

2017 General Notebook

Author: Andrew D. Nguyen, [Evolutionary Physiologist](#)

Affiliation: Biology Department, University of Vermont ; University of Florida, Department of Entomology and Nematology

Contact: anbe642@gmail.com

Date started: 2017-01-01

Date end (last modified): 2018-01-01



This work is licensed under a [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](#).

Introduction:

Notebook for 2017 new year. It'll log the rest of my dissertation and potentially new post doc ideas and/or projects

List of projects and description

- Hsp rxn norm: Understanding how the local thermal environment shapes thermal tolerance and stress response (using Hsps as a proxy for stress) in forest ants of the genus *Aphaenogaster*. CTmax and rxn norm of Hsp expression measured across forest ants from Fl to Maine.
- Range limits: Identifying the factors/forces that set range limits in common forest ants (*Aphaenogaster picea*). Modelling + measured their cold physiology in forest ants of Maine and Vt.
- Multiple stressors: Understanding how progressive desiccation and starvation impacts thermal tolerances in *A. picea*. We pre-treated with desiccation and starvation and then measured KO-time. We also measured the stress response.
- Thermal niche paper: Collaborative paper understanding how the environment shapes the ability to withstand cold and hot temperatures. In field and in a common garden, we measured upper and lower thermal limits of ants from GA-Maine (2 species).
- Stress in nature: Are ants stressed under experimental warming that projects climate change? Ants were collected from warming chambers (0-5 C increase from ambient) and we measured their stress response.
- Biological rhythms in *Rhagoletis*: Determining the relationship between behavioral rhythms in adult *Rhagoletis* and diapause exit timing + depth(eclosion and mass specific metabolic rate).

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

- [Page 1: 2017-01-01](#). SICB meeting

- [Page 2: 2017-01-03](#). Yearly goals
- [Page 3: 2017-01-10](#). Status of projects
- [Page 4: 2017-01-17](#). Displaying hierarchical structure of a repo or set of directories in the terminal with tree command
- [Page 5: 2017-01-17](#). Project idea: The impact of temperature variation on ant colony level performance.
- [Page 6: 2017-01-17](#). Tutorial for making fancy documents(Cover letters) in latex and rendering into pdf!
- [Page 7: 2017-01-19](#). Making sense of reviewer comment; multiple stressors ms
- [Page 8: 2017-01-19](#). Notes from SICB
- [Page 9: 2017-01-20](#). Status of projects
- [Page 10: 2017-01-26](#). **Hoffmann et al. 2013**, Functional Ecology Paper notes ; Constraints on upper thermal limits
- [Page 11: 2017-01-26](#). Notes on teaching statements
- [Page 12: 2017-01-26](#). SHC lab meeting, organizational
- [Page 13: 2017-01-26](#). **Bubliy & Loeschke 2005** paper notes; comparisons of selection on stress resistance
- [Page 14: 2017-01-26](#). **Mitchell & Hoffmann 2010** paper notes, Functional Ecology; Thermal ramping influences evolutionary potential
- [Page 15: 2017-01-27](#). Embedding a calendar into your webpage
- [Page 16: 2017-01-31](#). **Sørensen et al. 2003**; Ecology letters, The Evolutionary and ecological role of heat shock proteins; paper notes
- [Page 17: 2017-01-31](#). List of Adaptive Variation in Hsps lit table.
- [Page 18: 2017-01-31](#). Statuf of projects: writing

- [Page 19: 2017-01-31](#). SHC lab meeting reading: Ranga et al. 2017; Evol Ecol: Sibling Drosophila species (*Drosophila leontia* and *Drosophila kikkawai*) show divergence for thermotolerance along a latitudinal gradient
- [Page 20: 2017-02-02](#). SHC lab meeting: reading Ranga et al. 2017; lab discussion
- [Page 21: 2017-02-03](#). **Todgham et al. 2017**; JEB, The effect of temperature adaptation on the ubiquitin-proteasome pathway in notothenioid fishes
- [Page 22: 2017-02-06](#). **Tomanek & Somero 2002**; JEB, Paper notes
- [Page 23: 2017-02-06](#). **Tomanek 2010**; JEB, paper notes
- [Page 24: 2017-02-06](#). **Parcel & Lindquist 1994**; cold spring harbor , paper notes
- [Page 25: 2017-02-07](#). How far down do Aphaenogaster colonies go in the winter?
- [Page 26: 2017-02-07](#). A definition of a trade-off
- [Page 27: 2017-02-09](#). SHC lab meeting: Practice talk for biolunch and dissertation defense
- [Page 28: 2017-02-14](#). Status of projects + PHD Progress
- [Page 29: 2017-02-15](#). **Aphaenogaster, Hsp rxn norm paper stats revisited**: regression without open habitat; Sampling info ; PCA effective number of dimensions calculation for Hsps
- [Page 30: 2017-02-20](#). prep for committee meeting 2017-02-24, 2:00PM
- [Page 31: 2017-02-21](#). Measuring evolutionary rates in darwins and haldanes

- [Page 32: 2017-02-22](#). Resistance vs Tolerance strategies to stress; paper notes-**Núñez-Farfán et al. 2007**; The Evolution of Resistance and Tolerance to Herbivores; Annu. Rev. Ecol. Evol. Syst.
- [Page 33: 2017-02-23](#). Multiple regression models testing effect of Hsp parameters on Ctmax
- [Page 34: 2017-02-23](#). SHC lab meeting: Going over hxp rxn norm paper.
- [Page 35: 2017-02-23](#). What are Evolutionary Innovations?
- [Page 36: 2017-02-24](#). Last biolunch and committee meeting notes
- [Page 37: 2017-02-27](#). Mappies in R for Zamira
- [Page 38: 2017-02-28](#). Questions to think about from SHC
- [Page 39: 2017-02-28](#). More analyses for hsp rxn norm paper
- [Page 40: 2017-03-04](#). Visit at UF; feedback on dissertation talk
- [Page 41: 2017-03-07](#). To do.
- [Page 42: 2017-03-09](#). General reading notes
- [Page 43: 2017-03-09](#). Dissertation format check with Sean M.
- [Page 44: 2017-03-10](#). General phylogenetics workflow
- [Page 45: 2017-03-10](#). SHC lab schedule
- [Page 46: 2017-03-13](#). Meeting with SHC for dissertation timeline
- [Page 47: 2017-03-13](#). SHC analysis suggestion hsp rxn norm paper
- [Page 48: 2017-03-14](#). Reading notes- Armstrong et al. 2011; hsp70 and neurophysiology
- [Page 49: 2017-03-16](#). Writing notes
- [Page 50: 2017-03-17](#). Writing notes
- [Page 51: 2017-03-20](#). Writing notes

- [Page 52: 2017-03-21](#). Writing notes
- [Page 53: 2017-03-22](#). Prepping for oral defense, reading notes
- [Page 54: 2017-03-23](#). To do list
- [Page 55: 2017-03-24](#). multiple stressors reviewer comments
- [Page 56: 2017-03-28](#). Definition of phenotypic plasticity
- [Page 57: 2017-03-28](#). Meeting with new lab tech, KB
- [Page 58: 2017-03-31](#). Making thesis fixes, notes
- [Page 59: 2017-03-31](#). To do list
- [Page 60: 2017-04-07](#). Project updates
- [Page 61: 2017-04-10](#). To do and project updates
- [Page 62: 2017-04-11](#). Hsp rxn norm project: Species efficiency curves
- [Page 63: 2017-04-13](#). Hsp rxn norm proj: amplicon quant
- [Page 64: 2017-04-14](#). Hsp rxn norm project: Species efficiency curves cont'd
- [Page 65: 2017-04-14](#). Proteome stability project: Meeting with Wai
- [Page 66: 2017-04-19](#). Proteome stability project: Data analysis
- [Page 67: 2017-04-20](#). Proteome stability project: Re-analysis with normalizing with gapdh
- [Page 68: 2017-04-21](#). Proteome Stability project: More data analysis; fitting curves separately to each species
- [Page 69: 2017-04-24](#). Post doc project ideas: Diapause timing and circadian clocks
- [Page 70: 2017-04-25](#). Proteome stability, individual colony variation and other stats

- [Page 71: 2017-04-26](#). Climate cascade meeting: Project updates; to do list
- [Page 72: 2017-04-28](#). Stressed in nature project: data analysis
- [Page 73: 2017-04-30](#). Stressed in nature project: Re-analysis
- [Page 74: 2017-05-01](#). Stressed in nature project: re-analysis
- [Page 75: 2017-05-03](#). Stressed in nature project: Analyses...again
- [Page 76: 2017-05-04](#). Stressed in nature project: Analyses, including technical replicates
- [Page 77: 2017-05-05](#). Stressed in nature project: Analyses with recoded data and splitting by site
- [Page 78: 2017-05-08](#). quantifying gene stability with geNorm
 - [Linear mixed effects models for all genes.](#)
 - [More precise delta delta CT calculations](#)
 - [Corresponding plots](#)
 - [Meeting with SHC](#)
 - [Linear mixed effects models for only hsps](#), excluding housekeeping genes
 - [Separating linear mixed effects models](#) for each HSP gene
- [Page 79: 2017-05-09](#). Proteome stability project: addressing notes from 2017-04-26 climate cascade meeting
- [Page 80: 2017-05-11](#). Stress in nature project: parsing out chamber data
- [Page 81: 2017-05-11](#). Stress in nature project: Parsing chamber data
- [Page 82: 2017-05-16](#). Removing the history of large files in github through terminal
- [Page 83: 2017-05-16](#). Stress in nature project: meeting with SHC
 - [2017-05-18](#) Meeting with SHC, stressed in nature project

- [Page 84: 2017-05-22](#). Meeting with SHC and NJG, stressed in nature project
- [Page 85: 2017-05-24](#). Harvard + Duke Forest chamber info
- [Page 86: 2017-05-30](#). Multi Multi-panel figures in R
- [Page 87: 2017-06-01](#). Protein stability thoughts
- [Page 88: 2017-06-06](#). Hsp gxp rxn norm paper: writing notes
- [Page 89: 2017-06-20](#). Project idea: Viability selection on ant founding queens
- [Page 90: 2017-07-10](#). Meeting with Dan Hahn
- [Page 91: 2017-07-11](#). Reading rhagoletis papers
- [Page 92: 2017-07-12](#). Meeting prep with Dan: To do/progress list
- [Page 93: 2017-07-13](#). Project status; dissertation
- [Page 94: 2017-07-14](#). Meeting with Dan, 2pm
- [Page 95: 2017-07-17](#). Rhagoletis reading batch 2
- [Page 96: 2017-07-17](#). UF communication tips
- [Page 97: 2017-07-18](#). Qinwin ECB training
- [Page 98: 2017-07-19](#). UF computing center, tutorial/workshop
- [Page 99: 2017-07-19](#). Prep to do list, meeting Friday
- [Page 100: 2017-07-19](#). Skype meeting with TPowell, GRagland.
- [Page 101: 2017-07-21](#). Meeting with Dan 2pm friday
- [Page 102: 2017-07-21](#). notes on rhagoletis paper genomic structure
- [Page 103: 2017-07-21](#). Diapause reading papers
- [Page 104: 2017-07-24](#). Dambroski and Feder 2012; Journal of evo bio paper

- [Page 105: 2017-07-26](#). diapause paper by Wilsterman et al.
- [Page 106: 2017-07-27](#). Meeting with Dan; progress and to do list
- [Page 107: 2017-07-28](#). Dissertation work progress/ list
- [Page 108: 2017-07-31](#). Dissertation work progress/list
- [Page 109: 2017-08-01](#). Paper notes: Hormonal control of diapause by Denlinger, Yocom, and Rinehart. Book chapter 10; 2012; and Thoughts
- [Page 110: 2017-08-07](#). Reading notes
- [Page 111: 2017-08-07](#). Getting oriented on respirometry
- [Page 112: 2017-08-08](#). Reading on circadian clocks
- [Page 113: 2017-08-14](#). Two new papers on niche breadth, related to range limits, thermal tolerance breadth projects
- [Page 114: 2017-08-14](#). Meeting with Dan 2pm, update, to do's
- [Page 115: 2017-08-17](#). Meeting with Dan 12PM
- [Page 116: 2017-08-24](#). Apple maggot collecting trip log and day 1 sampling
- [Page 117: 2017-08-25](#). Apple sampling day 2
- [Page 118: 2017-08-26](#). Sampling apple maggots day 3
- [Page 119: 2017-08-27](#). Sampling apple maggots day 4
- [Page 120: 2017-08-28](#). Calibrating Li-cors (measures co2)
- [Page 121: 2017-08-28](#). Sampling apple maggots day 5
- [Page 122: 2017-08-29](#). Apple maggot sampling day 6
- [Page 123: 2017-08-29](#). Rhagoletis respirometry- purging
- [Page 124: 2017-08-30](#). Apple maggot sampling day 7
- [Page 125: 2017-09-18](#). Rhagoletis project update

- [Page 126: 2017-09-22](#). field notes from 2017-09-19 hawthorne collection
- [Page 127: 2017-11-10](#). Overall project status, planning, to do 's
- [Page 128: 2017-11-14](#). Prepping meeting with Dan ; Thursday 2017-11-16
- [Page 129: 2017-11-17](#). Rhago notes , looked at Feder papers again; refs- [Page 91: 2017-07-11](#). Reading rhagoletis papers
- [Page 130: 2017-11-19 to 2017-11-20](#). Initial reading of Philosophical Transactions B review articles on [wild clocks](#) (<http://rstb.royalsocietypublishing.org/content/372/1734>)
- [Page 131: 2017-11-21](#). Figured out code to subset out a dataframe based on a time interval
- [Page 132: 2017-11-21](#). Field clocks reviews - methods in field chronobiology
- [Page 133: 2017-11-24](#). Prep for meeting with Dan, 2017-11-28 + some thoughts on diapause + to do list
- [Page 134: 2017-11-28](#). Reading paper for lab meeting Overgaard and MacMillan 2017; Ann Rev Phys
- [Page 135: 2017-11-29](#). Notes on how to analyze biological rhythm data
 - [2017-11-30 update](#). dates not ordered, affecting results for wavelet analyses, not periodogram analyses though
- [Page 136: 2017-11-30](#). Continuous Wavelet analyses for different time bins: 6 min, 15 min, 30 min, 1 hr
- [Page 137: 2017-12-03](#). Discrete wavelet analysis on h404; 15 min bins

- [Page 138: 2017-12-05](#). Prepping meeting with Dan on 2017-12-07, 10:30AM
 - [Page 139: 2017-12-08](#). behavioral data (trikinetics) data parsing update--fixing date order again; update from [2017-11-30 update](#).
 - [Page 140: 2017-12-11](#). Dimensions of biodiversity workshop: managing big data
 - [Page 141: 2017-12-12](#). paper notes- Bulla et al. 2017 on Marine Biorhythms and van der Veen et al. 2017 on flexible clock systems
 - [Page 142: 2017-12-13](#). Prepping meeting with Dan, 9AM , 2017-12-14
 - [Page 143: 2017-12-17](#). Douse 2013; MESA paper
 - [Page 144: 2017-12-17](#). [Time series analysis demo](#)
 - [Page 145: 2017-12-18](#). Costantini et al. 2010; *Ecology Letters*; **Ecological processes in a hormetic framework**
 - [Page 146: 2017-12-18](#). Trying spectral analysis with spectrum() function
 - [Page 147: 2017-12-18](#). Table of contents of physical notebook: NB#001; Apple Maggots; year 2017; grant # 108398
 - [Page 148: 2017-12-18](#). Chao leading lab meeting
 - [Page 149: 2017-12-20](#). 2017-12-20. continuous wavelet analysis revisted: 6 min bins and exploring power vs period to id multiple dominant peaks
 - [Page 150: 2017-12-21](#). Meeting with Dan
-

Page 1: 2017-01-01. Society for Integrative and Comparative Physiology (SICB) meeting in New Orleans

Table of talks I'd like to attend

Date	Time	Room	Speaker	Title
Wednesday, Jan4	19:30	Hilton Ballroom	Swalla	Follow the Yellow Brick Road: An Odyssey from Myoplasm to Marine Biology to Genomics
Thursday, Jan5	8:00	217	Kelly	Protein coding and regulatory variation contribute to heat adaptation in the copepod <i>Tigriopus californicus</i> .
Thursday, Jan5	8:15	217	Logan	The evolutionary potential of a global insect invader in the face of rapid environmental change
Thursday, Jan5	8:30	217	Campbell	Urban heat islands and temperature-mediated physiological shifts between populations of the Puerto Rican crested anole
Thursday, Jan5	8:30	221	Charbonneau	Who Are the 'Lazy' Ants? Inter-worker Variation Gives Insight into Potential Functions of Inactivity
Thursday, Jan5	9:00	217	Ferris	The genomics of rapid adaptation to climatic extremes in house mice across the Americas
Thursday, Jan5	9:15	214	Geyman	Temperature Effects on Parasite Larval Size Over Time and Across Multiple Life Stages
Thursday, Jan5	9:30	214	Genovese	Plasticity in thermal tolerance of early life history stages of marine invertebrate larvae
Thursday, Jan5	9:30	217	Jangjoo	Gene expression associated with dispersal ability under different temperature conditions in the alpine butterfly, <i>Parnassius smintheus</i>

