we can analyze the data as function-valued traits. Growth would be a good option.

[Genome paper](#)

Measuring physiology: IR gas exchange analyzer

**Measures photosynthetic rate and transpiration rate! **

Cool technique to QTL with function valued traits [here](#)

```
<div id='id-section51'>  
###Page 51: 2016-08-02; 2016-08-03. Climate cascade meeting
```

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)
- Multiple stressors ms: NJG gave me edits 2016-08-02, rework, then send to Sara. Aiming to submit next week?
- Range limits ms: Go over figures, meet with NJG 2016-08-03 to

go over intro, methods, and results.

- Figure suggestions:
 - recolor map, keep maps consistent
 - shift cold tolerance vs tmin legend from horizontal to vertical.
 - double checkt he interaction of tmin and pre treatment temp; the betas
 - create 2 panel fig for basal cold tolerance and hardening.
 - Thermal niche ms: Lacey's hands
 - HSP modulation paper: SHC's hands
 - Stressed in nature MS: Curtis' hands ; he was suppose to give me a timeline
 - Genome sequencing? Mlau's hands
 - Phylogenomics of common forest ants: SHC and Bernice assembling data matrix. ADN needs to send vouchers to Bernice.
2. Ask about post doc (**NJG and SHC think its ok to stay at same institution**)
3. Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper? Apply for funding? **Suitor Travel Grant**

Deadline is october 31

4. Biolunch: Should I talk about github?(**SHC and NJG are ok with this but I need to think about my delivery and what people can “handle”**) Range limits? Dissertation talk (I want to give this in the Spring (**SHC says yes**))?
-

<div id='id-section52'>

Page 52: 2016-08-04. Following up

stats, range limits project

**analysis of data with pre treatment
temperature as continuous within an
anova**

```
## anova model  
  
k.dat$pretreat_Temp<-as.numeric(as.character(k.dat$pretreat_Temp))  
  
cold.mod1<-aov(treatment_recovery_s~Tmin*pretreat_Temp+Colony,data=k.dat) # testing interaction between pre-treat  
temp and T min (both continuous)
```

Df	Sum Sq	Mean Sq	F value	Pr(>F)
----	--------	---------	---------	--------

Tmin	1	116145	116145	5.755	0.018765 *
------	---	--------	--------	-------	------------

pretreat_Temp	1	261310	261310	12.949	0.000553 **
---------------	---	--------	--------	--------	-------------

```
*  
Tmin:pretreat_Temp 1 162568 162568 8.056 0.005747 **  
  
Residuals 80 1614444 20181
```

analysis of data with pre treatment temperature as a factor within a linear model

```
##analysis of data with pre treatment temperature as a factor within a linear model  
k.dat$pretreat_Temp<-as.factor(as.character(k.dat$pretreat_Temp))  
cold.mod1<-lm(treatment_recovery_s~Tmin*pretreat_Temp+Colony,data=k.dat) #testing interaction between factors of pretreatment with Tmin(continuous)  
#summary(cold.mod1)  
#stepwise aic  
qc<-stepAIC(cold.mod1,direction="both")  
summary(qc)  
  
#output:  
summary(qc)
```

Call:

```
lm(formula = treatment_recovery_s ~ Tmin + pretreat_Temp  
+ Tmin:pretreat_Temp,  
  data = k.dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-292.69	-79.96	-10.13	69.04	355.98

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	210.58	363.71	0.579	0.56432
Tmin	-24.64	24.27	-1.015	0.31321
pretreat_Temp0	450.14	514.37	0.875	0.38426
pretreat_Temp25	1796.59	514.37	3.493	0.00080

pretreat_Temp5	1173.92	514.37	2.282	0.02527
*				
Tmin:pretreat_Temp0	40.73	34.33	1.186	0.23916
Tmin:pretreat_Temp25	114.57	34.33	3.338	0.00131
**				
Tmin:pretreat_Temp5	76.71	34.33	2.235	0.02837
*				

```
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1  
‘ ’ 1
```

Residual standard error: 124 on 76 degrees of freedom

Multiple R-squared: 0.4577, Adjusted R-squared: 0.40

78

F-statistic: 9.164 on 7 and 76 DF, p-value: 3.644e-08

More digestable table:

```
knitr::kable(summary(qc)$coefficients)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	210.58099	363.71495	0.5789726	0.564
Tmin	-24.64324	24.27295	-1.0152553	0.313
pretreat_Temp0	450.14412	514.37061	0.8751358	0.384
pretreat_Temp25	1796.59479	514.37061	3.4928022	0.000
pretreat_Temp5	1173.91549	514.37061	2.2822367	0.025
Tmin:pretreat_Temp0	40.72533	34.32714	1.1863889	0.239
Tmin:pretreat_Temp25	114.57348	34.32714	3.3376940	0.001
Tmin:pretreat_Temp5	76.71280	34.32714	2.2347566	0.028

###Hardening ability

```
cold.mod8<-aov(hardening~Tmin*PT+Colony,data=mew6)
```

```
qc8<-stepAIC(cold.mod8,direction="both")
```

```
summary(qc8)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Tmin	1	85850	85850	5.903	0.02055	*
PT	2	550143	275071	18.915	3.01e-06	***
Colony	17	1435781	84458	5.808	6.88e-06	***
Tmin:PT	2	179795	89897	6.182	0.00513	**
Residuals	34	494455	14543			

Good post to read for understanding interactions [here](#)

2016-08-11 updated analyses

Basal cold tolerance re-analyzed

	df	SS	MS	F-value	P-value	
Tmin	1	114575	114575	6.757	0.0122	
Pre-treatment	3	623523	207841	12.257	<0.001	
Tmin × Pre-treatment	3	189451	63150	3.724	0.016	
9						
Colony	17	228419	13436	0.792	0.6931	

Residuals 51 864771 16956

Cold hardening re-analyzed (double checked)

	df	SS	MS	F-value	P-value		
Tmin	1	411796	411796	26.318	<0.001		
Pre-treatment	2	363498	181749	11.616	<0.001		
Tmin × Pre-treatment	2	98308	49154	3.141	0.055		
986							
Colony	17	1285635	75626	4.833	<0.001		
Residuals	34	531992	15647				

Interaction non-significant; the change was caused by a mistake made by consolidating scripts.

<div id='id-section53'>

Page 53: 2016-08-08. Post doc ideas part 2

**1. How does selection operate on the life histories of poplar?
[Similar to this post doc listing]
(<http://evol.mcmaster.ca/~brian/evo.html>)**

Approach: **Identify and characterize how natural selection operates at different life stages of poplar**

- Measure selection gradients on age structured populations in the field
- **Is it possible to heat shock leaves out in the field?**
- Viability selection([Mojica & Kelly ref](#)) : One thing missing from selection studies is that organisms can die before expressing a trait (They confusingly call this the invisible fraction of variation). Can we test this by taking cuttings and planting them? Or does it have to be from seeds? (I think the latter)
- [Good natural history](#)

How does contemporary episodes of natural selection compare with past local adaptation to climate?

Approach: Compare selection in the field to common garden. There is a cool [paper](#) by Kingsolver et al. 2012 that suggests we account for environmental covariation with selection gradient analyses. If we have a relatedness matrix, we can see if individuals are spatially clustered with environment.

2. How does selection operate on populations of monkeyflowers (*Mimulus guttatus*) with different modes of reproduction?

Approach: **Identify and characterize how natural selection operates on perennials and annals** Which one is more susceptible? Are there shifts between one or the other?

- Measure selection gradients across a whole cline (whole west coast of US) for perrenials and annuals.
- Perrenials experience greater within generation variation than among-so they may harbor greater plasticity than annuals.

3. Gladecresses *Leavenworthia alabamica*

- measure selection gradients between self compatible vs self incompatible for populations in Alabama.
- Low adaptive potential in self compatible vs self incompatible.

4. Identifying specific genotypes for optimal growth in the Shrub willow (*Salix pupurea*)

Approach1: **Evaluate growth as a function valued trait across latitudinal cline**

- Mate and plot genotypes in the field. Or take clippings and plant?
- Measure growth across latitude.

Approach2: Evaluate growth as a function valued trait within a common garden

- Possible to have them reared at 6 temperatures and 3 moisture levels?
-

Analysis: Determine shifts in growth reaction norms.

<div id='id-section54'>

###Page 54: 2016-08-10. Climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)
- **Multiple stressors ms: SHC's hands- discussion is too disjointed, reworking organization**
- **Range limits ms: Fixed figures, go over!**
- Thermal niche ms: Lacy and I are working on it. Discussion left to do
- HSP modulation paper: SHC's hands
- Stressed in nature MS: Curtis' hands ; he was suppose to give me a timeline
- Genome sequencing? Mlau's hands
- Phylogenomics of common forest ants: ADN to send Bernice samples this week.

2. Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - Apply for funding. **Suitor Travel Grant** [Deadline is october 31](#)
 3. Biolunch, working title: **Strategies for achieving reproducible research** ; get picture of the meeting
-

<div id='id-section55'>

Page 55: 2016-08-11. Overlaying raster files in a map in R

Good link to show how to overlay [here](#). I've had to use this to plot climate cut offs (example: [here](#))

Some code:

Cropping world map, I set coords to region I'm interested in: Maine

```
w2 <- getData('worldclim', var='bio', res=.5, lat=45, lon=-68) # grab worldclim data; with .5 res you need to specify coordinates  
  
extent<-c(-72, -65, 42, 48)
```

```
bew<-crop(w2,extent)
```

Here is the code to make cut offs: designate extreme values and then plotting it will be easy

You have to get rid of NAs and assign to variable.

```
Tm<-na.omit(bew[[5]])  
Tm[bew[[5]] < 246.5] <- 100 # absent  
Tm[bew[[5]] > 246.5] <- 1
```

###Here is plotting the cut off

```
dbio2$coco<-ifelse(dbio2$Found_Notfound=="1","red","black") # specify color of points base don presence absence  
  
plot(lar[[5]],col=c("white","grey75"),legend=F)  
map("worldHires",c("USA","Canada"),add=TRUE)  
map("state", c('maine','vermont','new hampshire'), add = TRUE)  
points(dbio2$Lon,dbio2$Lat,pch=16,col=dbio2$coco)
```

```
<div id='id-section56'>
```

```
###Page 56: 2016-08-16 range limits paper, data analysis of chill  
coma recovery time (CCRT) revisited
```

From my G matrix analysis, I find variation in the cooler-warmer axis. So for my statistics for relating CCRT to local environment (to see if they're locally adapted), I used an ANCOVA:

```
CCRT ~ pre-treatment temp * Tmin
```

This just says whether the relationship between CCRT and Tmin at each pre-treatment temperature are *different or not*. But what I may want, is an estimate of those relationships. So I should run a regression or mixed effect model to generalize to the whole population.

Mixed effect model with pretreatment * Tmin interaction, random intercept and slope? for every colony measured at each pretreatment temp

```
mod5.r<-lmer(formula=inv_c~pretreat_Temp*Tmin+(1+pretreat  
_Temp|Colony),REML=TRUE,data=test)
```

I'll compare this model to:

Mixed effect model with fixed effect of Tmin, random intercept and slope? for every colony measured at each pretreatment temp

```
mod3<-lmer(formula=inv_c~Tmin+(1+pretreat_Temp|Colony),RE  
ML=TRUE,data=test)
```

and also compare it to:

Mixed effect model with fixed effect of Tmin and pretreatment temp, random intercept and slope? for every colony measured at each pretreatment temp

```
mod4<-lmer(formula=inv_c~pretreat_Temp+Tmin+(1+pretreat_T  
emp|Colony),REML=TRUE,data=test)
```

my “comparison” using anova function:

```
refitting model(s) with ML (instead of REML)  
Data: test
```

Models:

```
mod3: inv_c ~ Tmin + (1 + pretreat_Temp | Colony)
mod2: inv_c ~ pretreat_Temp + (1 + pretreat_Temp | Colony)
)
mod4: inv_c ~ pretreat_Temp + Tmin + (1 + pretreat_Temp |
Colony)
mod5.r: inv_c ~ pretreat_Temp * Tmin + (1 + pretreat_Temp |
Colony)
```

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	P
r(>Chisq)								
mod3	13	555.36	602.00	-264.68	529.36			
mod2								
mod2	15	544.10	597.91	-257.05	514.10	15.2606	2	
0.0004855 ***								
mod4	16	543.62	601.02	-255.81	511.62	2.4798	1	
0.1153190								
mod5.r	19	540.10	608.26	-251.05	502.10	9.5188	3	
0.0231317 *								

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1								
' ' 1								

mod5.r is stat diff from the other more simple models

Let's look at the output:

Linear mixed model fit by REML [`'lmerMod'`]

Formula: `inv_c ~ pretreat_Temp * Tmin + (1 + pretreat_Tem`
`p | Colony)`

Data: test

REML criterion at convergence: `525.8`

Scaled residuals:

Min	1Q	Median	3Q	Max
<code>-1.9347</code>	<code>-0.5625</code>	<code>-0.1789</code>	<code>0.4116</code>	<code>5.4326</code>

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
Colony	(Intercept)	<code>0.03646</code>	<code>0.1909</code>	
	pretreat_Temp0	<code>0.15330</code>	<code>0.3915</code>	<code>-0.23</code>
	pretreat_Temp25	<code>0.20398</code>	<code>0.4516</code>	<code>-0.92 -0.13</code>
	pretreat_Temp5	<code>0.26667</code>	<code>0.5164</code>	<code>-0.17 -0.50</code>
	Residual	<code>0.32402</code>	<code>0.5692</code>	

Number of obs: `267`, groups: Colony, `18`

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	3.59188	1.03188	3.481
pretreat_Temp0	-2.90454	1.68322	-1.726
pretreat_Temp25	-4.39184	1.80599	-2.432
pretreat_Temp5	-4.47550	1.96022	-2.283
Tmin	0.11598	0.06922	1.675
pretreat_Temp0:Tmin	-0.23723	0.11283	-2.102
pretreat_Temp25:Tmin	-0.28354	0.12088	-2.346
pretreat_Temp5:Tmin	-0.30516	0.13104	-2.329

Correlation of Fixed Effects:

	(Intr)	prt_T0	pr_T25	prt_T5	Tmin	p_T0:T	p_T
25:							
pretrt_Tmp0	-0.519						
pretrt_Tmp25	-0.770	0.183					
pretrt_Tmp5	-0.443	-0.039	0.499				
Tmin	0.997	-0.517	-0.766	-0.442			
pretrt_Tm0:T	-0.518	0.997	0.184	-0.037	-0.520		
pretrt_T25:T	-0.768	0.184	0.997	0.498	-0.770	0.185	
pretrt_Tm5:T	-0.443	-0.037	0.498	0.997	-0.445	-0.036	0.
500							

Considering only the random effect of colony

```
mod2<-lmer(formula=treatment_recovery_s.x~pretreat_Temp+(1|Colony),REML=TRUE,data=test)

mod3<-lmer(formula=treatment_recovery_s.x~Tmin+(1|Colony),REML=TRUE,data=test)

mod4<-lmer(formula=treatment_recovery_s.x~pretreat_Temp+Tmin+(1+pretreat_Temp|Colony),REML=TRUE,data=test)

#mod5.r<-lmer(formula=inv_c~pretreat_Temp*Tmin+(1|Colony),REML=TRUE,data=test)

mod6<-lmer(formula=treatment_recovery_s.x~pretreat_Temp*Tmin+(1|Colony),REML=TRUE,data=test)

anova(mod3,mod4,mod2,mod6)
```

```
mod3: treatment_recovery_s.x ~ Tmin + (1 | Colony)
mod2: treatment_recovery_s.x ~ pretreat_Temp + (1 | Colony)
mod4: treatment_recovery_s.x ~ pretreat_Temp + Tmin + (1 | Colony)
mod6: treatment_recovery_s.x ~ pretreat_Temp * Tmin + (1 | Colony)

      Df     AIC     BIC   logLik deviance    Chisq Chi Df Pr(>Chisq)
mod3  4  3628.0  3642.4 -1810.0    3620.0
mod2  6  3583.1  3604.6 -1785.5    3571.1  48.9531    2  2.
```

344e-11 ***

mod4	7	3577.2	3602.3	-1781.6	3563.2	7.8337	1	0
.005128	**							
mod6	10	3564.4	3600.2	-1772.2	3544.4	18.8832	3	0
.000289	***							

###model output for mod6

Linear mixed model fit by REML [`'lmerMod'`]

Formula: treatment_recovery_s.x ~ pretreat_Temp * Tmin +
(1 | Colony)

Data: test

REML criterion at convergence: 3649.7

Scaled residuals:

Min	1Q	Median	3Q	Max
-3.2557	-0.6656	-0.1116	0.4587	3.8248

Random effects:

Groups	Name	Variance	Std.Dev.
Colony	(Intercept)	752.9	27.44
	Residual	33965.8	184.30

Number of obs: 280, groups: Colony, 19

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	207.92	291.37	0.714
pretreat_Temp0	439.58	396.48	1.109
pretreat_Temp25	1736.31	395.31	4.392
pretreat_Temp5	1215.86	399.33	3.045
Tmin	-24.52	19.57	-1.253
pretreat_Temp0:Tmin	39.34	26.65	1.476
pretreat_Temp25:Tmin	109.35	26.52	4.124
pretreat_Temp5:Tmin	79.62	26.73	2.979

Correlation of Fixed Effects:

	(Intr)	prt_T0	pr_T25	prt_T5	Tmin	p_T0:T	p_T
25:							
pretrt_Tmp0	-0.678						
prtrt_Tmp25	-0.681	0.500					
pretrt_Tmp5	-0.674	0.495	0.497				
Tmin	0.997	-0.676	-0.679	-0.672			
prtrt_Tm0:T	-0.676	0.997	0.498	0.493	-0.678		
prtrt_T25:T	-0.680	0.499	0.997	0.496	-0.682	0.501	
prtrt_Tm5:T	-0.674	0.496	0.497	0.997	-0.677	0.497	0.
499							

```
<div id='id-section57'>
```

Page 57: 2016-08-25. Hsp modulation follow up stats

```
summary(aov(log10(B_40)~axis3_desig,data=mergy))  
Df Sum Sq Mean Sq F value Pr(>F)  
axis3_desig 3 4.947 1.6490 7.154 0.000413 ***  
Residuals 52 11.986 0.2305  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' ' 1
```

I separated out groupings based on phylogenetic axes. The model anova is significant.

Now I'll do a post hoc test.

```
TukeyHSD(aov(log10(B_40)~axis3_desig,data=mergy))  
diff lwr upr p adj  
North-A. picea 0.1185330 -0.3739818 0.611047  
8 0.9189644  
South-A. picea -1.0921848 -1.7596714 -0.424698  
2 0.0003710
```

zAxis 2 A. picea-A. picea	0.2516912	-0.5104439	1.013826
3 0.8169654			
South-North	-1.2107178	-1.9910398	-0.430395
8 0.0007709			
zAxis 2 A. picea-North	0.1331582	-0.7295202	0.995836
6 0.9765503			
zAxis 2 A. picea-South	1.3438760	0.3706435	2.317108
5 0.0031663			

<div id='id-section58'>

Page 58: 2016-08-29 & 30. Climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)
- Multiple stressors ms: working on SHC edits
 - Send out Wednesday.
- Range limits ms: **Go over figure; SHC has ms**; eta? Not looked at it.
 - sampling map: make larger, points should be gray; sites

- that were used for common garden should have a gold outline
- fig 6, cold phys; get rid of “cold”, use different words.
 - Thermal niche ms: **Lacey and I working on discussion**
 - HSP modulation paper: SHC submitted
 - Stressed in nature MS: Samples to rerun.
 - update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis’ samples.

The DF 20140717 sample box was found when we dug through all the freezers

in the winter and I didn't have time to extract RNA and qPCR them all.

The HF 20140812 box was the box we weren't able to find anywhere.

There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.

- Genome sequencing? Mlau's hands
- Phylogenomics of common forest ants: status?
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range

limits paper.

- **construct talk; when to give practice talk ?**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing and then get everything signed.
 - Biolunch, working title: Strategies for achieving reproducible research Sept 2nd.
-

<div id='id-section59'>

Page 59: 2016-09-01. SHC lab meeting

Fall 2016

Room	Date	Activity	Person.in.Charge	Breakfast
124	Sept. 8	IDPs	Sara	Sara
124	Sept. 15	American Naturalist paper	Sara	Megna
122	Sept. 22	Experimental design	Megna	Katie
124	Sept. 29	Manuscript - A. picea range limits	Andrew	Laurel
124	Oct. 6	Proposal - NSF post-doc fellowship	Andrew	Delaney
124	Oct.	Experimental	Julia	Julia

	13	design		
122	Oct. 20	Research update	Bonnie	Bonnie
124	Oct. 27	Results presentation	Delaney	Delaney
124	Nov. 3	Paper discussion	Laurel	Sara
124	Nov. 10	Results discussion	Laurel	Laurel
122	Nov. 17	Manuscript - CNP in Aphaenogaster	Katie	Bonnie
NA	24- Nov	Thanksgiving		
124	Dec. 1	Meeting talk - range limits	Andrew	Sara
124	Dec. 8	Dimensions of Biodiversity new papers!!!	Everyone	Andrew

Note dietary requirements for breakfasts:

- Dairy-free options
- No coconut
- No nuts in baked goods
- No honeydew melon

Including notes from meeting (added 2016-09-02)

- LSO needs to check monthly eye wash, chemical inventory, lab safety
- Do your lab safety training.
-

Tuesday morning (2016-09-06): Schedule time to look for ants, collect ~ 20.

<div id='id-section60'>

2016-09-01: Paper notes: Paccard et al. 2016

ref:

Paccard A, Van Buskirk J, Willi Y, Eckert CG, Bronstein JL. 2016. Quantitative Genetic Architecture at Latitudinal Range Boundaries: Reduced Variation but Higher Trait Independence. *The American Naturalist*.

Quick and dirty: They compared variance-covariance G matrices among

**9 populations in a *Arabdopsis* species
that spans a cline. It was in a common
garden with 2 levels of moisture
treatments.**

Findings:

- Genetic variance was highest at the middle of their range and lowest at the edges (south and **north**)
- More trait independence at the **northern** part of their range

Making sense of the properties of G

Confusing sentence in methods: *We calculated four measures of multivariate evolutionary potential and G-matrix geometry (size, sphericity, and orientation) for each treatment and population.*

Separate or the same item?

1. **Size:** sum of genetic variances across all traits. I guess this means the total amount of genetic variance.
2. **Sphericity = # of dimensions:** Sum of all eigenvalues / first eigenvalue. It tells you how independent traits are. If it is 1 then gmax or the first pc explains most of the variation. But if it is a high number (# of dimensions of G), then it tells you many

traits are independent and variances are distributed among traits. It can also tell you whether genetic constraints exist in certain directions without specifying direction

3. **Orientation of G relative to common standard vector:**

Compare dmax (dominant eigenvector of variance-covariance matrix of population means for 10 traits across the 9 populations— **D matrix** describes population divergence). For each population they measured the orientation as the absolute value of the angle between dmax and gmax.

4. **Response to selection: Random skewers method:** They calculate change in phenotype by simulating Betas in the delta

$$Z = \mathbf{G} * \text{Beta}$$

<div id='id-section61'>

Page 61: 2016-09-06. Playing with rpart with range limit data

Using bioclim variables to classify presence-absence

Guidance for picking “best” tree

- Convention is to pick one with the lowest cross-validate relative error or smallest(simplest) tree within 1 standard error of best tree
-

```
###Full dataset layout
```

```
str(dbio2)

'data.frame':   102 obs. of  38 variables:

$ n            : int  1 2 3 4 5 6 7 8 9 10 ...
$ date         : int  19960507 20140709 20140709 201407
10 20050625 20030715 20050625 20130718 19910901 20050630
...
$ state        : Factor w/ 1 level "Maine": 1 1 1 1 1 1
1 1 1 1 ...
$ county       : Factor w/ 23 levels "", "cumberland", ...
: 23 2 8 8 6 6 6 21 7 6 ...
$ locality     : Factor w/ 84 levels "", "18-LP-4C", ...
81 42 17 17 6 3 4 76 61 67 ...
$ habitat       : Factor w/ 12 levels "", " ", "Behind di
ning hall", ...: 11 8 5 6 NA NA NA 3 12 NA ...
$ Lat          : num  43.6 43.9 43.9 43.9 44.3 ...
$ Lon          : num  -70.8 -70.2 -69.7 -69.7 -68.3 ...
$ masl         : num  158 NA NA NA 68 100 230 NA NA 105
...
$ subfamily    : Factor w/ 2 levels "", "Myrmicinae": 2
2 2 2 2 2 2 2 2 ...
$ ant.genus    : Factor w/ 2 levels "", "Aphaenogaster": 
2 2 2 2 2 2 2 2 ...
$ ant.species  : Factor w/ 2 levels "", "picea": 2 2 2 2
2 2 2 2 2 ...
$ code         : Factor w/ 2 levels "", "aphpic": 2 2 2
```

```
2 2 2 2 2 2 2 ...  
$ collection      : Factor w/ 75 levels "", "Aaron", "Acadia  
NP", ... : 5 1 4 1 3 3 3 1 6 7 ...  
$ collector       : Factor w/ 11 levels "Aaron", "Acadia Bi  
oBlitz", ... : 10 3 3 3 8 2 8 4 10 9 ...  
$ Found_Notfound: int 1 1 1 1 1 1 1 1 1 1 ...  
$ MAT              : num 7 7.6 7.8 7.8 6.9 6.6 6.3 6.6 6.6  
6.8 ...  
$ MDR              : num 129 108 105 105 107 107 107 106 109 1  
24 110 ...  
$ ISO              : num 32 28 28 28 28 28 28 29 30 28 ...  
$ SD               : num 94.2 92.7 90.5 90.5 90.7 ...  
$ Tmax             : num 27.1 26.3 26 26 25.5 25.2 24.8 24  
.9 27.1 25.9 ...  
$ Tmin             : num -132 -115 -107 -107 -117 -121 -12  
3 -120 -142 -125 ...  
$ TAR              : num 403 378 367 367 372 373 371 369 4  
13 384 ...  
$ TWQ              : num 24 33 37 37 -22 -25 -28 -23 20 25  
...  
$ TDQ              : num 179 186 192 192 183 180 177 177 -  
53 186 ...  
$ TwarmQ           : num 188 193 192 192 183 180 177 177 1  
89 186 ...  
$ TminQ            : num -57 -47 -42 -42 -52 -56 -59 -54 -  
66 -58 ...  
$ AP               : num 1195 1146 1157 1157 1261 ...  
$ PWM              : num 131 123 125 125 144 148 150 140 1
```

```
10 127 ...
$ PDM      : num  86 79 76 76 78 79 81 77 69 81 ...
$ PSD      : num  12 13 14 14 17 18 18 17 11 14 ...
$ PWQ      : num  349 335 341 341 388 401 407 385 3

01 343 ...
$ PDQ      : num  267 244 244 244 245 250 256 245 2

31 248 ...
$ PwarmQ   : num  275 248 244 244 245 250 256 245 2

68 248 ...
$ PminQ    : num  293 293 297 297 342 354 359 340 2

40 294 ...
$ var       : Factor w/ 2 levels "absent","present":
2 2 2 2 2 2 2 2 2 ...
$ color     : chr  "red" "red" "red" "red" ...
$ coco      : chr  "red" "red" "red" "red" ...
```

All bioclim variables

```
knitr:::kable(round(cor(dbio2[17:35]),3))
```

	MAT	MDR	ISO	SD	Tmax	Tmin	TAR
MAT	1.000	-0.273	0.352	-0.637	0.663	0.876	-0.512
MDR	-0.273	1.000	0.541	0.787	0.483	-0.674	0.913
ISO	0.352	0.541	1.000	-0.047	0.537	0.104	0.179
SD	-0.637	0.787	-0.047	1.000	0.133	-0.916	0.967

Tmax	0.663	0.483	0.537	0.133	1.000	0.229	0.299	.
Tmin	0.876	-0.674	0.104	-0.916	0.229	1.000	-0.860	.
TAR	-0.512	0.913	0.179	0.967	0.299	-0.860	1.000	.
TWQ	-0.740	0.137	-0.387	0.526	-0.506	-0.649	0.371	.
TDQ	0.620	-0.722	-0.027	-0.859	-0.031	0.845	-0.844	.
TwarmQ	0.852	0.168	0.402	-0.143	0.939	0.511	-0.009	.
TminQ	0.948	-0.519	0.249	-0.848	0.398	0.980	-0.752	.
AP	0.560	-0.606	0.119	-0.836	-0.056	0.771	-0.785	.
PWM	0.598	-0.647	0.060	-0.843	-0.015	0.809	-0.800	.
PDM	0.769	-0.437	0.218	-0.717	0.344	0.818	-0.622	.
PSD	-0.265	-0.587	-0.470	-0.341	-0.700	0.106	-0.471	.
PWQ	0.495	-0.733	-0.040	-0.861	-0.180	0.775	-0.854	.
PDQ	0.793	-0.399	0.303	-0.724	0.364	0.826	-0.619	.
PwarmQ	-0.878	0.525	-0.127	0.771	-0.395	-0.916	0.692	.
PminQ	0.684	-0.678	0.072	-0.898	0.041	0.884	-0.844	.

rpart predictive model: full bioclim

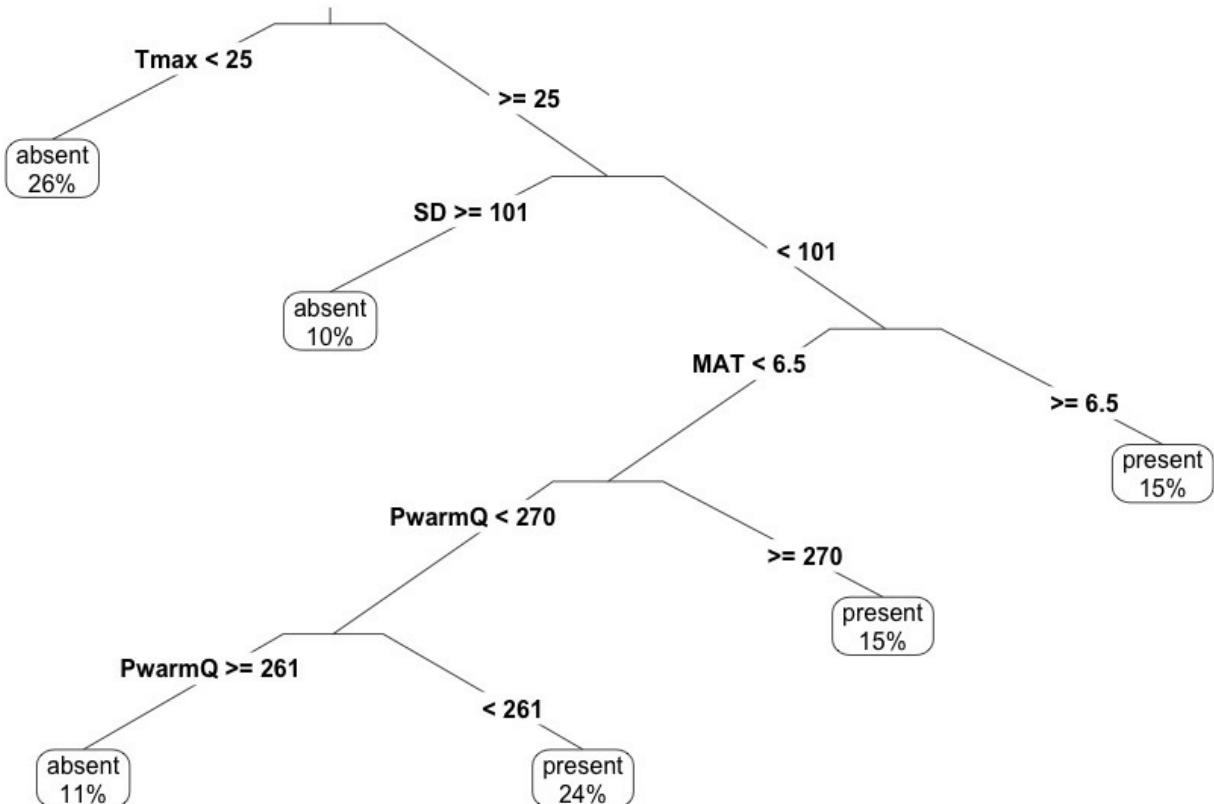
```
vars<-as.data.frame(cbind(dbio2[,17:35],V1=dbio2[,36])) #  
all bioclim variables  
  
form<-as.formula(V1~.)  
tree.1<-rpart(form,data=vars,control=rpart.control(minspl
```

```
it=20, cp=0), method="class")  
printcp(tree.1)  
plotcp(tree.1)  
  
rpart.plot(tree.1, type=3, extra=100)
```

classification tree

Table statistics of model:

CP	nsplit	rel error	xerror	xstd
0.42	0	1.00	1.26	0.0981595
0.12	1	0.58	0.82	0.0990346
0.06	2	0.46	0.76	0.0976589
0.00	5	0.28	0.66	0.0944956



model accuracy

```

m<-predict(tree.1,vars[-20])
m.pre<-ifelse(m[,1]< m[,2],"present","absent")
#confusion matrix
#following this tutorial
#http://eric.univ-lyon2.fr/~ricco/tanagra/fichiers/en_Tan
agra_Validation_Croisee_Suite.pdf
mc<-table(vars$V1,m.pre);mc
sum(ifelse(vars$V1== m.pre,1,0))/nrow(vars)
  
```

Confusion matrix indicating 86.2% accuracy:

	absent	present
absent	42	8
present	6	46

Subset of bioclim variables:

```
sub<-data.frame(cbind(dbio2$MAT,dbio2$Tmin,dbio2$SD,dbio2
$TAR,dbio2$ISO,dbio2$MDR,dbio2$AP,dbio2[,31]))
names(sub)<-c("MAT","Tmin","SD","TAR","ISO","MDR","AP","P
SD")
knitr:::kable(round(cor(sub),3))
```

	MAT	Tmin	SD	TAR	ISO	MDR	AP	PS
MAT	1.000	0.876	-0.637	-0.512	0.352	-0.273	0.560	-0.2
Tmin	0.876	1.000	-0.916	-0.860	0.104	-0.674	0.771	0.1
SD	-0.637	-0.916	1.000	0.967	-0.047	0.787	-0.836	-0.3
TAR	-0.512	-0.860	0.967	1.000	0.179	0.913	-0.785	-0.4
ISO	0.352	0.104	-0.047	0.179	1.000	0.541	0.119	-0.4
MDR	-0.273	-0.674	0.787	0.913	0.541	1.000	-0.606	-0.5
AP	0.560	0.771	-0.836	-0.785	0.119	-0.606	1.000	0.3
PSD	-0.265	0.106	-0.341	-0.471	-0.470	-0.587	0.381	1.0

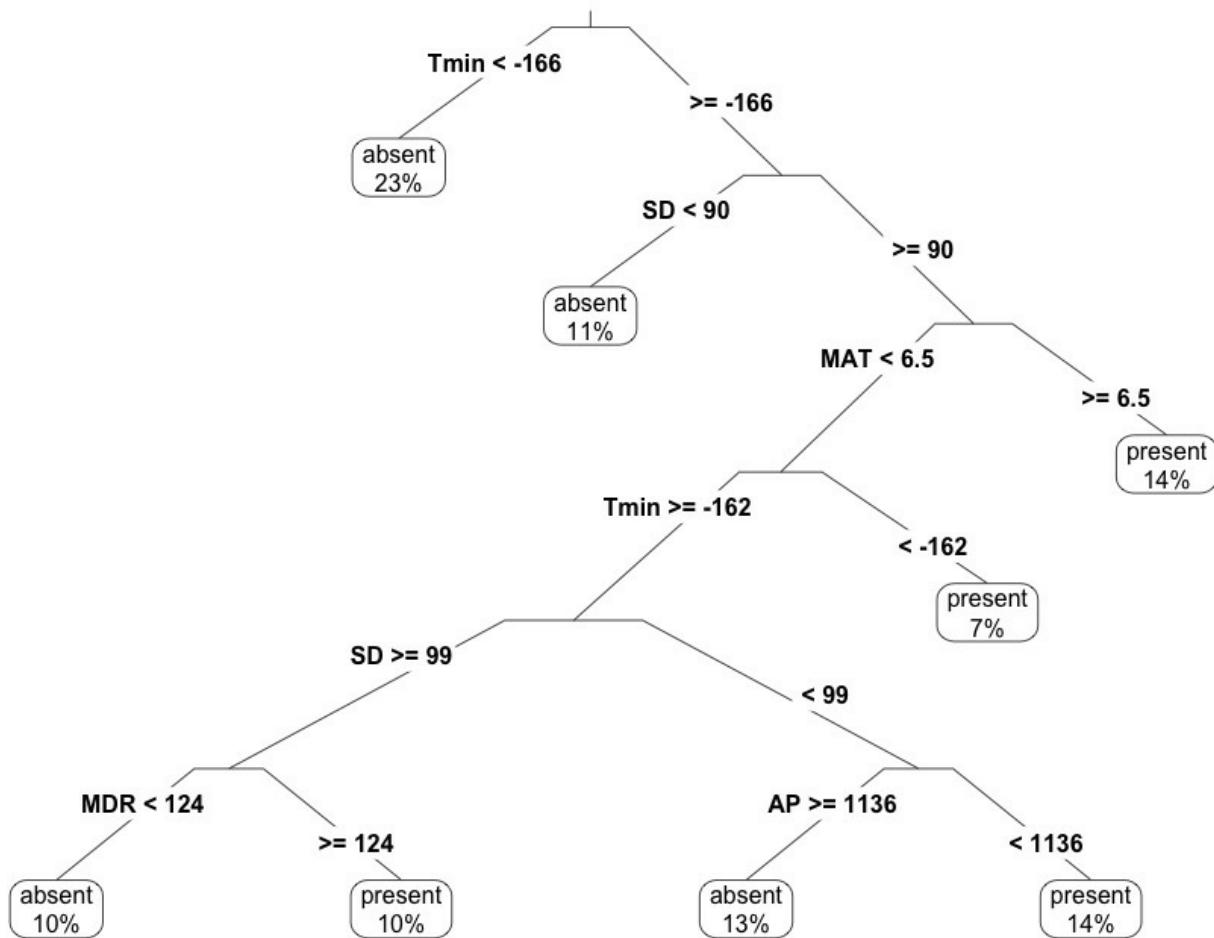
Classification tree with subset of bioclim

```
vars<-as.data.frame(cbind(sub,V1=dbio2[,36]))  
#names(vars)[1]<- "V1"  
  
form<-as.formula(V1~.)  
tree.1<-rpart(form,data=vars,control=rpart.control(minsplit=20, cp=0),method="class")  
printcp(tree.1)  
plotcp(tree.1)  
rpart.plot(tree.1,type=3,extra=100)
```

output of classification tree

Table statistics of model:

CP	nsplit	rel error	xerror	xstd
0.38	0	1.00	1.24	0.0986179
0.14	1	0.62	1.00	0.1009756
0.04	2	0.48	0.84	0.0994100
0.02	6	0.32	0.66	0.0944956
0.00	7	0.30	0.52	0.0880285



model accuracy

```

m<-predict(tree.1,vars[-9])
m.pre<-ifelse(m[,1]< m[,2],"present","absent")
knitr:::kable(mc)

```

Confusion matrix indicating 85.2% accuracy

	absent	present
absent		
present		

absent	46	4
present	11	41

<div id='id-section62'>

Page 62: 2016-09-06. Climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)

- Multiple stressors ms:
 - **SHC hands**
- Range limits ms: **Aaron made comments, go over with Nick**

- Thermal niche ms: **Lacey and I working on discussion**

- Stressed in nature MS: Samples to rerun.
 - update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.

There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.

- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **construct talk; when to give practice talk ?**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing and then get everything signed.

Notes: Only NJG and ANBE in attendance.

- **Go over thesis layout next time**
-

```
<div id='id-section63'>  
###Page 63: 2016-09-07. PCA update for range limit data ; see * Page 63: 2016-09-07. PCA update for range limit data
```

Aaron wants to explore PCA decomposition of bioclim variables

PCA of all bioclim

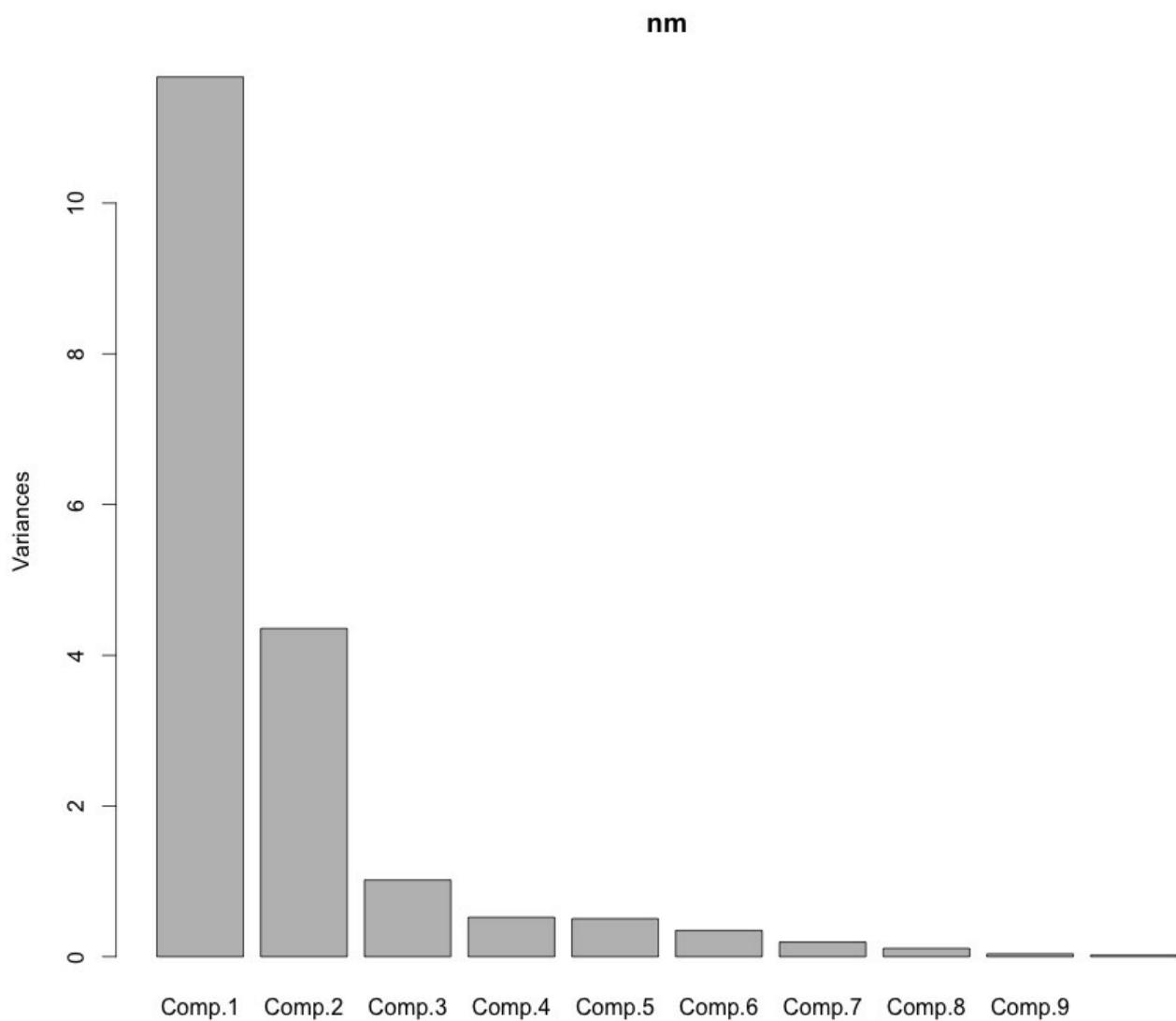
```
nm<-princomp(scale(dbio2[,17:35]))  
knitr:::kable(round(nm$loadings[,1:4],3))
```

Table of loadings

	Comp.1	Comp.2	Comp.3	Comp.4
MAT	0.238	-0.242	0.191	-0.079
MDR	-0.192	-0.307	-0.347	0.086
ISO	0.037	-0.309	-0.614	-0.515
SD	-0.267	-0.124	0.000	0.393
Tmax	0.052	-0.451	0.099	0.239
Tmin	0.281	-0.026	0.184	-0.206
TAR	-0.248	-0.211	-0.129	0.327
TWQ	-0.205	0.213	0.151	-0.155
TDQ	0.259	0.111	0.034	0.002
TwarmQ	0.128	-0.389	0.247	0.209
TminQ	0.274	-0.112	0.140	-0.205
AP	0.258	0.103	-0.324	0.158
PWM	0.268	0.100	-0.230	0.275
PDM	0.259	-0.108	-0.046	0.164
PSD	0.052	0.413	-0.107	0.240
PWQ	0.256	0.180	-0.215	0.198
PDQ	0.259	-0.124	-0.122	0.075
PwarmQ	-0.263	0.107	-0.228	-0.014

PminQ	0.282	0.065	-0.130	0.143
-------	-------	-------	--------	-------

Screeplot of PCA of all bioclim vars



Variance explained

```
summary(nm)
```

Importance of components:

	Comp.1	Comp.2	Comp.3
Comp.4			
Comp.5			
Standard deviation	3.4169139	2.0868333	1.00881816
	0.7		

2270248 0.71067369

Proportion of Variance 0.6205736 0.2314732 0.05409423 0.0
2776159 0.02684514
Cumulative Proportion 0.6205736 0.8520468 0.90614101 0.9
3390259 0.96074773

PC1 explains 62%, PC2 explains 23%, PC3 explains 5%.

##Statistical analysis: Using logistic regression, glm() function for first 3 PCs

```
dmo1<-glm(dbio2$Found_Notfound~pca.clm[,1]+pca.clm[,2]+  
pca.clm[,3],family="binomial")  
summary(dmo1)
```

Call:

```
glm(formula = dbio2$Found_Notfound ~ pca.clm[, 1] + pca.  
clm[,  
2] + pca.clm[, 3], family = "binomial")
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.6588	-0.9896	0.3712	0.9299	2.3119

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.11715	0.24828	-0.472	0.63702

```

pca.clim[, 1] 0.23114    0.08479   2.726  0.00641 ** 
pca.clim[, 2] -0.57836   0.15037   -3.846  0.00012 *** 
pca.clim[, 3] -0.19877   0.24715   -0.804  0.42126 

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 
' ' 1

```

(Dispersion parameter **for** binomial family taken to be **1**)

```

Null deviance: 141.36 on 101 degrees of freedom
Residual deviance: 112.62 on 98 degrees of freedom
AIC: 120.62

```

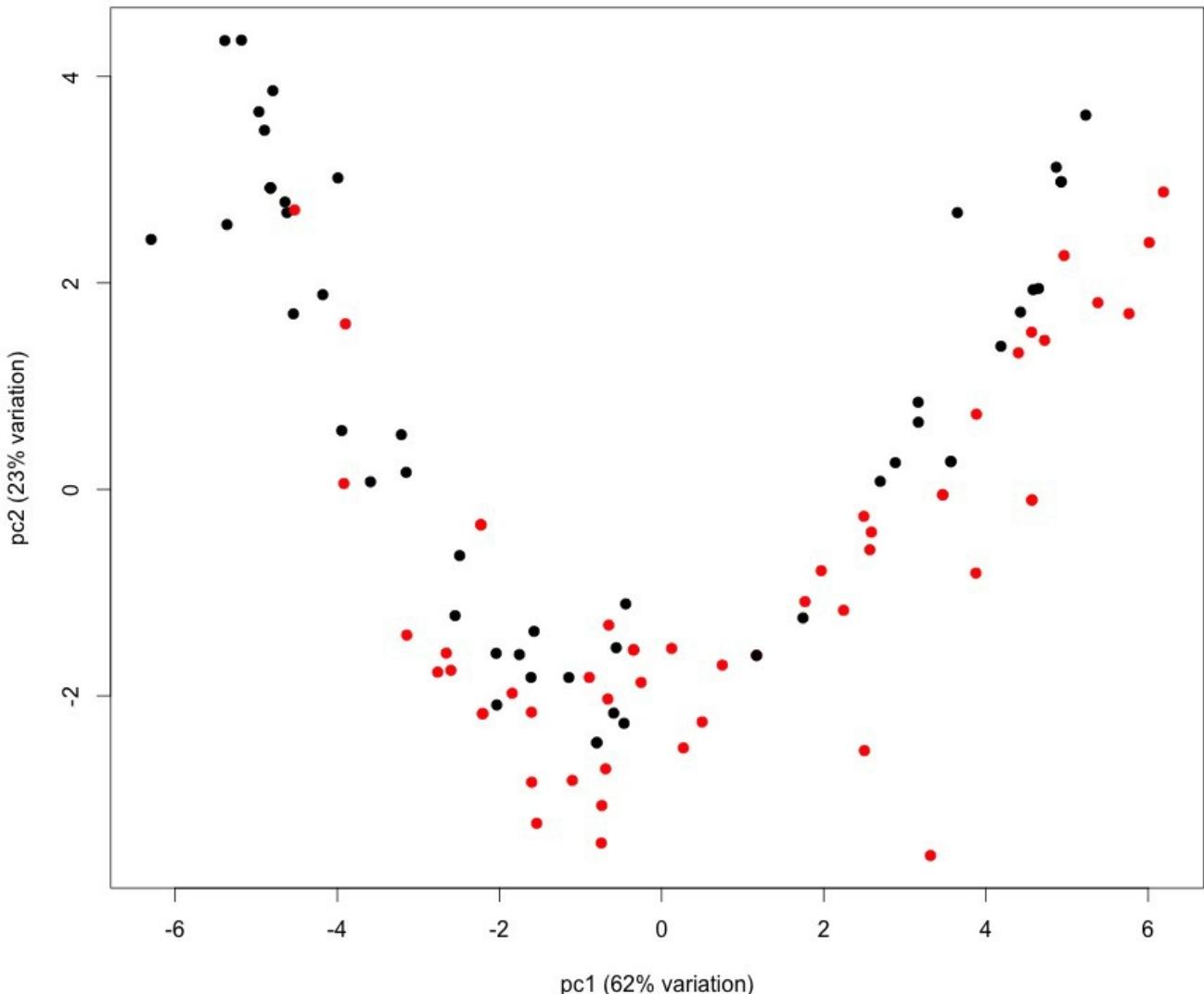
Number of Fisher Scoring iterations: 5

```
#more digestable table
knitr::kable(round(summary(dmo1)$coefficients,3))
```

Table output of logistic regression

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.117	0.248	-0.472	0.637
pc1	0.231	0.085	2.726	** 0.006**
pc2	-0.578	0.150	-3.846	0.000
pc3	-0.199	0.247	-0.804	0.421

Overlaying presence-absence onto climate space as represented by PCs



Aaron's thoughts

Hi Andrew,

The scree plot suggests both PC1 and maybe PC2, not definitely not PC3 are useful. The GLM supports this.

The loadings on PC2 are clear: MDR, ISO, Tmax, TwarmQ, PSD, none of which load heavily on PC1

But the loadings on PC1 are a mess. None exceed 0.3 in loading, and the 0.2-0.3 (absolute values) are: MAT, SD, Tmin, TAR, TDQ, TminQ, AP, PWM, PDM, PDQ, PWarmQ, and PminQ

.

Looks to me like a lot of min temps and precip on PC1 and maxima on PC2, but I don't know my bioclim vars.

But the "bowing" on the biplot is a common problem when you have more than 1 env. gradient working in the data that are working at cross-purposes. Which you described in text, and which you get out of the regression (or classification) tree (which I did get backwards – it's about the predictee, not the predictors, but not both).

So my suggestion would be to stick with the CART analysis. If you must do a GLM, you should only work with uncorrelated BioClim vars. You'll just have to choose the set a priori and defend it.

Best,

Aaron

```
<div id='id-section64'>
```

```
<div id='id-section65'>
```

```
<<<<<< HEAD
```

```
=====
```

Page 65:2016-09-12. variable importance

[Online tutorial](#)

[Youtube version](#)

```
<div id='id-section66'>
```

Page 66: 2016-09-13. climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)
- Multiple stressors ms:
 - **my hands, need to edit and send out by**

wednesday/thursday

- Range limits ms: **SHC's hands**
- Thermal niche ms: **Lacey and I working on discussion**
- Stressed in nature MS: Samples to rerun.
 - update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.

There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.

- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing and then get everything signed.
- **Go over thesis layout next time**
 - Introduction (> 3 pages), manuscripts, then

synthesis/conclusion (~3 pages) ; SHC and NJG agreee

<div id='id-section67'>

Page 68: 2016-09-14. SICB meeting

Venue: Hilton New Orleans Riverside

Address: Two Poydras Street, New Orleans, LA 70130, UNITED STATES

Closest airport: Louis Armstrong New Orleans Airport.

27 minutes away from hilton but there is discounted round trip airport trans: \$40/person

Budget:

- \$40 transportation (put 32 in budget)
 - \$388 flight
 - \$580 + taxes and fees hotel
 - \$ 99 registration to SICB
-

<div id='id-section68'>

Page 68: 2016-09-19; 2016-09-20. Climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)
- Multiple stressors ms:
 - **sent to SHC 2016-09-16**
 - Range limits ms: **SHC's hands**
 - Thermal niche ms: **Lacey and I working on discussion**
 - Stressed in nature MS: Samples to rerun.
 - update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
 - Proteome stability project: **ETA end of the week (5/6 done)**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range

limits paper.

- **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
- Apply for funding. Suitor Travel Grant Deadline is october 31
- **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.
- **Go over thesis layout next time**
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Abstract? I have one written up for NSF post doc fellowship

<div id='id-section69'>

Page 69: 2016-09-21. qPCR redos for 18s rRNA

Table of colonies with unstable HSG as determined by linear regression (18s ~ Temp).

	colony	Df	SS	MS	F	p_value
5	ALA1	1	2123420.91	2123420.91	8.054925	0.0218751
9	Avon19-	1	85577.02	85577.02	5.659013	0.0446244

		1				
15	CJ2	1	860194.07	860194.07	26.944017	0.0008317
55	GF34-1	1	9742336.46	9742336.46	45.449574	0.0001463
85	LPR4	1	2802821.86	2802821.86	14.940584	0.0047729

others: Yates3, Duke8

<div id='id-section70'>

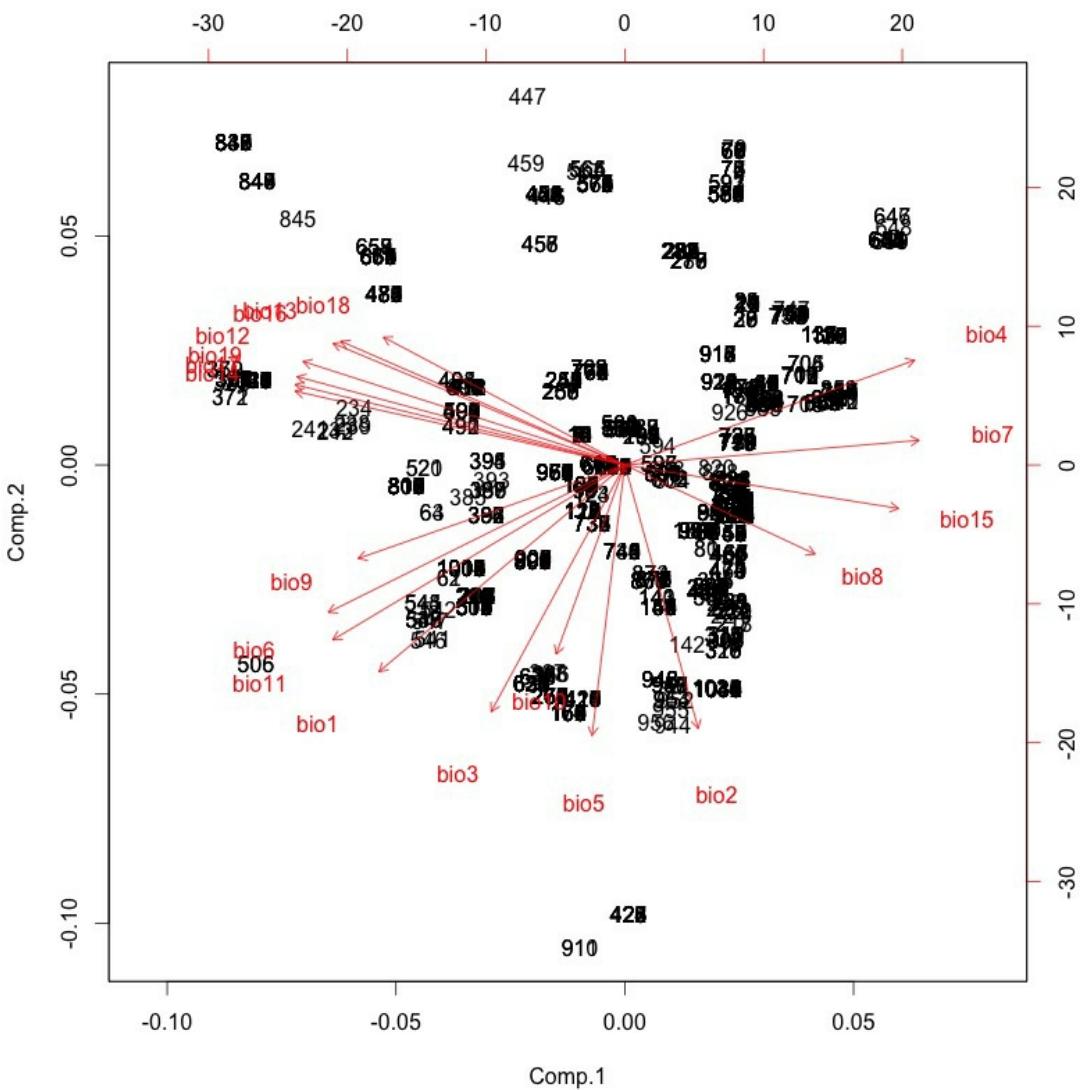
Page 70: 2016-09-26. selecting poplar clones

Overall goal: Make a map highlighting climate gradient and plotting potential sites to select clones from. The magnitude of the points will relate to the GSL.

General workflow

1. Grab climate data and plot all sites
2. Link previous dataset to a another dataset that has empirical GSL from either IH or Burlington.
3. Make map

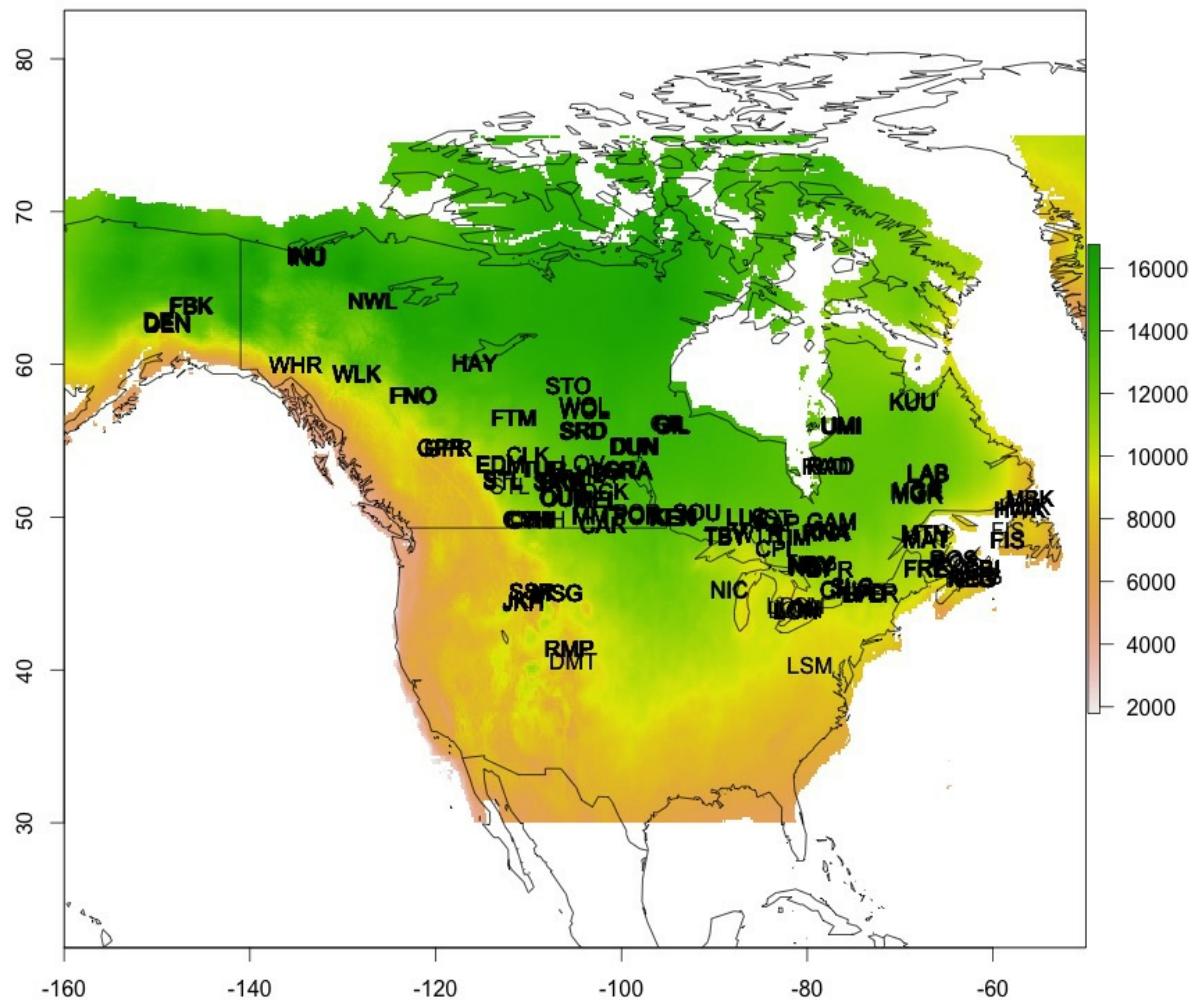
###Climate data



Looks like PC1 (~55%) represents precipitation to temperature seasonality axis and PC2 (19%) represents precipitation to overall temperature axis.

All possible sites

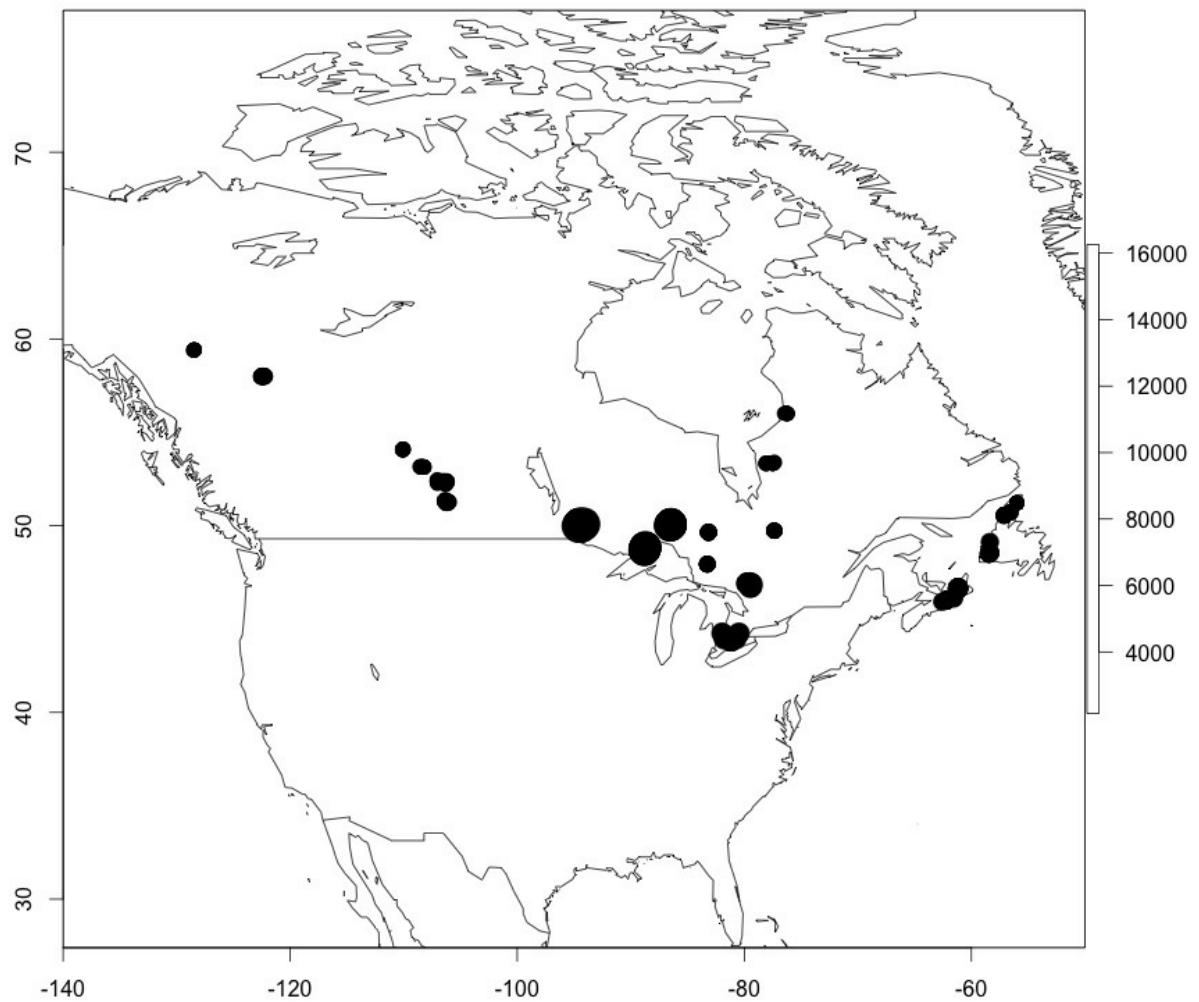
bio4



Subsetted sites

Looks like IH has both BF and BS data but Burlington doesn't

bio4



range of GSLs: 2.016667-4.833333 months

Table for previous fig

PopCode	GSL	BS	BF	months
CBI	84.20000	200.9000	116.7000	2.806667
CLK	62.00000	183.8889	121.8889	2.066667
CPL	67.57143	190.3571	122.7857	2.252381
CYP	61.85714	184.8571	123.0000	2.061905

FIS	77.80000	194.0000	116.2000	2.593333
FNO	70.11111	181.0000	110.8889	2.337037
GAM	64.76000	186.6400	121.8800	2.158667
HWK	68.80000	189.8500	121.0500	2.293333
KAP	68.75000	193.3125	124.5625	2.291667
KEN	145.00000	256.6000	111.6000	4.833333
LLC	136.95000	248.1500	111.2000	4.565000
LON	91.72727	208.7273	117.0000	3.057576
MBK	61.77778	182.3333	120.5556	2.059259
NBY	91.23529	210.4706	119.2353	3.041177
NEG	76.88636	195.6591	118.7727	2.562879
OUT	69.40000	188.1000	118.7000	2.313333
RAD	63.77778	181.7037	117.9259	2.125926
SKN	63.38462	184.8462	121.4615	2.112821
TBY	137.60000	250.1333	112.5333	4.586667
TUR	63.40000	186.5000	123.1000	2.113333
UMI	61.00000	182.0000	121.0000	2.033333
WLK	60.50000	175.0000	114.5000	2.016667

<div id='id-section71'>

Page 71: 2016-09-26 and 2016-09-27.

Climate cascade meeting

1. Project updates:

- Gene expression project: on hold; focusing on 2 manuscripts (multiple stressors and range limits ms)

- **Present some analyses**

- Multiple stressors ms:

- **Working on SHC edits**

- Range limits ms: **SHC lab meeting to go over Thursday September 29th**

- Thermal niche ms: **Lacey and I working on discussion**

- Stressed in nature MS: Samples to rerun.

- update: Curtis can no longer work+ write on project

- in reference to missing samples

- Fit in time to process Curtis' samples.

- **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8**

plates total. 2 days for 8 plates.

- Proteome stability project: **ETA end of the week (5/6 done); database searching**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.
- **Thesis related**
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph

<div id='id-section72'>

Page 72: 2016-09-27. evolution of hsp gxp data analysis

Exploring different approaches

1. PCA decomp bioclim variables I think *a priori* are important and using that in regression vs. just bio5(Tmax)
2. And then building a global model with predictors I think are important (*a priori*) vs constructing a fully complex model.