Thursday, Jan5	11:45	211-213	Gough	Physical Properties and Anisotropy in the Central Tissue Layer of Cetacean Tail Flukes
Thursday, Jan5	11:45	214	Debiasse	Testing the effect of ocean acidification on a sponge-coral species interaction
Thursday, Jan5	13:00	215-216	Flynn	Response of Amphibian Gut Microbiome to Coal Combustion Waste
Thursday, Jan5	13:30	224	Tangwancharoen	Divergence in cis-regulatory elements and HSPB1 gene expression along a temperature cline in the copepod <i>Tigriopus californicus</i>
Thursday, Jan5	13:45	219	Kingsolver	Inconstancy is informative: Estimating performance curves in fluctuating environments
Thursday, Jan5	14:00	219	Howey	Effect of Temperature on Snake Locomotion and the Interpretation of Thermal Performance Curves
Thursday, Jan5	15:15	214	Tielens	Geological age and host polymorphism affect functional diversity and community composition in plant-insect interactions across a space-for-time chronosequence on the Hawaiian Islands.
Thursday, Jan5	19:00	208	Sheriff	Integrating physiology, behavior, and ecology to understand the mechanisms that regulate and limit animal populations
Thursday, Jan5	15:30- 17:30	Exhibit Hall (posters)	Brueggemann	The effect of cholesterol and -tocopherol on cold tolerance, post-cold performance, and rapid cold hardening
Thursday, Jan5	15:30- 17:30	Exhibit Hall (posters)	Chan	Taking the heat: High thermal tolerance of larval and adult mangrove snails
Friday, Jan6	7:50	206	Ragland; Williams	Introduction to Evolutionary Impacts of Seasonality symposium
Friday, Jan6	8:15	211-213	Gilbert	Natural Selection on Thermal Preference and Performance over a Rapid Timescale
Friday, Jan6	8:30	215-216	Nourabadi	Fitness consequences of pH adaptation in an experimentally evolved beneficial

				symbiosis
Friday, Jan6	9:00	210	Obrien	The metabolic costs of animal weapons
Friday, Jan6	9:30	206	Buckley; Kingsolver	Insect Development, Thermal Plasticity and Fitness Implications in Changing, Seasonal Environments
Friday, Jan6	9:30	211-213	Riddell	Potential responses to climate change are improved by physiological acclimation of water loss
Friday, Jan6	10:30	221	Clay	Transcriptomics of salamander tail tips reveal potential biomarkers of stress.
Friday, Jan6	11:00	221	Holden	Preparing for winter dormancy: Early-life experience affects condition, metabolism, and hormonal response to cold temperatures in the checkered garter snake, <i>Thamnophis marcianus</i>
Friday, Jan6	11:30	215-216	Tanner; stillmann	Locally adapted <i>Phyllaplysia taylori</i> populations in Central California show higher thermal plasticity potential
Friday, Jan6	11:45	215-216	Oyen	Common garden experiments reveal local adaptation in critical thermal limits of bumblebees (<i>Apidae</i> , <i>Bombus</i>) over short geographic distances
Friday, Jan6	13:45	214	Norin	Plasticity, performance, and pace of life: individual differences in physiological and behavioural flexibility towards daily changes in temperature and oxygen availability
Friday, Jan6	14:00	214	Smith	Assessing the protein and metabolic costs of a trade-off between reproduction and immunity
Friday, Jan6	14:00	220	Crickenberger	Do temperature and competition interact to set a range limit?
Friday, Jan6	14:15	219	Kenny	How temperature influences the viscosity of hornworm hemolymph
Friday, Jan6	14:15	223	Refsnider	Plasticity in Behavioral Thermoregulation by Lizards on an Elevational Gradient: A Reciprocal Transplant Experiment
Friday, Jan6	14:45	214	Jaumann	Nutritional Stress Decreases Fecundity and Choosiness in a Butterfly

Friday, Jan6	15:00	206	Williams; Ragland	Evolutionary impacts of seasonality: synthesis and directions forward
Friday, Jan6	15:00	214	Triedel; Williams	The effect of diet nutrient composition on development and life history traits of a wing polymorphic cricket, <i>Gryllus lineaticeps</i>
Friday, Jan6	19:30	215-216	Sosik	Life in the Plankton, Stories from automated submersible microscopy and flow cytometry
Friday, Jan6	15:30-17:30	Exhibit Hall (posters)	Kornegay	Methylation and chromatin remodeling complex from sponges to humans
Saturday, Jan7	8:45	215-216	Robberts;Rank;Stillman; Williams	The effects of snow cover on overwinter physiology of a montane insect
Saturday, Jan7	9:30	211-213	Wada	A potential link between organismal adrenocortical responses and cellular heat shock responses
Saturday, Jan7	10:15	219	Stoehr	Temperature, Photoperiod and Nutrients Affect Phenotypically Plastic Wing Patterns in the Cabbage White Butterfly
Saturday, Jan7	10:45	215-216	Allen	Diet-by-temperature interactions on a sexually selected trait and sexual dimorphism
Saturday, Jan7	10:45	220	Bryant	Uncoupling Proteins and Thermal Acclimation and Adaptation in Atlantic killifish, <i>Fundulus heteroclitus</i>
Saturday, Jan7	11:00	215-216	Mikucki	Seasonal Differences in Diapause Induction in a Vermont Population of <i>Pieris rapae</i> Butterflies
Saturday, Jan7	11:00	220	Novarro	Geographic patterns of thermal tolerance in a widespread lungless salamander
Saturday, Jan7	11:30	211-213	Finger	The Effect of Heat Shock on Constitutive and Inducible Heat Shock Proteins and Corticosterone in the Zebra Finch
Saturday, Jan7	11:30	215-216	Lockwood	Molecular targets of thermal stress during early development in <i>Drosophila melanogaster</i>
				Daily Energy Expenditure, but Not Self-

Saturday, Jan7	11:45	215-216	Niedojadlo	Maintenance Costs, Are Related to Hematological Variables in Response to Temperature Acclimation
Saturday, Jan7	13:30	215-216	Lisovski	Biologically Significant Dimensions of Seasonality
Saturday, Jan7	13:45	217	Lozier	Population Genomics of Color Pattern Variation in a Widespread North American Bumble Bee
Saturday, Jan7	14:00	215-216	Betini	Fitness Trade-off Between Seasons Causes Multigenerational Cycles in Phenotype and Population Size
Saturday, Jan7	14:00	220	Mitchell	Do covariances between maternal behavior and embryonic physiology drive sex-ratio evolution under environmental sex determination?
Saturday, Jan7	14:15	225-226	Ceja	Ecologically modeling the distribution of an intertidal crab concerning global change
Saturday, Jan7	15:30- 17:30	Exhibit Hall (posters)	Pigg; Williams	Effects of ambient temperature on the organization of lipids of the avian stratum corneum
Sunday, Jan8	8:00	220	McGee	Machine learning predicts cichlid feeding kinematics from craniofacial morphology
Sunday, Jan8	8:30	217	Newman	The Influence of the Early-life Environment on Stress Physiology and Fitness in the Wild
Sunday, Jan8	8:30	224	Irvine	Proteomic Changes Due to Elevated Temperature in Ascidian Ovaries
Sunday, Jan8	8:45	214	Matoo; Montooth	Role of Genetic Variation on the Ontogeny of Metabolism during Development.
Sunday, Jan8	8:45	224	Hurley	The Heat is On: Decrease in Avian Sperm Functionality at High Ambient Temperatures
Sunday, Jan8	9:15	215-216	Soda; Slice	Vector autoregressive-moving average models as tools to visualize differences in shape trajectories
Sunday,	9:15	218	Diamond	Rapid evolution of ant thermal

Jan8				tolerance within an urban heat island
Sunday, Jan8	9:30	214	Neel; Mcbrayer	Thermal dependence of sprint performance and critical thermal limits in ecologically distinct populations of a small ectotherm
Sunday, Jan8	9:30	217	Graham	Evolutionary history matters: Maternal hormonal response to a natural stressor and effects on offspring growth and behavior
Sunday, Jan8	9:30	218	Hall; Warner	Thermal Spikes Caused by the Urban Heat Island Effect Result in Differential Egg Survival of a Non-native Lizard (<i>Anolis cristatellus</i>)
Sunday, Jan8	10:00	221	Dormio	Staying Active for Life: Investigating the Covariance Between Behavioral and Physiological Trade-offs in Treefrogs
Sunday, Jan8	10:15	217	Vitousek	Do Brief, Acute Stressors Have Lasting Effects on Phenotype?
Sunday, Jan8	10:30	214	Shah;Ghalambor	Does Climate Variability Predict Thermal Tolerance? A Comparison of Thermal Breadths in Aquatic Insects Across Elevation & Latitude
Sunday, Jan8	10:45	218	Balaban	Elastic energy storage and thermal performance in fence lizards
Sunday, Jan8	10:45	224	Mccue	Repeated exposure to food limitation earlier in life enables rats to spare lipid stores during prolonged starvation
Sunday, Jan8	11:00	224	Mcternan	Resting metabolism comparisons among populations of a subspecies of lizard differing in climate and vegetation types
Sunday, Jan8	11:30	217	Senner	The Stress Response of <i>Peromyscus</i> Mice to Experimental High Elevation Conditions
Sunday, Jan8	11:45	214	Nguyen	Constraints on cold tolerance and hardening ability limit the distribution of forest ants at its northern range boundary.
Sunday, Jan8	11:45	222	Crall	A Neonicotinoid Pesticide Disrupts Nest Behavior and Social Interactions in

				Bumblebee Colonies
Sunday, Jan8	13:30	210	Harrison	Hypometric scaling of metabolic rate arises from size-dependent natural selection on ATP demand
Sunday, Jan8	13:45	214	Clark; Williams	A Genetic Polymorphism for a Hormonal Circadian Rhythm is Associated With a Shift in Metabolic Fuel Use in Flight-Capable but not Flightless Crickets
Sunday, Jan8	13:45	210	Salin	How Does Mitochondrial Functioning Constrain Energy Efficiency?
Sunday, Jan8	14:00	222	Cirino	Effects of male quality and territory quality on female preference of varying condition
Sunday, Jan8	14:15	214	Braciszewski	Relatedness and differential disease resistance in eastern Pacific Haliotids
Sunday, Jan8	14:15	224	Boothby	How do Tardigrades Survive Extremes? Disordered Proteins as Mediators of Tardigrade Stress Tolerance
Sunday, Jan8	14:30	221	Miles	Desert tortoises race against climate change: past, present and future
Sunday, Jan8	14:45	220	Schoenle	Why Does Malaria Infection Reduce Fitness in Wild Birds?: A Test of Physiological Mechanisms
Sunday, Jan8	15:00	222	Leary	Acute Stress is a Target of Intra- and Intersexual Selection in the Green Treefrog, <i>Hyla cinerea</i> : Implications for Fitness, Honest Signals, and the Evolution of Endocrine-based Acoustic Armaments

Page 2: 2017-01-03. Yearly Goals

1. Submit and Publish 3 manuscripts: range limits, Hsp rxn norm, and multiple stressors (also thermal niche paper).

2. Get a post doc. This'll probably involve learning a new study system.
And also sending out tons of applications.
 3. Learn and build a shiny app.
 4. Learn and become more proficient in statistics (Machine learning?,
Baysian, predictive modelling, mixed effects modelling, eigentensor
analyses).
 - Quantitative genetics: more statistical genetics
 5. Form new collaborations? It'd be awesome to work with Brent
Sinclair, Brent Lockwood, Joel Kingsolver, Caroline Williams, Jon
Stillman, Alex Gunderson.
 6. Participate in a meta-analysis? Would be cool.
 7. Learn more physiology: Q10, metabolism related topics, lipid
membranes, metabolites.
-

Page 3: 2017-01-10. Status of Projects

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:**
 - In SHC's hands, due 2017-02-14
 - **Range limits ms:** SHC lab gave verbal edit, still need to incorporate

- **Thermal niche ms:** Lacy to check intro, and add refs.
- **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. 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.

the terminal with tree command

1. In the terminal, [install the tree command](#)
 - You may have to install the [homebrew](#)
2. use the tree command in the terminal

```
1 2017_Ecological_Genomics andrewnguyen$ tree
2 .
3 └── 2017_Ecological_Genomics.Rproj
4   ├── Online_notebook.md
5   ├── README.md
6   ├── RasterPCA_demo.Rmd
7   ├── RasterPCA_demo.html
8   ├── index.Rmd
9   ├── index.html
10  └── index.pdf
11
12 0 directories, 8 files
```

Page 5: 2017-01-17. Project idea: The impact of temperature variation on ant colony level performance.

Background

Ant colonies experience temperature fluctuations throughout the day and season. In response to temperature variation, ants must be able to forage under cool and hot conditions. One way to achieve high performance in the face of temperature variation is for the colony to match individual level performance with shifting temperatures. For

example, for colonies with diverse genetic structures, some genotypes may forage better under cool or hot conditions. Multiply mated queens (*Pogonomyrmex*) will produce offspring with more phenotypic variation in traits, compared to singly mated queens (*Aphaenogaster*).

Notes: Need to actually verify that foraging efficiency differs throughout the day in a colony, aka need to know more natural history. Look up Deborah Gordon's work. How will you address the confounding issue of demographic effects? Meaning what if there is genetic clumping of sperm and it corresponds to age polyethism, leading to the covariance between age and genotype. Also, the age structure of the colony itself may covary with genotype.

Question

1. Is the amount of physiological variation in upper thermal limits reflected in the number of fathers?
2. Do ant colonies utilize different genotypes to forage throughout the day and season?

or

Do ants display seasonal adaptation in forager performance?

Hypotheses

1. The number of fathers increases the amount of phenotypic variation in upper thermal limits.
2. There is genetic structuring in the foragers throughout the day and season.

Predictions

1. Additive genetic variance in upper thermal limits will be proportional to the number of fathers.

2. Thermally tolerant genotypes perform during the warmest parts of the day and vice versa for more cool tolerant genotypes.
3. The more thermally tolerant genotypes start off at low frequency early on in the season, but increase in frequency at peak summer, then decrease again.

Experimental Design

I should do a power analysis to see how many ants and colonies I'd need to sample. Sam Scarpino has an R package for this I think.

1. Document the natural history of 10 pogo colonies, to tune for sampling
2. If there is like 10 foraging bouts, or 10 time slices through the day, sample ~30 ants for each colony at each time slice.
3. Phenotype each ant: measure morphology (head width, leg lengths, alitrunk length, etc), number of ovarioles?, hydrocarbons, desiccation resistance, measure upper thermal limit
4. Pool ants at each time slice, isolate DNA, and then pool-seq (capture-seq, but chatting with April, it sounds like biased sampling because only the portion of the first exon and promoter are sequenced.)

Critical Results or Alternative outcomes

1. Positive relationship between parentage and phenotypic variance.
2. Structuring of alleles throughout the day or season. Some start off low, increase, then decrease in frequency. This would indicate seasonal adaptation of foragers!

It is possible that we won't find anything at all. But maybe because we're pool-sequencing, we can pick up the microbiome? And maybe that is structured diurnally or throughout the season.

Another idea: Compare parasite load between ant colonies with high and low number of fathers.

Quick and dirty: One of the benefits of a diverse genetic background in a colony is disease resistance. So, test this by pool-seq experiment and measuring the parasite loads for ~30 individuals, potentially across a whole phylogeny of ants that vary in number of daddys and number of queens. Parasite load from different taxonomic groups (fungus, nemotodes, bacteria, viruses) can be mapped onto the phylogeny!

Page 6: 2017-01-17. Tutorial for making fancy documents(Cover letters) in latex and rendering into pdf!

I took this [cool template](#) for making a cover letter, which was in ".tex" format. Then converted the tex file into pdf using [this tutorial](#), which involve installing mactex, sublime text 2, and skim.

Page 7: 2017-01-19. Making sense of reviewer comment; multiple stressors ms

Is there any direct evidence that heat shock proteins prevent heat knockdown?

Calabria et al. 2012 find differences in basal hsp70 protein expression but not heat-induced hsp70 expression between 3 genotypes of fruit flies(they differ in chromosomal structure). These differences were reflected in the fast ramp (0.1 C/min ramp), but not slow ramp(0.6 C/min ramping).

3 genotypes:

1. O(3+4)/O(3+4): higher hsp70 basal expression and thermal tolerance under fast ramp;
2. O(3+4+8)/O(3+4+8)
3. O(st)/O(st): lower hsp70 basal expression and thermal tolerance under fast ramp;

References:

1. Calabria G, Dolgova O, Rego C, et al (2012) Hsp70 protein levels and thermotolerance in *Drosophila subobscura*: a reassessment of the thermal co-adaptation hypothesis. *Journal of Evolutionary Biology* 25:691–700. doi: 10.1111/j.1420-9101.2012.02463.x
-

Page 8: 2017-01-19. Notes from SICB

2017-01-05; Met with Joel Kingsolver

Met and talked about range limits project, selection gradient idea, and hsp rxn norm gxp project.

- For my ancestral trait reconstruction of CTmax, I excluded the outgroup, suggests to include it
- For multi-panel figure with parameters on x axis and ctmax on y

- axis; try partial regression
- If I don't get into PNAS, try Proc B
 - For looking at shifts in trade-offs between multiple stressors look up Nancy Emory, central cali, she is a plant evo ecologist
 - Selection gradient proj idea; JK asks what has been done in well known systems? ie. *Arapidopsis* (sp?), check Andy Schidth

2017-01-06; met with Ray Huey

Showed hsp rxn norm paper stuff:

- Are parameters correlated- he was worried about this.
- Try fitting splines and extracting parameters
- Selection on traits; linear or quadratic? non-parametric?

2017-01-07 Met with Dan Hahn and Greg Ragland

Went over data and Dan liked how I parsed out all the predictions for the shifts in reaction norms of hsp gxp expression.

- Finger example highlights the differences between resistance and tolerance
 - Pressing two fingers together without moving = **resistance**
 - When 1 finger moves the other finger but can move back = **tolerance**
-

Page 9: 2017-01-20. Status of projects

1. Project updates:
 - **Hsp gene expression + Ctmax project:**
 - Have intro, discussion mostly complete by today, meet with

NJG Monday (2017-01-23)

- Submit to PNAS in May?
- **Multiple stressors ms:**
 - In SHC's hands, due 2017-02-14
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
 - **Thermal niche ms:** Lacy to check intro, and add refs.
- **Stressed in nature MS: Samples to rerun.**
 - **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. Thesis related

- **Defense talk: Have first version by Jan 31st**
 - Required defense talk in biolunch: Feb 24 (Friday)
- Working title: Evolutionary Innovations of Ants to Thermally Stressful Environments
- 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---**done**
 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 10: 2017-01-26. Hoffmann et al. 2013, Functional Ecology Paper notes

Starting off intro with constraints on upper thermal limits.

What can constraint upper thermal limits?

1. low plasticity (stillmann paper shows this in crabs)
2. low adaptive potential (heritability)
 - low variation in upper compared to lower thermal limits
3. Phylogenetic constraints inferred from strong phylogenetic signal.
 - high lambda in PGLS for insects, lizards + snakes

Some definitions:

- heritability: degree of phenotypic variation in a trait that is genetic
- evolvability: extent to which the trait's mean can shift under selection, which depends on the amount of genetic variation and mean values.