Exploring hsp gxp parameters from boltzmann fits

Table of correlation between params

	FC_hsc70_1468_max	FC_hsc70_1468_slope
FC_hsc70_1468_max	1.000	0.569
FC_hsc70_1468_slope	0.569	1.000
FC_hsc70_1468_Tm	0.642	0.640
FC_hsp40_541_max	0.398	0.340
FC_hsp40_541_slope	0.104	0.189
FC_hsp40_541_Tm	0.076	0.174
FC_Hsp83_279_max	0.029	-0.154
FC_Hsp83_279_slope	-0.122	-0.079
FC_Hsp83_279_Tm	-0.207	-0.297

It doesn't have basal gxp; including basal and then doing a pca:

Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10	Comp.11	Comp.12
Standard deviation		2.0502371	1.4176264	1.2325728	1.1396205	0.84813044	0.74749858	0.68915615	0.60025005	0.4704591	0.36055794
Proportion of Variance		0.3612359	0.1727056	0.1305593	0.1116100	0.06181701	0.04801793	0.04081483	0.03096329	0.0190207	0.01117205
Cumulative Proportion		0.3612359	0.5339415	0.6645008	0.7761108	0.83792784	0.88594577	0.92676060	0.95772389	0.9767446	0.98791664
Standard deviation		0.294315885	0.232345681								
Proportion of Variance		0.007444064	0.004639294								
Cumulative Proportion		0.995360706	1.000000000								

	Comp.1	Comp.2	Comp.3	Comp.4
hsc70	0.366	0.117	-0.041	-0.400
hsp83	0.271	0.019	-0.238	-0.414
hsp40	0.141	0.309	-0.279	-0.433
FC_hsc70_1468_max	0.284	-0.006	0.529	-0.184
FC_hsc70_1468_slope	0.313	0.112	0.318	-0.110
FC_hsc70_1468_Tm	0.300	-0.063	0.495	0.234
FC_hsp40_541_max	0.210	-0.502	0.039	-0.185
FC_hsp40_541_slope	0.153	-0.521	-0.264	-0.048

FC_hsp40_541_Tm	0.232	-0.493	-0.175	0.173
FC_Hsp83_279_max	-0.321	-0.174	0.305	-0.324
FC_Hsp83_279_slope	-0.355	-0.249	0.168	-0.366
FC_Hsp83_279_Tm	-0.392	-0.124	0.124	-0.276

some stats with pcas of hsp gxp params to see how much it explains CTmax

```
summary(lm(h$K0_temp_worker~paramspc$scores[,1]+paramspc$scores[,2]+paramspc$scores[,3]+paramspc$scores[,4]))
```

Call:

```
lm(formula = h$K0_temp_worker ~ paramspc$scores[, 1] + paramspc$scores[, 2] + paramspc$scores[, 3] + paramspc$scores[, 4])
```

Residuals:

Min	1Q	Median	3Q	Max
-0.90448	-0.46768	0.02901	0.40598	1.08398

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	41.61667	0.10803	385.219	< 2e-16

```

***

paramspc$scores[, 1]  0.15861     0.05269    3.010   0.00548

**

paramspc$scores[, 2] -0.04312     0.07621   -0.566   0.57600

paramspc$scores[, 3]  0.40672     0.08765    4.640   7.41e-05

***

paramspc$scores[, 4]  0.05244     0.09480    0.553   0.58451

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.6206 on 28 degrees of freedom
Multiple R-squared:  0.5272,    Adjusted R-squared:  0.45
96
F-statistic: 7.805 on 4 and 28 DF,  p-value: 0.0002339

```

Mistake: I didnt control for housekeeping gene in basal gxp. redo

```

h<-read.csv("20160927_total_dataset_curated.csv")
basalxp<-h[,4:6]-h[,3]
paramspc<-princomp(scale(cbind(basalxp,h[,7:15])))
summary(paramspc)

```

Importance of components:

	Comp.1	Comp.2	Comp.3	C
Comp.4	Comp.5			
Standard deviation	1.8681865	1.5167796	1.3456615	1.1162675
	0.86282525			
Proportion of Variance	0.3002255	0.1979028	0.1557682	0.1071874
	0.06404021			
Cumulative Proportion	0.3002255	0.4981283	0.6538964	0.7610838
	0.82512400			
	Comp.6	Comp.7	Comp.8	
Comp.9				
Standard deviation	0.82407918	0.71977398	0.55118644	0.45902754
Proportion of Variance	0.05841776	0.04456556	0.02613389	0.01812527
Cumulative Proportion	0.88354177	0.92810732	0.95424121	0.97236649
	Comp.10	Comp.11	Comp.12	
Standard deviation	0.39039533	0.304603562	0.275767584	
Proportion of Variance	0.01311041	0.007981362	0.006541743	
Cumulative Proportion	0.98547690	0.993458257	1.000000000	

```
knitr:::kable(round(paramspc$loadings[,1:4],3))
```

	Comp.1	Comp.2	Comp.3	Comp.4
hsc70	-0.338	0.071	-0.410	-0.299

hsp83	-0.275	0.237	-0.295	-0.234
hsp40	-0.098	0.057	-0.476	-0.391
FC_hsc70_1468_max	-0.316	-0.358	0.172	-0.246
FC_hsc70_1468_slope	-0.195	-0.360	0.211	-0.172
FC_hsc70_1468_Tm	-0.289	-0.347	0.253	-0.044
FC_hsp40_541_max	-0.414	0.177	0.147	0.127
FC_hsp40_541_slope	-0.310	0.265	-0.087	0.461
FC_hsp40_541_Tm	-0.348	0.304	0.081	0.313
FC_Hsp83_279_max	-0.390	0.076	0.292	-0.183
FC_Hsp83_279_slope	0.053	0.439	0.418	-0.286
FC_Hsp83_279_Tm	0.193	0.406	0.290	-0.410

Stats

```
summary(stepAIC(lm(h$K0_temp_worker~paramspc$scores[,1]+paramspc$scores[,2]+paramspc$scores[,3]+paramspc$scores[,4]),direction="both"))
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
--	----------	------------	---------	----------

(Intercept)	41.62969	0.12246	339.931	< 2e-16

paramspc\$scores[, 1]	-0.17806	0.06555	-2.716	0.01119
*				

```
paramspc$scores[, 2] -0.23931      0.08074   -2.964   0.00614
**
paramspc$scores[, 3]  0.15733      0.09101    1.729   0.09486
.
.
.
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.6928 on 28 degrees of freedom
Multiple R-squared:  0.4062,    Adjusted R-squared:  0.34
25
F-statistic: 6.384 on 3 and 28 DF,  p-value: 0.00196
```

```
<div id='id-section73'>
```

Page 73: 2016-09-28. building ultrametric trees

I need to build ultrametric trees to do phylogenetic analyses. They need to be ultrametric to meet the assumptions of Homoscedasticity. I'll be using [BEAST 2.3.1](#). And I'll build 2 types; 1 with BL differences across whole phylogeny and another with species as polytomies.

1. I created a new folder in /Data/Phylogenetics/20160928_beast
2. It has 2 newick files:

20160927_phylogeny_aphaeno_BL_species.newick and

20160927_phylogeny_aphaeno_BL.newick

- 20160927_phylogeny_aphaeno_BL.newick has BL for each colony, and I previously added CJ10, LPR4, and Bing in there; so I have to take them out because there is no sequence data for those samples. New file:
20160927_phylogeny_aphaeno_BL_none.newick

3. It also has this fasta file that was previously parsed:

20160516_Andrew_SNP_sequences.fas

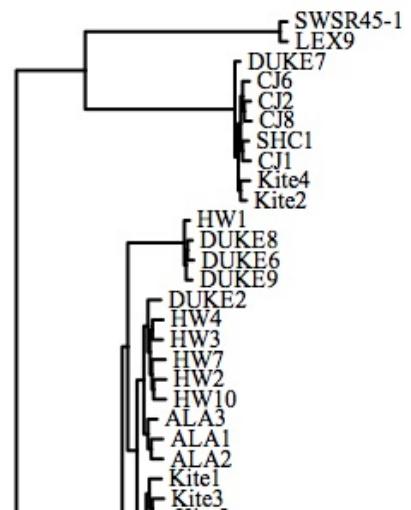
4. In downstream analyses, I got rid of novomessor which I'll do for this fasta too. New file:

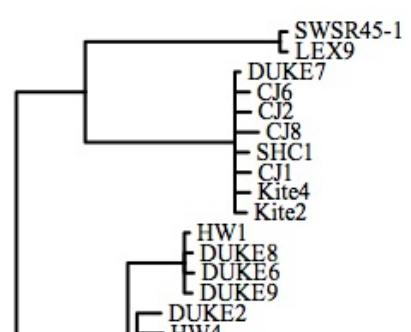
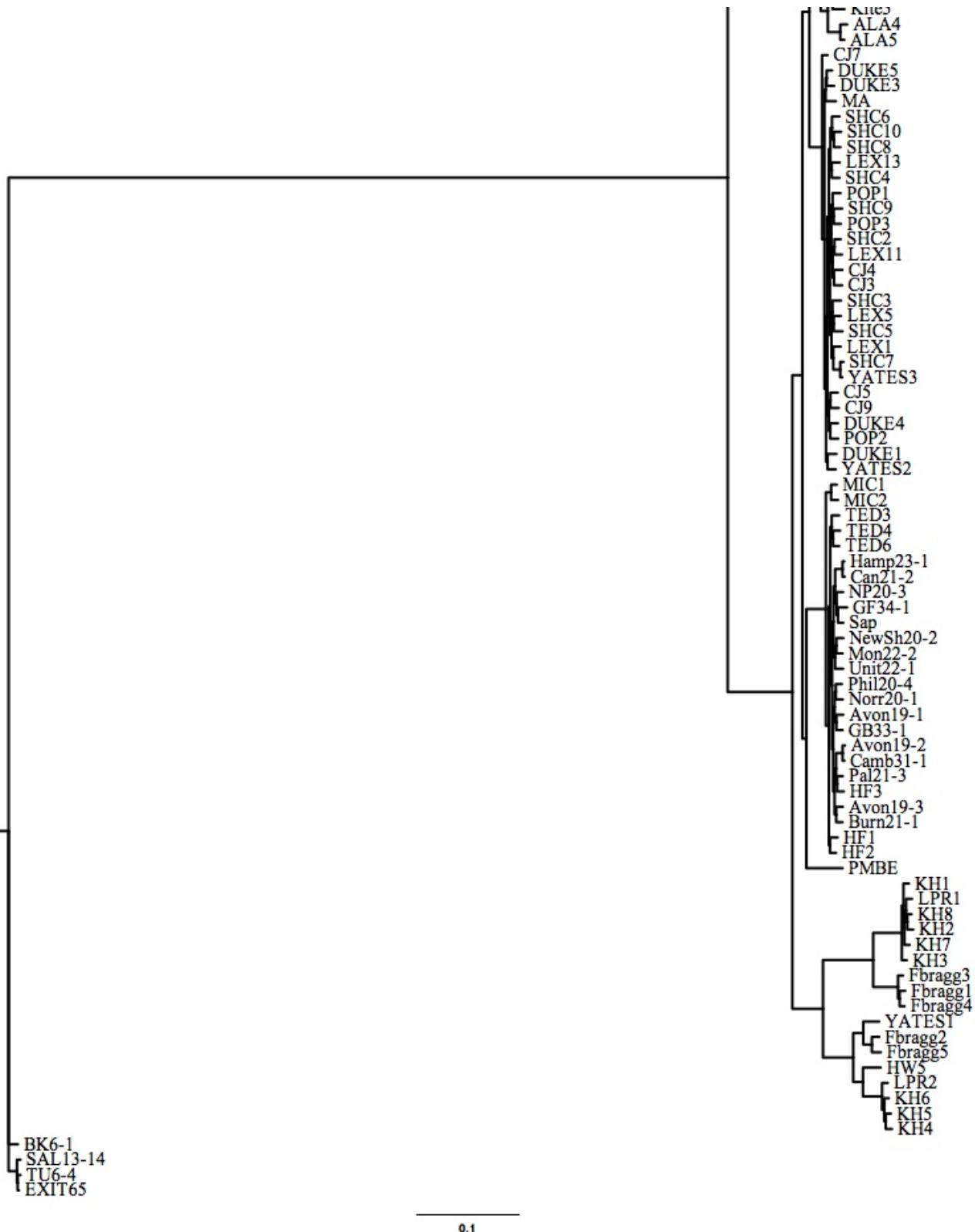
20160516_Andrew_SNP_sequences_nonov.fas

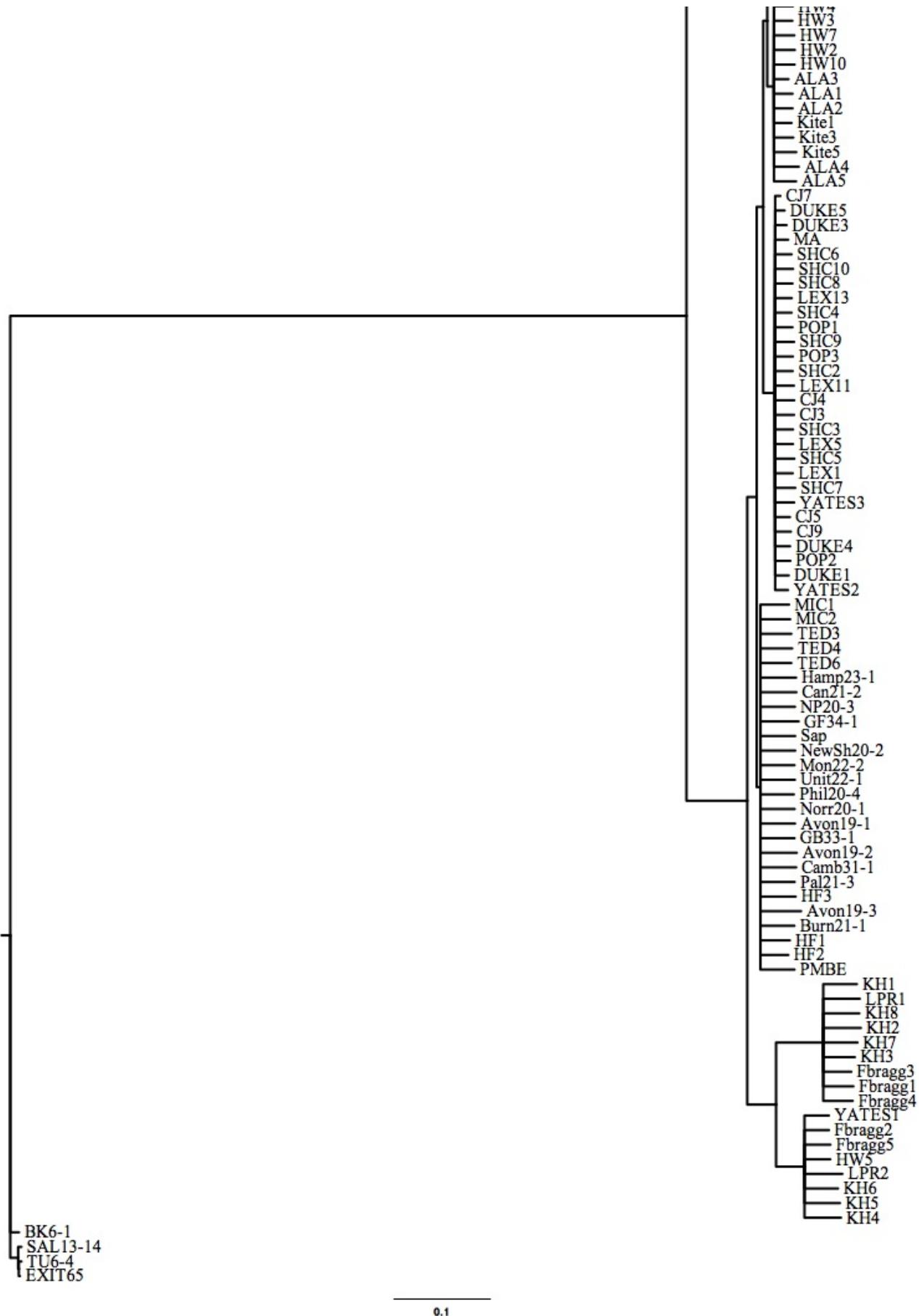
I'll use BEAST on cipres, but I'll need to set up Beauti which sets up the input for BEAST.

The two trees:

Pop level







Yes, so you can't put tree information

into a beast analysis I dont think.

Anyway, here are the settings:

NOte Need to convert fasta into nexus file in order for beauti to read as nucleotide, otherwise it'll read it as amino acids**

1. 1 partition,(1 SNP matrix)
2. tip dates specified as year and before the present
3. gammer site model, GTR+gamma
4. Relaxed clock log normal
5. priors: ingropu = aphaenogaster, outgroup= veromessor;
ucldMean set as mean = 50, sigma = 5 based on fossil data
6. MCMC chain length = 100,000,000

Cannot get it to work. YULE model best for species. But I have pop and species.

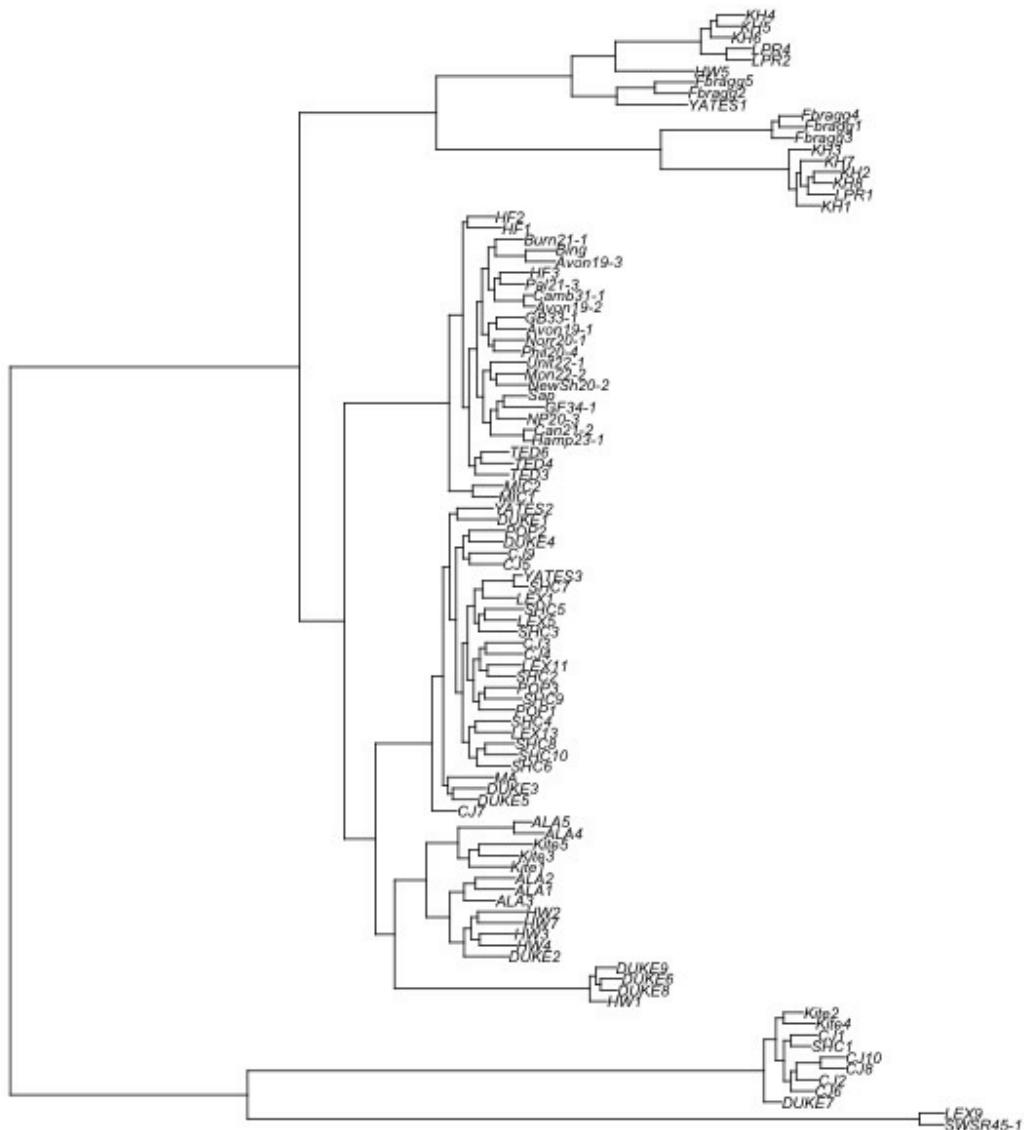
<div id='id-section74'>

Page 74: 2016-09-28. phylogenetic regressions (PGLS) and anovas

Did PGLS in 3 ways:

1. untransformed BL
2. transformed for all tips
3. forced polytomies for species

1. untransformed BL



PGLS 1. untransformed BL

```
pgmod<-gls(K0_temp_worker~ bio5*habitat_v2, correlation =
```

```
corBrownian(phy = aph_onlytree), data = aph_phylo, method  
= "ML")
```

```
summary(pgmod)
```

Generalized least squares fit by maximum likelihood

Model: K0_temp_worker ~ bio5 * habitat_v2

Data: aph_phylo

AIC	BIC	logLik
-----	-----	--------

289.458	302.4838	-139.729
---------	----------	----------

Correlation Structure: corBrownian

Formula: ~1

Parameter estimate(s):

numeric(0)

Coefficients:

	Value	Std. Error	t-value	p-
value				
(Intercept)	39.60864	3.41576	11.595866	0
.0000	.0000	.0000	.0000	.0000
bio5	0.00663	0.00968	0.685406	0
.4947	.4947	.4947	.4947	.4947
habitat_v2flat woods	9.03418	50.92693	0.177395	0
.8596	.8596	.8596	.8596	.8596
bio5:habitat_v2flat woods	-0.02718	0.15809	-0.171921	0
.8639	.8639	.8639	.8639	.8639

Correlation:

(Intr)	bio5	hbt_2w
--------	------	--------

```
bio5           -0.870
habitat_v2flat woods      -0.017  0.041
bio5:habitat_v2flat woods  0.016 -0.042 -1.000
```

Standardized residuals:

	Min	Q1	Med	Q3	M
ax	-1.52993175	-0.23380594	-0.04718187	0.06754746	0.450998
89					

Residual standard error: 2.995514

Degrees of freedom: 100 total; 96 residual

Phyl ANOVA 1. untransformed BL

```
phlaov<-phylANOVA(aph_onlytree,aph_phylo$habitat_v2,aph_p
hylo$K0_temp_worker,p.adj="hochberg")
phlaov
```

\$F

[1] 49.0392

\$Pf

[1] 0.135

\$T

deciduous forest flat woods

```
deciduous forest      0.000000 -7.002799
flat woods           7.002799  0.000000
```

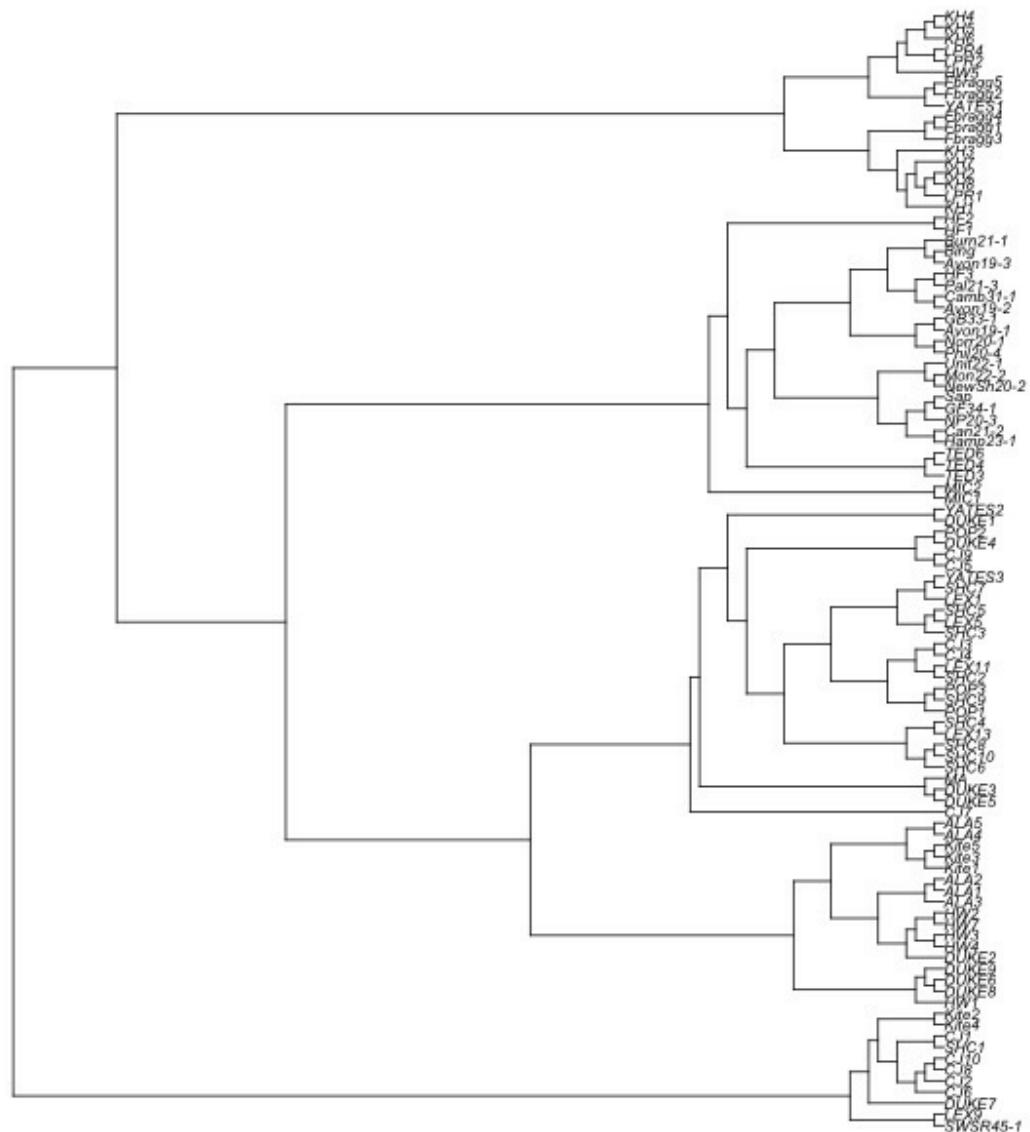
```
$method
```

```
[1] "hochberg"
```

```
$Pt
```

	deciduous forest	flat woods
deciduous forest	1.000	0.135
flat woods	0.135	1.000

2. transformed for all tips



PGLS 2. transformed for all tips

```
ult.tree1<-compute.brlen(aph_onlytree)
plot(ult.tree1,cex=.5)
aph phylo1<-ant dat clim[match(ult.tree1$tip.label,ant da
```

```
t_clim$colony.id2),]  
pgmod1<-gls(K0_temp_worker~ bio5*habitat_v2, correlation  
= corBrownian(phy = ult.tree1),data = aph_phylo1, method  
= "ML")  
summary(pgmod1)
```

Generalized least squares fit by maximum likelihood

Model: K0_temp_worker ~ bio5 * habitat_v2

Data: aph_phylo1

AIC BIC logLik

335.9159 348.9418 -162.958

Correlation Structure: corBrownian

Formula: ~1

Parameter estimate(s):

numeric(0)

Coefficients:

	Value	Std.Error	t-value	p-
value				
(Intercept)	39.94706	4.77385	8.367890	0 .0000
bio5	0.00486	0.01258	0.386220	0 .7002
habitat_v2flat woods	14.08505	51.06703	0.275815	0 .7833
bio5:habitat_v2flat woods	-0.04342	0.15883	-0.273386	0 .7851

Correlation:

	(Intr)	bio5	hbt_2w
bio5	-0.806		
habitat_v2flat woods	-0.008	0.025	
bio5:habitat_v2flat woods	0.008	-0.025	-1.000

Standardized residuals:

Min	Q1	Med	Q3	
				Max
-0.836519826	-0.092391337	0.004278385	0.080275482	0.34
7423662				

Residual standard error: 5.332277

Degrees of freedom: 100 total; 96 residual

PHYLO ANOVA 2. transformed for all tips

```
phlaov2<-phylANOVA(ult.tree1,aph_phylo$habitat_v2,aph_phylo$K0_temp_worker,p.adj="hochberg")
phlaov2
$F
[1] 49.0392

$Pf
[1] 0.234
```

```
$T
```

```
deciduous forest flat woods
```

```
deciduous forest 0.000000 -7.002799
```

```
flat woods 7.002799 0.000000
```

```
$method
```

```
[1] "hochberg"
```

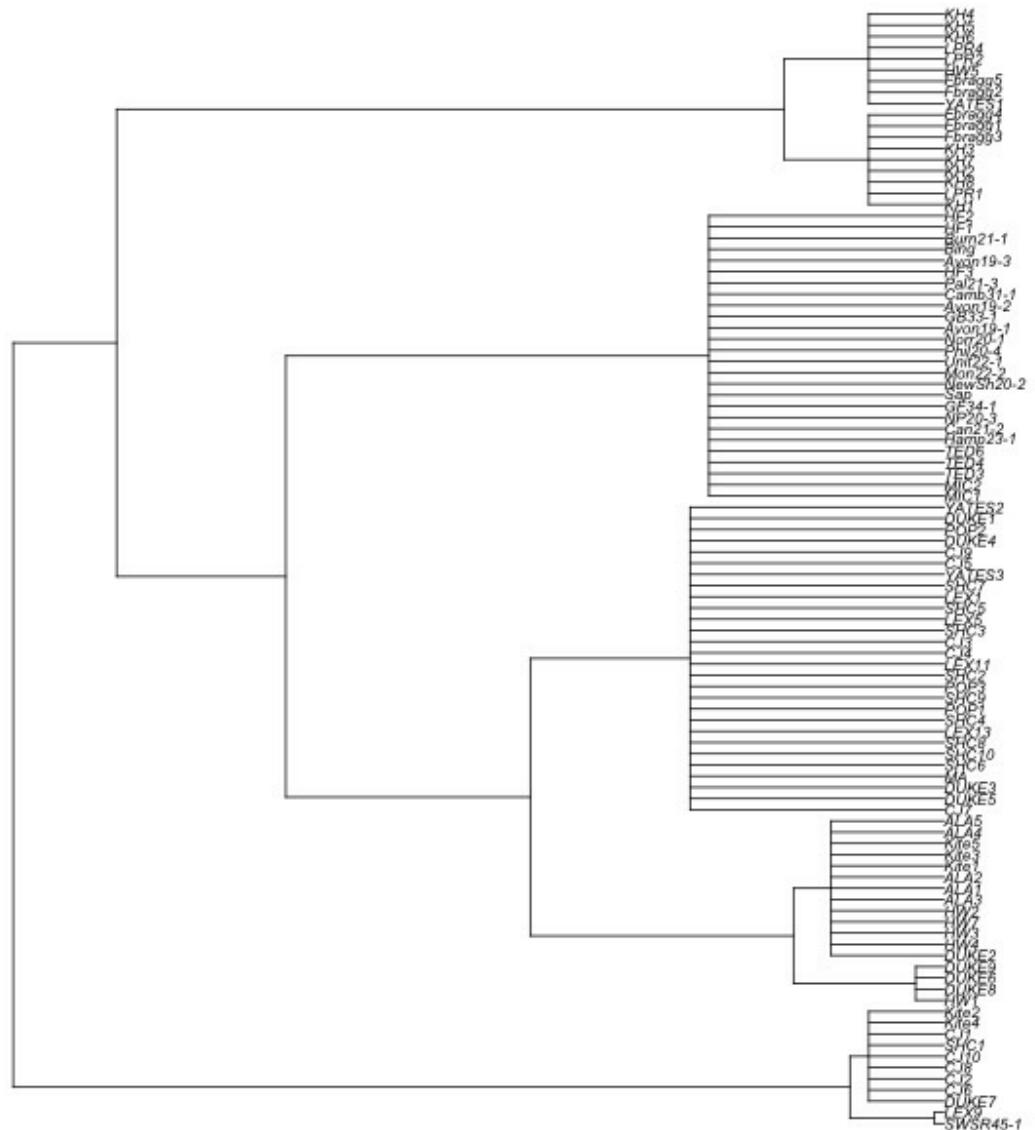
```
$Pt
```

```
deciduous forest flat woods
```

```
deciduous forest 1.000 0.234
```

```
flat woods 0.234 1.000
```

3. forced polytomies for species



PGLS 3. forced polytomies for species

```
plot(aph_onlytree)
nodelabels(cex=.5)
ant_tree_root1<-collapse.to.star(ant_tree_root,192) # flo
```

```
r  
ant_tree_root2<-collapse.to.star(ant_tree_root1,184) #ash  
ant_tree_root3<-collapse.to.star(ant_tree_root2,158) #pic  
ea  
ant_tree_root4<-collapse.to.star(ant_tree_root3,131)# rud  
is  
ant_tree_root5<-collapse.to.star(ant_tree_root4,119) # mi  
amiana  
ant_tree_root6<-collapse.to.star(ant_tree_root5,116) #lam  
ellidens  
ant_tree_root7<-collapse.to.star(ant_tree_root6,104) # fu  
lva  
ant_tree_root8<-collapse.to.star(ant_tree_root7,103) # te  
nn  
#ant_tree_root9<-collapse.to.star(ant_tree_root8) # outgr  
oup  
plot(ant_tree_root8)  
ult2.tree<-compute.brlen(ant_tree_root8)  
plot(ult2.tree)
```

```
aph_phylo2<-ant_dat_clim[match(ult2.tree$tip.label,ant_da  
t_clim$colony.id2),]  
pgmod2<-gls(K0_temp_worker~bio5*habitat_v2, correlation =  
corBrownian(phy = ult2.tree),data = aph_phylo2, method =  
"ML")  
summary(pgmod2)
```

Generalized least squares fit by maximum likelihood

Model: K0_temp_worker ~ bio5 * habitat_v2

Data: aph_phylo2

AIC BIC logLik

255.776 268.8019 -122.888

Correlation Structure: corBrownian

Formula: ~1

Parameter estimate(s):

numeric(0)

Coefficients:

	Value	Std. Error	t-value	p-value
(Intercept)	37.82400	2.043758	18.507082	0 .0000
bio5	0.01175	0.005942	1.978037	0 .0508
habitat_v2flat woods	22.58447	12.917075	1.748420	0 .0836
bio5:habitat_v2flat woods	-0.06971	0.039823	-1.750585	0 .0832

Correlation:

	(Intr)	bio5	hbt_2w
bio5	-0.881		
habitat_v2flat woods	-0.132	0.148	
bio5:habitat_v2flat woods	0.132	-0.149	-0.999

Standardized residuals:

	Min	Q1	Med	Q3	M
ax					
67	-2.24865470	-0.26276358	0.05811258	0.26246427	0.990708

Residual standard error: 1.836591

Degrees of freedom: 100 total; 96 residual

PHYLO ANOVA 3. forced polytomies for species

```
aph_phylo<-ant_dat_clim[match(ult2.tree$tip.label,ant_dat  
_clim$colony.id2),]  
aph_phylo$habitat_v2<-droplevels(aph_phylo$habitat_v2)
```

```
phlaov3<-phylANOVA(ult2.tree,aph_phylo$habitat_v2,aph_phy  
lo$K0_temp_worker,p.adj="hochberg")
```

phlaov3

\$F

[1] 49.0392

\$Pf

[1] 0.183

\$T

	deciduous forest flat woods	
deciduous forest	0.000000	-7.002799
flat woods	7.002799	0.000000

\$method

[1] "hochberg"

\$Pt

	deciduous forest flat woods	
deciduous forest	1.000	0.183
flat woods	0.183	1.000

Intepretting a phylogenetic ANOVA [here](#)

The way the phylogenetic ANOVA (sensu Garland et al. 1993 ; Syst. Biol.) works is by first computing a standard ANOVA, and then comparing the observed F to a distribution obtained by simulating on the tree under a scenario of no effect of x on y. This "accounts for" the tree in the sense that it attempts to account for the possibility that species may have similar y conditioned on x because x influences y; or because they share common history and are thus similar by virtue of th

is history (and
not at all due to x)

It is not particularly surprising that your P-value was lower in the phylogenetic ANOVA than in your regular ANOVA. In general, the effect of the phylogenetic ANOVA on P depends on the distribution of the factor, x. If x is clumped on the tree, than the P-value of a phylogenetic ANOVA will tend to be higher than a regular ANOVA. By contrast, if x is overdispersed phylogenetically, the P-value of the phylogenetic ANOVA will tend to be lower than the regular ANOVA.

I hope this is of some help.

- Liam

--

Liam J. Revell

University of Massachusetts Boston

web: <http://faculty.umb.edu/liam.revell/>

email: ***@umb.edu

blog: <http://phytools.blogspot.com>

```
<div id='id-section74.1'>
```

2016-09-28. SHC suggestion: ancestral trait reconstruction -> regressions/anovas

[summary pdf figs](#)

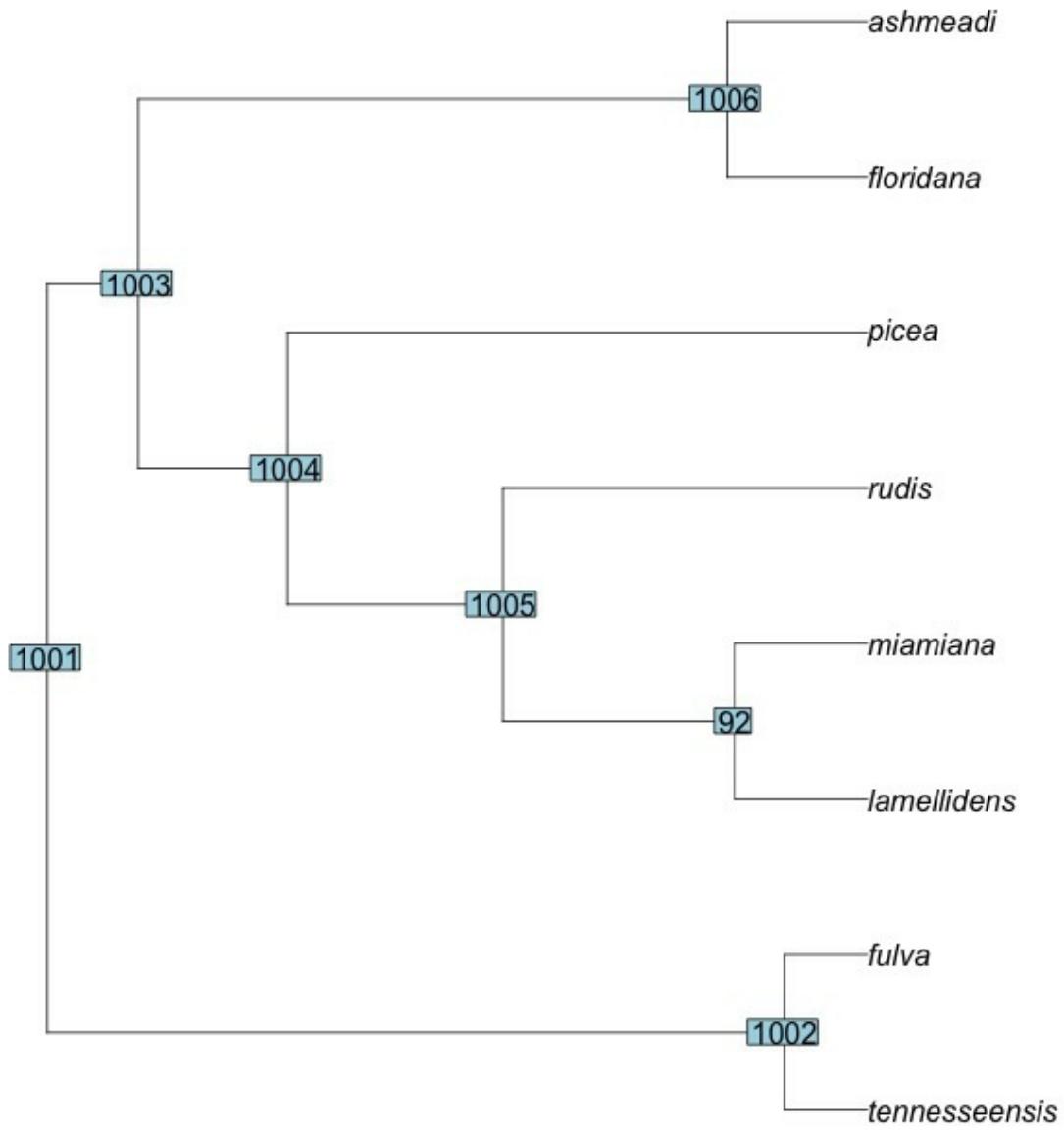
```
<div id='id-section74.2'>
```

2016-09-29. PIC

Dataset

	Species	CTmax	Tmax	Habitat
1	ashmeadi	42.80833	324.0000	FW
2	floridana	42.76852	323.7778	FW
6	picea	40.50096	262.9615	DF
7	rudis	41.33808	300.3214	DF
5	miamiana	40.95128	329.3846	DF
4	lamellidens	42.09375	318.2500	DF
3	fulva	41.01222	310.5556	DF
8	tennesseensis	40.75000	311.0000	DF

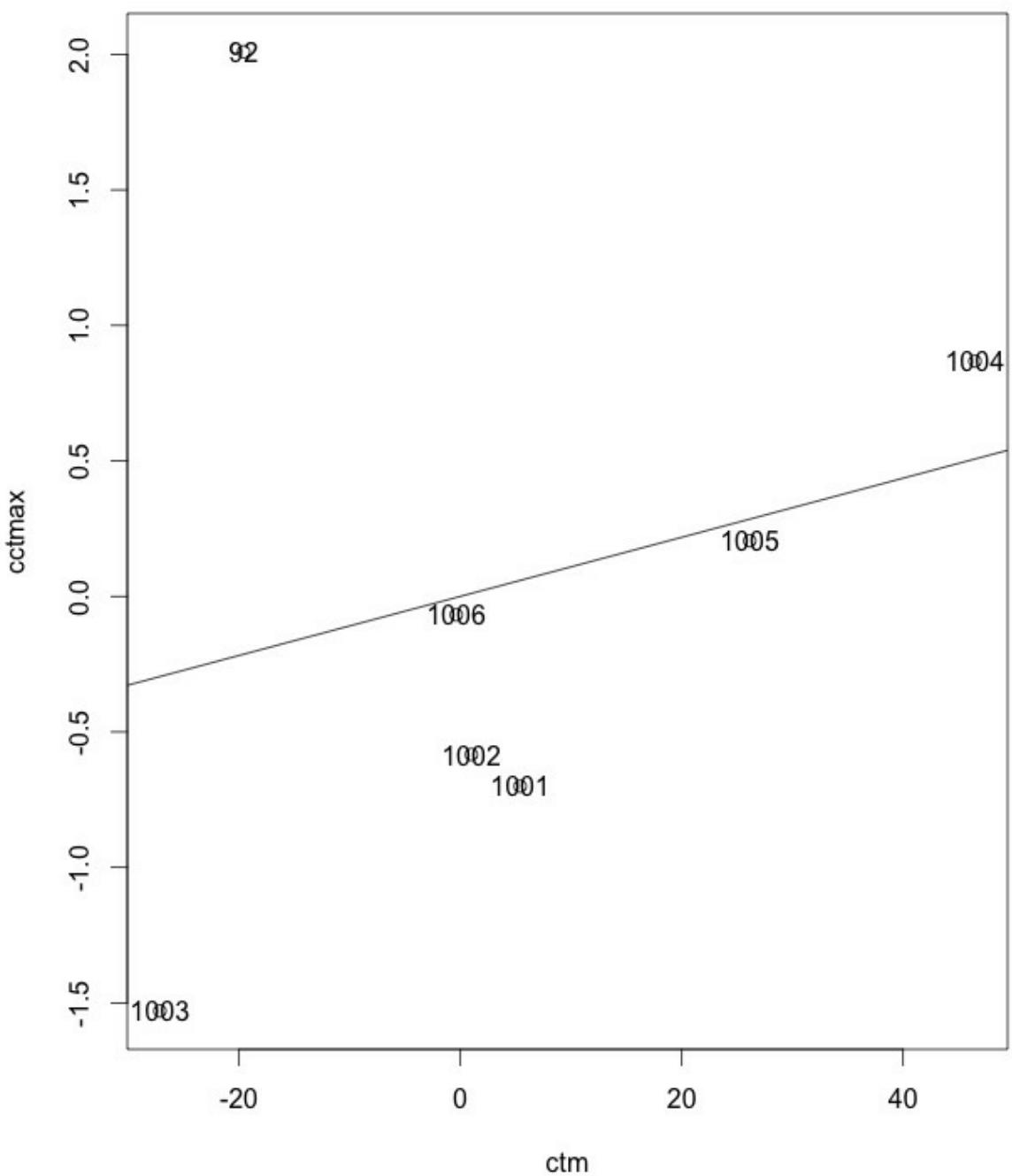
NOdes of phylogeny



Independent contrast estimates for
CTMAX

Node	cctmax
1001	-0.7004417
1002	-0.5834076
1003	-1.5296702
1004	0.8678094
1005	0.2051669
92	2.0095026
1006	-0.0679396

**Better fig with contrsts of CTmax v
Tmax with points labeled by nodes**



<div id='id-section75'>

Page 75: 2016-10-03 and 2016-10-04.Climate cascade meeting

1. Project updates:

- Gene expression project:
 - **Go over analyses:**
 - Phylo anova, PGLS, ancestral trait reconstruction
 - GXP: basal expression, PGLS with CTmax and gxp parameters
 - **Go over figure layout for ms**
 - Left to do: QC and analyze hsp83 and hsp40
- Multiple stressors ms:
 - **Ask about SHC comments on confusion of mismatch membrane stability**
- Range limits ms: **SHC lab gave verbal edits:**
 - focus on 1 end of thermal niche breadth(although it is nice to mention it because CTmin decreases across lat)-CTmin.
 - Discussion needs to talk about cold adaptation; why trade-offs?
 - Walk through results better
- Thermal niche ms: **Lacey and I working on discussion**
- Stressed in nature MS: Samples to rerun.

- update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- Proteome stability project: **no clue what status is**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - **Talk title: Northern range limits of a common forest ant is associated with trade-offs in cold physiology**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.
- **Thesis related FORMS FOUND HERE**
 - Formatting:
 - Introduction (> 3 pages), manuscripts, then

synthesis/conclusion (~3 pages) ; SHC and NJG

agree

- Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph

- Deadlines:

1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.
-

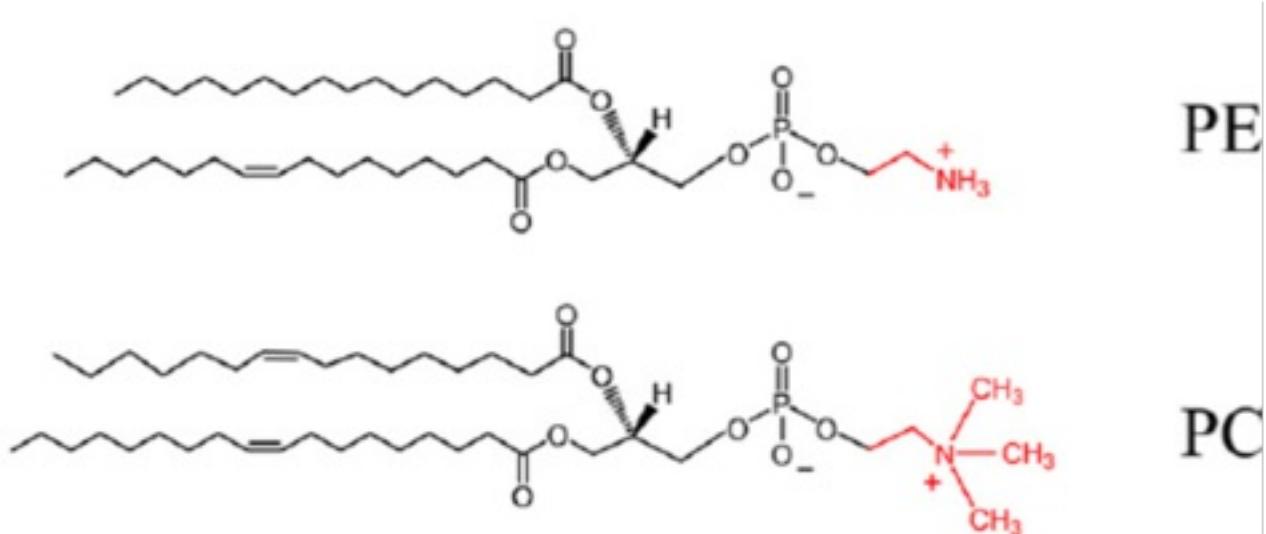
<div id='id-section76'>

Page 76: 2016-10-03. Membrane stability

Trying to get how membrane stability is altered among different stressors. **Two things can change/alter membrane fluidity; glycero-phospholipid head groups**

(phosphatidylethanolamine, PE; phosphatidylcholine PC) and lipid saturation(saturated vs unsaturated). In warmer environments, higher PC and lipid saturation confer homeostasis. Cooler environments = PE and unsaturated lipids. Membrane fluidity for desiccation resistance usually covaries with cold acclimation/adaptation.

PC bind 10-12 water molecules and PE binds 7-8. PE binds less water and it should be enriched under desiccation stress.



Going through some of the literature and what they found.

Hayward et al. 2014 has a nice explanation

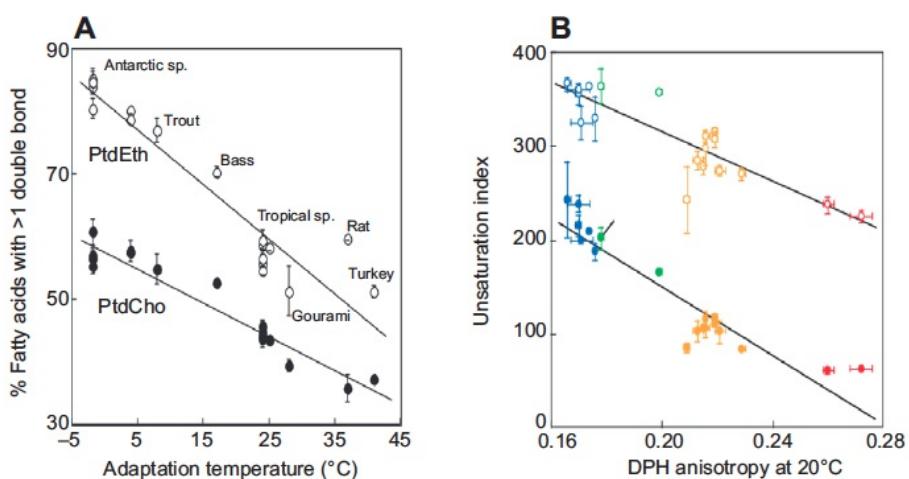


Fig. 1. A comparative analysis of brain synaptic membrane lipid composition and physical structure in relation to habitat or thermo-regulated body temperature. (A) A linear, orderly relationship for five fish species, one bird and one mammal in the percentage of fatty acids that were unsaturated in phosphatidylethanolamine (PtdEth) and phosphatidylcholine (PtdCho). (B) The unsaturation index of membrane phospholipids, which is a measure of the number of unsaturated bonds in fatty acids of membrane phosphoglycerides, is related to a measure of membrane static physical structure, namely the fluorescence anisotropy of the membrane probe 1,6-diphenyl-hexatriene (DPH). Lower values of anisotropy indicate reduced constraint on intra-membrane molecular mobility and vice versa. High values of the unsaturation index were correlated with reduced membrane rigidity or high membrane 'fluidity'. Species were divided into four groups: Antarctic (blue), temperate (green), tropical (orange) and homeothermic (red). For graphs in both A and B, the correlation coefficients are all highly significant ($P<0.0001$) (adapted from Logue et al., 2000).

In A, those that have greater unsaturated fatty acids, are more cold tolerant (operative body temperature to be exact). More fatty acid content negatively correlated with DPH anisotropy at 20 C (something that distorts light). DPH related to membrane rigidity and fluidity; high values = reduced constraint on intra-membrane molecular mobility. **So High unsaturated fatty acid index is related to reduced membrane rigidity or high membrane fluidity (lower values of DPH anisotropy)**

Cossins and Prosser 1978 PNAS shows:

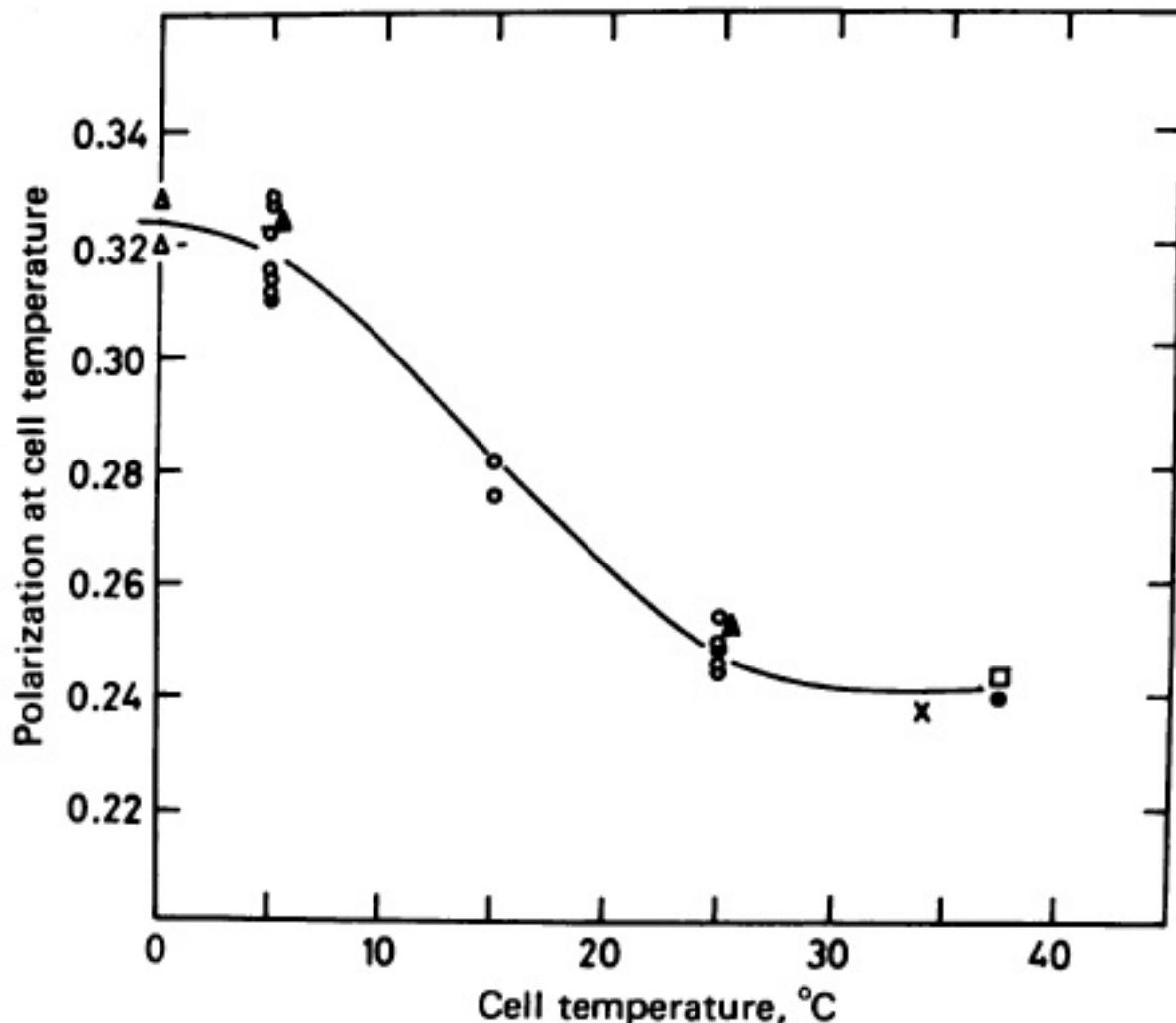


FIG. 2. Effect of adaptation or acclimation at different temperatures upon membrane viscosity expressed as polarization measured at their respective acclimation, environmental, or body (i.e., cellular) temperatures. Arctic sculpin (Δ), goldfish (\circ), green sunfish-bluegill hybrid (\blacktriangle), desert pupfish (\times), rat (\bullet), and hamster (\square). Each point represents an individual animal.

High membrane fluidity (polarization) is higher in more cold acclimated fish. and...

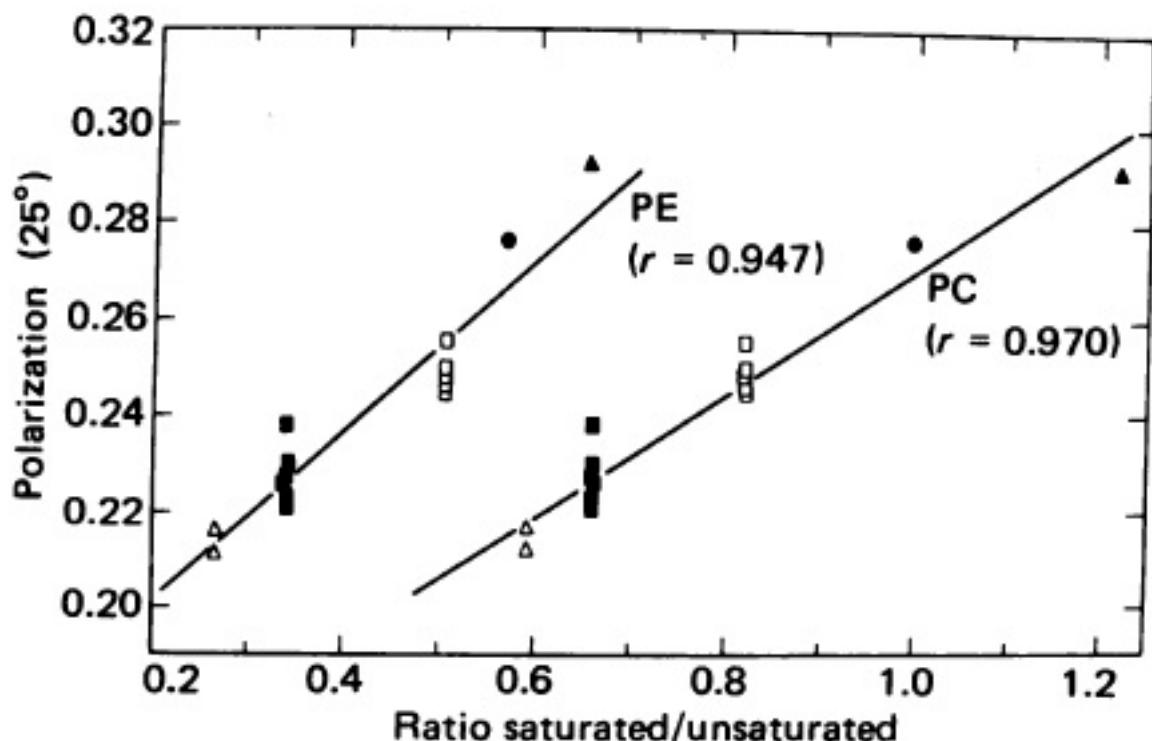


FIG. 3. Relationship between viscosity of synaptosomal membranes of various fish species and rat, and the ratio of saturated to unsaturated fatty acids for choline phosphoglycerides (PC) and ethanolamine phosphoglycerides (PE). Each point represents an individual animal. Membrane viscosity is expressed as polarization of diphenylhexatriene measured at 25° and fatty acid data are from Table 2. Arctic sculpin (Δ), goldfish acclimated at 5° (\blacksquare), goldfish from 25° (\square), desert pupfish (\bullet), rat (\blacktriangle).

High membrane fluidity is related to higher saturated:unsaturated. I think this makes sense, high unsat FA makes the polarization smaller.

Cooper et al. 2014; Functional ecology finds that acclimation influences PE/PC ratios

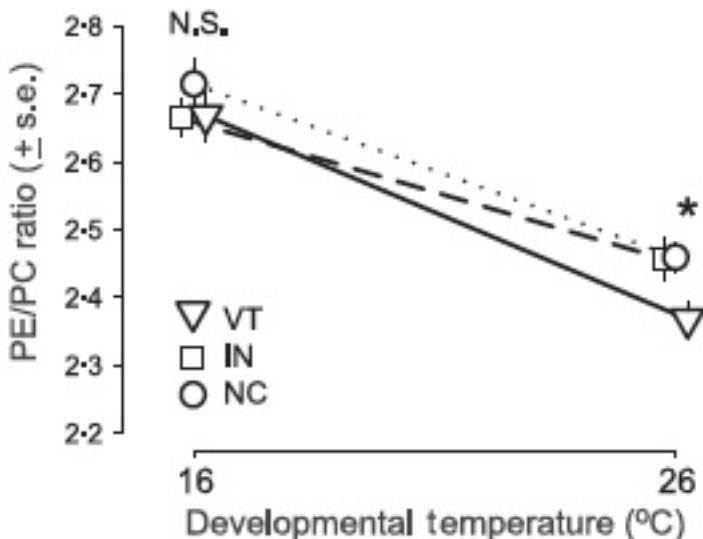


Fig. 1. Flies from all populations have a strong plastic response of phosphatidylethanolamine (PE)/phosphatidylcholine (PC) to developmental temperature, but flies from the most variable thermal environment (VT) respond more strongly to warm developmental conditions. Following development at 16 °C, flies from the three populations did not differ significantly in the mean PE/PC ($F_{2,36} = 0.679$, $P = 0.5134$), but when developed at 26 °C, flies from the three populations differed significantly in PE/PC ($F_{2,36} = 3.461$, $P = 0.04$), with VT having lower PE/PC than Indiana or North Carolina flies.

Summary table of directions of effects for stressors on membrane fluidity

Stress.type	Heat	Cold	Desiccation	pH
membrane.fluidity	decrease	increase	increase	
membrane.rigidity	increase	decrease	decrease	
PC	increase	decrease	decrease	
PE	decrease	increase	increase	
PE.PC.ratio	decrease	increase	increase	
saturated.FA	increase	decrease	decrease	

unsaturated.FA	decrease	increase	increase	
saturated.unsaturated.ratio	increase	decrease	decrease	

refs:

Hazel, J. R., and E. Eugene Williams. 1990. The role of alterations in membrane lipid composition in enabling physiological adaptation of organisms to their physical environment. *Progress in Lipid Research* 29:167-227.

Hazel, J. R., S. J. McKinley, and M. F. Gerrits. 1998. Thermal acclimation of phase behavior in plasma membrane lipids of rainbow trout hepatocytes. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 275:R861-R869.

Cooper, B. S., L. A. Hammad, and K. L. Montooth. 2014. Thermal adaptation of cellular membranes in natural populations of *Drosophila melanogaster*. *Functional Ecology* 28:886-894.

Cossins, A. R., and C. L. Prosser. 1978. Evolutionary adaptation of membranes to temperature. *Proceedings of the National Academy of Sciences of the United States of America* 75:2040-2043.

Hayward, S. A. L., B. Manso, and A. R. Cossins. 2014. Molecular basis of chill resistance adaptations in poikilothermic animals. *The Journal of Experimental Biology* 217:6-15.

Holmstrup, M., K. Hedlund, and H. Boriss. 2002. Drought acclimation and lipid composition in *Folsomia candida*: implications for cold shock, heat shock and acute desiccation stress. *Journal of Insect Physiology* 48:961–970.

<div id='id-section77'>

Page 77: 2016-10-04 Lab Safety Officer (LSO) meeting.

Department of Risk Management and Safety- Francis Churchill mainly speaking

Agenda:

1. News and updates

- staff changes- new lab safety coordinator
- lab fires at uvm
 - Chemistry- no blame; removing syringe that had fire . no evac, not a big fire
 - Votey building - small fire; no damage no hurt; alcohol near a burner-fire
 - faculty said not to leave in 1 class; that is bad.

You should leave if fire alarm goes off.

- Explosion at U Hawaii
 - **post doc in lab; working with pressure vessel (creating fuel for bacteria to make biofilms and biofuels); mixing hydrogen and oxygen and some carbon dioxide. Did over and over, and had minor issues; but in march it blew up. Took her arm off. Lab had good safety; but regulatory agents don't know how stuff get mixed; we all need to get better at hazard assessment. Fined \$115,000; \$750,000 building damage. Brought up issue of coverage of insurance for post doc researchers**
 - Violations: Failed to provide a safe workplace; failed to ensure employees to follow proper procedures. Chemical Hygiene plan did not include SOPs for relevant safety.
- Fine at Oregon
 - \$275,000 by EPA for mismanagement of chemicals; did not get rid of their chemicals; no labeled; every bottle out there should be labeled.

We're going to be inspected by the US EPA and the state department(DEC)

risk control governance: 22% of safety trainings are not being

completed; high for lab supervisors!!! Lab safety notebooks need to be updated.

2. uvm police services

Office Sue Roberts: Work place violence. Active shooters? Training to safeguard to active shooters. How to respond?

Showing a video: Run, Hide, Fight. Know how to exit your building(how to get in or out). First responders don't tend to the injured; secondary responders will.

Systems in place:

1. own police agency on campus with master keys and card access; allowing quick response times.
2. CAT Alerts.
3. Emergency blue lights- direct connect to UVM police.

Violence in the workplace

1. Detect early, to get resources to person with alarming behavior.
2. 2 teams on campus meet weekly and monthly; safety response team (discuss faculty and staff on campus) and care team (focus on needs of students). There is an anonymous care form (please give tons of info).

Phone systems

- for lan line; 911 goes to uvm police and they know where you are, send office to location
- for cell phones it goes to 911 call center in williston or it could go to brattoboro. Pay attention to where you are because phones don't give you pinpoint accuracy. Know street address.
- **Put UVM police into contact list: 802-656-3473**

3. summary of audits

There are top 10 audit deficiencies: **FILL OUT DATES; use yellow waste label**

1. safety training incomplete
2. **chemical waste is older than 6 months (we need a sticker and they need to collect the waste)**
3. mislabelling in chemical waste containers (completely fill out tags!!)
4. reports of hazard assessments are not available
5. lab online inventory (HCOC) has not been updated wtihin 6 months
6. chemical containers not fully labeled (**Waste and non waste need labels**)
7. research samples not albeled properly: sample ID, hazards, date material made
8. info on emergency contact door is not current
9. lab monthly inspection not done

10. eyewash flush log not visible and current.

Creating corrective actions: Stuff for you to fix.:

4. lab safety basics

UVM lab safety; monthly self inspection: Policy, all labs must do monthly inspection. Document on checklist.

If you don't have one, they are distributed out to departments.

1. Defrost freezers. check website so that our freezer is not ruined.
2. Label samples
3. annual refresher training (everybody complete it?)
4. can write discrepancies.

Labels: You need manufacturer's label and don't need anything else, just sign and date it.

Safety Audits at UVM: LabCliq. LSO can do corrective actions but the PI has to use Labcliq to verify online. Then PI gets email.

What trainings do you need? [HERE](#)

- Take all things that are applicable to your laboratory!!
- Green section 6 classes+ Annual refresher training. 4 online

safety trainings and 2 classroom safety trainings.

- **Red section** Fire safety training.

Lab safety notebook webpage [HERE](#)

5. CITI training opportunity

6. Q & A

- Fraudulent calls: Target international faculty and staff referring to immigration status, healthcare, taxes. If you get calls, notify police services to set up trap on that phone. Check for scams on UVM police website

<div id='id-section78'>

Page 78: 2016-10-05. Hsp gxp function valued trait fig

Boltzmann function and fit to dataset

```
Boltz<-function(data=x){  
  B<-nls(gxp ~ (1+(max-1)/(1+exp((Tm-T)/a))),data=data, s  
tart=list(max=80,Tm=35,a=1.05), trace=TRUE,control=nls.co  
ntrol(warnOnly = TRUE, tol = 1e-05, maxiter=1000))  
  #summary(B)
```

```

    return(summary(B)$parameters)
}

T<-c(25,28,30,31.5,33,35,36.5,38.5,40,41)
gxp<-c(1.050927323,
1.795269722,
2.394945916,
2.025719648,
5.995719441,
12.75328258,
35.0828896,
44.80226791,
63.64704198,
67.607218)

FB1<-as.data.frame(cbind(T,gxp));FB1

Boltz(FB1)
knitr::kable(Boltz(FB1))

```

function that estimates values based on Boltzmann parameters