Paper argues that mid-latitude species are most prone to heat stress (presently and to the future). Why? And what is mid-latitude?

Page 11: 2017-01-26. Notes on teaching statements

This science [article](#) has interesting advice. The teaching statement is usually a writing filter to weed out applicants who have not thought stuff through. Illuminates the character of the writer. It should be 1-2 pages.

- 1. Cater teaching statement to the institution**
- 2. Demonstrate a real commitment to teaching**
 - Low expectations on unique ways of teaching, but just show commitment (usually a sentence or two)
 - cite evidence, but be brief
3. Avoid presenting teaching as a 2nd priority
4. *Show us you care, but make it short and to the point*
5. echo teachign interests in the cover letter, 1-2 sentences
6. Write about courses you'd like to teach! For me, it'd be evolutionary physiology, evolution, computational biology?, ecological genomics

- cater this to the institution
 - display ideas, but don't be too ambitious
7. Teaching statement also indicates to the search committee how the applicant sees themselves in
 - don't limit yourself to advanced courses, but also include beginners!
 8. Display willingness to learn, pay attention, and change to reflect open mindedness and eagerness; shows you're not limited by what you are now
 9. **Draw on your experiences as a student, a scholar, and human being**
 10. Don't promise too much

A good quote:

I used to assume that any student who did not get an "A" on an exam wasn't trying, until it occurred to me that no matter how hard I try in dance classes, I rarely excel. Becoming less critical of my students--while still maintaining high standards--has made me a better teacher.

**Page 12: 2017-01-26. SHC lab meeting,
organizational**

**Write down 3 concrete goals and timeline
for the semester:**

1. Publish the multiple stressors manuscript
 - deadline is Feb 12 to submit revisions
2. Get a post doctoral position or real job by end of semester
 - NSF, should know by Feb 9th
 - Hibbitt fellow, no clue, but should be soon
 - prepping application for data scientist position
3. Have submittable version of hsp rxn norm paper to PNAS
 - Part of thesis, so should be accomplished by April 7th

Benchmark goals

1/3- Wait for SHC to give back feedback and take 2 days to submit it, prior to Feb12. Keep writing on PNAS paper, have introduction, discussion done. Results and methods are mostly complete but could be tweaked. Post doc stuff is passive

2/3- get notified from NSF, hibbit. For Hibbitt, if they're interested, they'll interview on site. For PNAS paper, shape up ms.

3/3- Constant revisions on PNAS paper

Set schedule

I have a biolunch talk (dissertation requirement talk) February 24th. Give practice talk February 10th.

Tentative: Thursdays 4-5PM

**Page 13: 2017-01-26. Bubliy & Loeschke 2005
paper notes**

Correlated response to selection for stress resistance.

Fruit fly paper doing a selection experiment on:

1. cold (CS) in units of survival rate
 - chilled 5 C, recover 24 hours
 2. heat (HS) in units of survival rate
 - hardened 30 min at 36C, recover 20 hours at 25 C, then heat shocked 38C 1 hour
 3. desiccation (DS)
 4. starvation (SS)
 5. Heat knockdown time (KS) -in units of time (minutes)
 - KO at 40C
 6. Lifespan (LS)
 - And then they checked all of the traits from each selected line relative to controls.
 - 21 generations
1. Cold shock resistance (CS)
 - CS, HS, and DS all increased
 2. Heat shock resistance (HS)
 - HS, KS, DS all responded
 3. Heat knockdown resistance (KS)
 - only KS responded

KS and SS increased developmental time.

Upper and lower thermal limits can equally respond to selection after X generations in fruit flies. IN this experiment, there isn't necessarily a trade off between the two (except in cold shock resistance).

Page 14: 2017-01-26. Mitchell & Hoffmann 2010, Functional Ecology; Thermal ramping influences evolutionary potential

In fruit flies, they measured upper thermal limits as knockdown time in a static and ramping experiment. They did this in a quantitative genetic design, so they could estimate the narrow-sense heritability (partition out variation to the additive genetic component). They also measured this for ~11 species. (looks like a lot of work)

Methods:

Slow ramp: 0.06 C/min

Fast, static HS: 38C preset temp

Fast and slow show a positive relationship.

Set up a model to partition out variation Va (additive genetic variance), Vp (phenotypic variance), and Ve (environmental variance) and presented models for their coefficient of variation.

Results:

- Ramping had low heritability(not sig different than 0), but static had high heritability (sig different than 0)
- Ramping had more environmental variance, but do the data need to be standardized first, such that the mean is 0 and variance is 1? No, the units are the same for both traits
- Ramping had higher phenotypic variance

THoughts

This paper contrasts a little bit with van Heraawarden & Sgrò paper, where they focus on one species. But they find that ramping, fast HS, and hardening all significantly correlate. But there is an additive genetic axis where

Page 15: 2017-01-27. Embedding calendar into webpage

[Followed this tutorial:](#)

1. Create a calendar in google (gmail)
2. Go to settings
3. click calendars

4. go down to where it says embed this calendar, it is in html code, so copy that
5. tutorial goes through yml, but I just used the html code to make a new webpage with my calender in it and created a link in the navigation bar

[Here it is](#), useful for me to access my calender wherever I am now!

2017-01-31 Update

Calendar doesn't show up on my iphone or ipad (safari), what is the point? Taking the calendar out.

Page 16: 2017-01-31. Sørensen et al. 2003; Ecology letters, The Evolutionary and ecological role of heat shock proteins

Trying to get a handle on how this paper frames the role of Hsps with upper thermal limits. From the abstract, they mention it as a resistance mechanism.

1. **Defining stress:** a condition that disturbs the normal function of the biological system or a condition that decreases fitness (Hoffmann & Parsons 1991; Bijlsma & Loeschke 1997)
 - I like part of the definition. I'd define stress **as a perturbation**

to a biological system that decreases fitness. Or, a force acting on a biological system that decreases fitness.

- These forces can be *extrinsic* (environmental) and *intrinsic* (genetics, inbreeding, deleterious mutations, ageing).

2. **Hsp natural history** (cite Ritossa 1962 as first discovering heat shock response through chromosome puffs(which had Hsps on it))
3. **Protein quality control system (PQC)**: importance increases upon exposure to environmental /genetic stresses
 - function is 2 fold: correct folding and to assist in degradation of denatured or aggregated proteins
 - not very informative
4. Stress as an ecological and evo force (Don't agree with the framing)
 - types of responses:
 - move, alter physiological state through hibernation or diapause
 - adapt or acclimate
 - fail and die
 - Adaptive change in Hsp over days (Nguyen et al. 1994; Ferguson et al. 1998) or over seasons (Fader et al. 1994; Hofmann & Somero 1995; Pyza et al. 1997; Miner et al. 2000) and natural populations(no citations, what?)
5. Effects of stress on rates of evolution
 -
6. **Costs of Hsp expression**
 - unclear whether reduced fitness is due to cost of acclimation or reared under poor conditions. Why would we tease these apart?

- Benefits of acclimation (stress /longevity) can be separated from costs
 - ex: Hercus et al. 2003: repeated mild stress lowered fertility and fecundity in the short term but not long term
 - One way to separate out costs and benefits is by altering acclimation treatments (Scott et al. 1997; Hoffmann & HewaKapuge 2000; Thomson et al. 2001; Wilson & Franklin 2002)
- lowers fertility fecundity, energy, development, and survival
- Direct costs measured by Krebs & Feder 1998: hardened flies at different stages (1-3rd instar) in 4isofemale lines. Multiple heat exposures reduced survival but did not affect development time. Expression of Hsp70 was not correlated with survival, suggesting differences in expression cannot explain survival effects.
- said this already but high hsps lower growth and cell division (Feder et al. 1992; Krebs & Feder 1997, reduced reproduction (Krebs & Loeschke 1994; Silbermann & Tatar 2000))
 - Silbermann & Tatar 2000 showed heat induced hsp expression reduced egg hatching of moms in fruit flies.
 - Krebs & Loeschke 1994 found reduced fecundity

7. The role of Hsps for adaptation

- Improved heat resistance of insects after hsp expression in **insects** ((Gehring & Wehner 1995; Dahlgaard et al. 1998), **fish**(Basu et al. 2002) , plants (Sun et al. 2002), **mammals** (Ulmasov et al. 1993; Matz et al. 1996a).

- Some of the first data on the possible ecological relevance of Hsp expression comes from selection studies. In different species of *Drosophila*, it was shown that (contrary to predictions at that time) expression of Hsp70 was lower in lines frequently, or continuously exposed to severe stress (Bettencourt et al. 1999; Sørensen et al. 1999; Lansing et al 2000)
 - interpretation = costs of Hsp in pops exposed to stress outweighs benefits and adaptation is achieved through some other means
- Density dependence might drive hsp expression
- Age will influence hsp expression
- Life stages may differ in stressors exposed
 - Stage specific Hsp expression and resistance - no citation
 - not much support (Sørensen et al. 1999)
- In adults, Hsp70 down regulated with decreasing heat stress resistance (Sørensen & Loeschcke 2002a)

8. Conclusions and future prospectives

- However, new results show that Hsp expression is highly fine-tuned (not being only an on-off mechanism) and that Hsps are also continuously expressed after mild chronic stress exposure.
- Hsps as biomarkers of stress
- However, local adaptation and selection for other kinds of adaptive mechanisms may disturb the evaluation of the results. The results by Sørensen et al. (1999, 2001) and Koehler et al. (2000) showing that there is selection against Hsp expression in populations being exposed to chronic stress clearly demonstrates this problem.

This is a pretty old study. Looking up papers that cite it that look cool/interesting:

1. T. Esperk, A. Kjærsgaard, R. J. Walters, D. Berger, W. U. Blanckenhorn, **Plastic and evolutionary responses to heat stress in a temperate dung fly: negative correlation between basal and induced heat tolerance?**, Journal of Evolutionary Biology, 2016, 29, 5
2. Ramadoss Dineshram, Kondethimmanahalli Chandramouli, Ginger Wai Kuen Ko, Huoming Zhang, Pei-Yuan Qian, Timothy Ravasi, Vengatesen Thiagarajan, **Quantitative analysis of oyster larval proteome provides new insights into the effects of multiple climate change stressors**, Global Change Biology, 2016, 22, 6
3. L.Ye. Kozeko, D.B. Rakhmetov, **Variation in dynamics of the heat shock proteins HSP70 synthesis in Malva sylvestris and M. pulchella (Malvaceae) in connection with tolerance to high temperature, flooding and drought**, Ukrainian Botanical Journal, 2016, 73, 2, 194
 - constitutive and inducible; for thermal and drought tolerant plants
4. Delphine Cottin, Natacha Foureau, Frédéric Hervant, Christophe Piscart, **Differential regulation of hsp70 genes in the freshwater key species Gammarus pulex (Crustacea, Amphipoda) exposed to thermal stress: effects of latitude and ontogeny**, Journal of Comparative Physiology B, 2015, 185, 3, 303
 - good ref, added to my table

5. Jennifer A. Jost, Emily N. Soltis, Marshall R. Moyer, Sarah S. Keshwani, **Linking zebra mussel growth and survival with two cellular stress indicators during chronic temperature stress**, Invertebrate Biology, 2015, 134, 3
6. D Porcelli, R K Butlin, KJ Gaston, D Joly, R R Snook, **The environmental genomics of metazoan thermal adaptation**, Heredity, 2015, 114, 5, 502

Overall thoughts:

Not very informative in terms of what Hsps are actually doing and how they relate to upper thermal limits. What is the difference between protection and resistance? Tolerance? Some consistent wording would be nice. The paper focused too much on costs and benefits rather than adaptive variation. This is nice to know but what aspect of the thermal ecology do Hsps relate to? I would have hoped there would be a better discussion fo basal vs induced hsps.

Page 17: 2017-01-31. List of Adaptive Variation in Hsps lit table

Author	Year	Journal	Taxa	Life.Stage	Hsp	Expression.type	Comparison	Basal.warm.vs.cold.adapted.	Indu
Dietz & Somero	1992	PNAS	fish	adults	hsp90	protein	interspecific	NA	indu
Gehring & Wehner	1995	PNAS	ants	adults	hsp70	protein	interspecific	NA	high
Feder	1996	JEB	Fruit flies	larvae, pupae	hsp70	protein	intraspecific	hard to tell	high
Feder	1997	Functional Ecology	Fruit flies	larvae	hsp70	protein	intraspecific		
Krebs &	1997	Evolution	fruit flies	larvae,	hsp70	protein	intraspecific		

Feder				adult						
Bosch	1998	PNAS	Hydra	Polyps	hsp60	protein	intraspecific	NA		high indu
Krebs	1999	Cell Stress & Chaperones	Fruit flies	larvae	hsp70	protein	interspecific	NA		right
Dahlhoff & Rank	2000	PNAS	beetles	adults	hsp70	gene or protein	intraspecific	lower		indu
Tomanek & Somero	2000	Physiological and Biochemical Zoology	marine snails	adult	hsp83	protein	interspecific	NA		high
Tomanek & Somero	2000	Physiological and Biochemical Zoology	marine snails	adult	hsp70	protein	interspecific	NA		high
Boshoff	2000	CMLS	humans	blood cells		protein	blood cells			
Zatsepina	2001	JEB	Fruit flies	adult	hsp70	protein	intraspecific	NA		lower
Tomanek & Somero	2002	JEB	marine snails	adult	hsp70	protein	interspecific	NA		high
Tomanek & Somero	2002	JEB	marine snails	adult	hsp90	protein	interspecific	NA		high
Hofmann	2002	Integ. & Comp. Biol.	urchins	tube feet	hsp70	protein				high
Dahlgaard	2002	Functional Ecology	Fruit flies	adults	hsp70	proteins	intraspecific			
Garbuz	2003	JEB	Fruit flies	adults	hsp70	protein	interspecific	NA		high
Place	2005	Polar biology	Antarctic fish	adult	hsp70	gene				
Huang	2007	Journal of Insect Physiology	pea leafminer	adults	all	gene	intraspecific			
Laayouni	2007	BMC evo bio	Drosophila subobscura	larvae	hsp26	gene	intraspecific	NA		
Laayouni	2007	BMC evo bio	Drosophila subobscura	larvae	hsp68	gene	intraspecific	NA		
Bettencourt	2008	BMC Biology	Fruit flies	Larvae		gene	intraspecific			
Dong	2008	Biol Bull	Limpets	adults	hsp70	protein	interspecific	higher		high
Elekonich	2009	Cell stress and chaperones	honey bees	adults	hsp70	gene	tissues	higher in thorax of older bees		
Jensen	2009	Journal of Experimental Zoology	Fruit flies	adult	hsp70	protein	intraspecific	NA		no d
Lockwood	2010	JEB	Mytilus	gill	hsp70	gene	interspecific	no difference		no d
Lockwood	2010	JEB	Mytilus	gill	hsp24	gene	interspecific	NA		high
Mizrahi	2010	Cell Stress & Chaperones	land snails	adult	hsp70	protein	interspecific	revisit		
Tomanek	2010	JEB	mussels	adult	hsp70	protein	interspecific	NA		right
Franssen	2011	PNAS	seagrass	grass	27 hsp	gene	intraspecific			2/27
Carmel	2011	Heredity	Fruit flies	adults	hsp40	protein	intraspecific	higgher		high
Carmel	2011	Heredity	Fruit flies	adults	sHsps	protein	intraspecific	no difference		no d

Calabria	2012	J. Experimental Biology	Drosophila subobscura	adults	hsp70	protein	intraspecific	higher	no d
Madeira	2012	Cell Stress & Chaperones	crab	haemolymph	hsp70	protein	intraspecific	NA	high
Graham	2012	Journal of Heredity	Drosophila pseudoobscura	adults					
Bedulina	2013	Molecular Ecology	Amphipods	Adults	hsp70	gene	interspecific	higher; hsp70	lower
Cottin	2014	Journal of Comparative Physiology B	Amphipods	adult	hsp70	gene	intraspecific	lower	lower
Cottin	2014	Journal of Comparative Physiology B	Amphipods	adult	hsc70	gene	intraspecific	lower	no d
Cottin	2014	Journal of Comparative Physiology B	Amphipods	early	hsp70	gene	intraspecific	no difference	no d
Cottin	2014	Journal of Comparative Physiology B	Amphipods	early	hsc70	gene	intraspecific	no difference	no d
Franssen	2014	Marine Genomics	seagrass	shoots	28	gene	interspecific	higher	no d
Madeira	2015	Comparative Biochemistry and Physiology, Part A	shrimp	muscle	hsp70	protein	interspecific	no difference	no d
Nguyen	2016	BMC Evo Bio	ants	adults	hsp70 (hsc70-4 h1)	gene	interspecific	higher	high
Nguyen	2016	BMC Evo Bio	ants	adults	hsp70 (hsc70-4 h2)	gene	interspecific	no difference	high
Nguyen	2016	BMC Evo Bio	ants	adults	hsp83	gene	interspecific	no difference	high
Nguyen	2016	BMC Evo Bio	ants	adults	hsp40	gene	interspecific	lower	high

Page 18: 2017-01-31. Status of projects: writing

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - Met with NJG yesterday, revised intro, reread, should be

- ready to review with NJG 2017-02-06; Monday
- Meeting with NJG 2017-02-03, Friday 1PM to go over results, figure legends, figures.
- Submit to PNAS in May?
- **Multiple stressors ms:**
 - In SHC's hands, due 2017-02-14
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
 - **Thermal niche ms:** Lacy to check intro, and add refs.
- **Stressed in nature MS: Samples to rerun.**
 - **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. Thesis related

- **Defense talk:**
 - Giving practice talk **February 9th** (Thursday in lab meeting, 4PM)
 - Need to write out script by **Feb 7th**
 - Required defense talk in biolunch: **Feb 24 (Friday)**
- Working title: Evolutionary Innovations of Ants to Thermally Stressful Environments
- Formatting:

- Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - started filling in introduction, eta? No clue, working on in the background
 - Deadlines:
 1. Intent to graduate: February 1st for May. —**done**
 2. Send defense committee form to grad college—**done**
 3. **Graduate college format check March 4th**
 4. Defense notice 3 weeks before defense (oral defense by March 24th—*flexible*).
 5. Final thesis April 7th.
-

Page 19: 2017-01-31. SHC lab meeting reading: Ranga et al. 2017; Evol Ecol: Sibling *Drosophila* species (*Drosophila leontia* and *Drosophila kikkawai*) show divergence for thermotolerance along a latitudinal gradient

This study compares the different aspects of lower and upper thermal limits between two closely related species (*Drosophila*) in India across a wide climate gradient (8 -32 degrees north):

1. *D. leontia*
2. *D. kikkawai*

Measurements:

1. Upper thermal limits

- Heat knockdown time: 39 C treatment in water bath, measured time where they could not stand
- Survival: 24 hour survey of flies 100% dead when treated at 39C over time (10-80min)
- Hardening:

2. Lower thermal limits

- Chill coma recovery time
 - 0C treatment for 8 hour then recover at room temp (22C); measured time they could stand upright
- Cold survival
 - 0 C treatment for 8 -96 hours, measured survival after 24hours

3. Absolute hardening capacity (AHC)

- $AHC = KT - C$
- KT = thermotolerance after hardening
 - warm: (33, 35 or 37 C)
 - cool: (0, 2 or 4 C)
- C = basal thermo-tolerance

4. Relative hardening capacity

- $RHC = (KT - C)/C$

5. Egg-to-adult viability

- Hot: eggs subjected to 39 C 1 hour
- Cold: eggs subjected to 0 C 12 hours
- transferred to 22C 65% RH

Results

(Fig 2)

1. *D. leontia* lower plasticity

- Clinal variation in Upper thermal limits (KD time, heat survival)
 - No benefit from hardening
- No clinal variation in lower thermal limits (CCRT, cold survival)
 - No benefit from hardening

2. *D. kikkawai* —higher plasticity

- Clinal variation in Upper thermal limits (KD time, heat survival)
 - Benefit from hardening
- Clinal variation in lower thermal limits (CCRT, cold survival)
 - Benefit from hardening

Table 2: They did anova for each trait, testing the effect of each population

Why not do an ANCOVA:

```
1 | aov(Thermal trait ~ Latitude * Species)
```

Fig 3: Egg-to-Adult viability

1. *D. leontia* lower plasticity

- Southern pop had higher viability under controls, heat stressed, but not cold stress.
- **Notes makes sense because they have no relationship to cold**

o

2. *D. kikkawai* higher plasticity

- o Southern pop had higher viability under heat stress but lower under controls and cold stress.