```

fud<- function(T=seq(25,70,.1),Tm=40,slope=1.8,max=50){
  y<- 1+ (max-1)/(1+exp(((Tm-T)/slope)))
  return(y)
}

```

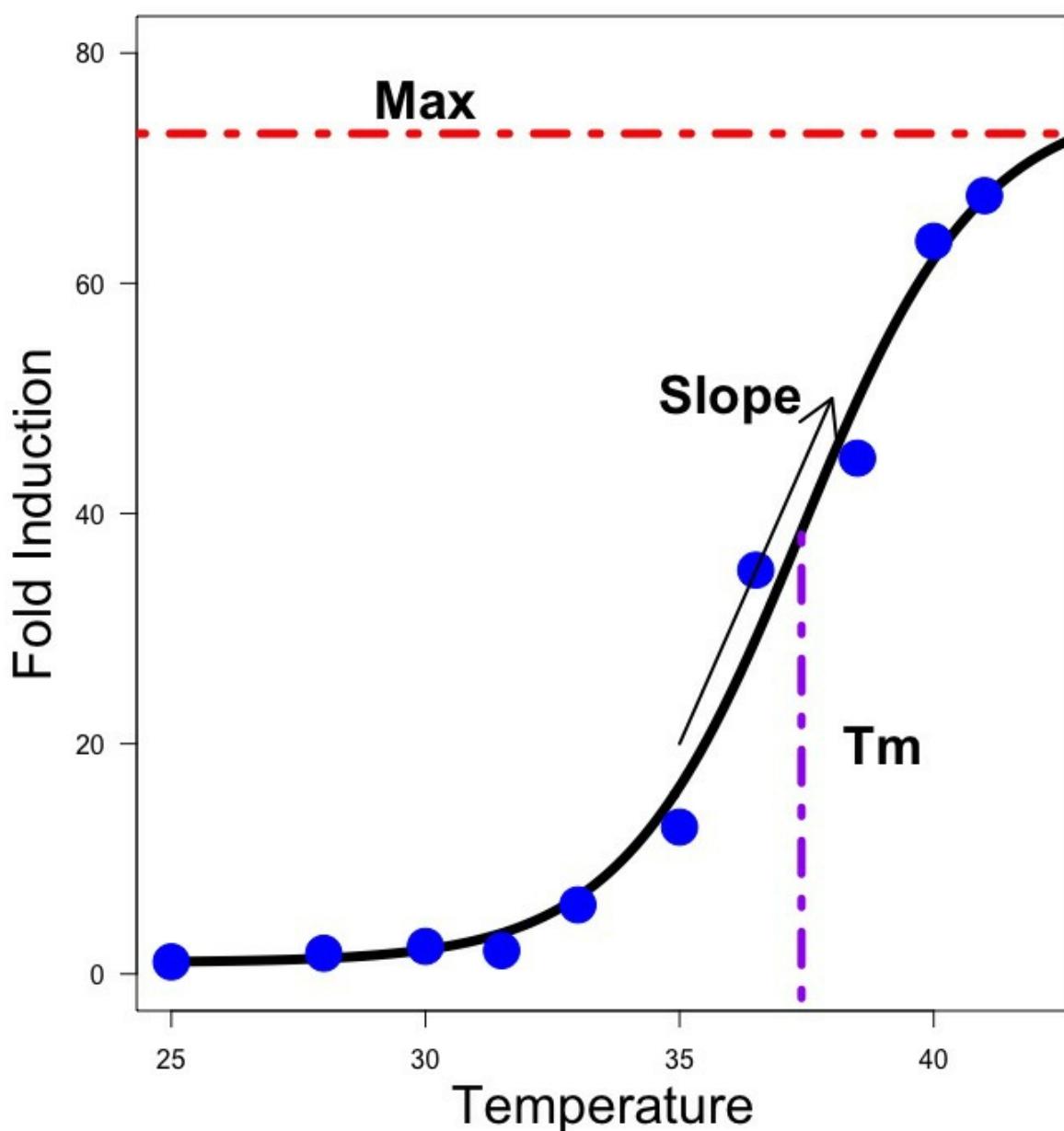
parameter fits

	Estimate	Std. Error	t value	p value
max	76.179606	8.0617514	9.449511	0.0000310
Tm	37.432787	0.5585165	67.021804	0.0000000
a	1.765851	0.3248254	5.436310	0.0009701

With units and real data

```
plot(seq(0,70,.1),fud(T=seq(0,70,.1)),col="blue",type="n"
,ylim=c(0,80),las=1,xlab="",ylab="",xlim=c(25,42))

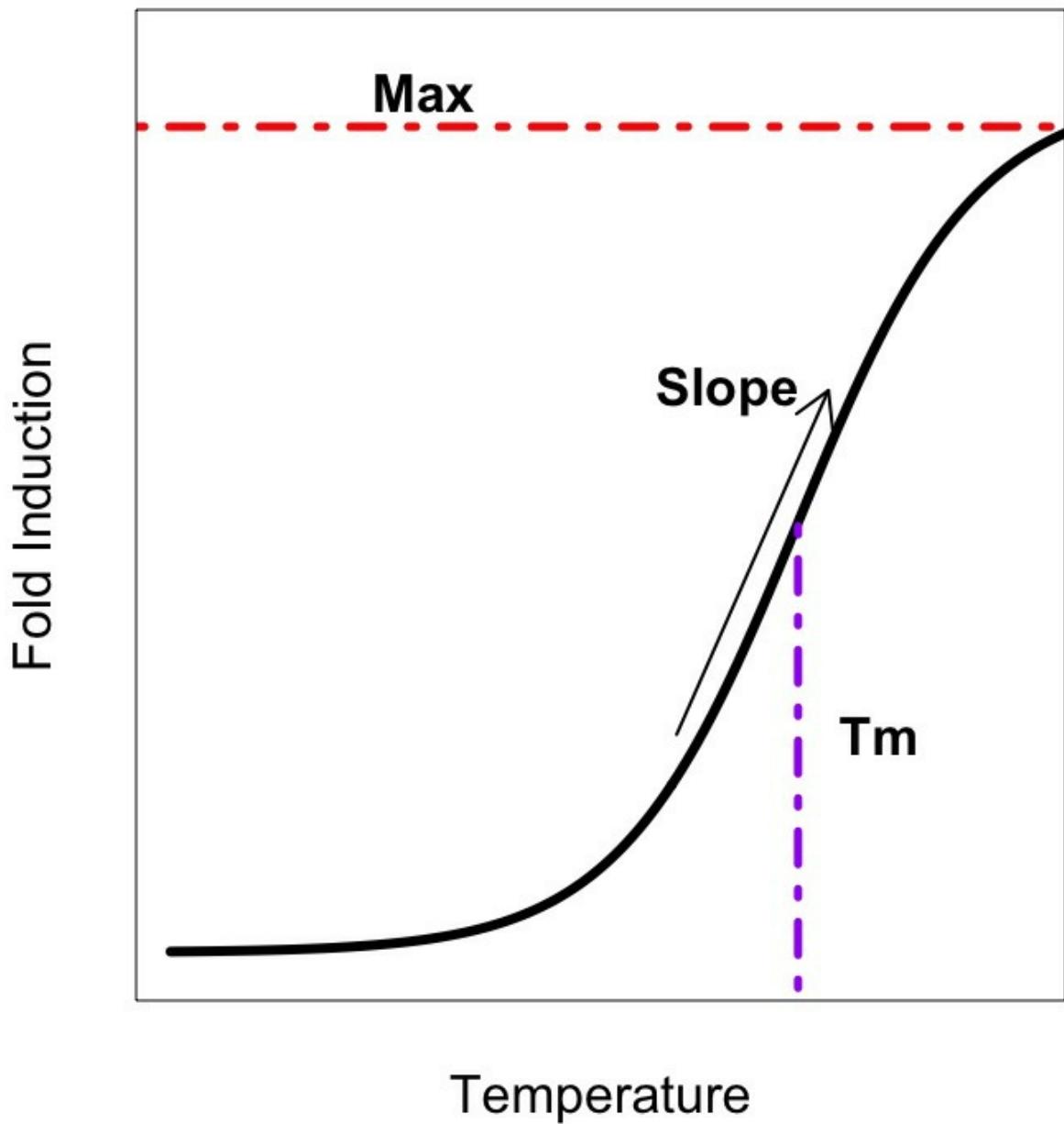
mtext("Fold Induction", side=2, line=2.5, cex=2)
mtext("Temperature", side=1, line=2.7, cex=2)
lines(seq(25,70,.1),fud(Tm=37.4,slope=1.76,max=76),lwd=6)
points(FB1$T,FB1$gxp,pch=19,col="blue",cex=3)
lines(c(37.4,37.4),c(-10,39),lwd=5,lty="dotdash",col="purple")
abline(h=73,lty="dotdash",col="red",lwd=5)
arrows(35,20,38,50,code=2,lwd=2, )
text(c(39,30,36),c(20,76,50),c("Tm","Max","Slope"),font=2
,cex=2)
```



No units or real data

```
plot(seq(0,70,.1),fud(T=seq(0,70,.1)),col="blue",type="n"
,ylim=c(0,80),las=1,xlab="",ylab="",xlim=c(25,42),axes=FALSE)
mtext("Fold Induction", side=2, line=2.5, cex=2)
mtext("Temperature", side=1, line=2.7, cex=2)
```

```
lines(seq(25,70,.1),fud(Tm=37.4,slope=1.76,max=76),lwd=6)
#points(FB1$T,FB1$gxp,pch=19,col="blue",cex=3)
lines(c(37.4,37.4),c(-10,39),lwd=5,lty="dotdash",col="purple")
abline(h=73,lty="dotdash",col="red",lwd=5)
arrows(35,20,38,50,code=2,lwd=2,)
text(c(39,30,36),c(20,76,50),c("Tm","Max","Slope"),font=2
,cex=2)
box()
```



<div id='id-section79'>

**Page 79: 2016-10-06. SHC lab
meeting: NSF post doc app**

Lab safety stuff:

1. Do trainings online
2. Check waste and dispose it, ethidium bromide gels
3. Do monthly inspections

Newar works on Fridays; works up to 6 hours.

Notes:

- use performance curves or reaction norm instead of function-valued traits
- separate out terms, performance for fitness proxy and then reaction norm for physiology or any traits-phenology GxE = reaction norm; generate performance curve-growth over season
- context dependent expression of traits drive relative performance
- who cares about separating out photoperiod vs temp
- env can shape relationship between traits and performance in non-linear and unexpected ways or in ways that influence the process of adaptation, adaptive potential.
- look at many gxp traits-relating those to each other and to performance

- integrate all of these traits and overlay them on a complex environmental background
- stoichiometry: give ratios not just %
- expand on QG of gene expression
- selection may act in context-dependent manner
- be careful about constraints and trade-offs
- Think about training objective # 3; goal of grant? reword to make sure its a goal
- **certain clones:** does not tell you a whole lot. how should poplar be selected? Talk about general principals that you can lead to suggest to growers. What kind of outreach . prescribe based on environmental variables I am measuring.
- more info that is concrete on what the patterns are; feels adrift; not tied tightly between sections
- introduction- get rid of 2nd paragraph. maybe 1 sentence to previous paragraph
- research objectives: clarify traits; response function; add a little bit or shift; clarify parts

- get the realized GSL ; using existing rad seq data; predict performance as a function of temperature
-

<div id='id-section80'>

Page 80: 2016-10-07. Prepping cliamte cascade meeting

1. Project updates:

- Gene expression project:
 - **Go over analyses:**
 - **Go over figure layout for ms**
- Multiple stressors ms:
 - **Ask about SHC comments on confusion of mismatch membrane stability**
- Range limits ms: **SHC lab gave verbal edits:**
 - focus on 1 end of thermal niche breadth(although it is nice to mention it because CTmin decreases across lat)-CTmin.
 - Discussion needs to talk about cold adaptation; why trade-offs?
 - Walk through results better

- Thermal niche ms: **Lacey and I working on discussion**
- Stressed in nature MS: Samples to rerun.
- update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- Proteome stability project: **no clue what status is**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - **Talk title: Northern range limits of a common forest ant is associated with trade-offs in cold physiology**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to

find better flight prices.

- Thesis related [FORMS FOUND HERE](#)

- Formatting:

- Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph

- [Deadlines](#):

1. Intent to graduate: February 1st for May.
2. Send defense committee form to grad college—now
3. Graduate college format check March 4th
4. Defense notice 3 weeks before defense (oral defense by March 24th).
5. Final thesis April 7th.

<div id='id-section81'>

Page 81: 2016-10-11. ANCOVA models for testing interaction of hsp gxp parameter and habitat on CTmax

```
apply(b2[,3:11],2,function(x){summary(aov(b2$K0_temp_work  
er~b2$habitat_v2*x))})
```

\$FC_hsc70_1468_max

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	20.902	20.902	81.798	1.1e-11	***
x	1	0.375	0.375	1.467	0.232	
b2\$habitat_v2:x	1	0.374	0.374	1.462	0.233	
Residuals	45	11.499	0.256			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

5 observations deleted due to missingness

\$FC_hsc70_1468_slope

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	20.902	20.902	84.903	6.33e-12	***
x	1	1.169	1.169	4.749	0.0346	*
b2\$habitat_v2:x	1	0.000	0.000	0.000	0.9999	
Residuals	45	11.078	0.246			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

5 observations deleted due to missingness

\$FC_hsc70_1468_Tm

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	20.902	20.902	89.676	2.79e-12	***
x	1	1.125	1.125	4.828	0.0332	*
b2\$habitat_v2:x	1	0.633	0.633	2.718	0.1062	
Residuals	45	10.489	0.233			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

5 observations deleted due to missingness

\$FC_hsp40_541_max

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	21.311	21.311	85.111	9.4e-12	***
x	1	0.360	0.360	1.440	0.2368	
b2\$habitat_v2:x	1	0.875	0.875	3.494	0.0684	.
Residuals	43	10.767	0.250			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

7 observations deleted due to missingness

\$FC_hsp40_541_slope

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	21.311	21.311	81.495	1.75e-11	***
x	1	0.605	0.605	2.312	0.136	
b2\$habitat_v2:x	1	0.153	0.153	0.585	0.449	
Residuals	43	11.245	0.262			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

7 observations deleted due to missingness

\$FC_hsp40_541_Tm

	Df	Sum Sq	Mean Sq	F value	Pr(>F)						
b2\$habitat_v2	1	21.311	21.311	104.527	4.39e-13	***					
x	1	1.642	1.642	8.052	0.00691	**					
b2\$habitat_v2:x	1	1.594	1.594	7.816	0.00771	**					
Residuals	43	8.767	0.204								

Signif. codes:	0	'***'	0.001	'**'	0.01	'*'	0.05	'. '	0.1	' '	1
7	observations deleted due to missingness										

\$FC_Hsp83_279_max

	Df	Sum Sq	Mean Sq	F value	Pr(>F)						
b2\$habitat_v2	1	23.226	23.226	95.284	8.72e-13	***					
x	1	0.063	0.063	0.260	0.612						
b2\$habitat_v2:x	1	0.330	0.330	1.355	0.250						
Residuals	46	11.213	0.244								

Signif. codes:	0	'***'	0.001	'**'	0.01	'*'	0.05	'. '	0.1	' '	1
4	observations deleted due to missingness										

\$FC_Hsp83_279_slope

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	23.226	23.226	95.648	8.22e-13	***
x	1	0.156	0.156	0.641	0.428	
b2\$habitat_v2:x	1	0.281	0.281	1.157	0.288	
Residuals	46	11.170	0.243			

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' ' 1
```

```
4 observations deleted due to missingness
```

```
$FC_Hsp83_279_Tm
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	23.226	23.226	95.177	8.88e-13	***
x	1	0.068	0.068	0.279	0.600	
b2\$habitat_v2:x	1	0.313	0.313	1.283	0.263	
Residuals	46	11.225	0.244			

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

```
' ' 1
```

```
4 observations deleted due to missingness
```

Summary table of each parameter and its interaction with habitat on CTmax:

summary.table	max70	slope70	Tm70	max40	slope40	Tm40
habitat	yes	yes	yes	yes	yes	yes
parameter	no	yes	yes	no	no	yes
habitat * parameter	no	no	no	no	no	yes

Effect of habitat type on hsp gxp parameters

```
apply(b2[,3:11],2,function(x){summary(aov(x~b2$habitat_v2))})
```

```
$FC_hsc70_1468_max
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	4819	4819	30.98	1.21e-06	***
Residuals	47	7312	156			

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' '
```

```
5 observations deleted due to missingness
```

```
$FC_hsc70_1468_slope
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	2.562	2.5621	12.99	0.000754	***
Residuals	47	9.270	0.1972			

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' '
```

```
5 observations deleted due to missingness
```

```
$FC_hsc70_1468_Tm
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
b2\$habitat_v2	1	18.41	18.409	25.53	7.03e-06	***
Residuals	47	33.89	0.721			

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

5 observations deleted due to missingness

\$FC_hsp40_541_max

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	110.7	110.69	5.018	0.0301 *
Residuals	45	992.5	22.06		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

7 observations deleted due to missingness

\$FC_hsp40_541_slope

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	2.683	2.683	4.294	0.044 *
Residuals	45	28.123	0.625		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

7 observations deleted due to missingness

\$FC_hsp40_541_Tm

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	39.38	39.38	14.2	0.000476 ***
Residuals	45	124.81	2.77		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

7 observations deleted due to missingness

\$FC_Hsp83_279_max

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	149.4	149.43	5.649	0.0215 *
Residuals	48	1269.8	26.45		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

4 observations deleted due to missingness

\$FC_Hsp83_279_slope

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	1.92	1.9227	2.345	0.132
Residuals	48	39.35	0.8198		

4 observations deleted due to missingness

\$FC_Hsp83_279_Tm

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
b2\$habitat_v2	1	42.56	42.56	9.229	0.00385 **
Residuals	48	221.37	4.61		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

4 observations deleted due to missingness

Summary table of effect of habitat on hsp gxp parameter

param	habitat
max70	yes
slope70	yes
Tm70	yes
max40	yes
slope40	yes
Tm40	yes
max83	yes
slope83	no
Tm83	yes

<div id='id-section82'>

Page 82: 2016-10-11. variance partitioning in CTmax of aphaeno

- Phylogenetic axes = first 9
- Ecology = MAT, TMax, and habitat type

```
#model construction
```

```
var2<- varpart(Aph.dat$K0_temp_worker, ~ Axis.1 + Axis.2+  
Axis.3+ Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9, ~biol  
+bio5+habitat_v2,data=Aph.dat)
```

```
$part
```

```
$SS.Y
```

```
[1] 121.5443
```

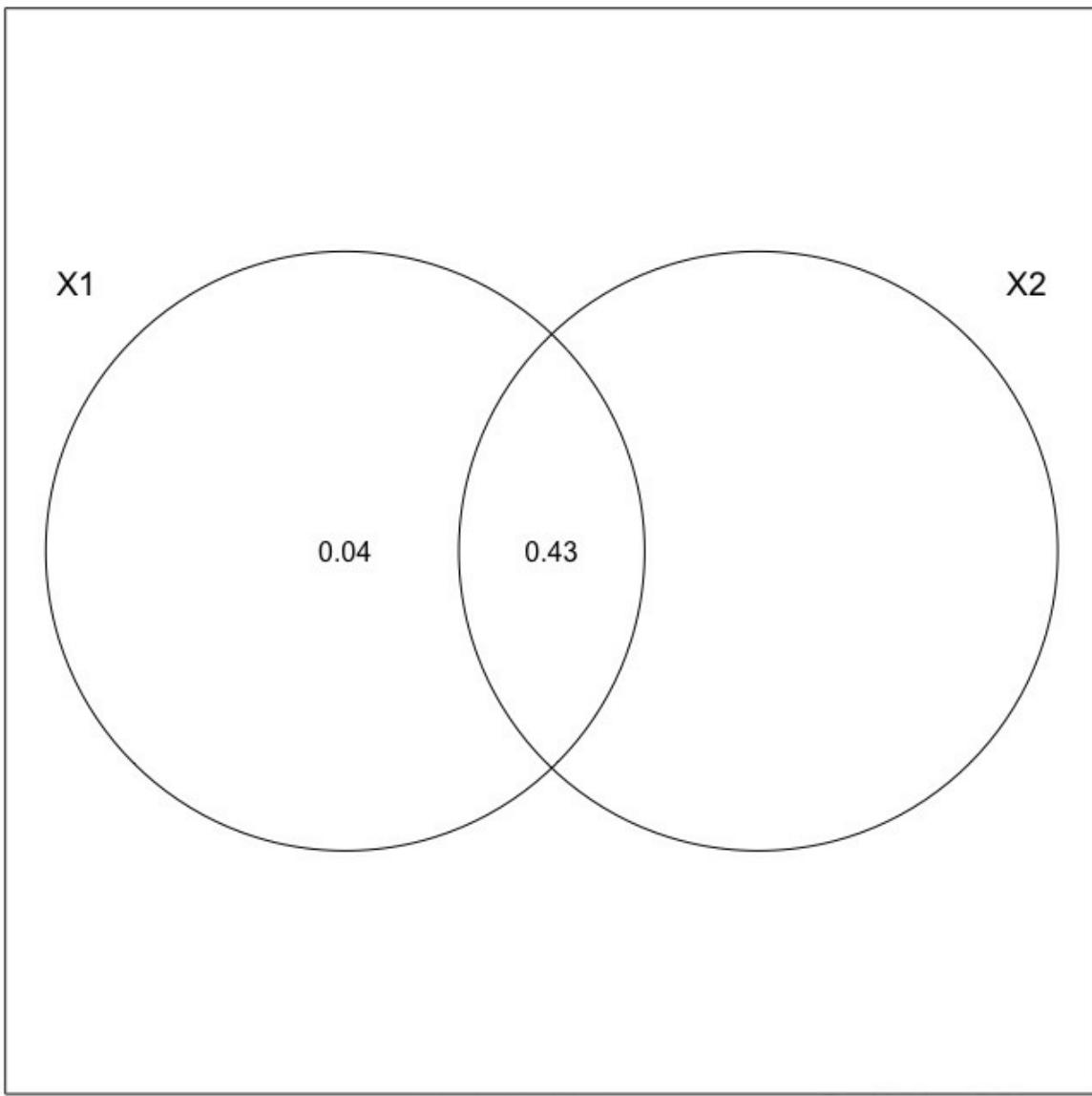
```
$fract
```

	Df	R.squared	Adj.R.squared	Testable
[a+b] = X1	9	0.5199228	0.4719151	TRUE
[b+c] = X2	3	0.4388392	0.4213030	TRUE
[a+b+c] = X1+X2	12	0.5288496	0.4638634	TRUE

```
$indfract
```

	Df	R.squared	Adj.R.squared	Testable
[a] = X1 X2	9	NA	0.042560390	TRUE
[b]	0	NA	0.429354679	FALSE
[c] = X2 X1	3	NA	-0.008051705	TRUE
[d] = Residuals	NA	NA	0.536136636	FALSE

Figure with different components



Values <0 not shown

<div id='id-section83'/>

**Page 83: 2016-10-12. Testing effect of
MAT on Hsp gxp and looking at
correlations between phylogeny and**

climate.

```
> apply(mergy[,38:43],2,function(x){summary(lm(log10(x)~mergy$biol))})  
$FC_83
```

Call:

```
lm(formula = log10(x) ~ mergy$biol)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.69315	-0.17367	-0.02182	0.16945	0.66741

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.993238	0.115874	8.572	1.19e-11 ***
mergy\$biol	-0.000497	0.001227	-0.405	0.687

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.2879 on 54 degrees of freedom

Multiple R-squared: 0.003028, Adjusted R-squared: -0.01543

F-statistic: 0.164 on 1 and 54 DF, p-value: 0.6871

\$FC_70

Call:

```
lm(formula = log10(x) ~ mergy$biol)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.63143	-0.12966	0.02354	0.18406	0.45652

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.571710	0.105899	14.842	<2e-16 ***
mergy\$biol	0.000679	0.001122	0.605	0.547

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.2631 on 54 degrees of freedom

Multiple R-squared: 0.006742, Adjusted R-squared: -0.01165

F-statistic: 0.3666 on 1 and 54 DF, p-value: 0.5474

\$FC_40

Call:

```
lm(formula = log10(x) ~ mergy$biol)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.87164	-0.16033	0.05806	0.23030	0.71656

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.8929016	0.1372969	6.503	2.63e-08 ***
mergy\$biol	0.0002741	0.0014540	0.188	0.851

Signif. codes:	0	'***'	0.001	'**'
		'0.01	'*' 0.05	'.' 0.1
		'	' 1	

Residual standard error: 0.3411 on 54 degrees of freedom

Multiple R-squared: 0.0006575, Adjusted R-squared: -0.01785

F-statistic: 0.03553 on 1 and 54 DF, p-value: 0.8512

\$B_83

Call:

```
lm(formula = log10(x) ~ mergy$biol)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.86395	-0.31896	-0.04139	0.33454	0.76906

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
--	----------	------------	---------	----------

(Intercept)	0.203307	0.186138	1.092	0.280
mergy\$biol1	-0.002098	0.001971	-1.064	0.292

Residual standard error: 0.4624 on 54 degrees of freedom

Multiple R-squared: 0.02054, Adjusted R-squared: 0.00
2405

F-statistic: 1.133 on 1 and 54 DF, p-value: 0.292

\$B_70

Call:

```
lm(formula = log10(x) ~ mergy$biol1)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.9569	-0.3399	-0.0464	0.3489	0.8581

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.199005	0.172676	1.152	0.254
mergy\$biol1	-0.002843	0.001829	-1.555	0.126

Residual standard error: 0.429 on 54 degrees of freedom

Multiple R-squared: 0.04284, Adjusted R-squared: 0.02
512

F-statistic: 2.417 on 1 and 54 DF, p-value: 0.1259

\$B_40

Call:

```
lm(formula = log10(x) ~ mergy$biol)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.68902	-0.28172	0.07947	0.31104	0.98014

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.300482	0.221888	1.354	0.181
mergy\$biol	-0.003086	0.002350	-1.313	0.195

Residual standard error: 0.5512 on 54 degrees of freedom

Multiple R-squared: 0.03096, Adjusted R-squared: 0.01301

F-statistic: 1.725 on 1 and 54 DF, p-value: 0.1946

Summary: none are significant

Correlation between Mean Annual Temperature (MAT), Tmax, and 4 phylogenetic axes

	MAT	Tmax	Axis.1	Axis.2	Axis.3	Axis.4
MAT	1.000	0.910	0.857	0.197	0.182	0.132
Tmax	0.910	1.000	0.836	0.128	0.204	0.110
Axis.1	0.857	0.836	1.000	0.002	0.000	0.008
Axis.2	0.197	0.128	0.002	1.000	0.000	-0.002
Axis.3	0.182	0.204	0.000	0.000	1.000	0.000
Axis.4	0.132	0.110	0.008	-0.002	0.000	1.000

20161013 follow up: checking 18s HKG stability

If there is an effect of rearing temperature, Tmax, and/or heat shock treatment, phylo axes, then the HKG is not stable.

```
ct<-read.csv("../Data/20150810_raw_CT_values.csv")

z<-inner_join(ct,mergy,by="Colony")
z$qpcr_block<-as.factor(z$qpcr_block)

#different 18s ct among treatments?
#different 18s ct
summary(stepAIC(lm(log2(X18)~bio5*treatment+qpcr_block+Axis.1+Axis.2+Axis.3+Rearing_Temp,data=z2)),direction="forward")
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	3.176814	0.056303	56.424	< 2e-16	***
qpcr_block2	0.107059	0.017592	6.086	1.84e-08	***
qpcr_block3	0.163280	0.018586	8.785	2.83e-14	***
Axis.1	-0.136572	0.072299	-1.889	0.0616	.
Axis.2	0.204421	0.112195	1.822	0.0712	.
Axis.3	-0.278600	0.165081	-1.688	0.0944	.
Rearing_Temp	-0.003763	0.002393	-1.573	0.1187	

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1
' ' 1					
Residual standard error:	0.07137	on 107 degrees of freedom			
Multiple R-squared:	0.5648	, Adjusted R-squared:	0.5404		
F-statistic:	23.15	on 6 and 107 DF,	p-value:	< 2.2e-16	

20161013 Taking out Axis1 because it covaries with bio5(Tmax)

```
apply(mergy[,38:43],2,function(x){summary(stepAIC(lm(log10(x)~mergy$bio5+mergy$Rearing_Temp+mergy$Axis.2+mergy$Axis.3)),direction="forward")})
```

Start: AIC=-142.41

```
log10(x) ~ mergy$bio5 + mergy$Rearing_Temp + mergy$Axis.2  
+ mergy$Axis.3
```

	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.3	1	0.002884	4.1926	-144.37
- mergy\$Axis.2	1	0.008699	4.1984	-144.29
- mergy\$bio5	1	0.017061	4.2068	-144.18
<none>			4.1897	-142.41
- mergy\$Rearing_Temp	1	0.257200	4.4469	-140.96

Step: AIC=-144.37

```
log10(x) ~ mergy$bio5 + mergy$Rearing_Temp + mergy$Axis.2
```

	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.2	1	0.009219	4.2018	-146.25
- mergy\$bio5	1	0.021070	4.2137	-146.08
<none>			4.1926	-144.37
- mergy\$Rearing_Temp	1	0.254448	4.4471	-142.96

Step: AIC=-146.25

```
log10(x) ~ mergy$bio5 + mergy$Rearing_Temp
```

	Df	Sum of Sq	RSS	AIC
- mergy\$bio5	1	0.01849	4.2203	-147.99
<none>			4.2018	-146.25
- mergy\$Rearing_Temp	1	0.29906	4.5009	-144.26

Step: AIC=-147.99

$\log_{10}(x) \sim \text{mergy\$Rearing_Temp}$

	Df	Sum of Sq	RSS	AIC
<none>			4.2203	-147.99
- mergy\$Rearing_Temp	1	0.30548	4.5258	-145.94

Start: AIC=-151.28

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2}$
+ mergy\$Axis.3

	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.3	1	0.006133	3.6020	-153.18
- mergy\$bio5	1	0.014353	3.6102	-153.05
- mergy\$Axis.2	1	0.125441	3.7213	-151.29
<none>			3.5959	-151.28
- mergy\$Rearing_Temp	1	0.211236	3.8071	-149.97

Step: AIC=-153.18

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2}$

	Df	Sum of Sq	RSS	AIC
- mergy\$bio5	1	0.011172	3.6132	-155.00
<none>			3.6020	-153.18
- mergy\$Axis.2	1	0.128482	3.7305	-153.15
- mergy\$Rearing_Temp	1	0.218797	3.8208	-151.76

Step: AIC=-155

$\log_{10}(x) \sim \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2}$

	Df	Sum of Sq	RSS	AIC
<none>			3.6132	-155.00
- mergy\$Axis.2	1	0.13788	3.7510	-154.83
- mergy\$Rearing_Temp	1	0.22616	3.8393	-153.48
Start:	AIC=-127.73			
log10(x) ~ mergy\$bio5 + mergy\$Rearing_Temp + mergy\$Axis.2 + mergy\$Axis.3				
	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.3	1	0.03867	5.4351	-129.32
- mergy\$bio5	1	0.10859	5.5051	-128.58
<none>			5.3965	-127.73
- mergy\$Axis.2	1	0.42509	5.8216	-125.33
- mergy\$Rearing_Temp	1	0.64013	6.0366	-123.23
Step:	AIC=-129.32			
log10(x) ~ mergy\$bio5 + mergy\$Rearing_Temp + mergy\$Axis.2 + mergy\$Axis.3				
	Df	Sum of Sq	RSS	AIC
- mergy\$bio5	1	0.14392	5.5791	-129.80
<none>			5.4351	-129.32
- mergy\$Axis.2	1	0.41361	5.8488	-127.06
- mergy\$Rearing_Temp	1	0.67128	6.1064	-124.56
Step:	AIC=-129.8			
log10(x) ~ mergy\$Rearing_Temp + mergy\$Axis.2 + mergy\$Axis.3				
	Df	Sum of Sq	RSS	AIC

<none>		5.5791	-129.80
- mergy\$Axis.2	1	0.47047	6.0495 -127.11
- mergy\$Rearing_Temp	1	0.63445	6.2135 -125.56
Start:	AIC=- 88.85		
log10(x) ~	mergy\$bio5 + mergy\$Rearing_Temp + mergy\$Axis.2		
+ mergy\$Axis.3			

	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.2	1	0.02655	10.576	-90.709
- mergy\$bio5	1	0.27432	10.824	-89.365
<none>			10.549	-88.854
- mergy\$Axis.3	1	0.47944	11.029	-88.277
- mergy\$Rearing_Temp	1	0.48666	11.036	-88.239

Step:	AIC=- 90.71		
log10(x) ~	mergy\$bio5 + mergy\$Rearing_Temp + mergy\$Axis.3		

	Df	Sum of Sq	RSS	AIC
- mergy\$bio5	1	0.29726	10.873	-91.101
<none>			10.576	-90.709
- mergy\$Rearing_Temp	1	0.46041	11.036	-90.237
- mergy\$Axis.3	1	0.49173	11.068	-90.073

Step:	AIC=- 91.1		
log10(x) ~	mergy\$Rearing_Temp + mergy\$Axis.3		

	Df	Sum of Sq	RSS	AIC
- mergy\$Axis.3	1	0.36201	11.235	-91.201

<none>		10.873	-91.101	
- mergy\$Rearing_Temp	1	0.50260	11.376	-90.480

Step: AIC=-**91.2**

$\log_{10}(x) \sim \text{mergy\$Rearing_Temp}$

	Df	Sum of Sq	RSS	AIC
--	----	-----------	-----	-----

<none>		11.235	-91.201
--------	--	---------------	----------------

- mergy\$Rearing_Temp	1	0.56062	11.796	-90.377
-----------------------	----------	----------------	---------------	----------------

Start: AIC=-**126.78**

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2}$
+ mergy\$Axis.3

	Df	Sum of Sq	RSS	AIC
--	----	-----------	-----	-----

- mergy\$Axis.2	1	0.0042	5.4901	-128.735
-----------------	----------	---------------	---------------	-----------------

- mergy\$Axis.3	1	0.0404	5.5262	-128.354
-----------------	----------	---------------	---------------	-----------------

- mergy\$bio5	1	0.1532	5.6391	-127.182
---------------	----------	---------------	---------------	-----------------

<none>		5.4859	-126.780
--------	--	---------------	-----------------

- mergy\$Rearing_Temp	1	4.5602	10.0461	-93.689
-----------------------	----------	---------------	----------------	----------------

Step: AIC=-**128.74**

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
--	----	-----------	-----	-----

- mergy\$Axis.3	1	0.0392	5.5292	-130.323
-----------------	----------	---------------	---------------	-----------------

- mergy\$bio5	1	0.1609	5.6509	-129.060
---------------	----------	---------------	---------------	-----------------

<none>		5.4901	-128.735
--------	--	---------------	-----------------

- mergy\$Rearing_Temp	1	4.8078	10.2978	-94.254
-----------------------	----------	---------------	----------------	----------------

Step: AIC=-130.32

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp}$

	Df	Sum of Sq	RSS	AIC
<none>		5.5292	-130.323	
- mergy\$bio5	1	0.204	5.7332	-130.221
- mergy\$Rearing_Temp	1	4.770	10.2992	-96.246

Start: AIC=-80.6

$\log_{10}(x) \sim \text{mergy\$bio5} + \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2}$
+ $\text{mergy\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
- mergy\$bio5	1	0.1822	12.346	-81.733
<none>		12.164	-80.595	
- mergy\$Rearing_Temp	1	0.7613	12.925	-79.074
- mergy\$Axis.2	1	1.1960	13.360	-77.156
- mergy\$Axis.3	1	3.4308	15.595	-68.185

Step: AIC=-81.73

$\log_{10}(x) \sim \text{mergy\$Rearing_Temp} + \text{mergy\$Axis.2} + \text{mergy\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
<none>		12.346	-81.733	
- mergy\$Rearing_Temp	1	0.8276	13.174	-79.970
- mergy\$Axis.2	1	1.3181	13.664	-77.849
- mergy\$Axis.3	1	3.9458	16.292	-67.648

\$FC_83

Call:

```
lm(formula = log10(x) ~ mergy$Rearing_Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.6121	-0.1422	-0.0417	0.1399	0.7465

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.39083	0.27720	1.410	0.1641
mergy\$Rearing_Temp	0.02473	0.01228	2.013	0.0489 *