Fig 4: Mortality

1. *D. leontia* lower plasticity

- o **HS:** No diff in mortality between hardening and controls (consistent with fig 2)
 - o **HS:** Southern had lower mortality than north
 - o **CS:** No diff between hardening and controls and pops
- #### 2. *D. kikkawai* higher plasticity
- o **HS** higher mortality in control than hardened, makes sense; southern has lower mortality
 - o **CS:** Northern have lower mortality than south ; hardening has lower mortality

Fig 5: AHC (Absolute hardening capacity) and RHC relative hardening capacity: Same direction

1. *D. leontia* lower plasticity

- o no sig hardening between northern and southern pops

2. *D. kikkawai* higher plasticity

- o Upper thermal limits: positive plasticity
 - Northern pops have higher upper thermal limits
- o Lower thermal limits:
 - CCRT: northern have faster recovery time
 - survival: northern have higher survival

Fig 6: Hardening vs pretreatment temperatures

1. Upper thermal limits:
 - *D. kikkawai* higher plasticity had higher hardening than *D. leontia* lower plasticity for both KD time and survival
2. Lower thermal limits:
 - *D. kikkawai* higher plasticity has higher cold tolerance than *D. leontia* lower plasticity

Fig 7 looks the same, wtf

Figure 8 Focus on the climate variables each species experiences

- North is more variable, south less variable (A)
 - Tave and RH negatively related to latitude
1. *D. leontia* lower plasticity
 - higher abundance at the warm end
 2. *D. kikkawai* higher plasticity
 - higher abundance at the cold end

Thoughts

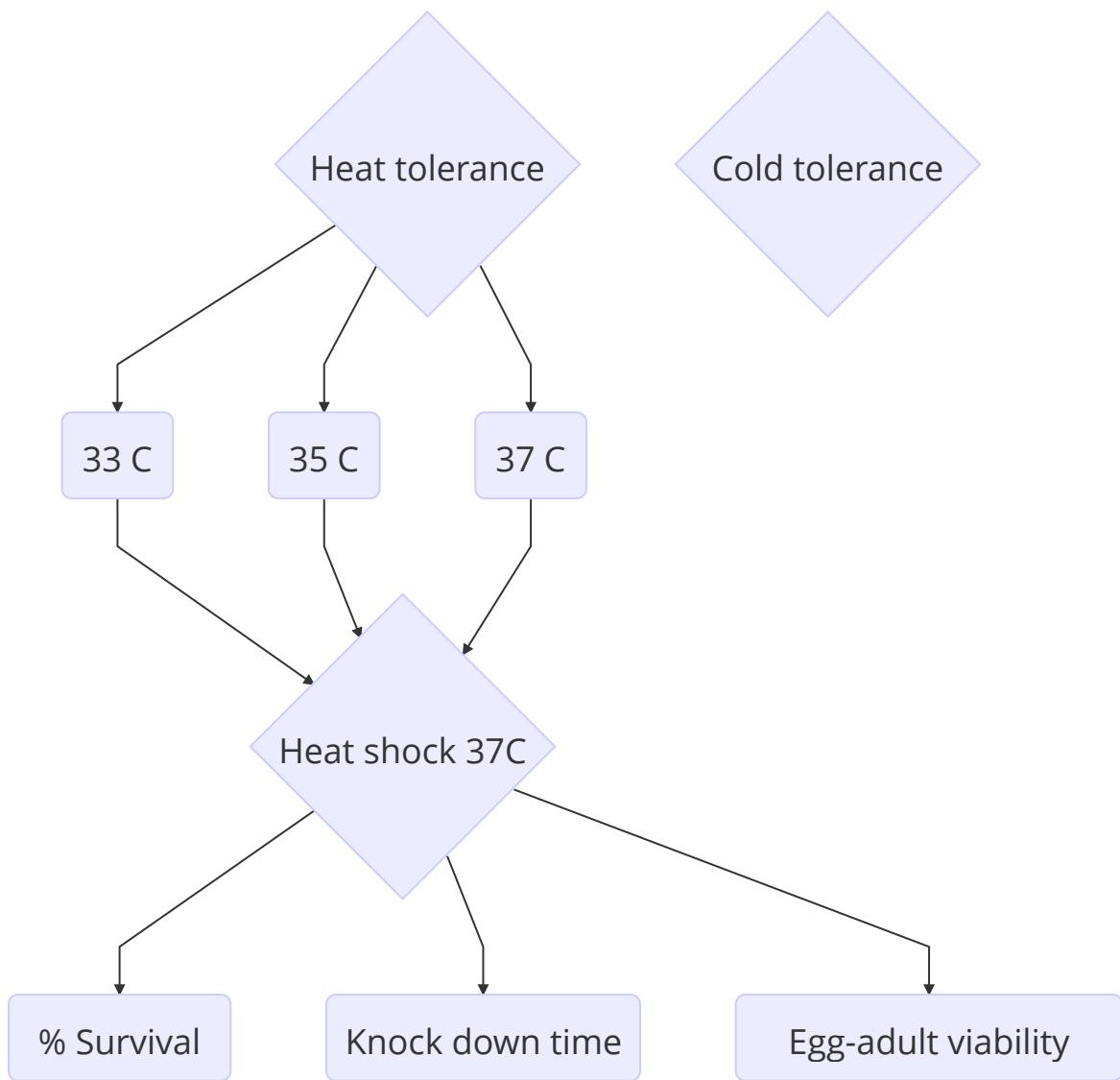
- What is the demographic history? I ask because it can explain some of the results.
- Optimal temp for *D. leontia* is the south and *D. kikkawai* is in the north. So *D. leontia* could be moving south to north and *D. kikkawai* could be moving north to south.
- Painful read, could have analyzed the data differently.

Page 20: 2017-02-02. SHC lab meeting: reading Ranga et al. 2017; lab discussion

Preliminary stuffY:

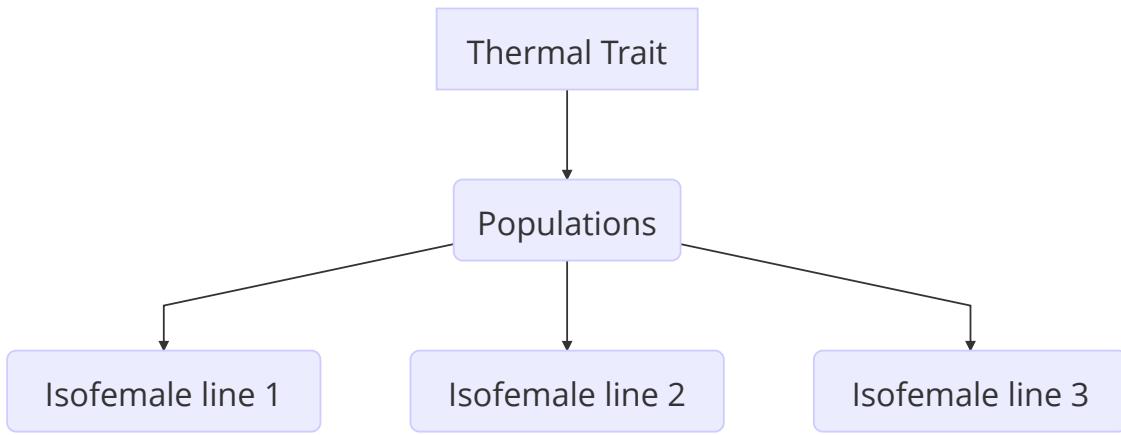
- SHC needs to quickly read multiple stressors ms

Katie miller wants to draw out stuff



Not exactly the experimental design....

Model construction for testing the effects of lines and populations on thermal traits: Nested ANOVA



**Page 21: 2017-02-03. Todgham et al. 2017;
JEB, The effect of temperature adaptation
on the ubiquitin-proteasome pathway in
notothenioid fishes**

Ref: Todgham AE, Crombie TA, Hofmann GE. 2016. The effect of temperature adaptation on the ubiquitin-proteasome pathway in notothenioid fishes. The Journal of Experimental Biology;jeb.145946.

Background

Ubiquitin tags proteins for degradation through the proteasome(protein chomper). In the face of heat stress, different species can adaptively modulate this pathway to cope with the deleterious effects of protein damage. Antarctic species live in -1.9C while new zealand species live in 10C .

Objective: Understand the thermal compensation of Ub-proteasome pathway in antarctic fish vs ones from warmer climates.

Approach and workflow

Compared ubiquitin ,proteasome activity and expression between 2 tissues (gill and liver) for 4 different fish species.

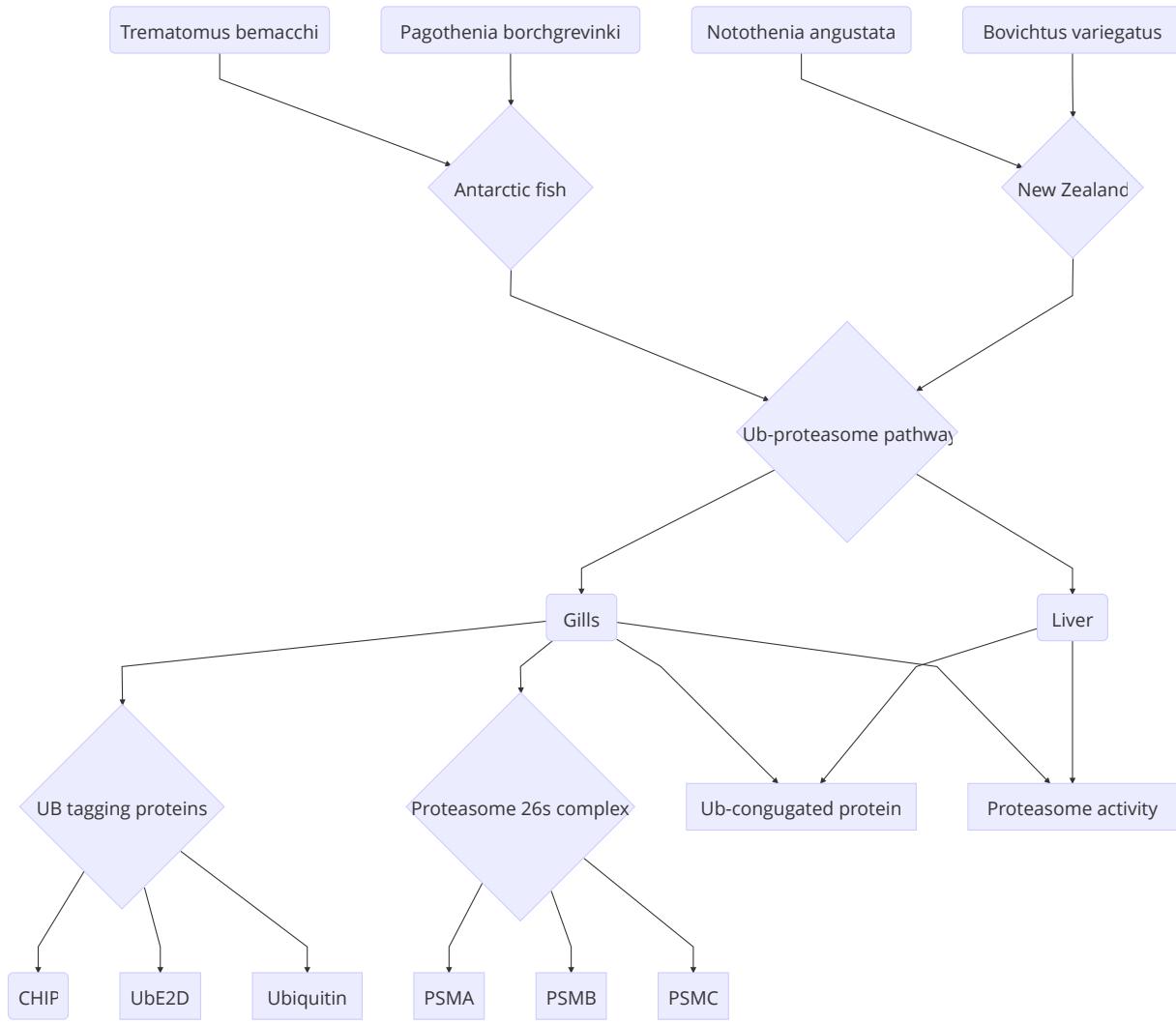


Figure 1: UB conjugated proteins

Antarctic and *N. angustata* have higher ubiquitin tagged proteins than *B. variegatus* for both liver and gill.

Figure 2. Protasome activity

Antarctic have higher proteasome activity at 0 and 10 C for both gill and liver. The magnitude of differences were higher in the gills. This may be why they focused only on the gills for gene expression

Figure 3: Gill gxp for UB related genes

General pattern as Fig 1. where antarctic fish and *N. angustata* have higher expression of UB related gene expression than *B. variegatus*. But *N. angustata* had higher expression than the antarctic fish. There could be different dynamics that would be captured with a reaction norm approach.

Figure 4: Gil gxn for proteasome complex related genes.

Similar result as figure 1. (except top panel) . But here we see categorical differences between Antarctic and new zealand. Antarctic fish have higher expression.

Thoughts:

It looks like cold tolerant species up-regulate their protein degradation pathway. It fits the model that if proteins are too damaged, they're clear those proteins.

Cool method to test for proteasome activity: (Coux et al., 1996)

1. chymotrypsin-like activities, which cleaves after large hydrophobic residues
 2. trypsin-like activities, which cleaves after basic residues
 3. peptidylglutamylpeptide hydrolyzing activities, which cleaves after acidic residues .
-

Page 22: 2017-02-06. Tomanek & Somero 2002; JEB Paper notes

Reference:

Tomanek L, Somero GN. 2002. Interspecific- and acclimation-induced variation in levels of heat-shock proteins 70 (hsp70) and 90 (hsp90) and heat-shock transcription factor-1 (HSF1) in congeneric marine snails (genus *Tegula*): implications for regulation of hsp gene expression. *J Exp Biol* 205:677–685.

Rational/Objectives/Questions

How do species adaptively vary in their stress response?

Species differ in vertical distribution

1. *Tegula brunnea* = subtidal
2. *T. montereyi* = subtidal to low intertidal zones
3. *T. funerals* = low to mid intertidal zone (with wider latitudinal distribution)
 - experiences the highest heat stress

Methods

1. Collected mid July and treated 13, 18 and 23 °C for 30-34 days
2. Dissected tissues and isolated proteins
3. They measured protein levels of Hsp70/90/40, HSF with western

blots.

Take homes

Our results indicate that the cellular thermometer model can account for intraspecific but not interspecific variation in T_{on} and that the relative levels of expression of different isoforms of hsp70 may provide a better measure of heat stress than the total expression of all hsp70 isoforms.

Figures are uninterpretable.

Page 23: 2017-02-06. Tomanek 2010; JEB

Reference:

Tomanek L. 2010. Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. *J Exp Biol* 213:971–979.

This is one of the few papers that compare the thermal reaction norms of Hsp protein expression among 3 different species. These species vary in the thermal environment

Species: *Chlorostoma* formerly *Tegula*

1. *C. funebralis* = warm tolerance
2. *C. brunnea* = intermediate tolerance
3. *C. montereyi* = low to intermediate tolerance

Figure 1.

These species shift the reaction norm of Hsp protein expression to the right, in support for the tolerance mechanism. And *C. funebralis* shifts their max expression upwards at the higher temperature.

Some species lack HSR:

1. Hydra lack a HSR (Bosch et al. 1988)
 - mRNA gets rapidly degraded
2. antarctic marine organisms have little HSR
 - lacks inducibility of HSPs

Argues that the variation in the thermal reaction norms of Hsp expression is related to the temperature variation different taxonomic groups experience.