Signif. codes:	0 *** 0.001 ** 0.01 * 0.05 . 0.1			
	' ' 1			

Residual standard error: 0.2745 on 56 degrees of freedom

Multiple R-squared: 0.0675, Adjusted R-squared: 0.05
084

F-statistic: 4.053 on 1 and 56 DF, p-value: 0.0489

\$FC_70

Call:

```
lm(formula = log10(x) ~ mergy$Rearing_Temp + mergy$Axis.2  
)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.67832	-0.16434	0.02663	0.17901	0.38810

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.12428	0.26812	7.923	1.16e-10	*
**					
mergy\$Rearing_Temp	-0.02197	0.01184	-1.855	0.0689	.
mergy\$Axis.2	0.81467	0.56233	1.449	0.1531	

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1
	' '	1			

Residual standard error: 0.2563 on 55 degrees of freedom

Multiple R-squared: 0.07529, Adjusted R-squared: 0.04167

F-statistic: 2.239 on 2 and 55 DF, p-value: 0.1162

\$FC_40

Call:

```
lm(formula = log10(x) ~ mergy$Rearing_Temp + mergy$Axis.2)
```

)

Residuals:

Min	1Q	Median	3Q	Max
-0.80408	-0.10662	0.07152	0.25390	0.55421

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.08767	0.33317	0.263	0.7934
mergy\$Rearing_Temp	0.03680	0.01471	2.501	0.0154 *
mergy\$Axis.2	1.50486	0.69876	2.154	0.0357 *

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 . 0.1
	' '	1		

Residual standard error: 0.3185 on 55 degrees of freedom

Multiple R-squared: 0.2085, Adjusted R-squared: 0.17

97

F-statistic: 7.242 on 2 and 55 DF, p-value: 0.001614

\$B_83

Call:

lm(formula = log10(x) ~ mergy\$Rearing_Temp)

Residuals:

Min	1Q	Median	3Q	Max

-0.90287 -0.32839 0.03175 0.37027 0.81465

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.73480	0.45228	-1.625	0.11
mergy\$Rearing_Temp	0.03350	0.02004	1.672	0.10

Residual standard error: 0.4479 on 56 degrees of freedom

Multiple R-squared: 0.04753, Adjusted R-squared: 0.03052

F-statistic: 2.794 on 1 and 56 DF, p-value: 0.1002

\$B_70

Call:

lm(formula = log10(x) ~ mergy\$bio5 + mergy\$Rearing_Temp)

Residuals:

Min	1Q	Median	3Q	Max
-0.80424	-0.20413	-0.03442	0.25526	0.81219

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.479502	0.641632	-2.306	0.0249*
mergy\$bio5	-0.002799	0.001965	-1.424	0.1600

```
mergy$Rearing_Temp  0.097784   0.014196   6.888 5.75e-09
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' '
```

Residual standard error: 0.3171 on 55 degrees of freedom

Multiple R-squared: 0.4779, Adjusted R-squared: 0.4589

F-statistic: 25.17 on 2 and 55 DF, p-value: 1.734e-08

\$B_40

Call:

```
lm(formula = log10(x) ~ mergy$Rearing_Temp + mergy$Axis.2  
+ mergy$Axis.3)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.32627	-0.32412	0.04458	0.31258	0.90367

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.92089	0.50132	-1.837	0.071726 .

```

mergy$Rearing_Temp  0.04213     0.02214    1.903 0.062428 .

mergy$Axis.2        -2.51976    1.04941   -2.401 0.019819 *
                                        

mergy$Axis.3        -6.14160    1.47835   -4.154 0.000117 *
                                        

**


---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1


Residual standard error: 0.4782 on 54 degrees of freedom
Multiple R-squared:  0.3049 ,   Adjusted R-squared:  0.26
63
F-statistic: 7.897 on 3 and 54 DF,  p-value: 0.0001852

```

<div id='id-section83.5'>

2016-11-01 adding full models with automated stepAIC

```

apply(merg[,38:43],2,function(x){summary(stepAIC(lm(log10
(x)~merg$bio5+merg$Rearing_Temp+merg$Axis.1+merg$Axis.2+m
erg$Axis.3)),direction="forward")})

Start: AIC=-135.83

log10(x) ~ merg$bio5 + merg$Rearing_Temp + merg$Axis.1 +
merg$Axis.2 +
merg$Axis.3

```

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.3	1	0.00006	4.2616	-137.82
- merg\$Axis.2	1	0.00563	4.2671	-137.75
- merg\$Axis.1	1	0.03032	4.2918	-137.42
- merg\$bio5	1	0.05267	4.3142	-137.12
<none>			4.2615	-135.83
- merg\$Rearing_Temp	1	0.32622	4.5877	-133.62

Step: AIC=-137.82

$\log_{10}(x) \sim \text{merg$bio5} + \text{merg$Rearing_Temp} + \text{merg$Axis.1} + \text{merg$Axis.2}$

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.2	1	0.00557	4.2671	-139.75
- merg\$Axis.1	1	0.03288	4.2944	-139.39
- merg\$bio5	1	0.05995	4.3215	-139.03
<none>			4.2616	-137.82
- merg\$Rearing_Temp	1	0.32790	4.5895	-135.60

Step: AIC=-139.75

$\log_{10}(x) \sim \text{merg$bio5} + \text{merg$Rearing_Temp} + \text{merg$Axis.1}$

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.1	1	0.02927	4.2964	-141.36
- merg\$bio5	1	0.05486	4.3220	-141.02
<none>			4.2671	-139.75
- merg\$Rearing_Temp	1	0.35722	4.6243	-137.17

Step: AIC=-141.36

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.02771	4.3241	-142.99
<none>			4.2964	-141.36
- merg\$Rearing_Temp	1	0.33717	4.6336	-139.05

Step: AIC=-142.99

log10(x) ~ merg\$Rearing_Temp

	Df	Sum of Sq	RSS	AIC
<none>			4.3241	-142.99
- merg\$Rearing_Temp	1	0.3481	4.6722	-140.58

Start: AIC=-147.19

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.1 +
merg\$Axis.2 +
merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.1	1	0.009107	3.5000	-149.05
- merg\$Axis.3	1	0.009894	3.5008	-149.03
- merg\$bio5	1	0.016701	3.5076	-148.92
- merg\$Axis.2	1	0.046939	3.5379	-148.43
<none>			3.4909	-147.19
- merg\$Rearing_Temp	1	0.215627	3.7065	-145.78

Step: AIC=-149.05

$\log_{10}(x) \sim \text{merg\$bio5} + \text{merg\$Rearing_Temp} + \text{merg\$Axis.2} + \text{merg\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.3	1	0.005260	3.5053	-150.96
- merg\$bio5	1	0.008554	3.5086	-150.91
- merg\$Axis.2	1	0.057491	3.5575	-150.12
<none>			3.5000	-149.05
- merg\$Rearing_Temp	1	0.210727	3.7107	-147.71

Step: AIC=-150.96

$\log_{10}(x) \sim \text{merg\$bio5} + \text{merg\$Rearing_Temp} + \text{merg\$Axis.2}$

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.006235	3.5115	-152.86
- merg\$Axis.2	1	0.059127	3.5644	-152.01
<none>			3.5053	-150.96
- merg\$Rearing_Temp	1	0.218048	3.7233	-149.52

Step: AIC=-152.86

$\log_{10}(x) \sim \text{merg\$Rearing_Temp} + \text{merg\$Axis.2}$

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.2	1	0.065809	3.5773	-153.80
<none>			3.5115	-152.86
- merg\$Rearing_Temp	1	0.225290	3.7368	-151.31

Step: AIC=-153.8

log10(x) ~ merg\$Rearing_Temp

	Df	Sum of Sq	RSS	AIC
<none>		3.5773	3.5773	-153.8
- merg\$Rearing_Temp	1	0.18654	3.7639	-152.9

Start: AIC=-122.77

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.1 +
merg\$Axis.2 +
merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.3	1	0.01759	5.0640	-124.58
- merg\$Axis.1	1	0.03695	5.0833	-124.37
- merg\$bio5	1	0.09873	5.1451	-123.69
- merg\$Axis.2	1	0.14349	5.1899	-123.20
<none>		5.0464	5.0464	-122.77
- merg\$Rearing_Temp	1	0.61137	5.6577	-118.37

Step: AIC=-124.58

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.1 +
merg\$Axis.2

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.1	1	0.06171	5.1257	-125.90
- merg\$Axis.2	1	0.13474	5.1987	-125.11
- merg\$bio5	1	0.15531	5.2193	-124.89
<none>		5.0640	5.0640	-124.58

- merg\$Rearing_Temp 1 0.62522 5.6892 -120.06

Step: AIC=-125.9

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.2

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.11746	5.2431	-126.63
- merg\$Axis.2	1	0.17282	5.2985	-126.04
<none>			5.1257	-125.90
- merg\$Rearing_Temp	1	0.66713	5.7928	-121.05

Step: AIC=-126.63

log10(x) ~ merg\$Rearing_Temp + merg\$Axis.2

	Df	Sum of Sq	RSS	AIC
<none>			5.2431	-126.63
- merg\$Axis.2	1	0.21853	5.4617	-126.35
- merg\$Rearing_Temp	1	0.63456	5.8777	-122.23

Start: AIC=-85.77

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.1 +
merg\$Axis.2 +
merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.2	1	0.09471	10.350	-87.247
- merg\$bio5	1	0.14357	10.399	-86.979
- merg\$Axis.3	1	0.17560	10.431	-86.803
- merg\$Rearing_Temp	1	0.34221	10.597	-85.900

<none>		10.255	-85.771
- merg\$Axis.1	1	0.51791	10.773 -84.963

Step: AIC=-87.25

$\log_{10}(x) \sim \text{merg\$bio5} + \text{merg\$Rearing_Temp} + \text{merg\$Axis.1} + \text{merg\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.09885	10.449	-88.705
- merg\$Axis.3	1	0.20541	10.555	-88.127
- merg\$Rearing_Temp	1	0.28656	10.636	-87.690
<none>			10.350	-87.247
- merg\$Axis.1	1	0.45249	10.802	-86.808

Step: AIC=-88.71

$\log_{10}(x) \sim \text{merg\$Rearing_Temp} + \text{merg\$Axis.1} + \text{merg\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
- merg\$Rearing_Temp	1	0.30750	10.756	-89.052
<none>			10.449	-88.705
- merg\$Axis.3	1	0.37408	10.823	-88.700
- merg\$Axis.1	1	0.60533	11.054	-87.495

Step: AIC=-89.05

$\log_{10}(x) \sim \text{merg\$Axis.1} + \text{merg\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
<none>			10.756	-89.052

- merg\$Axis.3 1 0.42229 11.178 -88.857
- merg\$Axis.1 1 0.71553 11.472 -87.381

Start: AIC=-122.03

log10(x) ~ merg\$bio5 + merg\$Rearing_Temp + merg\$Axis.1 +
merg\$Axis.2 +
merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.0001	5.4282	-124.032
- merg\$Axis.2	1	0.0329	5.4610	-123.689
- merg\$Axis.1	1	0.0409	5.4690	-123.605
- merg\$Axis.3	1	0.0666	5.4947	-123.338
<none>			5.4281	-122.033
- merg\$Rearing_Temp	1	4.5125	9.9406	-89.546

Step: AIC=-124.03

log10(x) ~ merg\$Rearing_Temp + merg\$Axis.1 + merg\$Axis.2 +
merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.2	1	0.0357	5.4639	-125.659
- merg\$Axis.3	1	0.0798	5.5080	-125.200
- merg\$Axis.1	1	0.1695	5.5977	-124.279
<none>			5.4282	-124.032
- merg\$Rearing_Temp	1	4.5125	9.9407	-91.545

Step: AIC=-125.66

log10(x) ~ merg\$Rearing_Temp + merg\$Axis.1 + merg\$Axis.3

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.3	1	0.0784	5.5423	-126.847
- merg\$Axis.1	1	0.1733	5.6372	-125.879
<none>			5.4639	-125.659
- merg\$Rearing_Temp	1	4.5377	10.0016	-93.197

Step: AIC=-126.85

$\log_{10}(x) \sim \text{merg\$Rearing_Temp} + \text{merg\$Axis.1}$

	Df	Sum of Sq	RSS	AIC
- merg\$Axis.1	1	0.1750	5.7173	-127.075
<none>			5.5423	-126.847
- merg\$Rearing_Temp	1	4.4787	10.0209	-95.087

Step: AIC=-127.07

$\log_{10}(x) \sim \text{merg\$Rearing_Temp}$

	Df	Sum of Sq	RSS	AIC
<none>			5.7173	-127.075
- merg\$Rearing_Temp	1	4.7398	10.4571	-94.659

Start: AIC=-78.04

$\log_{10}(x) \sim \text{merg\$bio5} + \text{merg\$Rearing_Temp} + \text{merg\$Axis.1} + \text{merg\$Axis.2} + \text{merg\$Axis.3}$

	Df	Sum of Sq	RSS	AIC
- merg\$bio5	1	0.1636	11.382	-79.225

<none>			11.219	-78.036
- merg\$Axis.1	1	0.4666	11.685	-77.754
- merg\$Rearing_Temp	1	0.6847	11.903	-76.718
- merg\$Axis.2	1	0.9679	12.186	-75.402
- merg\$Axis.3	1	3.9432	15.162	-63.168

Step: AIC=-79.23

`log10(x) ~ merg$Rearing_Temp + merg$Axis.1 + merg$Axis.2
+ merg$Axis.3`

	Df	Sum of Sq	RSS	AIC
<none>		11.382	-79.225	
- merg\$Axis.1	1	0.4311	11.813	-79.144
- merg\$Rearing_Temp	1	0.6969	12.079	-77.897
- merg\$Axis.2	1	0.8346	12.217	-77.263
- merg\$Axis.3	1	3.9224	15.305	-64.643

\$FC_83

Call:

`lm(formula = log10(x) ~ merg$Rearing_Temp)`

Residuals:

Min	1Q	Median	3Q	Max
-0.61666	-0.14861	-0.03988	0.14529	0.74191

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.36231	0.28484	1.272	0.2087

```
merg$Rearing_Temp  0.02638      0.01254    2.104    0.0399 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.2804 on 55 degrees of freedom
Multiple R-squared:  0.07451,   Adjusted R-squared:  0.05
768
F-statistic: 4.428 on 1 and 55 DF,  p-value: 0.03995
```

\$FC_70

Call:

```
lm(formula = log10(x) ~ merg$Rearing_Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.6417	-0.1415	0.0238	0.1711	0.3910

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.06670	0.25908	7.977	9.51e-11 **
*				
merg\$Rearing_Temp	-0.01931	0.01140	-1.694	0.096 .

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.255 on 55 degrees of freedom

Multiple R-squared: 0.04956, Adjusted R-squared: 0.03

228

F-statistic: 2.868 on 1 and 55 DF, p-value: 0.09601

\$FC_40

Call:

lm(formula = log10(x) ~ merg\$Rearing_Temp + merg\$Axis.2)

Residuals:

Min	1Q	Median	3Q	Max
-0.80392	-0.10073	0.07339	0.22020	0.55569

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.09071	0.32917	0.276	0.7839
merg\$Rearing_Temp	0.03680	0.01453	2.533	0.0143 *
merg\$Axis.2	1.24166	0.83541	1.486	0.1431

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

‘ ’ 1

Residual standard error: 0.3145 on 53 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.166, Adjusted R-squared: 0.1345
F-statistic: 5.275 on 2 and 53 DF, p-value: 0.008145

\$B_83

Call:

```
lm(formula = log10(x) ~ merg$Axis.1 + merg$Axis.3)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.89374	-0.32249	0.03374	0.32440	0.77433

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.01073	0.05911	0.182	0.8566
merg\$Axis.1	-1.09010	0.57516	-1.895	0.0634 .
merg\$Axis.3	2.00468	1.37680	1.456	0.1512

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	' '	1		

Residual standard error: 0.4463 on 54 degrees of freedom

Multiple R-squared: 0.09566, Adjusted R-squared: 0.06

217

F-statistic: 2.856 on 2 and 54 DF, p-value: 0.06621

\$B_70

Call:

```
lm(formula = log10(x) ~ merg$Rearing_Temp)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.7507	-0.1789	-0.0132	0.2067	0.7046

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2.24217	0.32753	-6.846	6.75e-09	**
*					
merg\$Rearing_Temp	0.09734	0.01442	6.753	9.59e-09	**
*					

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1
' '	1				

Residual standard error: 0.3224 on 55 degrees of freedom

Multiple R-squared: 0.4533, Adjusted R-squared: 0.44

33

F-statistic: 45.6 on 1 and 55 DF, p-value: 9.589e-09

\$B_40

Call:

```
lm(formula = log10(x) ~ merg$Rearing_Temp + merg$Axis.1 +  
  merg$Axis.2 +  
  merg$Axis.3)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.38234	-0.22276	-0.00071	0.25240	0.84201

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.84687	0.49824	-1.700	0.09527	.
merg\$Rearing_Temp	0.03887	0.02200	1.767	0.08319	.
merg\$Axis.1	-0.85399	0.61446	-1.390	0.17062	
merg\$Axis.2	-2.42734	1.25523	-1.934	0.05870	.
merg\$Axis.3	-6.12416	1.46081	-4.192	0.00011	**
*					

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1
	' '	1			

Residual standard error: 0.4724 on 51 degrees of freedom

```
(1 observation deleted due to missingness)

Multiple R-squared:  0.3278,    Adjusted R-squared:  0.27
51

F-statistic: 6.218 on 4 and 51 DF,  p-value: 0.0003732
```

<div id='id-section84'>

Page 84: 2016-10-12]. Updating climate cascade to do list.

1. Project updates:

- Hsp gene expression + Ctmax project:
 - **Go over updated figures**
 - Starting to write: working title-**Shifts in the reaction norms of heat shock protein gene expression accompany evolutionary innovations in thermal tolerance of forest ants**
- Multiple stressors ms:
 - **Sent SHC another version**
- Range limits ms: **SHC lab gave verbal edits:**
 - focus on 1 end of thermal niche

breadth(although it is nice to mention it because
CTmin decreases across lat)-CTmin.

- Dicussion needs to talk about cold adaptation;
why trade-offs?
 - Walk through results better
- Thermal niche ms: **Lacey and I working on discussion**
- Stressed in nature MS: Samples to rerun.
- update: Curtis can no longer work+ write on project
 - in reference to missing samples
 - Fit in time to process Curtis' samples.
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- Proteome stability project: **should be getting data soon**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - **Talk title: Northern range limits of a common**

forest ant is associated with trade-offs in cold physiology

- Apply for funding. Suitor Travel Grant Deadline is october 31
- **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.

- **Thesis related** FORMS FOUND HERE

- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph
- Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.

<div id='id-section85'>

Page 85: 2016-10-14. Paper note:

Puentes, A., G. Granath, and J. Ågren.

2016. Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*. Evolution 70:2370-2386.

Similar paper here: [Page 60: 2016-09-01](#). Paper notes: Paccard A, Van Buskirk J, Willi Y, Eckert CG, Bronstein JL. 2016. Quantitative Genetic Architecture at Latitudinal Range Boundaries: Reduced Variation but Higher Trait Independence. *The American Naturalist*.

Method differences:

But Puentes et al. focus on *A. lyrata* in their native range (Norway-Sweden) and in the field, and Paccard et al. 2016 use populations from USA-Canada in lab conditions. Puentes uses 5 traits, Paccard used 10 traits. One of Cortlett Wood's papers suggests that the number of traits can alter how G changes among environments. Check what traits were used between studies.

Summary of findings: G is stable between Norway and Sweden populations

Page 86: 2016-10-14. Wiley House

Style Guide

I'll need to follow these general writing rules for submitting a ms to Evolution.

Use of *that* and *which*

- * That is used for defining or restrictive clauses:
 - * The patient made a list of the symptoms that were most troublesome

A defining clause is specific (limiting) to a particular person or thing; i.e., the patient had to list only those particular symptoms that were most troublesome

- * Which is used in nondefining or nonrestrictive clauses:
 - * The patient made a list of the symptoms, which were most troublesome

A nondefining clause is general (nonlimiting); it provides additional information, and the use of commas is often important. In this example, all the sympt

oms were very troublesome.

Redundancy

Avoid using a modifying word when the intended meaning is inherent in a word already used.

Redundancy is obvious in examples such as the results were plotted graphically, past history, bright blue in color, inactivates its activity, and completely filled. Does the term careful monitoring suggest that the alternative is careless monitoring?

Balancing a sentence:

It is important to ensure that a sentence balances on either side of certain words (correlatives) that emphasize similarity or contrast and that are used in parallel: both and and; either and or; neither and nor; not only and but; between and and; whether and or . For example, “I swam both in the morning and afternoon” should be “I swam both in the morning and in the afternoon” or “I swam in both the morning and the afternoon.” Note the position of the preposition in. (See also the section “Editing for Sense.”)

Key Points

- It is now acceptable to use the active or the passive voice.
- Use the past tense for the author's methods and results, and the present tense for interpretation and generally accepted "facts."
- The subject and verb must agree in number.
- "That" is defining; "which" is not.
- Check that articles ("a," "an," and "the") are used correctly.
- Sentences must balance (e.g., with "both ... and ...").
- In comparisons (e.g., with lower/higher/less/more), make sure it is clear what is being compared with what.
- Avoid sexist, dehumanizing, and stereotypical language.

Punctuation

Semi colons and Colons

SEMICOLONS

- The semicolon is stronger than a comma but not as decisive as a full point. It can be used to separate sentences (whereas a comma cannot).
- Use a semicolon before, and a comma after, the conjunctive adverbs however, that is,

nevertheless, etc.

COLONS

Colons are used to introduce material that restates, explains, enlarges upon, or summarizes previous material. They also introduce items in a list set off from text (but a colon is not needed in run-on lists introduced by the words for example, namely, including, etc.; e.g., in the sentence “The dessert looks nice with fruit on it, for example: strawberries, raspberries, and blueberries” the colon should not be there).

- In US spelling, if the material introduced by a colon consists of more than one sentence, or if it is a formal statement, quotation, or speech in dialogue it should take a capital after the colon. In UK spelling, a capital letter is not used after a colon (except in titles and subtitles).
- Ratios containing words should have a thin space on each side of the colon (e.g., the light : dark cycle) but ratios containing numbers should be closed up (e.g., 16:8 h).

Key Points:

- Use commas to clarify sentences.
- Do not use a comma to separate sentences; use a semicolon

on (this is a particularly common error before “however” and “nevertheless”).

- Do not use apostrophes with plural abbreviations (e.g., ANOVAs, not ANOVA’s).
- For hyphenation, refer to your journal style sheet.
- Do not hyphenate adverbs ending in -ly (e.g., dermatologically tested soap).
- Use hyphens in compound terms to clarify meaning (e.g., much-needed clothing).
- Use en dashes, not hyphens, for associations (e.g., dose-response curve).

<div id='id-section87'>

Page 87: 2016-10-14. NSF post doc app meeting: Keller Lab

SK background to grant

NSF used to have bioinformatics post doc competition and replaced with narrowly defined one in bio. It has to fit into 1 or a couple bins: 1 of them is plant genome research program (PGRP; funds poplar). SK attend PGRP meetings as part of training missions seriously. They build a program and come in as cohort(post doc fellows) and they have extra training sessions with them. Post doc presents work and are well supported for 3 years. SK fits squarely into: economically important

plant, genome wide approaches to the problem of plant growth/yield and response to stress and other challenges.

Project Summary

Project description

Large communication issue

- What is new and novel? Kattia
- **Figure 3: analysis is of a single trait** Hammer it down, multiple times, outside of fig legend and make it more clear.
- **HAMMER DOWN novelty is non-linear GxE interactions**
- Cant predict performance readily from 1 environment to another environment (that span the current and future climate)
- Say you'll measure wood traits
- **Bring more genomics more important:** Bring in population genetics into the proposal.
- Add path analysis
- Come up with precise alleles of what is adaptive.

- Fig 1 C. put an ellipse for central population:
- Set margins to 1 inch around.
- heavy lifting (SK): bring emphasis on gene expression way up (genetic variation among genotypes in their transcriptional response to that variation); ID genes or networks of genes that show differences in expression or organization. What parts of the transcription? GO, pathways? Genes in trade-offs in few networks or overdispersed across a network, relative to the total transcriptome. Stress response genes (Hsps)? Phenology associated genes (circadian clock). How can that be pulled out using the kingsolver method. (Not just as a tool that is cool to use, but as a question with an appropriately matched tool).
- **Look at the SNPs.** include in
 - There is gxp from fairbanks and indian head. “Timing for success title”
 - Karl: pair down first paragraph; reduce in length
 - SK, focus on the major ideas
- **Be more explicit about what the trait is used for Gmatrix.**
 - genetically based differences to identify GxE

Dissertation Abstract

Data management

<div id='id-section88'>

Page 88: 2016-10-18. Climate cascade meeting

1. Project updates:

- Hsp gene expression + Ctmax project:
 - **Go over updated figures**
 - figure 3, SHC says to switch back branches
 - figure 4, color code by habitat type, NJG:don't use dot dash, use dash
 - Starting to write: working title-**Shifts in the reaction norms of heat shock protein gene expression accompany evolutionary innovations in thermal tolerance of forest ants**
 - need to start writing methods and results
- Multiple stressors ms:
 - **Sent SHC another version** ; should submit soon

- Range limits ms: **SHC lab gave verbal edits:**
 - focus on 1 end of thermal niche
breadth(although it is nice to mention it because CTmin decreases across lat)-CTmin.
 - Discussion needs to talk about cold adaptation;
why trade-offs?
 - Walk through results better
- Thermal niche ms: **Lacey and I working on discussion**
- Stressed in nature MS: Samples to rerun.
- update: Curtis can no longer work+ write on project
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- Proteome stability project: **should be getting data soon**
- Attending SICB - Jan 4-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**

- **Talk title: Northern range limits of a common forest ant is associated with trade-offs in cold physiology**
- Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.
 - Application submitted today 2016-10-18
- **Thesis related** FORMS FOUND HERE
 - Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph
 - Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.

Page 89: 2016-10-25. Climate cascade updated list

1. Project updates:

- Hsp gene expression + Ctmax project:
 - figure 3, SHC says to switch back branches
 - Starting to write: working title-**Shifts in the reaction norms of heat shock protein gene expression accompany evolutionary innovations in thermal tolerance of forest ants**
 - need to start writing methods and results; submit to **MBE**
- Multiple stressors ms:
 - submitted **2016-10-24**
- Range limits ms: **SHC lab gave verbal edits:**
 - focus on 1 end of thermal niche breadth(although it is nice to mention it because CTmin decreases across lat)-CTmin.
 - Discussion needs to talk about cold adaptation; why trade-offs?
 - Walk through results better
- Thermal niche ms: **Lacey and I working on discussion**
 - Stressed in nature MS: Samples to rerun.
 - update: Curtis can no longer work+ write on project

- **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- Proteome stability project:
 - **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)
- Modulation of Hsp ms:
 - make fig 2 without spline curves with just points
 - grab elevation data for each sampling point in R
- Attending SICB - Jan 3-8 New Orleans, Give a talk about range limits paper.
 - **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
 - **Talk title: Northern range limits of a common forest ant is associated with trade-offs in cold physiology**
 - Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app.** I need to find out pricing (~ \$1000) and then get everything signed. Waiting to find better flight prices.

- Application submitted today 2016-10-18

- Thesis related [FORMS FOUND HERE](#)

- Formatting:

- Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - Dissertation Abstract is in multiple paragraphs, but for dissertation itself, make 1 paragraph

- [Deadlines](#):

1. Intent to graduate: February 1st for May.
2. Send defense committee form to grad college—now
3. Graduate college format check March 4th
4. Defense notice 3 weeks before defense (oral defense by March 24th).
5. Final thesis April 7th.

<div id='id-section90'>

Page 90: 2016-10-25. Meeting with M Pespeni

Meeting time, Wednesday 2-4; 2016-10-26

Things to discuss

- Potential post doc opportunity at MBL(Marine Biological Laboratory)

- Previous email pitch with prospective post doc mentor

A question that excites me is how organisms persist and respond to environmental change in natural populations?

(This falls into 2 strategic themes of MBL: comparative evolution and genomics, and organismal adaptation and resiliency to climate change) Well, response to selection depends on their quantitative genetic architecture (variance-covariance G matrix) and selection gradient.

Monogonont rotifers seem like a really great system to explore this question with a combination of field surveys, and lab studies. For example, since their lifespan is relatively short, there should be a lot of evolutionary responses within a season.

So this would involve sampling rotifers throughout the season (4-6 times), then genotyping (GBS or rad-seq, maybe whole genome sequencing if its not too large) and establishing clones each time. Genotyping would detect shifts in allele frequencies with respect to the environment that changes, signature of evolution.

Establishing clones would allow one to assess the evolutionary potential at each point in the season by estimating the variance-covariance G matrix. Selection should erode genetic variation, so G should be altered throughout the season that may hinder or facilitate future responses.

And evolutionary potential is really unique in rotifers because they can be clonal or mate. So I'd be interested in comparing G between these life strategies.

The problem with a G matrix, is that we have no clue what the key molecular players are: so to tackle this problem, one could leverage the collected data into a qtl analysis too.

I think it is fun to think about the evolutionary potential to environmental change for organisms that can switch from asexual to sexual reproduction. If you compare the G matrix between them, sexually produced offspring populations should have more genetic variance than clonal offspring populations. These animals are resilient to environmental change because of this! So it'd be cool to compare G between asexual vs sexual and whether trade-offs can shift among traits.

Melissa advice; write down questions, hypotheses and aims that will help facilitate the discussion

- sequencing for ecological genomics? : multieplex individuals , you'll need 1-2x coverage: or pool individuals and estimate allele frequencies (sequence RNA or DNA); if RNA, then you'll have potential for allele-biased expression influencing allele freq estimates. If DNA genomics from a pooled sample, then

playing field is level, but genomes are big. 2 ways to do it:

HARP(genotype parents(known)-then subsequent genotype larvae; needs and requires low coverage—then reconstruct allele frequencies).

- How many individuals per pop (10-100?) depends on how large your pop size (only need a few individuals)? If small pop—need more and there will be more random chance. Look up Christian Slaughter (experimental evolution). Look up papers ; power analyses.
- **GTA for ecological genomics**

2016-10-27 Brent's thoughts

Ask about Isofemale lines

- look up genome size (it's .35 pg)
 - What is changing G? What is the predictive power? model it.
 - Try to talk to Mike Angiletta, Rus Lande. (Genetic accomodation and assimilation)
-

<div id='id-section91'>

Page 91: 2016-10-26 SICB meeting talk

details for my talk

<div id='id-section92'>

Page 92: 2016-10-27. Proteome stability project update

- reminder: generated unfolding reaction norms for 6 ant colonies (3 colonies per species).
- received data from Bethany 2016-10-26
 - excel sheets wit relative abundance to first sample is in: 2016_Protein_stability_evolution/Data/2016/10_Oct
 - in this path, you'll see 3 folders, 1 set of samples queried against 18 species (it actually has a combo of ants/microbes because Bethany just took the top 18 searches) from uniprot. The other folder queries the NCBI database. And the last folder contains raw mass spec files.
- Bethany is going to run more of the sample to see if we can ID more proteins.