**Page 24: 2017-02-06. Parcel & Lindquist 1994;
cold spring harbor , paper notes**

Title: Heat Shock proteins and stress tolerance

-
- Protection (resistance) and tolerance are assumed to be the same (conflated).
- Induced thermotolerance = heat hardening; when exposed to sublethal stress and Hsps increase , it can confer resistance to future temperature threats
- *Nutrient availability, oxygen tension, diurnal rhythms, and a host of other variables exert highly reproducible effects on thermotolerance.*
- *In general terms, hsps function by preventing the accumulation of stress-damaged proteins.*
-

How does heat stress damage cells?

- Heat damages a wide variety of cellular structures and metabolic processes (for detailed reviews, see Nover 1991; Laszlo 1992).*
- *In higher eukaryotes, one of the most immediate effects of heat shock is extensive disruption of the cytoskeleton (Falkner et al. 1981; Coss et al. 1982; Glass et al. 1985; Welch and Suhan 1985; Iida et al. 1986).*
- The golgi gets fragmented and number of lysosomes increase.
- Mitochondria swell and decrease in number— decrease in oxidative phosphorylation
- influence gene expression; reduced
- DNA synthesis is slowed; inhibit chromatin assembly; DNA becomes unstable
- *The total protein content of the nucleus also increases with heat shock (Laszlo 1992).*
- changes lipid membranes; increased fluidity of bilayer and aggregation of integral membrane proteins

What is the response?

- *In many cases, mild heat pretreatments, which induce hsp synthesis,*

either reduce the extent of these perturbations or speed their repair.

- *Unfortunately, although we know that hsp70s have a central role in thermotolerance and have learned a great deal about the general biochemical functions of these proteins, we still do not know which heatinduced cellular perturbations are responsible for lethality or which cellular lesions are the most susceptible to repair by hsp70s.*

Hsp70 function in stress tolerance (so this paper frames in term of tolerance)

- *early experiments showed a close correlation between their induction and the induction of tolerance to high temperatures (Li and Werb 1982; Li and Laszlo 1985; Subjeck and Shyy 1986).*
 - Li & Werb 1982 compare thermal tolerance with Hsp protein expression

Where is hsp70 localized during HS?

- concentrated in membrane, nuclei, and nucleoli (*Pelham 1984; Pelham et al. 1984; Velazquez and Lindquist 1984; Welch and Feramisco 1984*)

In a few experiments, differences in thermotolerance among cultivars have been correlated with differences in hsp synthesis (Ougham and Stoddart 1986; Howarth 1989; Krishnan et al. 1989).

In one case, however, where the cosegregation of hsp polymorphisms and thermotolerance polymorphisms was examined, no simple correlation was observed (Fender and O'Connell 1989).

Page 25: 2017-02-07. How far down do Aphaenogaster colonies go in the winter?

Lubertazzi 2012 has the answer.

1 4.1.4.Winter. In winter months colonies avoid freezing
2 temperatures
3 by maintaining their nests below ground. Talbot
4 [24] found the average depth of 5 winter colonies
5 in Missouri
6 to be 25 cm. Colonies in Connecticut appear to prefer
7 deeper
8 nests, to a depth of at least 50 cm. Developmental
processes enter a diapause and worker activity within the nest is
minimal.

Depending on where they are, 25-50 cm!!!

reference:

Lubertazzi D. 2012. The Biology and Natural History of *Aphaenogaster rudis*. *Psyche: A Journal of Entomology* [Internet] 2012. Available from: <http://www.hindawi.com/journals/psyche/2012/752815/abs>

Page 26: 2017-02-07. A definition of a trade-off.

Reading Saltz et al. 2017, Trait Correlations in the Genomics Era; published in *TREE*

I like this:

- 1 However, predictions about the evolutionary dynamics of trait correlations go beyond heritability: often we
- 2 are interested in why traits are correlated, whether the correlation evolved under selection, and
- 3 whether it is possible for selection or drift to change the magnitude or direction of trait
- 4 correlations.

Trade-off definition:

Trade-offs occur when functional relationships among traits prevent evolution of optimum values for all traits simultaneously.

or...when the relationships among traits prevent the evolution of optimal values for all traits simultaneously.

or

when the correlational structure of functional traits are at odds with selection such that the suite of traits are not optimized.

Thoughts

Should there be a distinction between a negative correlation among traits as a **functional trade off**, vs how selection operates on traits to limit trait values in the next generation as an **evolutionary trade off**? Because if selection operates in the same direction of traits that are negatively correlated, then trait values are optimized. If selection operates perpendicular, then that'd be a trade off.

I like this too:

- 1 Quantitative genetics theory suggests that identifying genetic loci underlying trait correlations can contribute to answering these questions. Specifically, quantitative trait loci can produce trait correlations through pleiotropy, in which a single locus causally affects two or more traits,
- 2 or through linkage, in which two or more loci each affect different traits, but are in linkage disequilibrium (LD) and therefore are inherited together (Box 1). This distinction is important because trait correlations caused by pleiotropy are expected to evolve (adaptively and by drift) differently than trait correlations caused by LD. In general, because LD is expected to erode
- 3 through recombination, trait correlations generated by LD are expected to be transient and therefore have a limited scope for contributing to evolutionary change. Similarly, if a trait correlation is caused by LD, each trait might be produced by distinct functional mechanisms

reference

Saltz JB, Hessel FC, Kelly MW. 2017. Trait Correlations in the Genomics Era. *Trends in Ecology & Evolution* [Internet] 0. Available from: [http://www.cell.com/trends/ecology-evolution/abstract/S0169-5347\(16\)30242-7](http://www.cell.com/trends/ecology-evolution/abstract/S0169-5347(16)30242-7)

Page 27: 2017-02-09. SHC lab meeting: Practice talk for biolunch and dissertation defense

SHC : overall intro is fine, you'll give it better every time you do it. you'll get smoother.

SHC has problem with question:

What types of adaptations are needed to evolve into different environments?

What types of adaptations are needed to colonize different environments? (might want to be more specific, "thermally extreme environments")

SHC wants me to avoid, performance is physiology/ behavior. performance relates to fitness. They are not the same thing.

In slide 8; I'm showing TPC's across a temperature gradient. The whole thing is shifting, but what im saying is that only ctmax is changing. thinka bout this. display diff

slide 9: there is an issue with me only presenting 1 option. There are other ways!!! don't necessarily talk about other ways.

Slide 11: Engage with the figure: Here are the 3 heat shock proteins, but if i have a pathway, i better show how it works. Rethink.

Slide 17: this is vague. I mean that the "dynamics of how they are used". This is what i want to stress. "Temperature selection" is too vague and I can say something more specific.

Slide 18: this is confusing: too fast, both slide 18 and 19 have positive slopes and how ar they different?

SHC suggestion: I want to get across, how do i use hsps to be better protected? Have them there all the time! Get across in plain english

Slide 19: Tolerance: when you do get damage, you repair it , turn on when needed and turn on stronger.

Went too fast; and with graphs, it is hard for audience to get it. Bar graph instead to give people the words (SHC is not sure)

If HSPs are contributing to thermal limits, how are they doing it? SHC confused as to why upper thermal limits is on y axis

get rid of slide 20

Slide 21: title is bad, instead: ants live in different environments. colonized environments from extremely cold to extremely hot!

Slide 22: Question 2 needs adjustments. How implies that it assumed they already adaptively vary. "relate" is vague, be more precise.

Question 3 is awkward. SHC is not sure, think about how to frame it . How do upper thermal limits interact with other elements of physiology that relates to or influence species survival. or where species occur or performance

slide 24: NOT clear that i used drosophila for blast search. Which ones in fruit flies are heat inducible?

SHC suggestion: Save housekeeping business for this slide. Some are for heat response and some are for housekeeping function. Use fruit flies as a reference. Adds to story about shifts taht have gone on in these proteins used in ants as opposed to other insects. it is dynamic.

KM suggestions

conclusions for each chapter was unclear:

multiple stressors: symbols for desiccation and starved were the same. change colors!

Slide 25: give a road map for the first time when i present hsp83 gene trees and regulatory elements. diverse group....some might not understand how you can infer duplication on a gene tree.

Slide 27: call it hsc70-4 instead of hsp70.

What is going on with bombus?

Slide 30: have contrasting slide, meaning show hsp orthologues not heat inducible!

Slide 31: Flat title. Conservation and innovation—say something about the evolutionary dynamics of hsps....they are conserved in function, but there is a lot of opportunity for innovation. Switch between housekeeping to inducible. There is more to be explored....i.e. subfunctionalization. Not a simple answer, there is a complex answer. Conserved function but the individual players is not static!

One time to make general statements about the implications of the work I'm doing. Generalize so you can make general comments about biology. don't overstate but think critically!

Slide 33: both hypotheses don't belong on the same slide

Ask...whether there are adaptive shifts in upper thermal limits when ants colonize novel environments.

This is what I want to know....

Extension or enhancement of thermal limits. If you live in a warmer place, can you withstand higher temperatures.

If so, how do you do it? Are hsps involved, how!?

important

Then go to the clade. species are structured along the cline. colonize east coast over X million years. and they have colonized open environments.

Bonnie comments

When I go into the predictions, make it explicit that they are predictions.

Four main questions was confusing. I only have 3 points, but bonnie was expecting 4.

Make sure im' clear on DW and FW

Mistake : I'm assuming that everybody will understand a phylogenetically informed approach. Give overview of approach! What is the phylogeny doing? Set up the audience to understand. More important to be understandable than length.

Slide 35-40 I need to use titles!1

Conclusion to CTmax part: I set up conflict, but my conclusion does not resolve the conflict. Needs to be in plain english.

Are the differences in CTmax explained by differences in how Hsps are used?

How does graded protection help/work?

Tolerance mechanism is not a driver, it's a response to something else that is reducing the incidence of protein denaturing.

Max is more about repair.

three possibilities:

1. protection - active
2. repair - active
3. tolerance - passive

Frame your question! Think about this to tell a precise story.

RXN norm slides: connect it to something. Issue with how people study this. ANOVA approaches misses elements of adaptive changes in responses

DW buffer thermal environment compared to FW. Use data to decide whether i focus on habitat differences or to include cline in ctmax.

For chapter 2 conclusion: wrap up the whole chapter...thermal limits and hsp. Evolutionary lability or flexibility in how hsps are used can lead or ability to colonize new habitats. Link hsps back to habitat. integrate the data better

Natural shift to last question: The data i have might suggest, in the framework of climate change....DF species might be more protected because they dont get exposed to extremes. HOWEVER, climate change is not about change in temperature....but the constellation of changes associated with climate change . Increases in temperature, other things are changing....other stressors can lead species to become more vulnerable.

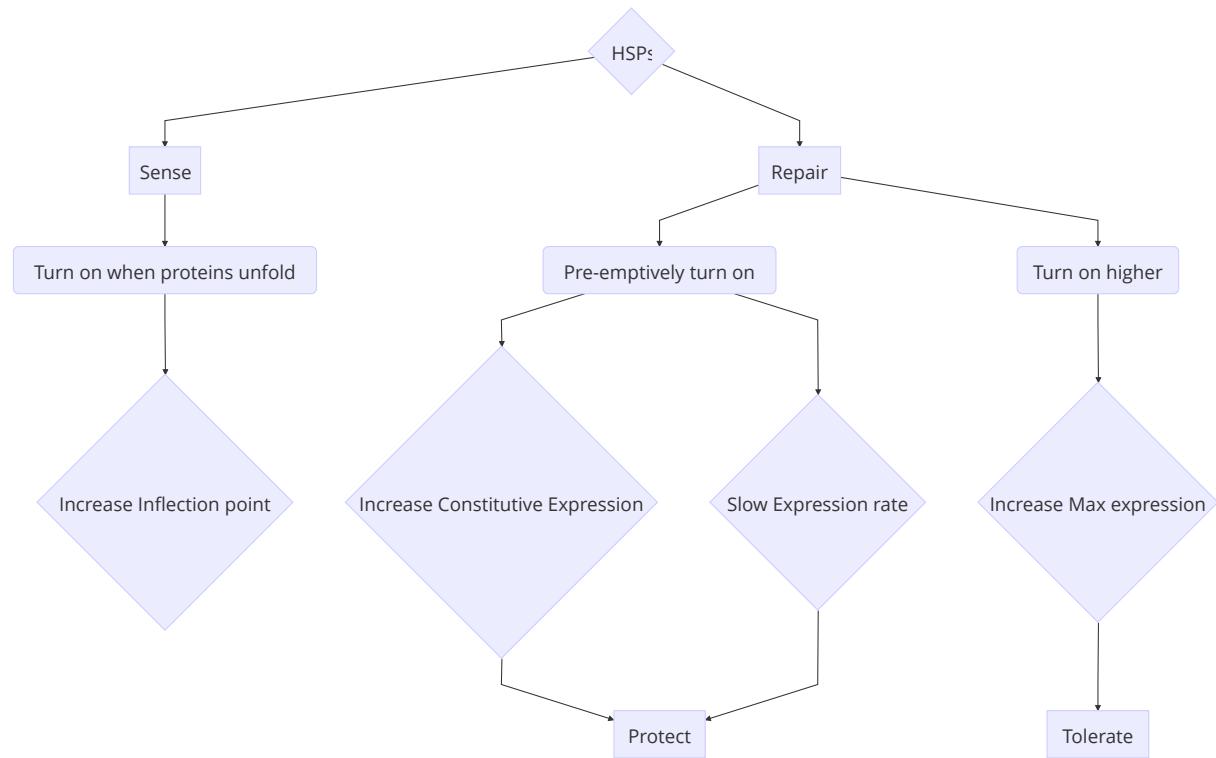
Goals instead of topics. Last part could be integrating my results with current climate change patterns.

Slide 68: hypothesis slide for multiple stressors section

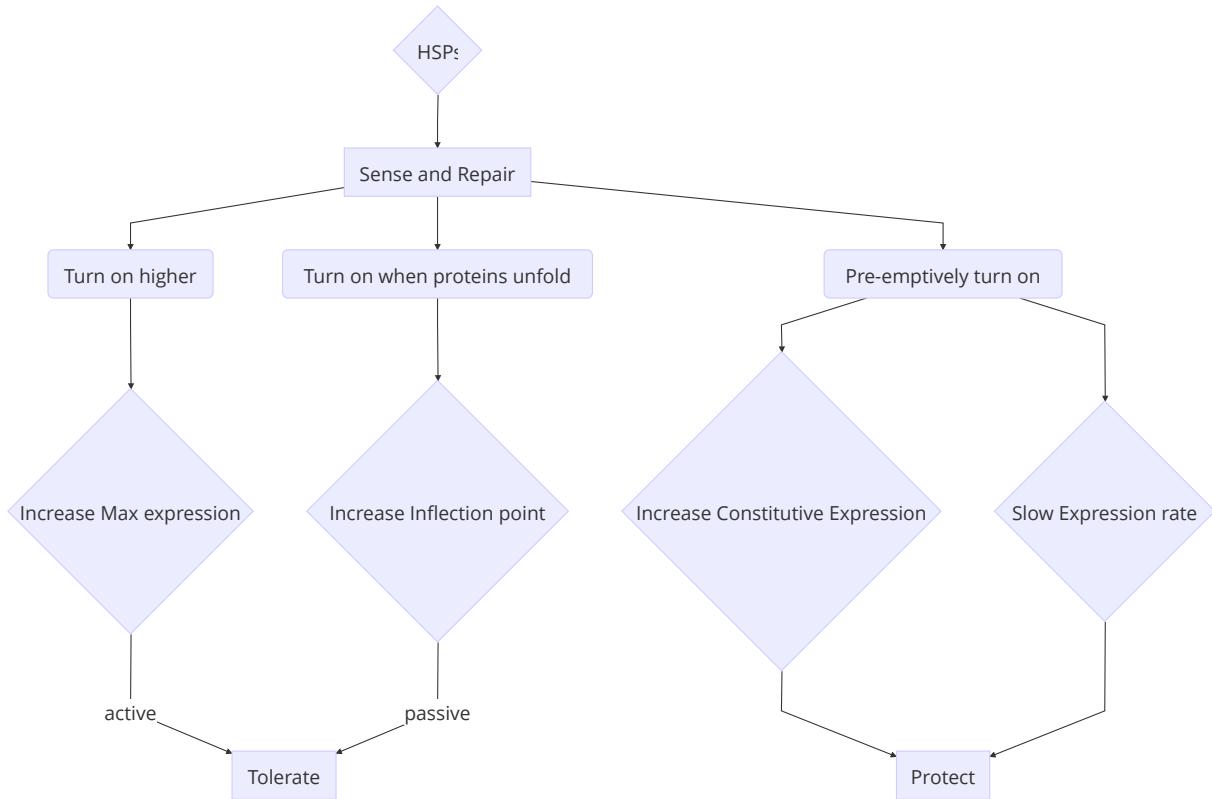
I have "or", split up the hypothesis

Slide 70: x axis in predictions could be severity. Overall, predictions are abstract. Explain cross protection or susceptibility mechanisms. ie extra stressor improves or reduces them. NO need for complex prediction slides. Why would another stressor can ever help you? (link back to HSPs...but don't need to add the data;; but you need the explanation)

You need methods: treatments, knockdown time explanation. All results need statistics. Play with overall conclusion.