<div id='id-section93'>

Page 93: 2016-10-31. CTmax and Hsp reaction norm stats

Stats overview:

1. Effect of local environment (Tmax and habitat type) on basal xp and other parameters.

Basal xp

```
summary(lm(b70~bio5+habitat_v2,data=b70))
```

Call:

```
lm(formula = b70 ~ bio5 + habitat_v2, data = b70)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.10674	-0.34255	0.07049	0.44475	1.56186

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	10.915110	1.957855	5.575	1.74e-06
bio5	0.005714	0.006543	0.873	0.38
habitat_v2flat woods	-0.124177	0.365522	-0.340	0.73

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.87 on 41 degrees of freedom

Multiple R-squared: 0.01836, Adjusted R-squared: -0.02952

F-statistic: 0.3835 on 2 and 41 DF, p-value: 0.6839

summary(lm(b83~bio5+habitat_v2,data=b83))

Call:

lm(formula = b83 ~ bio5 + habitat_v2, data = b83)

Residuals:

Min	1Q	Median	3Q	Max
-2.16408	-0.49336	0.03001	0.64313	1.96466

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	12.751247	2.132030	5.981	2.34e-07 ***
bio5	-0.002689	0.007140	-0.377	0.708
habitat_v2flat woods	-0.480410	0.362781	-1.324	0.191

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

' ' 1

Residual standard error: 0.9811 on 50 degrees of freedom
Multiple R-squared: 0.06047, Adjusted R-squared: 0.02289
F-statistic: 1.609 on 2 and 50 DF, p-value: 0.2103

summary(lm(b40~bio5+habitat_v2,data=b40))

Call:

lm(formula = b40 ~ bio5 + habitat_v2, data = b40)

Residuals:

Min	1Q	Median	3Q	Max
-1.7137	-0.6858	-0.1241	0.3196	3.0774

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	13.721471	2.714855	5.054	1.36e-05 ***
bio5	0.004703	0.009120	0.516	0.609
habitat_v2flat woods	-0.381890	0.509208	-0.750	0.458

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
' ' 1				

```
Residual standard error: 1.123 on 35 degrees of freedom  
Multiple R-squared:  0.01669,   Adjusted R-squared: -0.0  
395  
F-statistic: 0.297 on 2 and 35 DF,  p-value: 0.7449
```

Hsp70 (hsc70-4 h2) params (slope,Tm,max)

```
apply(b[,3:5],2,function(x){summary(lm(x~b$bio5+b$habitat  
_v2))})
```

```
$FC_hsc70_1468_max
```

Call:

```
lm(formula = x ~ b$bio5 + b$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-20.536	-8.414	-1.652	4.839	30.045

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.90802	28.24097	-0.032	0.9744
b\$bio5	0.13378	0.09449	1.416	0.1635

67

```
b$habitat_v2flat woods 20.35661     4.86449    4.185 0.0001
27 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1
```

Residual standard error: 12.34 on 46 degrees of freedom

Multiple R-squared: 0.4224, Adjusted R-squared: 0.39

73

F-statistic: 16.82 on 2 and 46 DF, p-value: 3.288e-06

\$FC_hsc70_1468_slope

Call:

```
lm(formula = x ~ b$bio5 + b$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.91667	-0.22656	0.08771	0.27554	0.87662

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.213328	1.023087	0.209	0.835
b\$bio5	0.002091	0.003423	0.611	0.544

75

31

```
b$habitat_v2flat woods 0.494706 0.176226 2.807 0.007  
31 **  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' '
```

Residual standard error: 0.4471 on 46 degrees of freedom

Multiple R-squared: 0.2228, Adjusted R-squared: 0.189

F-statistic: 6.595 on 2 and 46 DF, p-value: 0.003032

\$FC_hsc70_1468_Tm

Call:

```
lm(formula = x ~ b$bio5 + b$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.23057	-0.46633	-0.00151	0.62405	1.24574

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	35.043684	1.956972	17.907	< 2e-16 ***
b\$bio5	0.003766	0.006548	0.575	0.568

014

b\$habitat_v2flat woods 1.372953 0.337088 4.073 0.000

181 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.8552 on 46 degrees of freedom

Multiple R-squared: 0.3566, Adjusted R-squared: 0.32

87

F-statistic: 12.75 on 2 and 46 DF, p-value: 3.931e-05

Hsp83 params (slope,Tm,max)

```
apply(u[,9:11],2,function(x){summary(lm(x~u$bio5+u$habitat_v2))})
```

\$FC_Hsp83_279_max

Call:

```
lm(formula = x ~ u$bio5 + u$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.8432	-2.7507	-0.7032	2.3143	11.2074

Coefficients:

Estimate	Std. Error	t value	Pr(>
----------	------------	---------	------

```
t|)

(Intercept)          8.059606   8.941439   0.901   0.37
208
u$bio5              -0.002729   0.029897   -0.091   0.92
766
u$habitat_v2flat woods 4.720030   1.550712   3.044   0.00
386 **

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1
```

Residual standard error: 4.06 on 46 degrees of freedom
Multiple R-squared: 0.2054, Adjusted R-squared: 0.1709
F-statistic: 5.947 on 2 and 46 DF, p-value: 0.005045

\$FC_Hsp83_279_slope

Call:

```
lm(formula = x ~ u$bio5 + u$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.8619	-0.5948	0.1370	0.6879	1.3637

Coefficients:

Estimate	Std. Error	t value	Pr(>
----------	------------	---------	------

```
t|)  
(Intercept) -1.056652 1.865514 -0.566 0.  
574  
u$bio5 0.008211 0.006238 1.316 0.  
195  
u$habitat_v2flat woods 0.301698 0.323536 0.933 0.  
356
```

Residual standard error: 0.8471 on 46 degrees of freedom
Multiple R-squared: 0.09876, Adjusted R-squared: 0.05
957
F-statistic: 2.52 on 2 and 46 DF, p-value: 0.09148

\$FC_Hsp83_279_Tm

Call:

```
lm(formula = x ~ u$bio5 + u$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.4767	-0.7621	0.1731	0.9167	2.6581

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	31.80124	3.37214	9.431	2.54e-12 ***

```

u$bio5           0.01076    0.01128    0.955  0.3446
99
u$habitat_v2flat woods  2.16554    0.58483    3.703  0.0005
69 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 1.531 on 46 degrees of freedom
Multiple R-squared:  0.3423,    Adjusted R-squared:  0.31
37
F-statistic: 11.97 on 2 and 46 DF,  p-value: 6.533e-05

```

Hsp40 (hsc70-4 h2) params

(slope,Tm,max)

```
apply(n[,6:8],2,function(x){summary(lm(x~n$bio5+n$habitat_v2))})
```

\$FC_hsp40_541_max

Call:

```
lm(formula = x ~ n$bio5 + n$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.8615	-3.3291	-0.6736	1.7653	10.5454

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.4754213	10.9534314	0.865	0.3917
n\$bio5	-0.0009401	0.0367220	-0.026	0.9797
n\$habitat_v2flat woods	3.6490726	1.8969491	1.924	0.0609

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	'	'	'	' 1

Residual standard error: 4.749 on 44 degrees of freedom

Multiple R-squared: 0.1003, Adjusted R-squared: 0.05945

F-statistic: 2.454 on 2 and 44 DF, p-value: 0.09765

\$FC_hsp40_541_slope

Call:

```
lm(formula = x ~ n$bio5 + n$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.4300	-0.5157	0.2182	0.6412	1.3309

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.295834	1.816631	-0.163	0.871
n\$bio5	0.005677	0.006090	0.932	0.356
n\$habitat_v2flat woods	0.413173	0.314610	1.313	0.196

Residual standard error: 0.7877 on 44 degrees of freedom

Multiple R-squared: 0.1048, Adjusted R-squared: 0.06411

F-statistic: 2.576 on 2 and 44 DF, p-value: 0.08755

\$FC_hsp40_541_Tm

Call:

```
lm(formula = x ~ n$bio5 + n$habitat_v2)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.7066	-1.0076	0.2038	0.9873	3.5691

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
--	----------	------------	---------	----------

```

| )
(Intercept)           39.14520   3.84815  10.172 3.93e-
13 ***
n$bio5                -0.01175   0.01290  -0.911  0.3674
44
n$habitat_v2flat woods 2.46904   0.66643   3.705  0.0005
88 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 1.669 on 44 degrees of freedom
Multiple R-squared:  0.2539,    Adjusted R-squared:  0.2
2
F-statistic: 7.487 on 2 and 44 DF,  p-value: 0.00159

```

**Summary: no sig effect of Tmax (bio5)
on parameters, but habitat type does
in some cases:**

Table summary:

Parameter	hsp83	hsc70.4.h2	hsp40
basal	no	no	no
slope	no	yes	no

Tm	yes	yes	yes
max	yes	yes	no

<div id='id-section94'>

Page 94: 2016-10-31; 2016-11-01.

Climate cascade meeting setup and notes

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - figure 3, SHC says to switch back branches
 - Wrote up methods and results
 - Submit to? MBE, evolution, Goerge Somero and Brent think PNAS is a good fit. SHC and NJG thoughts?
 - reference for rad-seq:HF3-picea,fbragg2-floridana,KH4-ashmeadi,Duke6-mariae,ala2-miamiana, Lex13-rudis
- **Multiple stressors ms:**
 - submitted **2016-10-24**
 - in review **2016-11-01**
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate

- **Thermal niche ms:** Lacey and I working on discussion
- **Stressed in nature MS: Samples to rerun.**
 - update: Curtis can no longer work+ write on project
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- **Proteome stability project:**
 - **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)
- **Modulation of Hsp ms:**
 - make fig 2 without spline curves with just points (done)
 - grab elevation data for each sampling point in R (done)

2. Attending SICB - Jan 3-8 New Orleans, Give a talk about range limits paper.

- **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
- Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app** Application

submitted today 2016-10-18

- Bought hotel, rooming with Emily M.,
need to buy airplane tickets

3. Thesis related FORMS FOUND HERE

- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
- Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.

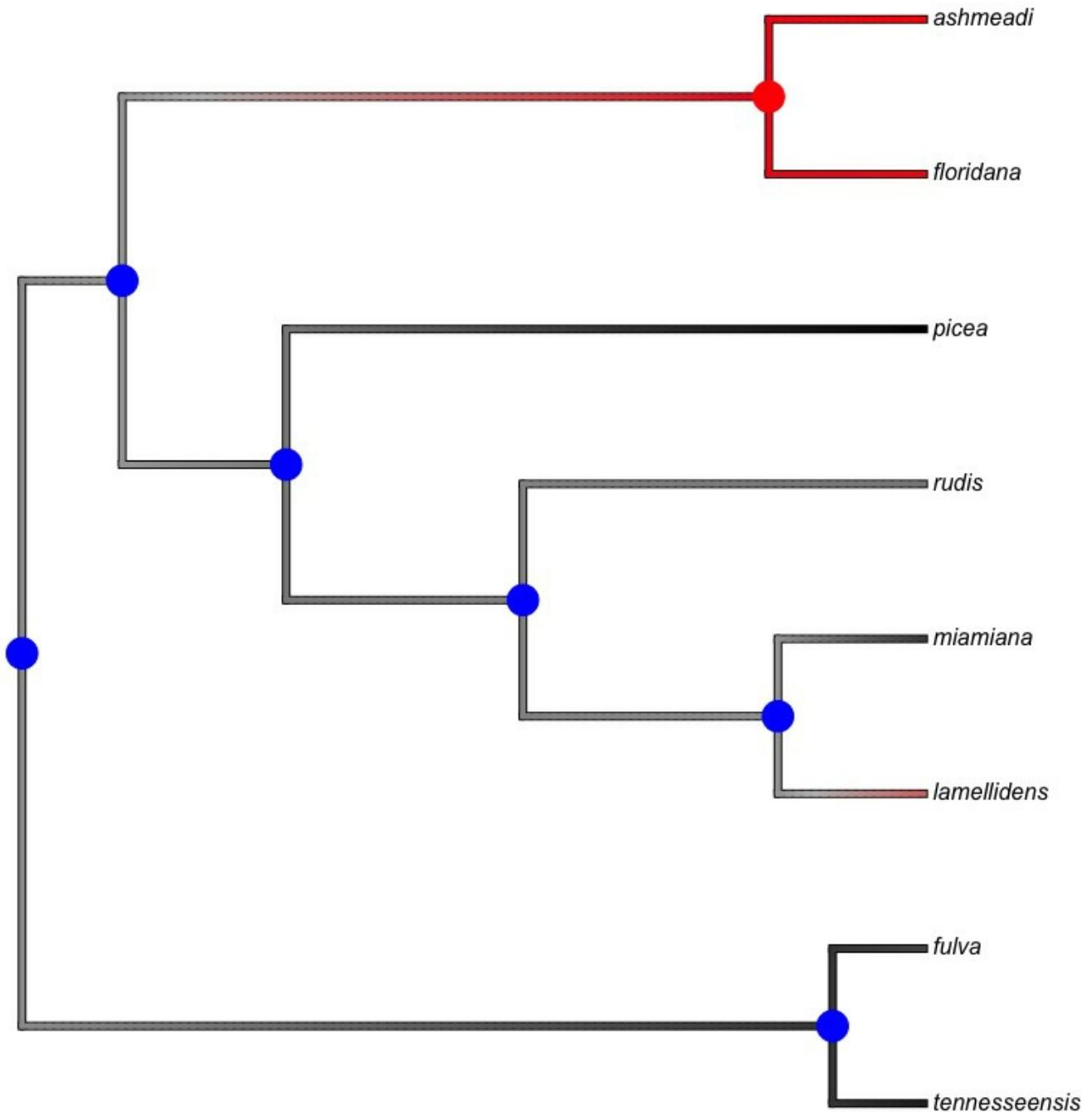
<div id='id-section95'>

Page 95: 2016-11-02. Ancestral trait reconstruction and CTmax PGLS ANBE common garden

Ancestral trait reconstruction

```
cols<-ifelse(esthab[,1]>esthab[,2],"blue","red")
par(mar=c(1,1,1,1))
plot(ult.tree1,cex=.5)
nodelabels(pch=19,cex=.75,col=cols)
```

```
#obj<-contMap(ult.tree1,trait,plot=FALSE,fsize=.1,method=
"fastAnc")
#obj$cols[ ]<-
obj<-setMap(obj,colors=colorRampPalette(c("black","gray",
"red"))(length(obj$cols)))
plot(obj,legend=FALSE)
nodelabels(pch=19,cex=3,col=cols)
```



Using ancTHRESH: [Paper here](#); troubleshooting error

Species level ancestral state reconstruction

```
####using ancThresh (revell 2014 evolution)
habitat<-as.character(sm.dat2$Habitat)
names(habitat)<-spec.tree$tip.label
```

```
er<-ancThresh(spec.tree,habitat,model="BM",ngen=20000)
```

```
$ace
```

	DF	FW
--	----	----

9	0.9411765	0.05882353
---	-----------	------------

10	1.0000000	0.00000000
----	-----------	------------

11	0.9411765	0.05882353
----	-----------	------------

12	1.0000000	0.00000000
----	-----------	------------

13	1.0000000	0.00000000
----	-----------	------------

14	1.0000000	0.00000000
----	-----------	------------

15	0.1176471	0.88235294
----	-----------	------------

```
$mcmc
```

	9	10	11	12	13	14	15
--	---	----	----	----	----	----	----

1	FW						
---	----	----	----	----	----	----	----

2	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

3	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

4	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

5	FW	DF	FW	DF	DF	DF	FW
---	----	----	----	----	----	----	----

6	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

7	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

8	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

9	DF	DF	DF	DF	DF	DF	FW
---	----	----	----	----	----	----	----

10	DF	DF	DF	DF	DF	DF	FW
----	----	----	----	----	----	----	----

11	DF	DF	DF	DF	DF	DF	FW
----	----	----	----	----	----	----	----

12	DF						
----	----	----	----	----	----	----	----

13	DF	DF	DF	DF	DF	DF	FW
----	----	----	----	----	----	----	----

14	DF						
----	----	----	----	----	----	----	----

15	DF	DF	DF	DF	DF	DF	FW
----	----	----	----	----	----	----	----

```
16 DF DF DF DF DF FW  
17 DF DF DF DF DF FW  
18 DF DF DF DF DF FW  
19 DF DF DF DF DF FW  
20 DF DF DF DF DF FW  
21 DF DF DF DF DF FW
```

\$par

	gen	DF	FW	logLik
1	0	0	Inf	-20.447477
2	1000	0	Inf	-15.497509
3	2000	0	Inf	-7.454956
4	3000	0	Inf	-7.583405
5	4000	0	Inf	-12.443948
6	5000	0	Inf	-7.824642
7	6000	0	Inf	-18.411281
8	7000	0	Inf	-8.957609
9	8000	0	Inf	-10.366720
10	9000	0	Inf	-16.032346
11	10000	0	Inf	-10.064034
12	11000	0	Inf	-12.232852
13	12000	0	Inf	-11.150292
14	13000	0	Inf	-10.575092
15	14000	0	Inf	-11.406253
16	15000	0	Inf	-18.842795
17	16000	0	Inf	-13.001441
18	17000	0	Inf	-10.961662
19	18000	0	Inf	-10.054596

```
20 19000  0 Inf -13.251417
```

```
21 20000  0 Inf -13.367340
```

```
$liab
```

	ashmeadi	floridana	picea	rudis	miamiana
--	----------	-----------	-------	-------	----------

```
lamellidens
```

1	0.4256323	0.41631348	-0.9982783	-0.7611795	-0.2870708
---	-----------	------------	------------	------------	------------

			-0.7228374		
--	--	--	------------	--	--

2	1.7161387	2.01928329	-0.5807896	-2.7040955	-2.1537732
---	-----------	------------	------------	------------	------------

			-1.3879036		
--	--	--	------------	--	--

3	0.2283514	0.33912611	-1.4167395	-0.8156619	-2.1513080
---	-----------	------------	------------	------------	------------

			-1.5626112		
--	--	--	------------	--	--

4	0.2245830	0.04840176	-0.1672812	-0.5182768	-1.4955228
---	-----------	------------	------------	------------	------------

			-1.2845957		
--	--	--	------------	--	--

5	2.8412873	2.52791039	-0.9945787	-0.1217164	-1.0169444
---	-----------	------------	------------	------------	------------

			-0.8199481		
--	--	--	------------	--	--

6	0.1611044	0.07884604	-1.6873462	-1.9551489	-2.5062990
---	-----------	------------	------------	------------	------------

			-1.8735545		
--	--	--	------------	--	--

7	0.6062956	0.67119993	-1.7010454	-3.1098352	-3.5942080
---	-----------	------------	------------	------------	------------

			-3.6599400		
--	--	--	------------	--	--

8	0.2781314	0.90142051	-0.9775805	-1.4564663	-2.0262664
---	-----------	------------	------------	------------	------------

			-1.9955650		
--	--	--	------------	--	--

9	0.4831741	0.32616809	-1.2045168	-1.4714718	-1.8546322
---	-----------	------------	------------	------------	------------

			-1.9443720		
--	--	--	------------	--	--

10	0.9545092	0.91442789	-1.9678349	-2.8803130	-2.0902628
----	-----------	------------	------------	------------	------------

			-2.1420066		
--	--	--	------------	--	--

11	0.5334539	0.45214518	-0.6975138	-1.7053550	-1.0307576
----	-----------	------------	------------	------------	------------

			-1.3671555		
--	--	--	------------	--	--

12	0.3455715	0.36069694	-0.8319524	-1.3187262	-0.3870102	
	-0.6778155					
13	1.0041606	0.14710811	-1.5117681	-1.1249621	-2.0614504	
	-1.6515543					
14	0.3898680	0.07794064	-2.5767746	-2.2195374	-2.0482449	
	-2.5311433					
15	0.1438718	0.01582491	-0.6168032	-1.8867342	-2.1514162	
	-2.2116893					
16	1.5174549	1.25039739	-0.3146283	-0.6646803	-2.9459244	
	-2.4065327					
17	0.2949082	0.79497182	-2.3475672	-1.3544484	-1.7933900	
	-1.0633168					
18	0.1754723	0.07905511	-1.7017423	-2.8025226	-2.3548623	
	-2.7766376					
19	0.6090604	0.62077613	-2.8309318	-2.2609481	-2.4802131	
	-2.9229599					
20	0.3999724	0.84360557	-3.1013779	-2.7228427	-3.7007126	
	-3.2687614					
21	0.7346061	0.86140836	-2.1879653	-2.5777420	-3.3467673	
	-3.9166340					
		fulva tennesseensis		9	10	
11		12				
1	-0.1323211	-0.4694033	0.1077790	0.05401888	0.143	
9697	0.5549107					
2	-1.3488986	-0.9985338	-1.0447443	-0.89772529	-1.372	
6393	-1.6255353					
3	-1.0540075	-0.8284915	-0.2338892	-1.07612349	-0.574	
0115	-1.0216461					

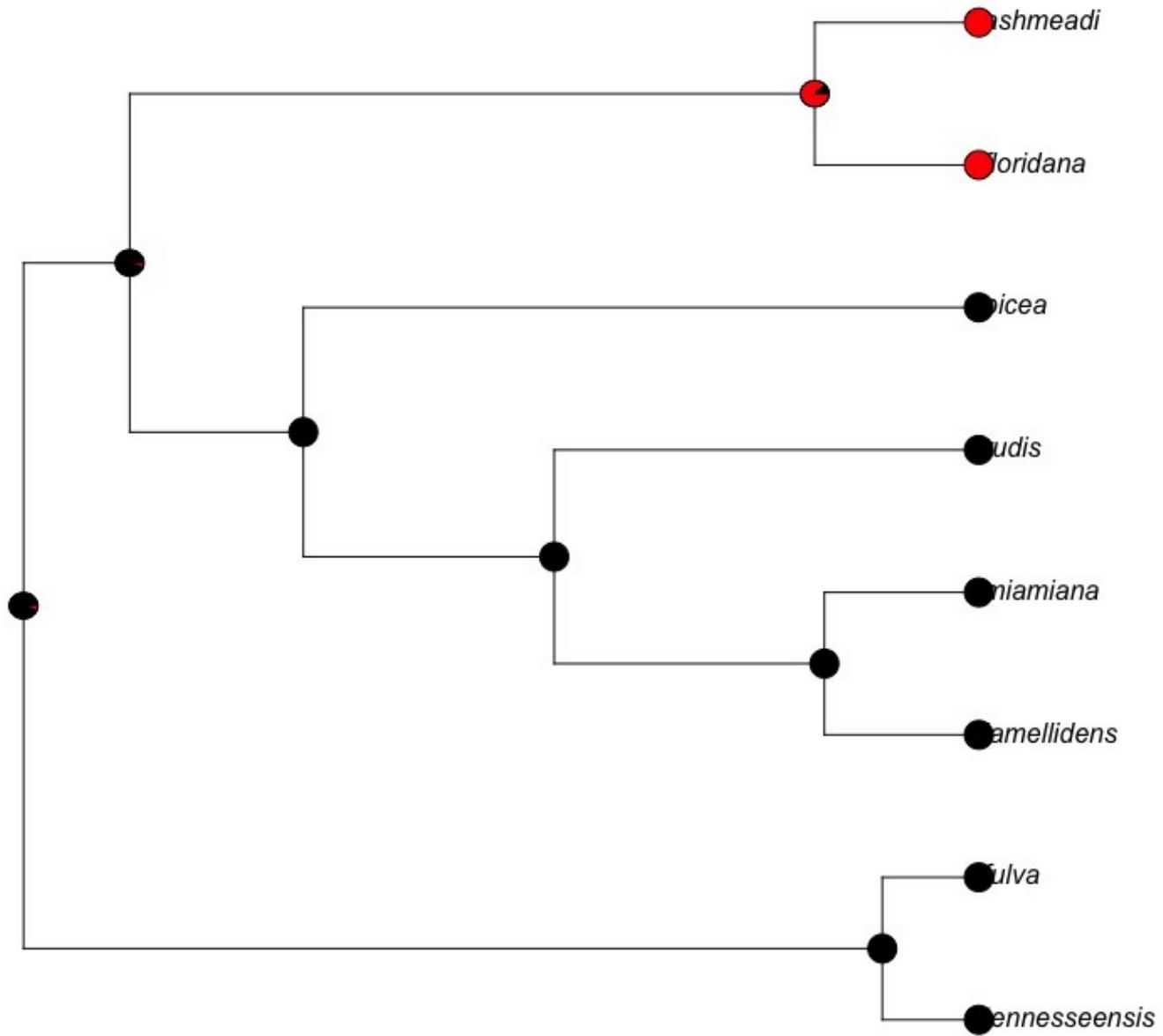
4	-0.8706067	-1.3873548	-0.4220714	-1.45708985	-0.188
5166	-0.4826553				
5	-0.4714702	-0.3575748	0.3448237	-0.79814696	0.417
9758	-0.1106286				
6	-2.0032554	-1.5914376	-1.1376738	-1.68649095	-0.922
9725	-1.5736772				
7	-2.8018002	-3.0551338	-1.8844310	-2.27322487	-1.717
0136	-2.4081014				
8	-2.3101088	-2.1159034	-1.4001121	-2.09323137	-1.459
9777	-1.8363667				
9	-2.9641505	-2.4055435	-1.0906425	-2.44676880	-0.855
6195	-0.3708003				
10	-0.8302079	-1.8660746	-1.5512891	-1.57882060	-0.911
1552	-0.7840139				
11	-1.1688755	-0.5233950	-1.0694070	-1.12396533	-1.392
7681	-0.8758065				
12	-0.8464924	-1.0075126	-1.4279378	-1.32894559	-0.802
3390	-1.1667292				
13	-0.5218728	-0.5705261	-0.0381062	-0.92343185	-0.474
4984	-1.3058079				
14	-0.6052192	-0.3901746	-1.3237021	-0.60220033	-0.699
8386	-1.6929037				
15	-1.2232572	-1.3633033	-0.3002129	-1.09482578	-1.040
4010	-0.8156088				
16	-1.6465250	-2.5813912	-1.9776983	-2.28317185	-2.025
9641	-1.1658657				
17	-3.3390536	-3.0821085	-1.7921216	-3.21916257	-2.036
0532	-2.6290063				

18	-2.8353118	-2.5476584	-2.4669142	-2.85649615	-2.284
2481	-2.4311806				
19	-2.2636787	-2.3547350	-1.7062219	-1.86053981	-1.675
8183	-1.9448948				
20	-2.2528318	-2.1913204	-1.6877972	-1.97671417	-2.175
3948	-2.6389801				
21	-0.9965233	-0.5414665	-0.4247154	-0.76066224	-0.667
4432	-1.9868298				
	13	14	15		
1	0.9519656	0.7427504	0.24005246		
2	-1.3613162	-1.6237741	1.61824721		
3	-1.0603488	-1.7249034	0.51079667		
4	-0.5200048	-1.3974409	0.28723904		
5	-0.4768141	-1.2742356	2.25125896		
6	-2.3110323	-2.0833774	0.16007229		
7	-3.6709435	-3.6396813	0.98393423		
8	-1.9847607	-2.2242295	0.69042714		
9	-1.3939044	-1.9765742	0.31119880		
10	-1.5328568	-1.8498664	1.16199792		
11	-1.5364744	-1.0508674	0.50063895		
12	-0.3373547	-0.1329506	-0.16251270		
13	-1.2442089	-1.5792052	0.10020199		
14	-1.6824202	-2.2044414	-0.04600028		
15	-1.7298317	-2.0001838	0.55174349		
16	-1.4318329	-2.2727758	1.01123041		
17	-2.2424728	-1.5377536	0.31713522		
18	-2.8320154	-2.1532767	0.21849413		
19	-2.2843315	-2.4451599	0.48568731		

20 -3.4562176 -3.2262318 0.26999207

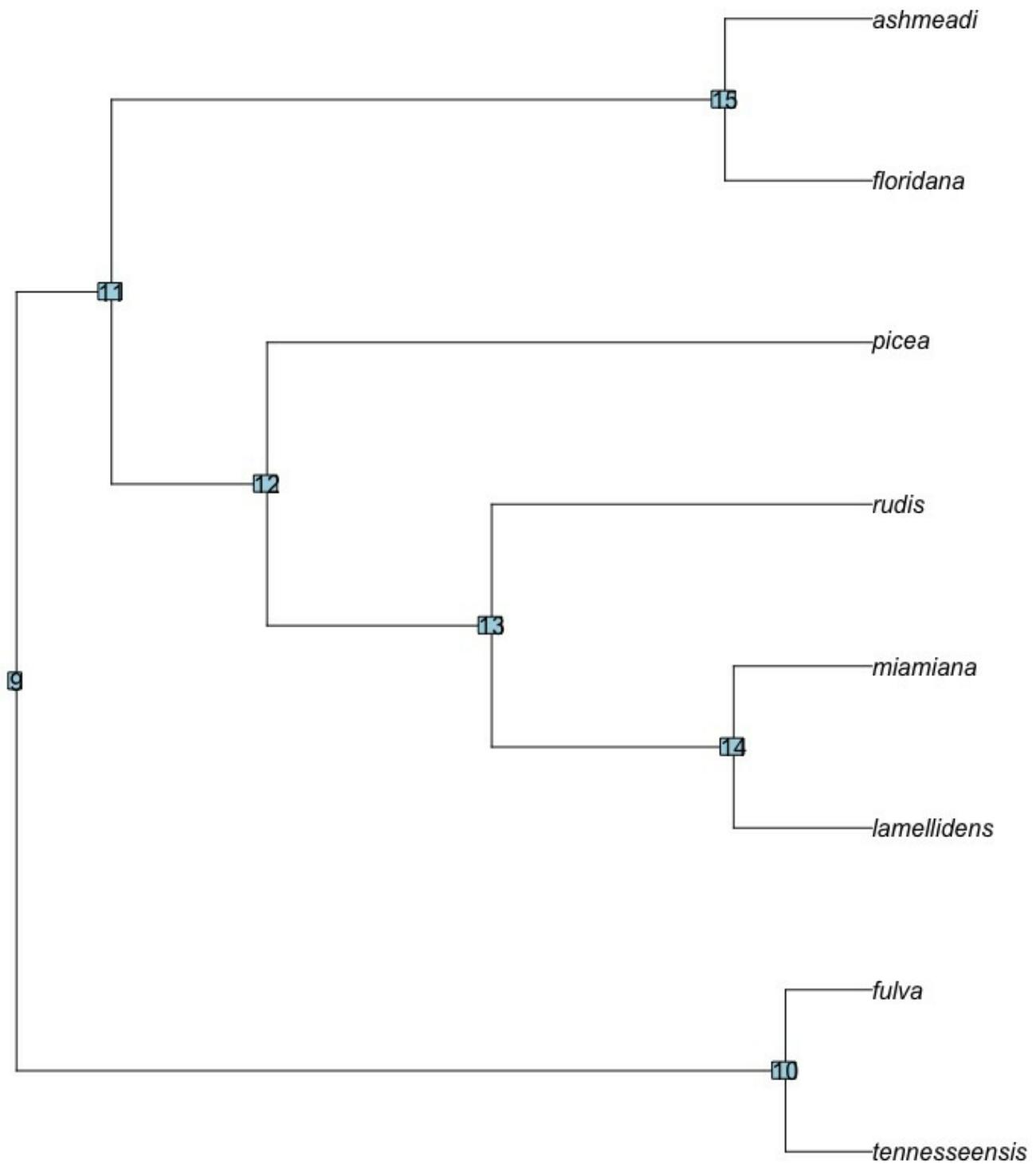
21 -2.8093678 -3.7144363 0.53795324

It automatically plots the results:



■ BMW

Reference for tree with node labels



Doing pgls in 3 ways:

1. Using colonies as tips (breaks assumptions because of reticulate evolution)

2. Forcing polytomies with species as replicates
3. Just doing species themselves (8)

1. Using colonies as tips (breaks assumptions because of reticulate evolution)

```
library(caper)
aph_phylo1$colony.id2<-as.character(aph_phylo1$colony.id2)
ult.tree1<-makeLabel(ult.tree1)
aph_phylo1$habitat_v2<-droplevels(aph_phylo1$habitat_v2)
pp<-comparative.data(phy=ult.tree1,data=aph_phylo1,names.col=colony.id2, vcv = TRUE, na.omit = FALSE, warn.dropped = TRUE)
```

```
momo<-pgls(K0_temp_worker~bio5+habitat_v2,data=pp,lambda=
"ML",bounds=list(lambda=c(0.001,1)))
summary(momo)
```

Call:

```
pgls(formula = K0_temp_worker ~ bio5 + habitat_v2, data =
pp,
lambda = "ML", bounds = list(lambda = c(0.001, 1)))
```

Residuals:

Min	1Q	Median	3Q	Max
-----	----	--------	----	-----

-2.3636 -0.6161 -0.1511 0.3177 2.8311

Branch length transformations:

kappa [Fix] : 1.000

lambda [ML] : 0.001

lower bound : 0.001, p = 1

upper bound : 1.000, p = < 2.22e-16

95.0% CI : (NA, 0.517)

delta [Fix] : 1.000

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	37.2440770	1.0902152	34.1621	< 2.2e-16 ***
bio5	0.0128318	0.0036686	3.4978	0.0007098 ***
habitat_v2flat woods	1.3750216	0.2575557	5.3387	6.157e-07 ***

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	'	'	'	1

Residual standard error: 0.8605 on 97 degrees of freedom

Multiple R-squared: 0.4051, Adjusted R-squared: 0.3929

F-statistic: 33.03 on 2 and 97 DF, p-value: 1.147e-11

It looks like the PGLS is using lambda of 0. So I tried estimating lambda and then plugging it in the PGLS model

```
#phylogenetic signal
x<-aph_phylo1$K0_temp_worker
names(x)<-aph_phylo1$colony.id2
phylosig(ult.tree1,x,test=TRUE,method="lambda")
$lambda
[1] 0.4833368

$logL
[1] -128.4395

$logL0
[1] -151.6493

$P
[1] 9.5454e-12

#phylosig(ult.tree1,x,test=TRUE,method="K",nsim=1000)

#redoing pglss with lambda from phylosig
momo3<-pglss(K0_temp_worker~habitat_v2+bio5,data=pp,lambda
=0.4833368)
summary(momo3)

Call:
pglss(formula = K0_temp_worker ~ habitat_v2 + bio5, data =
```

```
pp,  
lambda = 0.4833368)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.3928	-0.3833	0.1074	0.8404	3.3408

Branch length transformations:

```
kappa [Fix] : 1.000  
lambda [Fix] : 0.483  
delta [Fix] : 1.000
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	38.4681831	2.4673203	15.5911	<2e-16 ***
habitat_v2flat woods	0.5009582	0.5160753	0.9707	0.3341
bio5	0.0093601	0.0080294	1.1657	0.2466

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'
	' '	1		

Residual standard error: 1.082 on 97 degrees of freedom

Multiple R-squared: 0.02726, Adjusted R-squared: 0.007

F-statistic: 1.359 on 2 and 97 DF, p-value: 0.2617

2. Forcing polytomies with species as replicates

```
aph_phylo2$colony.id2<-as.character(aph_phylo2$colony.id2)
ult2.tree<-makeLabel(ult2.tree)
aph_phylo2$habitat_v2<-droplevels(aph_phylo2$habitat_v2)
pp<-comparative.data(phy=ult2.tree,data=aph_phylo2,names.
col=colony.id2, vcv = TRUE, na.omit = FALSE, warn.dropped
= TRUE)
```

```
momo<-pgls(K0_temp_worker~bio5+habitat_v2,data=pp,lambda=
"ML",bounds=list(lambda=c(0.001,1)))
summary(momo)
```

Call:

```
pgls(formula = K0_temp_worker ~ bio5 + habitat_v2, data =
pp,
lambda = "ML", bounds = list(lambda = c(0.001, 1)))
```

Residuals:

Min	1Q	Median	3Q	Max
-5.2426	-1.0208	-0.0880	0.9807	5.7995

Branch length transformations:

kappa [Fix] : 1.000
lambda [ML] : 0.991
lower bound : 0.001, p = 0.1627
upper bound : 1.000, p = 0.69339
95.0% CI : (NA, NA)
delta [Fix] : 1.000

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	38.9471677	2.2565730	17.2594	<2e-16 ***
bio5	0.0080535	0.0067643	1.1906	0.2367
habitat_v2flat woods	-0.0036539	0.4282500	-0.0085	0.9932

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'
	'.'	1		

Residual standard error: 1.839 on 97 degrees of freedom

Multiple R-squared: 0.01442, Adjusted R-squared: -0.005903

F-statistic: 0.7095 on 2 and 97 DF, p-value: 0.4944

Again, try to estimate lambda and then plug and chug

```
x<-aph_phylo2$K0_temp_worker  
names(x)<-aph_phylo2$colony.id2  
phylosig(ult2.tree,x,test=TRUE,method="lambda")  
$lambda  
[1] 0.9759065
```

```
$logL  
[1] -124.9107
```

```
$logL0  
[1] -151.6493
```

```
$P  
[1] 2.616073e-13
```

```
momo3<-pgls(K0_temp_worker~habitat_v2+bio5,data=pp,lambda  
=0.9759065)  
summary(momo3)
```

Call:

```
pgls(formula = K0_temp_worker ~ habitat_v2 + bio5, data =  
pp,  
lambda = 0.9759065)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.9077	-1.1334	0.0055	1.0044	5.3637

Branch length transformations:

kappa [Fix] : 1.000
lambda [Fix] : 0.976
delta [Fix] : 1.000

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	39.2879592	2.4115458	16.2916	<2e-16 ***
habitat_v2flat woods	0.0118194	0.4401407	0.0269	0.9786
bio5	0.0069203	0.0073845	0.9371	0.3510

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	'	'	'	' 1

Residual standard error: 1.766 on 97 degrees of freedom

Multiple R-squared: 0.009022, Adjusted R-squared: -0.0141

F-statistic: 0.4416 on 2 and 97 DF, p-value: 0.6443

3. Just doing species themselves (8)

```
#PGLS with caper  
spec.tree<-makeLabel(spec.tree)  
smp<-comparative.data(phy=spec.tree,data=sm.dat2,names.co  
l=Species, vcv = TRUE, na.omit = FALSE, warn.dropped = TR  
UE)
```

```
spmod<-pgls(CTmax~Habitat+Tmax,data=smp,lambda="ML")  
#spmod<-pgls(CTmax~Habitat,data=smp,lambda=0.885536,bound  
s=list(lambda=c(0.001,1)))  
summary(spmod)  
Call:  
pgls(formula = CTmax ~ Habitat + Tmax, data = smp, lambda  
= "ML")
```

Residuals:

Min	1Q	Median	3Q	Max
-0.7707	-0.1147	0.0567	0.3244	0.5081

Branch length transformations:

```
kappa [Fix] : 1.000  
lambda [ ML] : 0.000  
lower bound : 0.000, p = 1  
upper bound : 1.000, p = 0.0072118  
95.0% CI : (NA, 0.738)
```

```
delta [Fix] : 1.000
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	37.500342	2.914324	12.8676	5.048e-05 ***
HabitatFW	1.462473	0.435376	3.3591	0.02013 *
Tmax	0.011812	0.009520	1.2407	0.26975

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

Residual standard error: 0.4878 on 5 degrees of freedom

Multiple R-squared: 0.7947, Adjusted R-squared: 0.7125

F-statistic: 9.674 on 2 and 5 DF, p-value: 0.01911

```
profile_lambda=pgls.profile(spmod, which="lambda")  
plot(profile_lambda)
```

```
n<-sm.dat2$CTmax
```

```
names(n)<-sm.dat2$Species
```

```
phylosig(spec.tree,n,method="lambda",test=TRUE)
```

```
$lambda
```

```
[1] 0.885536
```

```
$logL
```

```
[1] -8.958222
```

```
$logL0  
[1] -10.06035
```

```
$P  
[1] 0.1376303
```

```
spmod<-pgls(CTmax~Habitat+Tmax,data=smp,lambda=0.885536)  
summary(spmod)
```

Call:

```
pgls(formula = CTmax ~ Habitat + Tmax, data = smp, lambda  
= 0.885536)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.93013	-0.02903	0.07964	0.38357	1.47947

Branch length transformations:

```
kappa [Fix] : 1.000  
lambda [Fix] : 0.886  
delta [Fix] : 1.000
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	39.3419794	4.4119919	8.9171	0.0002954 ***

```
HabitatFW      1.6291565  0.9278322  1.7559  0.1394628
Tmax          0.0055482  0.0145875  0.3803  0.7193135
...
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.8271 on 5 degrees of freedom
Multiple R-squared:  0.467,  Adjusted R-squared:  0.2538
F-statistic:  2.19 on 2 and 5 DF,  p-value: 0.2074
```

<div id='id-section96'>

Page 96: 2016-11-03. notes from skype meeting with KG, potential post doc opp

Marine Biological Labs, Hibbitt Early Career Fellows Program

about fellowship: its brand new for MBL; its trying to bring in new investigators early in their careers

MBL; resident and visiting scientists; there are a lot of courses in the summer (10 days to 6 weeks); teachers come from all over the world; Whitman scholars are fellowships that PIs can establish labs; groups of researchers meet here;

Other foundations:

1. Charles King foundation/trust
2. Life sciences research foundation
3. Ford Foundation
4. Hell and Hay whitney foundation?

check on deadlines

There is a genome for 15 different species. huge range in genome sizes, why?