2017-02-15: Alternative way of displaying to get at passive vs active for tolerance



Page 28: 2017-02-14. Status of projects + PHD Progress

Updated again: 2017-02-16

1. Project updates:

- **Hsp gene expression + Ctmax project:**
 - NJG made comments 2017-02-13, and I fixed them today
 - Send out for SHC lab meeting 2017-02-16 to go over next week (2017-02-23)
 - Submit to PNAS in May?
- **Multiple stressors ms:**
 - Resubmitted 2017-02-13; in review (today)
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
 - **Thermal niche ms:** Lacy to check intro, and add refs.
- **Stressed in nature MS: Samples to rerun.**
 - 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. Thesis related

- **Defense talk:**
 - Required defense talk in biolunch: **Feb 24 (Friday)**
 - Working title: Evolutionary Innovations of Ants to Thermally Stressful Environments
- **Formatting:**

- Introduction (> 3 pages), manuscripts, then synthesis/conclusion (~3 pages) ; SHC and NJG agree
 - CH1. Introduction — started filling in outline
 - **CH2. Hsp functional diversity paper — done (published)**; in thesis now
 - CH3. Hsp rxn norm paper -
 - **CH4. Multiple stressors paper — done (In review)**; in thesis now
 - CH5. Conclusion — have outline
 - Deadlines:
 1. Intent to graduate: February 1st for May. —**done**
 2. Send defense committee form to grad college—**done**
 3. **Graduate college format check March 4th**
 4. Defense notice 3 weeks before defense (oral defense by March 24th—*flexible*).
 5. Final thesis April 7th.
-

Page 29: 2017-02-15. Aphaenogaster, Hsp rxn norm paper stats revisited:

From previous analysis, I've found that there is a main effect of Tmax and Habitat type on CTmax in Aphaenogaster.

```

1 library(MASS)
2 #constructing model
3 umod<-
  lm(KO_temp_worker~bio5*habitat_v2+I(bio5^2),data=Aph.dat)
4 #model selection

```

```

5 summary(stepAIC(umod,direction="both"))
6
7 Coefficients:
8
9             Estimate Std. Error t value
10 (Intercept) -4.4102626 12.5885230 -0.350
11 bio5          0.2990131  0.0862737  3.466
12 habitat_v2flat woods 1.5151487  0.2472431  6.128
13 I(bio5^2)     -0.0004877  0.0001469 -3.320
14
15             Pr(>|t|)
16 (Intercept) 0.726851
17 bio5          0.000792 ***
18 habitat_v2flat woods 1.96e-08 ***
19 I(bio5^2)     0.001275 **
20
21 ---
22 Signif. codes:
23 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
24
25 Residual standard error: 0.8192 on 96 degrees of freedom
26 Multiple R-squared:  0.47, Adjusted R-squared:  0.4534
27 F-statistic: 28.37 on 3 and 96 DF,  p-value: 3.191e-13

```

Ok, what if I take out flat woods species, do I still get clinal variation?

```

1 #subset data
2 aphsub<-subset(Aph.dat,Aph.dat$habitat_v2!="flat
3   woods")
4 #construct model without habitat, since only focusing
5   on deciduous forest species
6 umod2<-lm(KO_temp_worker~bio5+I(bio5^2),data=aphsub)
7 #model selection
8 summary(stepAIC(umod2),direction="both")
9

```

```

8 Coefficients:
9                               Estimate Std. Error t value Pr(>|t|)
10 (Intercept) -3.1939844 13.5824985 -0.235 0.81468
11 bio5         0.2905107  0.0931214   3.120 0.00251 ** 
12 I(bio5^2)    -0.0004730  0.0001586  -2.982 0.00379 ** 
13 ---
14 Signif. codes:
15 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
16
17 Residual standard error: 0.8754 on 81 degrees of freedom
18 Multiple R-squared:  0.208, Adjusted R-squared:  0.1885
19 F-statistic: 10.64 on 2 and 81 DF,  p-value: 7.893e-05

```

General info on sampling

Species numbers

```

1 table(droplevels(Aph.dat$rad_seq_species))
2
3      ashmeadi      floridana       fulva
4          9             9             9
5      lamellidens     miamiana      picea
6          4            13            26
7      rudis tennesseensis
8          28             2

```

Habitat numbers

```
1 table(droplevels(Aph.dat$habitat_v2))  
2  
3 deciduous forest          flat woods  
4                 84                  16
```

Number of sites:

```
1 length(unique(Aph.dat$site))  
2 [1] 34
```

Range of Tmax * 10

```
1 range(Aph.dat$bio5)  
2 [1] 249 333
```

For pca of hsp parameters:

```
1 > pchsp$sdev  
2      Comp.1     Comp.2     Comp.3     Comp.4     Comp.5  
3 2.1385967 1.3517804 1.0759241 1.0023266 0.8465922  
4      Comp.6     Comp.7     Comp.8     Comp.9     Comp.10  
5 0.8464922 0.7749836 0.6587774 0.5320244 0.4468818  
6      Comp.11    Comp.12  
7 0.3442608 0.2738906
```

Here are the eigenvalues, and I want to get the number of dimensions
[Math Processing Error]

```
1 sum(pchsp$sdev) / pchsp$sdev[1]  
2 4.81275
```

Closer to 1 means that most of the variation among these traits is captured by first pc. The larger the number, more PCs capture most of the variation among these traits.

My session:

```
1 > sessionInfo()
2 R version 3.3.1 (2016-06-21)
3 Platform: x86_64-apple-darwin13.4.0 (64-bit)
4 Running under: OS X 10.12.2 (Sierra)
5
6 locale:
7 [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-
8/en_US.UTF-8
8
9 attached base packages:
10 [1] stats      graphics   grDevices utils      datasets
11 [6] methods    base
12
13 other attached packages:
14 [1] MASS_7.3-45   ggplot2_2.2.1 raster_2.5-8
15 [4] sp_1.2-4
16
17 loaded via a namespace (and not attached):
18 [1] Rcpp_0.12.9    lattice_0.20-34  assertthat_0.1
19 [4] grid_3.3.1     plyr_1.8.4    nlme_3.1-131
20 [7] gtable_0.2.0   scales_0.4.1   lazyeval_0.2.0
21 [10] tools_3.3.1   munsell_0.4.3  colorspace_1.3-2
22 [13] knitr_1.15.1   tibble_1.2
```

Page 30: 2017-02-20. prep for committee meeting 2017-02-24, 2:00PM

Layout and progress of my chapters

- Abstract:
(3/4 page)
- CH1: Introduction
(3-6 pages)
- CH2: The evolution of heat shock protein sequences, cis-regulatory elements, and expression profiles in the eusocial Hymenoptera
(published BMC Evolutionary Biology)
- Ch3: Molecular adaptations of protection and tolerance predict upper thermal limits in eastern forest ants
(aiming to submit to PNAS)
- writing in pnas format but will need to reformat for thesis layout?
- CH4: Effects of desiccation and starvation on thermal tolerance and the heat shock response in forest ants
(In 2nd round of reviews, Journal of Comparative Physiology B)
- CH5: Conclusions and future directions
(3 pages)

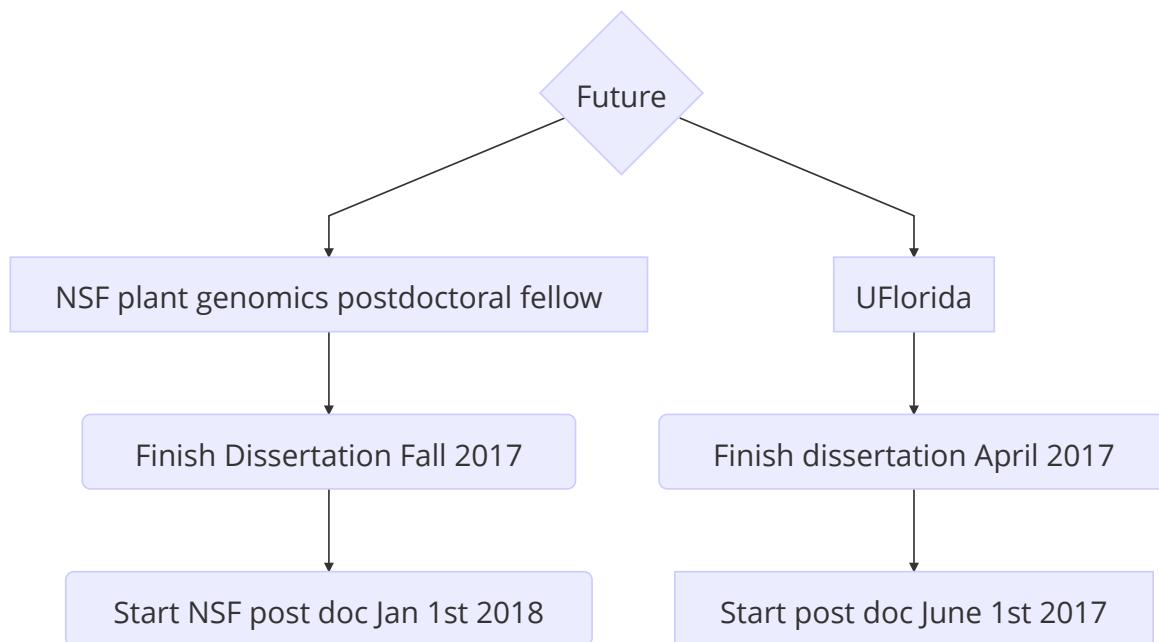
Timeline for completing my dissertation this term.

Important dates:

- Oral defense: March 24th is latest, but can be later, **March 27th-31st.**
 - Send out abstract and notification of oral defense 3 weeks prior
- Thesis due April 7th.

Set a date for my oral defense. Please bring your calendars.

Future post doc plans



Page 31: 2017-02-21. Measuring evolutionary rates in darwins and haldanes

I thought this was [cool](#):

a darwin ([Math Processing Error]): a change by a factor [Math Processing Error] per million years

[Math Processing Error]

- [Math Processing Error] and [Math Processing Error] are two sample means of natural logarithms of measurements
- [Math Processing Error] is the time difference between the two samples in millions of years

a haldane (h): a change by a factor of one standard deviation per generation

[Math Processing Error]

- [Math Processing Error] and [Math Processing Error] are two sample means of natural logarithms of measurements
- [Math Processing Error] is the pooled standard deviation of [Math Processing Error] and [Math Processing Error] values
- [Math Processing Error] is the time difference between two samples, measured in generations

Page 32: 2017-02-22. Resistance vs Tolerance strategies to stress; paper notes-Núñez-Farfán et al. 2007; The Evolution of Resistance and Tolerance to Herbivores; Annu. Rev. Ecol. Evol. Syst.

I want to get a sense of how the herbivore-plant literature uses certain words for responding to stress and potentially adopt that framework in temperature stress.

1. **Resistance:** Constitutive or induced response of plants against herbivory to avoid or reduce the amount of damage
2. **Tolerance:** Response of plants induced after consumption to buffer the negative fitness effect of damage.

These definitions seem inter-related. For example, an induced response that can avoid or reduce damage could also buffer the negative effects of damage. So, a molecule (secondary compound) can participate in both resistance and tolerance. I think the difference is the timing of herbivory. Meaning, if a molecule turns on early, it is acting to resist herbivory damage. But if a molecule turns on after herbivory, then its action will tolerate the damage already done.

In the coral literature, they like to use "[frontloading](#)" as another term for resistance.

1 ref: Barshis DJ, Ladner JT, Oliver TA, et al (2013)
Genomic basis for coral resilience to climate change.
PNAS 110:1387–1392. doi: 10.1073/pnas.1210224110

In Fineblum and Rausher 1995; they explain tolerance as the amount of reduced fitness for a given amount of damage, but resistance is the amount of damage a plant experiences for a given abundance of herbivores.

1 ref: Fineblum WL, Rausher MD (1995) Tradeoff between resistance and tolerance to herbivore damage in a morning glory. Nature 377:517–520. doi: 10.1038/377517a0

A diff expalantion of tolerance and resistance by Mauricio et al. 1997:

Resistance: Traits that reduce the amount of damage a plant experienced.

Tolerance: Ability of plant to sustain a fixed amount of herbivore damage without reduction in fitness.

A gem in this paper for the definition of a trade off:

For example, a negative correlation between two characters, both within and across species, is often taken to indicate the existence of tradeoffs between those characters (Cheverud 1984, Tilman 1990).

1 ref: Mauricio R, Rausher MD, Burdick DS (1997)
Variation in the Defense Strategies of Plants: Are Resistance and Tolerance Mutually Exclusive? Ecology
78:1301–1311. doi: 10.1890/0012-
9658(1997)078[1301:VITDSO]2.0.CO;2

In a diff system, but another [paper's definition](#):

Two routes to decreasing susceptibility to infection are resistance (the ability to clear pathogens) and tolerance (the ability to limit damage in response to pathogens).

1 ref: Ayres JS, Schneider DS (2009) The Role of Anorexia in Resistance and Tolerance to Infections in Drosophila. PLOS Biology 7:e1000150. doi: 10.1371/journal.pbio.1000150

Makes no sense....clearing pathogens limits the damage...And clearing a pathogen involves sensing and then responding to pathogens.

The first, resistance, is the ability of the host to reduce pathogen levels. The second, tolerance, is the ability to limit the impact of infections.

This makes more sense.

Thoughts:

I think there are two distinct parts of stress:

1. Whether you encounter it at all. This should involve resistance because there can be ways for organisms to avoid or reduce the amount of stress
2. What do you do when you encounter stress? You can limit the damage caused by it.

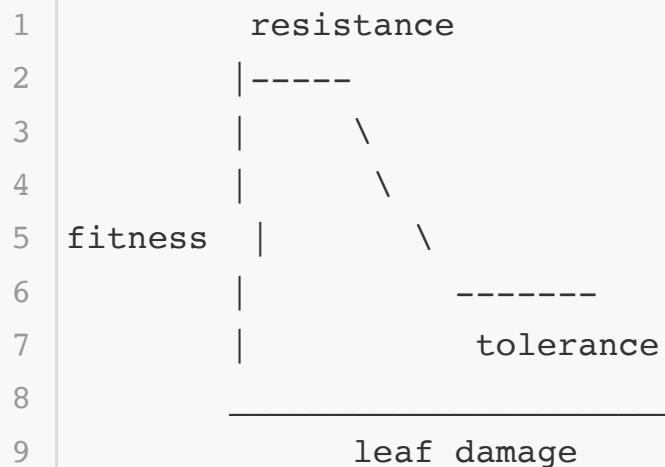
But there is a continuum right? Especially for temperature.

How would I define resistance or tolerance in terms of thermal stress biology? Or more generally?

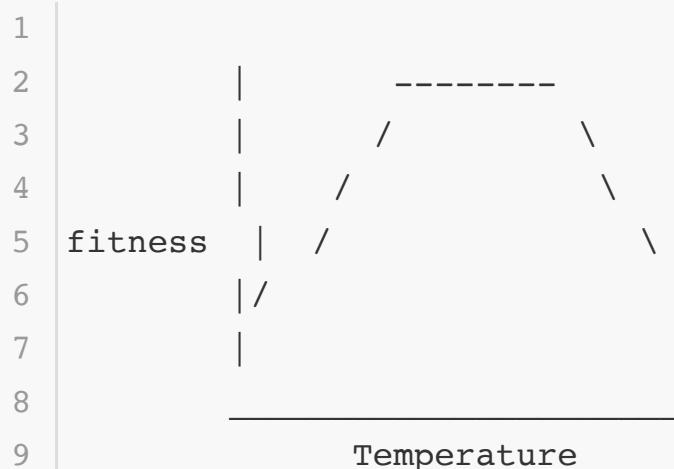
Resistance is a response (either induced or already present) that impedes disruption of a complex trait/biological system from a perturbation.

Tolerance is a response that mitigates the disruption of a complex trait/biological system from a perturbation.

Not sure about these definitions. In real biology, how would this work?



The amount of leaf damage that does not decrease fitness is resistance.
The part where the curve begins to change is tolerance. Can use a
breakpoint analysis.



The slope at the ends would indicate tolerance. But the width of topt
may be resistance.

Not sure this works.

Update: 2017-03-09

Another paper with a definition to think about:

Plant defence against any type of stress may involve resistance (traits that reduce damage) or tolerance (traits that reduce the negative fitness impacts of damage).

ref: Agrawal AA, Conner JK, Stinchcombe JR (2004) Evolution of plant resistance and tolerance to frost damage. *Ecology Letters*, **7**, 1199–1208.

Page 33: 2017-02-23. Multiple regression models testing effect of Hsp parameters on Ctmax

Doing a multiple regression as requested by SHC.

Could you do the straight-up regression? The PCA does a good job showing a single axis along which expression profiles vary, but the first most basic question is whether hsps are telling you the main story about CTmax.

Data structure:

```

1  'data.frame':   41 obs. of  81 variables:
2  $ n                  : int  1 2 3 4 5 6 7 8 9 10 ...
3  $ colony.id2         : Factor w/ 41 levels
4  "ALA1", "ALA4", ...
5  $ hsc70              : num  23 22.1 22.5 22.9 23.8 ...
6  $ hsp83              : num  22.4 22.1 22.3 22.4 21.6
7  ...
8  $ hsp40              : num  24.4 24 25.1 25.2 25.4 ...
9  $ FC_hsc70_1468_max : num  47.6 29.1 17 30.2 43.7 ...
10 $ FC_hsc70_1468_slope: num  1.026 1.071 1.01 0.391
11 0.909 ...
12 $ FC_hsc70_1468_Tm  : num  37.2 36.3 36.5 35.6 36.1
13 ...
14 $ FC_hsp40_541_max  : num  8.5 4.62 21.96 5.79 7.06
15 ...
16 $ FC_hsp40_541_slope: num  2.4311 1.3566 2.6054
17 0.8119 0.0455 ...
18 $ FC_hsp40_541_Tm   : num  37.3 35.3 40.9 35.3 33.1
19 ...
20 $ FC_Hsp83_279_max  : num  5.8 6.46 8.86 7.61 4.3 ...
21 $ FC_Hsp83_279_slope: num  2.0964 1.2968 1.757 0.0497
22 1.0885 ...
23 $ FC_Hsp83_279_Tm   : num  37 35 35.1 33.1 34.7 ...
24 $ Collection.date    : int  20141011 20141013 20150614
25 20150622 20140626 20140626 20140626 20140626 20140626
26 20140626 ...
27 $ site               : Factor w/ 16 levels "Alachua"
28 Co", "Avon", ...
29 1 1 2 3 4 4 4 4 4 4 ...
30 ...
31 ...
32 ...
33 ...
34 ...
35 ...
36 ...
37 ...
38 ...
39 ...
40 ...
41 ...

```

```

18 $ state           : Factor w/ 8 levels
  "FL", "MA", "MD", ...: 1 1 4 4 3 3 3 3 3 3 ...
19 $ queen_satus    : Factor w/ 4 levels
  "", "?", "QL", "QR": 4 3 1 1 4 4 4 4 4 4 ...
20 $ location        : Factor w/ 16 levels
  "ALA", "AVON", "BING", ...: 1 1 2 3 4 4 4 4 4 4 ...
21 $ lat             : num 29.7 29.7 44.8 45 39 ...
22 $ lon             : num -82.4 -82.4 -70.3 -69.9
  -77.2 ...
23 $ genus           : Factor w/ 1 level
  "Aphaenogaster": 1 1 1 1 1 1 1 1 1 1 ...
24 $ species          : Factor w/ 10 levels
  "?", "ashmeadi", ...: 7 7 8 8 4 9 9 4 9 4 ...
25 $ rad_seq_species : Factor w/ 8 levels
  "ashmeadi", "floridana", ...: 5 5 6 6 3 7 7 3 7 3 ...
26 $ colony.id       : Factor w/ 41 levels
  "Ala1", "Ala4", ...: 1 2 3 4 5 6 7 8 9 10 ...
27 $ Alt_names        : Factor w/ 41 levels
  "ALA1", "ALA4", ...: 1 2 3 4 5 6 7 8 9 10 ...
28 $ KO_temp_worker   : num 42 41.8 40.4 41.3 40.9 ...
29 $ mean_ind_weight_mg : num 0.664 0.509 NA NA 0.936
  ...
30 $ habitat          : Factor w/ 7 levels "deciduous
  forest", ...: 2 2 4 4 1 1 1 1 1 1 ...
31 $ habitat_v2        : Factor w/ 2 levels "deciduous
  forest", ...: 1 1 1 1 1 1 1 1 1 1 ...
32 $ nest.location     : Factor w/ 4 levels
  "log", "sand", "soil", ...: 3 3 1 3 1 1 1 4 4 4 ...
33 $ collectors         : Factor w/ 10 levels
  "ANBE", "ANBE, KM, Lloyd Davis", ...: 9 10 7 7 6 6 6 6 6 6
  ...
34 $ bio1              : int 204 204 47 53 125 125 125
  125 125 125 ...
35 $ bio2              : int 131 131 122 126 119 119
  119 119 119 119 ...

```

```
36 $ bio3 : int 47 47 29 29 33 33 33 33 33  
33 ...  
37 $ bio4 : int 5357 5357 9908 10071 8523  
8523 8523 8523 8523 8523 ...  
38 $ bio5 : int 329 329 249 259 304 304  
304 304 304 304 ...  
39 $ bio6 : int 55 55 -164 -164 -51 -51  
-51 -51 -51 -51 ...  
40 $ bio7 : int 274 274 413 423 355 355  
355 355 355 355 ...  
41 $ bio8 : int 268 268 1 7 212 212 212  
212 212 212 ...  
42 $ bio9 : int 173 173 -74 -70 12 12 12  
12 12 12 ...  
43 $ bio10 : int 268 268 170 178 233 233  
233 233 233 233 ...  
44 $ bio11 : int 132 132 -87 -84 12 12 12  
12 12 12 ...  
45 $ bio12 : int 1336 1336 1102 1052 1038  
1038 1038 1038 1038 1038 ...  
46 $ bio13 : int 202 202 113 106 102 102  
102 102 102 102 ...  
47 $ bio14 : int 56 56 69 64 70 70 70 70 70  
70 ...  
48 $ bio15 : int 41 41 12 13 12 12 12 12 12  
12 ...  
49 $ bio16 : int 547 547 305 291 292 292  
292 292 292 292 ...  
50 $ bio17 : int 189 189 231 211 221 221  
221 221 221 221 ...  
51 $ bio18 : int 547 547 296 286 291 291  
291 291 291 291 ...  
52 $ bio19 : int 277 277 243 229 221 221  
221 221 221 221 ...
```

53	\$ Axis.1	:	num	0.0182 0.0223 0.02 0.0241 -0.2567 ...
54	\$ Axis.2	:	num	0.0222 0.0285 0.0254 0.032 -0.0406 ...
55	\$ Axis.3	:	num	-0.00426 -0.00794 -0.00581 -0.00954 -0.03383 ...
56	\$ Axis.4	:	num	-2.54e-03 -2.59e-03 -6.49e-04 -5.45e-04 -6.64e-05 ...
57	\$ Axis.5	:	num	1.84e-02 2.04e-02 -4.40e- 02 -4.99e-02 2.88e-05 ...
58	\$ Axis.6	:	num	0.030309 0.030923 -0.002887 -0.00572 0.000598 ...
59	\$ Axis.7	:	num	-1.59e-02 -2.35e-02 6.21e- 05 -4.11e-03 7.34e-04 ...
60	\$ Axis.8	:	num	-4.06e-05 -6.81e-04 4.91e- 05 -3.62e-04 6.84e-05 ...
61	\$ Axis.9	:	num	2.55e-04 -2.88e-03 -2.03e- 05 -1.81e-03 1.54e-04 ...
62	\$ Axis.10	:	num	9.55e-03 -2.53e-02 2.81e- 04 1.53e-03 -1.75e-05 ...
63	\$ Axis.11	:	num	-5.18e-04 4.02e-03 1.38e- 03 2.43e-03 9.36e-05 ...
64	\$ Axis.12	:	num	4.51e-04 -5.99e-04 2.49e- 03 1.36e-02 -5.69e-06 ...
65	\$ Axis.13	:	num	-2.00e-04 1.58e-04 -3.37e- 05 -3.96e-05 -1.12e-06 ...
66	\$ Axis.14	:	num	3.96e-03 -2.98e-03 1.20e- 04 3.63e-06 -1.76e-05 ...
67	\$ Axis.15	:	num	1.55e-02 -5.38e-03 2.16e- 04 4.29e-04 -1.33e-06 ...
68	\$ Axis.16	:	num	6.07e-03 1.51e-02 -1.57e- 04 -1.69e-04 4.88e-06 ...
69	\$ Axis.17	:	num	-6.39e-04 4.58e-04 3.19e- 04 -2.48e-04 -2.03e-05 ...

```

70  $ Axis.18           : num  2.72e-03 -1.99e-03 -1.47e-
    03 1.22e-03 9.44e-05 ...
71  $ Axis.19           : num  0.001861 -0.000881
    -0.003553 0.006455 0.000168 ...
72  $ Axis.20           : num  1.30e-04 -4.88e-05 -5.60e-
    04 1.14e-03 1.80e-05 ...
73  $ Comp.1            : num  -3.711 -3.711 3.836 3.629
    0.736 ...
74  $ Comp.2            : num  -0.252 -0.252 0.376 0.955
    -0.848 ...
75  $ rRNA18s           : num  10.87 9.64 9.82 10.18 9.86
    ...
76  $ n.y               : int   1 2 4 5 7 8 9 10 11 12 ...
77  $ hsc70.1           : num  12.1 12.5 12.7 12.7 13.9
    ...
78  $ hsp83.1           : num  11.6 12.4 12.5 12.2 11.8
    ...
79  $ hsp40.1           : num  13.6 14.3 15.2 15 15.5 ...

```

Calculating basal expression. Subtracting ct values of hsp83,hsc70-4, and hsp40 from 18srRNA ct values. And then log2 transforming the data to get a measure of basal expression

```

1 | basalxp<-log2(jj[,3:5] - jj[,74])
2 | jj<-data.frame(jj,basalxp)

```

Constructing the multiple regression model:

```

1 regmod<-
  lm(KO_temp_worker~hsc70.1+hsp83.1+hsp40.2+FC_hsc70_1468_
max+FC_hsc70_1468_slope+FC_hsc70_1468_Tm+FC_hsp40_541_ma
x+FC_hsp40_541_slope+FC_hsp40_541_Tm+FC_Hsp83_279_max+FC
_Hsp83_279_slope+FC_Hsp83_279_Tm,data=jj)

```

Model selection based on AIC in both directions (forward and backwards):

```

1 summary(stepAIC(regmod,direction="both"))
2 #output
3 Call:
4 lm(formula = KO_temp_worker ~ hsp40.2 +
  FC_hsc70_1468_max + FC_hsp40_541_slope +
  FC_hsp40_541_Tm + FC_Hsp83_279_max +
  FC_Hsp83_279_slope,
  data = jj)
7
8 Residuals:
9   Min     1Q Median     3Q    Max
10 -0.92378 -0.28655  0.00972  0.25208  0.97126
11
12 Coefficients:
13                               Estimate Std. Error t value
14 (Intercept)            37.809870   2.777111 13.615 2.53e-
15 *** 
15 hsp40.2                -1.429154   0.673488 -2.122
16   0.04120 *
16 FC_hsc70_1468_max      0.042233   0.006797  6.213 4.56e-
17 07 *** 
17 FC_hsp40_541_slope -0.336009   0.127642 -2.632
18   0.01266 *

```

```

18 FC_hsp40_541_Tm      0.224272   0.049786   4.505 7.46e-
  05 ***
19 FC_Hsp83_279_max    -0.063139   0.022459   -2.811
  0.00813 **
20 FC_Hsp83_279_slope  0.251070   0.109254   2.298
  0.02784 *
21 ---
22 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
  0.1 ' ' 1
23
24 Residual standard error: 0.4913 on 34 degrees of
  freedom
25 Multiple R-squared:  0.69, Adjusted R-squared:  0.6353
26 F-statistic: 12.61 on 6 and 34 DF,  p-value: 1.926e-07

```

Model averaging with mumin package

```

1 selmod<-dredge(regmod)
2 selmod

```

- top 2 models

```

1 summary(model.avg(selmod[1:2]))
2 #output
3 (conditional average)
4                               Estimate Std. Error Adjusted SE z
5 value Pr(>|z|)
6 (Intercept)            36.527394   3.299388   3.371523
  10.834 < 2e-16 ***
7 FC_hsc70_1468_max    0.041943   0.006899   0.007150
  5.866 < 2e-16 ***
8 FC_hsp40_541_slope -0.329519   0.129646   0.134355
  2.453 0.01418 *

```

```

8  FC_hsp40_541_Tm      0.218131   0.051123   0.052921
  4.122 3.76e-05 ***
9  FC_Hsp83_279_max    -0.062754   0.022766   0.023598
  2.659  0.00783 **
10 FC_Hsp83_279_slope   0.252757   0.110742   0.114788
  2.202  0.02767 *
11 hsp40.2                 -1.429154   0.673488   0.698325
  2.047  0.04070 *
12 ---
13 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'.
  0.1 ' ' 1
14
15 Relative variable importance:
16
17 Importance:          FC_hsc70_1468_max
  FC_hsp40_541_slope
18 N containing models: 2
19
20 Importance:          FC_hsp40_541_Tm  FC_Hsp83_279_max
  FC_Hsp83_279_slope
21 N containing models: 2
22
23 Importance:          hsp40.2
24 N containing models: 1

```

- Top 10 models

```

1 summary(model.avg(selmod[1:10]))
2 #output

```

```

3 (conditional average)
4                               Estimate Std. Error Adjusted SE z
5 value Pr(>|z|)
6 (Intercept)            36.694206   3.447945   3.544068
7 10.354 < 2e-16 ***
8 FC_hsc70_1468_max    0.040852   0.007579   0.007834
9 5.215 2e-07 ***
10 FC_hsp40_541_slope -0.323397   0.136255   0.141024
11 2.293 0.021836 *
12 FC_hsp40_541_Tm     0.217875   0.055173   0.057078
13 3.817 0.000135 ***
14 FC_Hsp83_279_max    -0.059538   0.023868   0.024679
15 2.413 0.015842 *
16 FC_Hsp83_279_slope  0.252246   0.113812   0.117985
17 2.138 0.032522 *
18 hsp40.2              -1.446654   0.685537   0.711032
19 2.035 0.041893 *
20 FC_hsc70_1468_Tm    0.139557   0.128507   0.132953
21 hsc70.1              0.071127   0.087273   0.090593
22 0.785 0.432381
23 hsp83.1              0.056570   0.095056   0.098671
24 0.573 0.566426
25 FC_hsc70_1468_slope 0.103600   0.201892   0.209571
26 0.494 0.621065
27 FC_Hsp83_279_Tm     -0.030072   0.072190   0.074936
28 0.401 0.688195
29 FC_hsp40_541_max    -0.008143   0.019715   0.020465
30 0.398 0.690716
31 ---
32 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
33 0.1 ' ' 1
34
35 Relative variable importance:

```

```

22          FC_hsc70_1468_max
FC_hsp40_541_slope FC_hsp40_541_Tm FC_Hsp83_279_max
hsp40.2

23 Importance:      1.00      1.00      0.89
                 1.00      1.00      0.89

24 N containing models:   10      10      10
                 10      10      9

25          FC_Hsp83_279_slope
FC_hsc70_1468_Tm hsc70.1 hsp83.1 FC_hsc70_1468_slope

26 Importance:      0.87      0.15
                 0.09      0.07      0.07

27 N containing models:   8      2
                 1      1      1

28          FC_Hsp83_279_Tm FC_hsp40_541_max
29 Importance:      0.06      0.06
30 N containing models:   1      1

```

- Delta 4 AICc criterion

```

1 summary(model.avg(selmod, subset = delta < 4))
2
3 (conditional average)
4                               Estimate Std. Error Adjusted SE z
5 value Pr(>|z|)
6 (Intercept)            36.694206  3.447945  3.544068
7 10.354 < 2e-16 ***
8 FC_hsc70_1468_max    0.040852  0.007579  0.007834
9 5.215 2e-07 ***
10 FC_hsp40_541_slope -0.323397  0.136255  0.141024
11 2.293 0.021836 *
12 FC_hsp40_541_Tm     0.217875  0.055173  0.057078
13 3.817 0.000135 ***

```

```

 9 FC_Hsp83_279_max      -0.059538    0.023868    0.024679
 2.413 0.015842 *
10 FC_Hsp83_279_slope     0.252246    0.113812    0.117985
 2.138 0.032522 *
11 hsp40.2                  -1.446654    0.685537    0.711032
 2.035 0.041893 *
12 FC_hsc70_1468_Tm       0.139557    0.128507    0.132953
 1.050 0.293870
13 hsc70.1                  0.071127    0.087273    0.090593
 0.785 0.432381
14 hsp83.1                  0.056570    0.095056    0.098671
 0.573 0.566426
15 FC_hsc70_1468_slope     0.103600    0.201892    0.209571
 0.494 0.621065
16 FC_Hsp83_279_Tm       -0.030072    0.072190    0.074936
 0.401 0.688195
17 FC_hsp40_541_max       -0.008143    0.019715    0.020465
 0.398 0.690716
18 ---
19 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
 0.1 ' ' 1
20
21 Relative variable importance:
22
23          FC_hsc70_1468_max
24          FC_hsp40_541_slope  FC_hsp40_541_Tm  FC_Hsp83_279_max
25          hsp40.2
26          Importance:        1.00           1.00
27          1.00           1.00           0.89
28          N containing models: 10            10
29          10            10            9
30          FC_Hsp83_279_slope
31          FC_hsc70_1468_Tm  hsc70.1   hsp83.1  FC_hsc70_1468_slope
32          Importance:        0.87           0.15
33          0.09           0.07           0.07

```

```

27 N containing models:     8                      2
      1           1
28                               FC_Hsp83_279_Tm  FC_hsp40_541_max
29 Importance:                 0.06                  0.06
30 N containing models:       1                      1

```

Correlation between predictors:

```

1 knitr:::kable(round(cor(data.frame(jj[, 6:14], basalxp)), 3)
)
```

	FC_hsc70_1468_max	FC_hsc70_1468_slope	FC_hsc70_1468_Tm	FC_hsp40_541_max	FC_hsp40_541_slope	FC_hsp40_541_Tm
FC_hsc70_1468_max	1.000	0.522	0.599	0.425	0.289	0.291
FC_hsc70_1468_slope	0.522	1.000	0.507	0.331	0.378	0.365
FC_hsc70_1468_Tm	0.599	0.507	1.000	0.416	0.340	0.392
FC_hsp40_541_max	0.425	0.331	0.416	1.000	0.648	0.648
FC_hsp40_541_slope	0.289	0.378	0.340	0.648	1.000	0.648
FC_hsp40_541_Tm	0.295	0.368	0.617	0.665	0.609	0.609
FC_Hsp83_279_max	0.645	0.310	0.500	0.543	0.280	0.280
FC_Hsp83_279_slope	0.216	0.363	0.425	0.442	0.380	0.380
FC_Hsp83_279_Tm	0.453	0.195	0.783	0.396	0.223	0.223
hsc70	0.271	0.081	0.262	0.291	0.286	0.286
hsp83	0.165	0.126	0.119	0.261	0.312	0.312
hsp40	0.117	0.069	0.233	0.199	0.070	0.070

Another set of analyses:

Also, can you test the relationship between the PCs and CTmax without the open habitat species? This is to see if the same axis of variation also underlies CTmax variation without the big jump associated with the habitat shift.

Assembling dataset with pcs of hsp params with original data:

```
1 jj3<-data.frame(jj,hspparams)
2
3 summary(jj3$habitat_v2)
4 #deciduous forest      flat woods
5 #                      32          9
6
7 noop<-subset(jj3,jj3$habitat_v2!="flat woods")
8 #summary(lm(KO_temp_worker~pc1+pc2+pc3,data=noop))
9 summary(lm(KO_temp_worker~pc1,data=noop))
10
11 Coefficients:
12             Estimate Std. Error t value Pr(>|t|)
13 (Intercept) 41.36600   0.09696 426.630 < 2e-16 ***
14 pc1         -0.16944   0.06017  -2.816  0.00852 **
15 ---
16 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
17 0.1 ' ' 1
18 Residual standard error: 0.4823 on 30 degrees of freedom
19 Multiple R-squared:  0.2091,    Adjusted R-squared:
20 0.1827
21 F-statistic: 7.929 on 1 and 30 DF,  p-value: 0.008515
```

Rational for determining how many PCs to include in analysis

Found a source [here](#):

Because reduction of dimensionality is a goal of principal components analysis, several criteria have been proposed for determining how many PCs should be examined and how many should be ignored. One common criteria is to ignore principal components at the point at which the next PC offers little increase in the total explained variation. A second criteria is to include all those PCs up to a predetermined total percent explained variation, such as 90%. A third standard is to ignore components whose explained variation is less than 1 when a correlation matrix is used or less than the average variation explained when a covariance matrix is used, with the idea being that such a PC offers less than one variable's worth of information. A fourth standard is to ignore the last PCs whose explained variation is all roughly equal.