Bioinformatics; david, mark welsh (bay paul center); Lots of people do pool-seq; own Illumina hi-seq ; miseq; sanger sequencing. play up bioinformatics resource;

MBL are conveners; convening power

<div id='id-section97'>

Page 97: 2016-11-04. ms in prep

first authored

1. multiple stressors (submitted)
2. Curtis, stress in nature; submit to functional ecology
3. rxn norm of Hsps and CTmax; submit to PNAS
4. range limits paper with Jordan and Megan ; submit to American Naturalist

5. Modulation of Hsp ms (in review)
6. Proteome stability project (a stretch...)

with collaborators

1. Comparative ramp papers (CP lead?); submit to current biology?
 2. (co-lead author) thermal niche paper with LChick; submit molecular ecology?
 3. CNP work with katie miller (submit where?)
-

<div id='id-section98'>

Page 98: 2016-11-08. climate cascade meeting

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - figure 3, SHC says to switch back branches
 - Wrote up methods and results-- go over with Nick then send to SHC
 - Submit to PNAS
- **Multiple stressors ms:**
 - submitted **2016-10-24**
 - in review **2016-11-01**
- **Range limits ms:** SHC lab gave verbal edit, still need

to incorporate

- **Thermal niche ms:** Lacey and I working on discussion...eta?
- **Stressed in nature MS: Samples to rerun.**
 - update: Curtis can no longer work+ write on project
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- **Proteome stability project:**
 - **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)

2. **Attending SICB - Jan 3-8 New Orleans,** Give a talk about range limits paper.

- **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
- Apply for funding. Suitor Travel Grant Deadline is october 31
 - **Wrote up suiter award app** Application submitted today 2016-10-18
 - Bought hotel, rooming with Emily M., airplane tickets

3. Thesis related FORMS FOUND HERE

- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - started outline
- Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.

<div id='id-section99'>

Page 99: 2016-11-08. writing session with NJG

Writing Hsp reaction norm + CTmax ms in PNAS format

1. Someting to explore: variance among colony level means of CTmax in open vs closed habitats

- Narrow variance in warmer places could mean more stabilizing selection
 -
2. Try variance partitioning CTmax into Hsp, local environment, and phylogenetics
- Make CTmax vs Tmax figures with overlay of habitat type.
 - regress against latitude and PCA of climate variables too
 - try framing in terms of integrating *proximal* and *ultimate* explanations
3. put rxn norms in better context of theory; what is the alternative to hotter is better?
- Frazier et al. 2006, *AmNat*; the alternative is shifts in rxn norm horizontally, but not vertically= perfect-compensation hypothesis. In other words, biochemical adaptation can overcome rate-limiting effects of low temperature so that rmax is independent of Topt. Not mentioned in this explanation is that there can be constraints at higher temperatures that can potentially cause this pattern.
-

<div id='id-section99.1'>

1. among colony variance

```
ddply(Aph.dat,. (habitat_v2),summarize,CTmax=mean(K0_temp_
```

```
worker),var=var(K0_temp_worker))
```

	habitat_v2	CTmax	var
1	deciduous forest	41.04248	0.9443724
2	flat woods	42.77917	0.1750000

```
<div id='id-section99.2'>
```

PCA of climate variables

```
bclim<-princomp(scale(cbind(Aph.dat[,21:39])))
```

```
summary(bclim)
```

Importance of components:

	Comp.1	Comp.2	Comp.3
Comp.4	Comp.5	Comp.6	Comp.7
Standard deviation	3.6328923	1.7748683	1.19556867
	0.7430677	0.46454501	0.335626591
	0.215453516		
Proportion of Variance	0.7016431	0.1674725	0.07599067
	0.03187406	0.01147273	0.005988581
	0.002467848		
Cumulative Proportion	0.7016431	0.8691156	0.94510623
	0.97698029	0.98845302	0.994441598
	0.996909446		

```
knitr:::kable(round(bclim$loadings[,1:2],3))
```

	Comp.1	Comp.2
bio1	-0.269	-0.035
bio2	-0.144	-0.354

bio3	-0.268	-0.059
bio4	0.271	0.015
bio5	-0.249	-0.102
bio6	-0.267	-0.029
bio7	0.267	-0.013
bio8	-0.214	-0.040
bio9	-0.265	-0.073
bio10	-0.258	-0.061
bio11	-0.270	-0.034
bio12	-0.231	-0.123
bio13	-0.230	0.171
bio14	0.078	-0.495
bio15	-0.215	0.319
bio16	-0.238	0.148
bio17	0.058	-0.514
bio18	-0.248	0.145
bio19	-0.145	-0.385

<div id='id-section99.3'>

regression models; taking first two

pcas that explain 86% of variation

```
pcmod<-lm(K0_temp_worker~Comp.1*habitat_v2+Comp.2*habitat_v2 ,data=Aph.dat)
summary(stepAIC(pcmod,direction="both"))

Call:
lm(formula = K0_temp_worker ~ Comp.1 + habitat_v2 + Comp.2, data = Aph.dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.0136	-0.3372	0.1448	0.5228	1.5893

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	41.06999	0.10129	405.476	< 2e-16	

Comp.1	-0.04962	0.03006	-1.651	0.1020	
habitat_v2flat woods	1.56474	0.30657	5.104	1.68e-06	

Comp.2	-0.09366	0.05213	-1.797	0.0755	
.					

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ''''
	1				

```

Residual standard error: 0.8862 on 96 degrees of freedom
Multiple R-squared:  0.3797,    Adjusted R-squared:  0.36
F-statistic: 19.59 on 3 and 96 DF,  p-value: 5.466e-10

```

regressions with Tmax, habitat

```

umod<-lm(K0_temp_worker~bio5*habitat_v2 ,data=Aph.dat)
summary(stepAIC(umod,direction="both"))

Call:
lm(formula = K0_temp_worker ~ bio5 + habitat_v2, data = Aph.dat)

```

Residuals:

Min	1Q	Median	3Q	Max
-3.8297	-0.3348	0.2332	0.5586	1.4826

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
)				
(Intercept)	37.237343	1.084085	34.349	< 2e-16 ***
bio5	0.012855	0.003649	3.523	0.000652 ***
habitat_v2flat woods	1.376747	0.255980	5.378	5.2e-07 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
‘ ’ 1

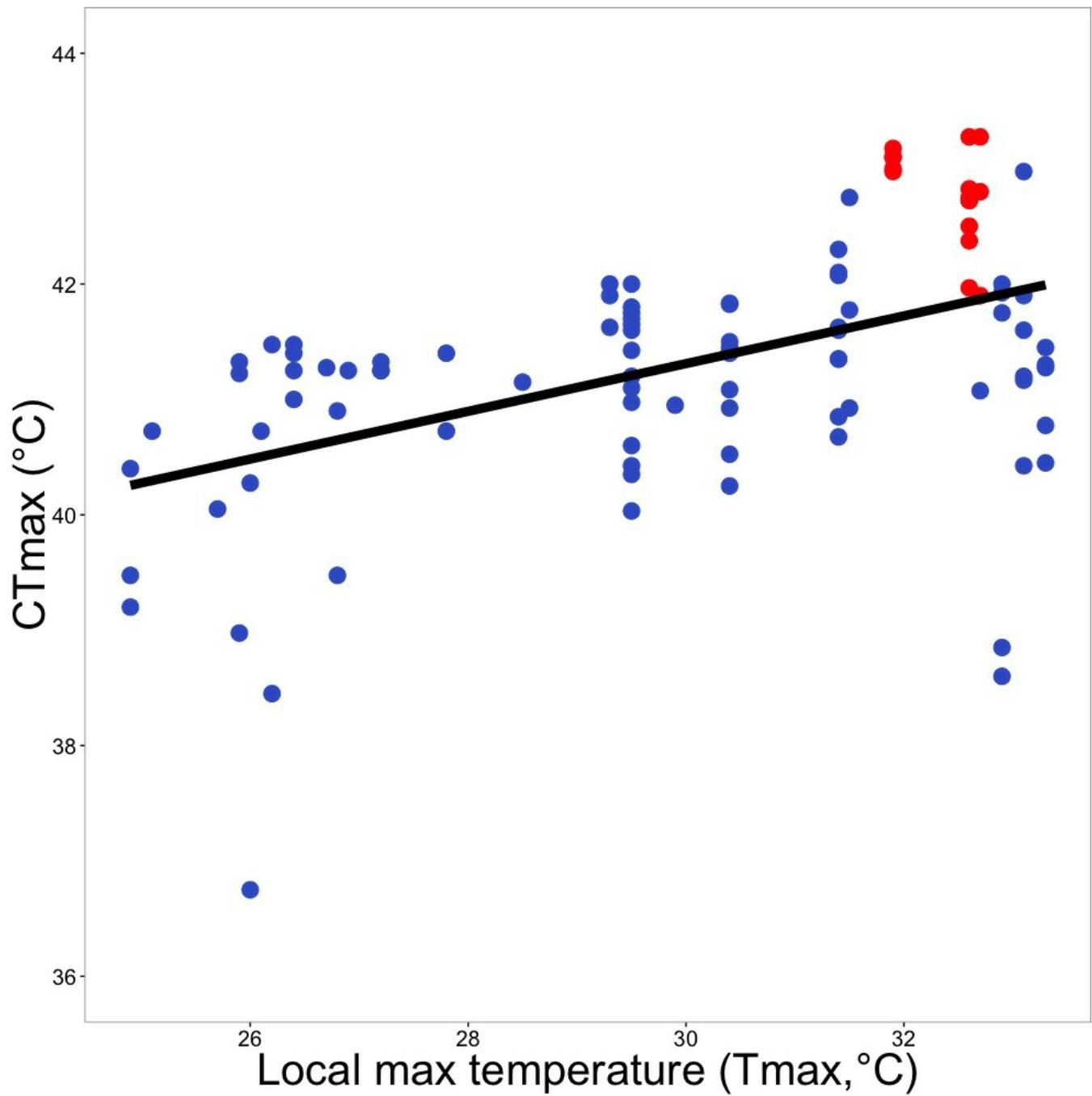
Residual standard error: 0.8605 on 97 degrees of freedom

Multiple R-squared: 0.4091, Adjusted R-squared: 0.39

69

F-statistic: 33.58 on 2 and 97 DF, p-value: 8.27e-12

Figure



regression with MAT

```
umod<-lm(K0_temp_worker~biol*habitat_v2 ,data=Aph.dat)
summary(stepAIC(umod,direction="both"))

lm(formula = K0_temp_worker ~ biol * habitat_v2, data = Aph.dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.8808	-0.2948	0.1394	0.5549	1.6231

Coefficients:

	Estimate	Std. Error	t value	Pr > t
(Intercept)	40.289262	0.266504	151.177	< 2e-16 ***
biol	0.006325	0.002090	3.027	0.00317 **
habitat_v2flat woods	4.264228	2.013656	2.118	0.03679 *
biol:habitat_v2flat woods	-0.015722	0.010713	-1.468	0.14549

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	''	'	'	'
	1			

Residual standard error: 0.8744 on 96 degrees of freedom

Multiple R-squared: 0.3962, Adjusted R-squared: 0.3773

F-statistic: 20.99 on 3 and 96 DF, p-value: 1.534e-10

regression with latitude

```
latmod<-lm(K0_temp_worker~lat*habitat_v2 ,data=Aph.dat)
summary(stepAIC(latmod,direction="both"))
```

Call:

```
lm(formula = K0_temp_worker ~ lat * habitat_v2, data = Aph.dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.9251	-0.2851	0.1050	0.5593	1.6421

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	43.26209	0.80838	53.517	< 2e-16 ***
lat	-0.05748	0.02079	-2.765	0.0682 **
habitat_v2flat woods	-2.95972	2.90928	-1.017	0.31155
lat:habitat_v2flat woods	0.13632	0.09109	1.497	0.13777

Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .
	'	'	'	' 1

Residual standard error: 0.8807 on 96 degrees of freedom

Multiple R-squared: 0.3874, Adjusted R-squared: 0.3682

F-statistic: 20.23 on 3 and 96 DF, p-value: 3.043e-10

```
<div id='id-section99.4'>
```

Hsps; pcas and variance partitioning of CTmax

```
summary(pchsp)
```

Importance of components:

	Comp.1	Comp.2	Comp.3
Comp.4	Comp.5	Comp.6	
Standard deviation	2.1385967	1.3517804	1.07592411
0232658	0.84659220	0.84649220	1.0
Proportion of Variance	0.3906613	0.1560828	0.09887942
8581459	0.06121969	0.06120523	0.0
Cumulative Proportion	0.3906613	0.5467441	0.64562350
3143809	0.79265778	0.85386301	0.7

```
knitr:::kable(round(pchsp$loadings[,1:7],3))
```

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6
hsc70	-0.073	-0.596	0.071	-0.224	-0.011	-0.001
hsp83	-0.023	-0.593	-0.008	0.098	0.001	0.001
hsp40	-0.023	0.008	0.461	0.803	0.001	-0.001
FC_hsc70_1468_max	-0.321	-0.160	0.404	-0.273	-0.001	-0.001
FC_hsc70_1468_slope	-0.280	-0.286	0.217	0.189	0.001	0.001

FC_hsc70_1468_Tm	-0.374	0.157	0.226	-0.133	-0.
FC_hsp40_541_max	-0.350	-0.082	-0.324	0.129	-0.
FC_hsp40_541_slope	-0.292	-0.149	-0.524	0.171	-0.
FC_hsp40_541_Tm	-0.368	0.063	-0.260	0.149	-0.
FC_Hsp83_279_max	-0.350	0.057	0.153	-0.213	0.
FC_Hsp83_279_slope	-0.290	0.171	-0.145	0.186	0.
FC_Hsp83_279_Tm	-0.351	0.310	0.171	-0.143	-0.

Some stats

```
summary(lm(jj$K0_temp_worker~pchsp$scores[,1]+pchsp$scores[,2]+pchsp$scores[,3]))
```

Call:

```
lm(formula = jj$K0_temp_worker ~ pchsp$scores[, 1] + pchs
p$scores[, 
    2] + pchsp$scores[, 3])
```

Residuals:

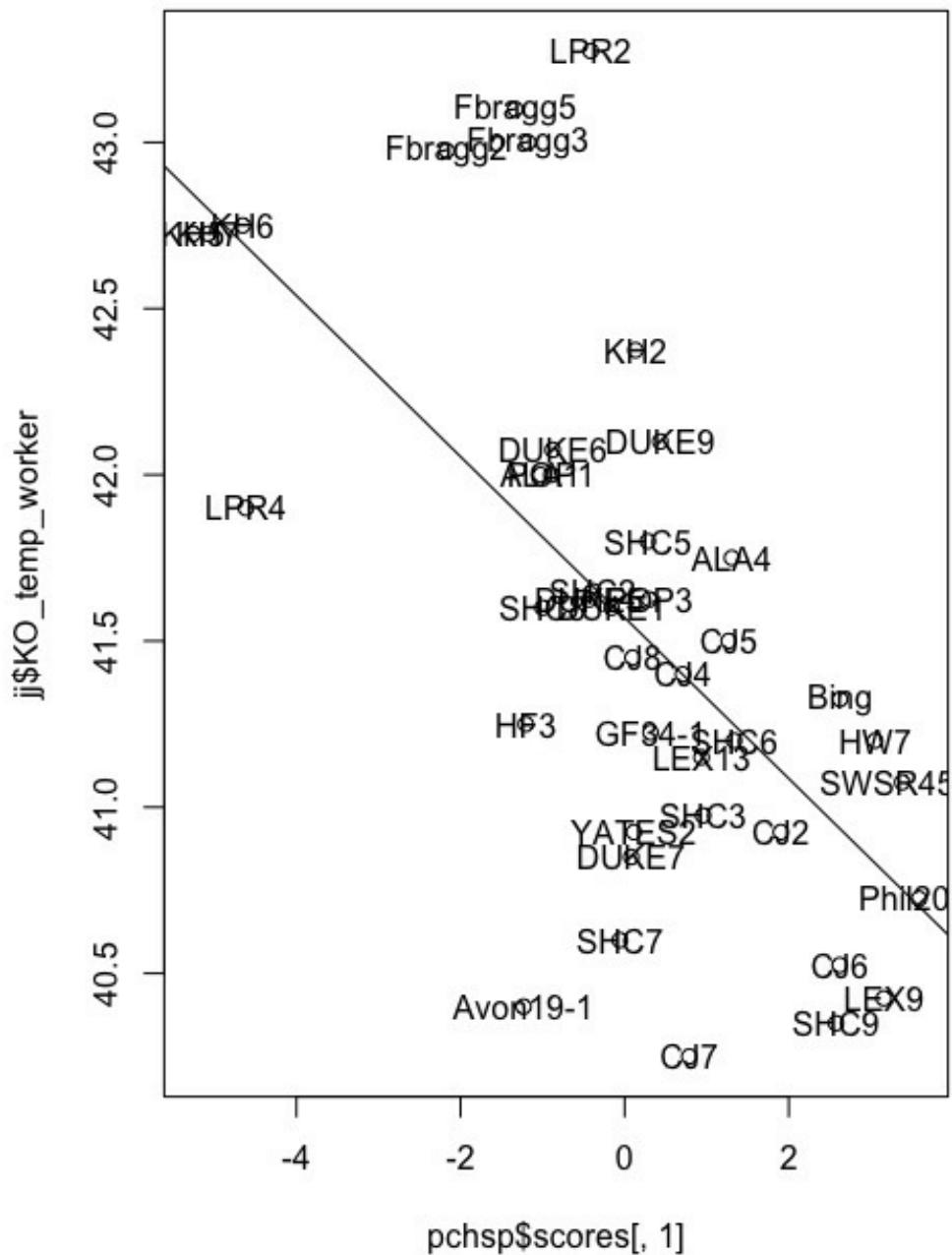
Min	10	Median	30	Max
-1.15358	-0.37044	0.04846	0.34646	1.54100

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
--	----------	------------	---------	----------

(Intercept)	41.570122	0.098692	421.211	< 2e-16 *
-------------	-----------	----------	---------	-----------

```
**  
pchsp$scores[, 1] -0.242155  0.046148 -5.247 6.55e-06 *  
**  
pchsp$scores[, 2] -0.001745  0.073009 -0.024    0.981  
  
pchsp$scores[, 3]  0.121858  0.091727  1.328    0.192  
  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1  
' ' 1  
  
Residual standard error: 0.6319 on 37 degrees of freedom  
Multiple R-squared:  0.4419,    Adjusted R-squared:  0.39  
67  
F-statistic: 9.767 on 3 and 37 DF,  p-value: 6.991e-05
```



<div id='id-section99.5'/>

Variance partitioning

```
var10<- varpart(jj$K0_temp_worker, ~ Axis.1 + Axis.2+ Axis.3+ Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9, ~biol+bio5+habitat_v2,~Hsppc1+Hsppc2,data=nw)

var10

plot(var10)
```

Partition of variance in RDA

```
Call: varpart(Y = jj$K0_temp_worker, X = ~Axis.1 + Axis.2 + Axis.3 + Axis.4 + Axis.5 + Axis.6 + Axis.7 + Axis.8 + Axis.9, ~biol + bio5 + habitat_v2, ~Hsppc1 + Hsppc2, data = nw)
```

Explanatory tables:

```
X1: ~Axis.1 + Axis.2 + Axis.3 + Axis.4 + Axis.5 + Axis.6 + Axis.7 + Axis.8 + Axis.9
X2: ~biol + bio5 + habitat_v2
X3: ~Hsppc1 + Hsppc2
```

No. of explanatory tables: 3

Total variation (SS): 26.477

Variance: 0.66191

No. of observations: 41

Partition table:

Df	R.square	Adj.R.square	Testable
----	----------	--------------	----------

[a+d+f+g] = X1	9	0.72027	0.63906	TRUE
[b+d+e+g] = X2	3	0.64967	0.62126	TRUE
[c+e+f+g] = X3	2	0.41531	0.38454	TRUE
[a+b+d+e+f+g] = X1+X2	12	0.78605	0.69435	TRUE
[a+c+d+e+f+g] = X1+X3	11	0.76028	0.66936	TRUE
[b+c+d+e+f+g] = X2+X3	5	0.67973	0.63398	TRUE
[a+b+c+d+e+f+g] = All	14	0.80893	0.70604	TRUE

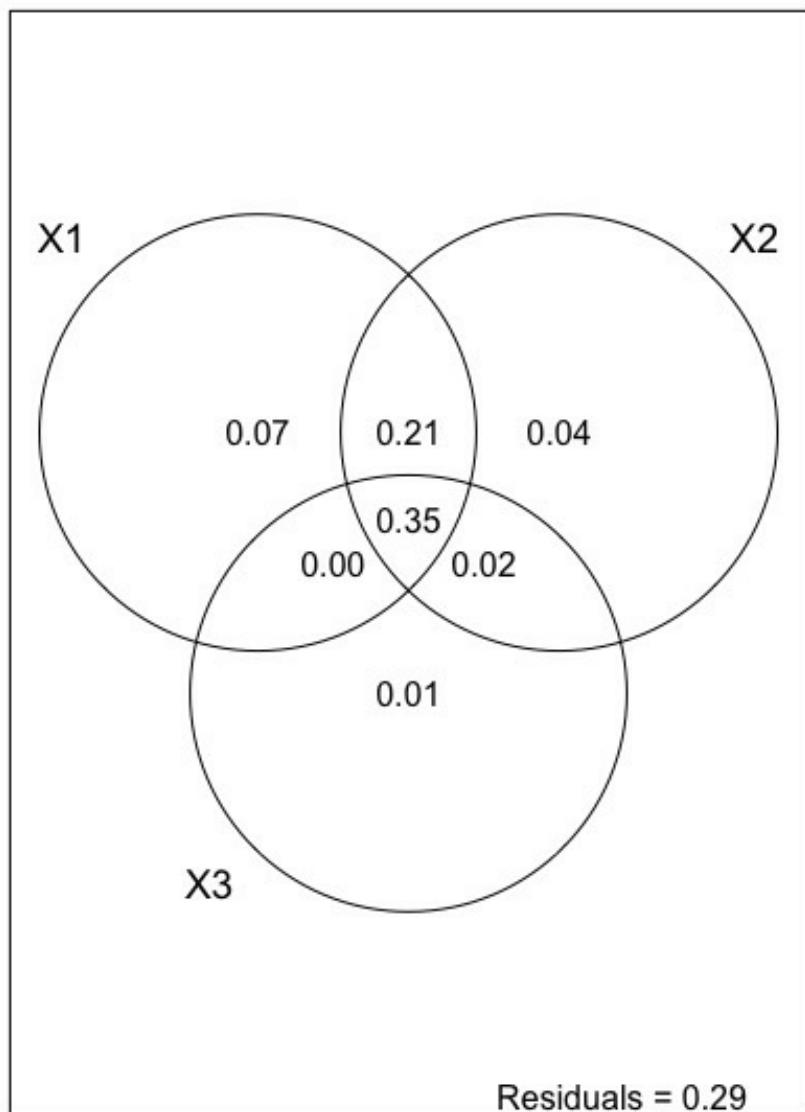
Individual fractions

[a] = X1 X2+X3	9	0.07206	TRUE
[b] = X2 X1+X3	3	0.03668	TRUE
[c] = X3 X1+X2	2	0.01169	TRUE
[d]	0	0.21275	FALSE
[e]	0	0.01861	FALSE
[f]	0	0.00103	FALSE
[g]	0	0.35322	FALSE
[h] = Residuals		0.29396	FALSE

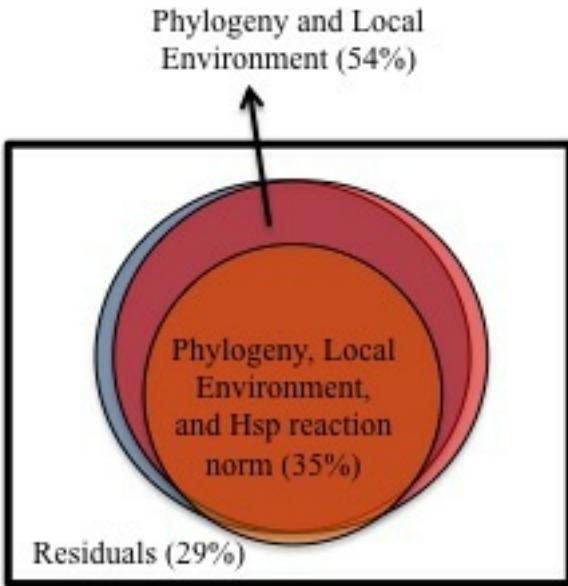
Controlling 1 table X

[a+d] = X1 X3	9	0.28482	TRUE
[a+f] = X1 X2	9	0.07309	TRUE
[b+d] = X2 X3	3	0.24944	TRUE
[b+e] = X2 X1	3	0.05529	TRUE
[c+e] = X3 X1	2	0.03029	TRUE
[c+f] = X3 X2	2	0.01271	TRUE

Use function 'rda' to test significance of fractions of interest



Slightly better figure



<div id='id-section100'>

Page 100: 2016-11-14 & 2016-11-15. climate cascade meeting

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - figure 3, SHC says to switch back branches
 - Wrote up methods and results-- go over with Nick then send to SHC
 - Submit to PNAS
- **Multiple stressors ms:**
 - submitted **2016-10-24** ; in review **2016-11-01**
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate

- **Thermal niche ms:** Lacey and I working on discussion...eta?
- **Stressed in nature MS: Samples to rerun.**
 - update: Curtis can no longer work+ write on project
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
- **Proteome stability project:**
 - **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)

2. **Attending SICB - Jan 3-8 New Orleans**, Give a talk about range limits paper.

- **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
- Support with Suiter Prize! \$1,000

3. **Thesis related** FORMS FOUND HERE

- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree

- started outline
 - Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.
-

<div id='id-section101'>

Page 101: 2016-11-16 Hsp reaction norm stats; adding quadratic term

```
lm(formula = K0_temp_worker ~ bio5 + habitat_v2 + I(bio5^  
2) ,  
  data = Aph.dat)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.6123	-0.3293	0.1297	0.4772	1.8485

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-4.4102626	12.5885230	-0.350	0.7268

```

51
bio5           0.2990131  0.0862737  3.466  0.0007
92 ***
habitat_v2flat woods  1.5151487  0.2472431  6.128  1.96e-
08 ***
I(bio5^2)      -0.0004877  0.0001469 -3.320  0.0012
75 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
' ' 1

Residual standard error: 0.8192 on 96 degrees of freedom
Multiple R-squared:  0.47, Adjusted R-squared:  0.4534
F-statistic: 28.37 on 3 and 96 DF,  p-value: 3.191e-13

```

<div id='id-section102'>

Page 102: 2016-11-22. climate cascade to do list

1. Project updates:

- **Hsp gene expression + Ctmax project:**

- rewrite results, intro and send out to NJG and

SHC

- Submit to PNAS

- **Multiple stressors ms:**

- **major revisions**

- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate

- **Thermal niche ms:** In my hands, get to it mid-december

- actionable items:

- recheck stats

- recheck figures

- make transitions between paragraphs in discussion

- **Stressed in nature MS: Samples to rerun.**

- update: Curtis can no longer work+ write on project

- **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**

- **Proteome stability project:**

- **~130 proteins for rudis, ~250 proteins for**

pogos(we got 500 proteins last time); labelling is ok

- Rerun mass spec, but loading more proteins (Bethany)

2. Attending SICB - Jan 3-8 New Orleans, Give a talk about range limits paper.

- **Practice talks: (December 1 2016 in SHC lab meeting ; Decemeber 7 2016 in EEEB)**
- Support with Suiter Prize! \$1,000

3. Thesis related FORMS FOUND HERE

- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - started outline
- Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committe form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.

Page 103: 2016-12-06. climate cascade update

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - rewrite results, intro and send out to NJG and SHC (methods done)
 - Submit to PNAS
- **Multiple stressors ms:**
 - **major revisions;** addressing now
 - go over figures
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
- **Thermal niche ms:** In my hands, get to it mid-december
 - actionable items:
 - recheck stats (are we using same dataset?)
 - recheck figures
 - make transitions between paragraphs in discussion (constructing outline)

- **Stressed in nature MS: Samples to rerun.**
 - update: Curtis can no longer work+ write on project
 - **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**
 - **Proteome stability project:**
 - **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)
2. **Attending SICB - Jan 3-8 New Orleans**, Give a talk about range limits paper.
- **Practice talks: (Decemeber 7 2016 in EEEB)**
 - Support with Suiter Prize! \$1,000
3. **Thesis related FORMS FOUND HERE**
- Formatting:
 - Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - started outline

- Deadlines:

1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.
-

<div id='id-section104'>

Page 104: 2016-12-19. climate cascade update

1. Project updates:

- **Hsp gene expression + Ctmax project:**

- rewrite results, intro and send out to NJG and SHC (methods done)
 - Submit to PNAS

- **Multiple stressors ms:**

- sent SHC revisions last week

- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate

- **Thermal niche ms:** Send new draft to Lacy tomorrow.

- **Stressed in nature MS: Samples to rerun.**

- update: Curtis can no longer work+ write on project
- **There are 74 samples: 3 days of RNA isolation + cDNA synthesis. 4 gene targets ran in duplicates is 2 plates per gene = 8 plates total. 2 days for 8 plates.**

- **Proteome stability project:**

- **~130 proteins for rudis, ~250 proteins for pogos**(we got 500 proteins last time); labelling is ok
 - Rerun mass spec, but loading more proteins (Bethany)

2. Attending SICB - Jan 3-8 New Orleans, on range limits paper.

- SICB talk Jan 8 2017, Sunday, 11:45AM.

3. Thesis related FORMS FOUND HERE

- Formatting:

- Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree

- started outline
 - Deadlines:
 1. Intent to graduate: February 1st for May.
 2. Send defense committee form to grad college—now
 3. Graduate college format check March 4th
 4. Defense notice 3 weeks before defense (oral defense by March 24th).
 5. Final thesis April 7th.
-

<div id='id-section105'>

Page 105: 2016-12-20. Reading a few papers

Reading some papers:

1. There is a cool paper by Gilchrist and Huey 2001, Evolution, that looks at the cross-generational effect of temperature on fitness in fruit flies. Ones reared from higher temperatures had offspring with higher fitness. This fitness benefit was gained by speeding up development.
 - Gilchrist GW, Huey RB (2001) PARENTAL AND DEVELOPMENTAL TEMPERATURE EFFECTS ON THE THERMAL DEPENDENCE OF FITNESS IN DROSOPHILA MELANOGASTER. *Evolution* 55:209-

2. Cool paper by [Huey and Slatkin 1976, The Quarterly Review of Biology](#) which developed the first thermoregulation model in lizards. They construct a mathematical model to quantify the costs and benefits of thermoregulation.
 - Huey RB, Slatkin M (1976) Cost and Benefits of Lizard Thermoregulation. *The Quarterly Review of Biology* 51:363-384.
 - Other follow up models:
 1. Vickers M, Manicom C, Schwarzkopf L (2011) Extending the cost-benefit model of thermoregulation: High-temperature environments. *Am Nat* 177(4):452-461.
 2. Christian KA, Tracy CR, Tracy CR (2006) Evaluating thermoregulation in reptiles: An appropriate null model. *Am Nat* 168(3):421-430.
 3. Sears MW, Angilletta MJ, Schuler MS, et al (2016) Configuration of the thermal landscape determines thermoregulatory performance of ectotherms. *PNAS* 201604824. doi: 10.1073/pnas.1604824113 (previous citations 1 and 2 found from this citation) [link](#)
3. One of [Huey's Science papers](#) that shows different populations from 3 continents track chromosomal changes with climate change.

Balanyá J, Oller JM, Huey RB, et al (2006) Global Genetic Change Tracks Global Climate Warming in *Drosophila subobscura*. *Science* 313:1773–1775. doi: 10.1126/science.1131002