Page 34: 2017-02-23. SHC lab meeting: Going over hxp rxn norm paper.

Announcements

2017 lab refresher

ssh in into the gotelli lab server

Overall feedback:

- Add Katie Miller in the acknowledgements
-

Introduction

Results

Discussion

Methods

Page 35: 2017-02-23. What are Evolutionary Innovations?

Reading this paper:

Hunter JP. 1998. Key innovations and the ecology of macroevolution. Trends in Ecology & Evolution 13:31–36.

Different definitions:

Miller 1949:

key adjustments in the morphological and physiological mechanism which are essential to the origin of new major groups

I like this definition because it describes the emergence of traits with a diversification of a clade.

Van Valen 1981:

A key character, in the adaptive sense, is a structure or element of physiology that makes a taxon more or less committed to a way of life different from, or appreciably more efficient than, that of its ancestors.

I like how this definition highlights the traits in a new species do better than ancestral ones. So there is a time aspect.

Levinton 1988:

key innovation is necessary, but not sufficient for a subsequent radiation

Baum & Larson 1991:

a trait that greatly modifies the selective regime of the lineage in which it evolves

Don't think this is exactly right. Modifies does not have a direction of effect.

Rosenzweig & McCord 1991:

a key adaptation is a change in the mathematical rule governing a trade-off constraint so that after the change, the trade-off is less severe

I wouldn't call this a key innovation. Instead, I'd call it breaking a trade off.

Erwin 1992:

Key Innovations characterize particular clades and are both necessary and sufficient to explain diversification within the clade

What the...

Heard and Hauser 1995:

An evolutionary change in individual traits that causally linked to an increased diversification rate in the resulting clade (for which it is a synapomorphy)

But this definition excludes new species(forms) that have been able to persist in new environments, but have low diversification rates.

This paper also proposes 3 mechanisms:

1. allowing escape from competition via invasion into a new adaptive zone
2. decreasing probability of extinction by increasing population density via increased individual fitness
3. favoring reproduction or ecological specialization

I think point 3 captures point 1 and 2 simultaneously.

My own words:

Evolutionary Innovations:

1. Critical adaptations that facilitate the emergence of new species.
 2. The evolution of critical traits that have enabled species to diversify
 3. Adaptations that have enabled species to diversify
 4. Evolutionary change in a characteristic that enhances persistence in a new species compared to ancestors
-

Page 36: 2017-02-24. Last biolunch and committee meeting notes.

Last biolunch

Nate:

- Ants are still here...problem with yesterday's adaptation is today's constraint
 - go through argument: yesterday's adaptation is today's constraint—generalizable to the whole clade? ; cross susceptibility,
- Ants should be resilient

Aimee:

- Range limits question; can't they move? Then, they'd be resilient.

Melanie Lloyd:

- I should change fold induction as delta delta CT

Brent's comments:

- semantics; evolutionary innovations makes it sound like something is novel, but it is tweaking an already established molecular mechanism.
 - careful if you're using evo inn for a paper
- fitting lines to HSR; what I'm looking at is the kinetics of the heat shock response. How does that change across evo time and diff time

- describing HSR; summarize it a little quicker(i had 3 slides) it could be more succinct so you can get into your experiments.
- motivation for importance: showing some data that actually substantiates that Hsps are important for thermal tolerance. (good example in drosophila). Hsps are not just a biomarker of stress. Higher hsps = confers greater heat tolerance. Link Hsp with upper thermal limits better.
- Setting up predictions for protective vs tolerance: show previously published data showing each protection /tolerance....but tell everybody that people dont systematically look at Hsps and its role in upper thermal limits. Protection (George's papers; Dong Millers in limpets 2008 or 2009 ish).
- Presenting Hsps gene phylogenies: make it clear which paralogues and orthologues correspond with each species
- 1st chapter is less well rounded; I went too fast.
- Why use thermal performance curves? Take it out?
- At the very end; safety margins but I never plot safety margins. I have my own data, plot it out! And describe it
- "I don't believe it". Small heat shock proteins , ants have small hsps. A matter of miscommunication

NJG:

- Change his picture
- Too many "um"s
- I can start with climate change and ants have adaptations to buffer its effects.

Melissa's comments:

- Ants are everywhere and can adapt to anything; but it doesn't show finer scale phylogenetic relationships on that
- There is constraint based on evo history and that may limit future adaptive potential
- Have broad map; and then have differences in species ranges and some species don't have a large range; some are secluded.
 - show map of ant species distributions (show number); they're not all everywhere
- set audience up for conclusion; by showing in differences in how species inhabit the world ; there is potential for specialization and diversification
- Spent a long time on Hsps and HSEs without us knowing where it is going. It is nice to have outline for 2 stories. Suggest brief outline slide
- Tighten up what I say about HSE and Hsps. Make it shorter.
- General comment: Innovations. Not the right word. Morphological trait based thing.
- "I'd like to tackle" in my dissertation: Say , the question that "I have addressed"
- Don't call them chapters: talk for 3 questions. "Call it my research" or refer to it that way.
- for multiple stressors part: there is a biotic and abiotic effect on species
- in multiple stressors part: say sequentially so that people know.
- When people interrupt me: Say "Great" instead of, where was I?
- Conclusion: These results suggest that if species didn't adapt in a warm environment they may face extinction.

- Set up why this idea of yesterday's adaptation is today's constraint is a novel conclusion
- Make habitat pictures up front. Make pictures the same size.
- describe the word passive in slide 63; be more explicit in what the Tm shift means
- Show the people what journey they're about to go on.
 - Tell them what you're about to tell them
 - Tell them
 - Tell them what you told them
-

Committee meeting:

- SHC wants me to add proteome stability work if I get NSF.
 - Send out multiple stressors manuscript
 - Date set as March 29th
-

Page 37: 2017-02-27. MapPies in R for Zamira

[mapPies syntax](#)

Libraries:

```

1 #for data parsing
2 library(dplyr)
3 library(tidyr)
4 library(reshape2)
5 library(plyr)
6 #phylogenetics packages
7 library(ape)
8 library(geiger)
9 library(ade4)
10 library(adephylo)
11 library(phytools)
12 library(MPSEM)
13 #for grabbing data from worldclim
14 library(raster)
15 libaray.maps)
16 library(sp)
17 library(dismo)
18
19 #Statistics
20 library(vegan) # variance partitioning function
21 library(MASS)
22 library(MuMIn)
23 options(na.action = "na.fail")
24 #Dealing with matrices
25 #library(spaa)

```

```

1 w <- getData('worldclim', var='bio', res=2.5)
2 dbiol1 <- extract(w, ant_dat[,c("lon","lat")])
3 ant_dat_clim <- cbind(ant_dat, dbiol1[,1:19])
4 str(ant_dat_clim) #data structure
5 >'data.frame': 115 obs. of 39 variables:

```

```
6 $ n : int 1 2 3 4 5 6 7 8 9 10 ...
7 $ Collection.date : int 20140520 20140520 20140520
20140616 20140616 20140616 20140616 20140616 20140616
20140617 ...
8 $ site : Factor w/ 45 levels "Alachua
Co", "Avon", ... : 36 36 36 22 22 22 22 22 22 25 ...
9 $ state : Factor w/ 17 levels
 "", "AZ", "CA", "FL", ... : 12 12 12 4 4 4 4 4 4 4 ...
10 $ queen_status : Factor w/ 4 levels
 "", "?", "QL", "QR": 4 4 4 3 3 4 3 4 4 4 ...
11 $ location : Factor w/ 42 levels
 "ALA", "AVON", "BF", ... : 34 34 34 20 20 20 20 20 20 23 ...
12 $ lat : num 40.4 40.4 40.4 29.8 29.8
...
13 $ lon : num -75.2 -75.2 -75.2 -82 -82
...
14 $ genus : Factor w/ 7 levels
 "Aphaenogaster", ... : 1 1 1 1 1 1 1 1 1 1 ...
15 $ species : Factor w/ 16 levels
 "", "?", "ashmeadi", ... : 14 14 15 7 7 7 3 3 3 7 ...
16 $ rad_seq_species : Factor w/ 15 levels
 "", "ashmeadi", ... : 14 14 14 7 7 7 2 2 2 7 ...
17 $ colony.id : Factor w/ 115 levels
 "Ala1", "Ala2", ... : 101 102 103 57 58 59 60 61 62 86 ...
18 $ colony.id2 : Factor w/ 115 levels "09-A", "10-
A", ... : 92 93 94 58 59 60 61 62 63 77 ...
19 $ Alt_names : Factor w/ 115 levels "09-A", "10-
A", ... : 103 104 105 61 62 63 64 65 66 90 ...
20 $ KO_temp_worker : num 42 41.9 41.6 42.5 42.4 ...
21 $ mean_ind_weight_mg: num 1.369 0.969 0.784 1.289
1.206 ...
22 $ habitat : Factor w/ 10 levels
 "", "city", "deciduous forest", ... : 3 3 3 9 9 9 8 8 8 6
...
```

```
23 $ habitat_v2           : Factor w/ 4 levels
  "city", "deciduous forest", ...: 2 2 2 4 4 4 4 4 4 4 ...
24 $ nest.location        : Factor w/ 5 levels
  "", "log", "sand", ...: 2 2 2 3 3 3 3 3 3 3 ...
25 $ collectors           : Factor w/ 18 levels
  "ANBE", "ANBE, KM, Lloyd Davis", ...: 1 1 1 2 2 2 2 2 2 2
...
26 $ bio1                  : num 108 108 108 202 202 202 202
  202 202 205 ...
27 $ bio2                  : num 114 114 114 125 125 125 125
  125 125 127 ...
28 $ bio3                  : num 31 31 31 46 46 46 46 46 46
  47 ...
29 $ bio4                  : num 8758 8758 8758 5361 5361
...
30 $ bio5                  : num 293 293 293 326 326 326 326
  326 326 327 ...
31 $ bio6                  : num -67 -67 -67 56 56 56 56 56
  56 61 ...
32 $ bio7                  : num 360 360 360 270 270 270 270
  270 270 266 ...
33 $ bio8                  : num 197 197 197 266 266 266 266
  266 266 268 ...
34 $ bio9                  : num -8 -8 -8 171 171 171 171
  171 171 174 ...
35 $ bio10                 : num 219 219 219 266 266 266 266
  266 266 268 ...
36 $ bio11                 : num -8 -8 -8 129 129 129 129
  129 129 134 ...
37 $ bio12                 : num 1162 1162 1162 1319 1319
...
38 $ bio13                 : num 116 116 116 196 196 196 196
  196 196 191 ...
39 $ bio14                 : num 74 74 74 58 58 58 58
  57 ...
```

```
40 $ bio15 : num 12 12 12 42 42 42 42 42 42 42
41 ...
42 $ bio16 : num 327 327 327 533 533 533 533 533
533 533 537 ...
43 $ bio17 : num 250 250 250 197 197 197 197
197 197 195 ...
44 $ bio18 : num 323 323 323 533 533 533 533 533
533 533 537 ...
45 dim(ant_dat_clim) # dimensions
46 >[1] 115 39
47 #grabbing summary of sampling
48 samples<-
  as.data.frame(summary(ant_dat_clim$rad_seq_species),row
  names=T);samples$species<-rownames(samples)
49 names(samples)<-c("number","species")
50 knitr::kable(samples)
51
```

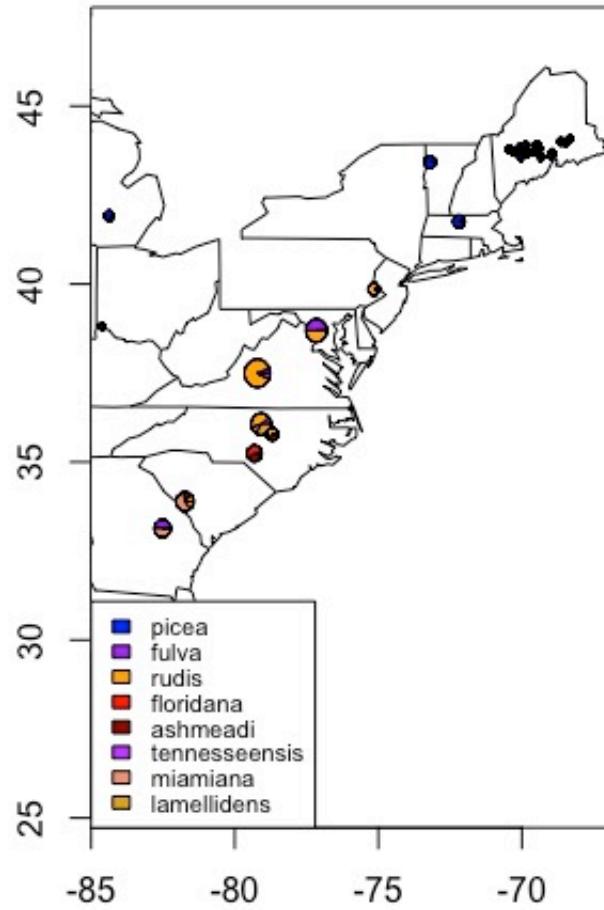
	number	species
	2	
ashmeadi	9	ashmeadi
barbatus	3	barbatus
caespeturn	1	caespeturn
Camponotus	1	Camponotus
Crematogaster	1	Crematogaster
floridana	9	floridana
Formica	1	Formica
fulva	10	fulva
lamellidens	4	lamellidens
miamiana	13	miamiana
pergandei	4	pergandei
picea	26	picea
rudis	29	rudis
tennesseensis	2	tennesseensis

MapPies

```
1 dpar<-ddply(ant_dat_clim,.
  (rad_seq_species,site),summarize,total=length(rad_seq_s
  pecies))
```

```
2
3 widepar<-dcast(dpar,site~rad_seq_species)[,-2]
4 gh<-ant_dat_clim[ !duplicated(ant_dat_clim$site), ]
5 gh1<-gh[order(gh$site), ]
6
7 widepar$lat<-gh1$lat
8 widepar$lon<-gh1$lon
9 widepar[is.na(widepar)]<-0
10 head(widepar)
11 widepar<-widepar[,-3:-6]
12 widepar<-widepar[,-4]
13 widepar<-widepar[,-7]
14
15 library(rworldmap)
16 plot(w, 5, xlim=c(-85,-67),
17 ylim=c(25,47.5),col="white", axes=TRUE,
18 legend=F,main="",box=FALSE)
19 map("state", c('florida', 'south carolina', 'north
carolina', 'georgia', 'virginia', 'west virginia',
'maryland', 'delaware', 'new jersey', 'rhode island',
'new york', 'connecticut', 'massachusetts',
'pennsylvania', 'vermont', 'new hampshire', 'maine',
'alabama', 'tennessee', 'kentucky',
'ohio','iowa','illinois','arkansas','missouri','minneso
ta','wisconsin','michigan','louisiana','mississippi',"t
exas","arizona","illinois","california","oregon","utah"
,"washington","kansas","new
mexico","montana","idaho","wyoming","north
dakota","south dakota","nebraska","oklahoma"), add =
TRUE)
```

```
19 mapPies(widepar,nameX="lon",nameY="lat",nameZs=c("picea  
","fulva","rudis","floridana","ashmeadi","tennesseensis  
","miamiana","lamellidens"),zColours=c("blue","purple",  
"orange","red","darkred","darkorchid1","darksalmon","go  
ldenrod"),xlim=c(-85,-67),  
ylim=c(25,47.5),landCol="gray50",addCatLegend =  
TRUE,add=TRUE,font=3)  
20 #mapBars(widepar,nameX="lon",nameY="lat",nameZs=c("picea  
","fulva","rudis","floridana","ashmeadi","tennesseensis  
","miamiana","lamellidens"),zColours=c("blue","purple",  
"orange","red","darkred","darkorchid1","darksalmon","g  
oldenrod"),xlim=c(-85,-67),  
ylim=c(25,47.5),landCol="gray50",addCatLegend =  
TRUE,add=TRUE)
```



Page 38: 2017-02-28. Questions to think about from SHC

In regards to a new paper in JEB finding that cataglyphis ants have higher constitutive protection than induced response.

SHC

A good opportunity for you to think about why you found the results that you did – desert Cataglyphis show a shift from induced to constitutive protection, while the Aphaenogaster you surveyed do not. What might be going on?

me:

One possibility is genetic assimilation where an inducible response evolved into a constitutive one. In other words, evolving into an extremely hot environment requires constant protection from protein damage, while Aphaenogaster turns on Hsps when they encounter protein damage.

SHC

That's a mechanistic process, not an explanation for *why* Cataglyphis would make this change while Aphaenogaster would not.

me:

Disagree. I explained the why. Maybe you're looking for this: Temperature variation shapes the plasticity in the heat shock response. If Aphaenogaster live in a more variable environment, then the HSR should be more plastic, while Cataglyphis lives in constant, extreme heat has a less plastic HSR and is constantly on to cope with heat stress.

SHC:

Yes, that is in fact better – your first answer was genetic assimilation, with the why partially implied in a follow-up sentence. But genetic assimilation is a how answer, not a why. Your PNAS discussion paragraphs will need to have the topics of discussion clearly separated out, not combined into one. Do you think you might have seen constitutively higher levels of Hsps if you had included the western members of the genus, or members of the outgroup? What is the value of looking at the set of species you've included instead of (or in addition to) looking at species inhabiting environmental extremes?

Page 39: 2017-02-28. More analyses for hsp rxn norm paper

From SHC:

Another question for you regarding the PNAS analyses. I am wondering whether the significant effect of "phylogeny" you are picking up is really just a significant effect of species (ie conspecifics are really genetically similar, much more so than species are to one another, and really phenotypically similar and live in really similar environments). I know the independent contrasts was pretty sparse, but perhaps you could color the points on your scatterplot

by species and add best-fit lines to show 1) whether the species differ, and 2) whether there are relationships between environment and CTmax within species. Just to have a look at it.

Constructed a model to test the interaction of Tmax and species id(factor), Tmax and habitat type:

```
1 umod2<-
  lm(KO_temp_worker~bio5*habitat_v2+bio5*rad_seq_species,
  data=Aph.dat)
2 summary(stepAIC(umod2,direction="both"))
3 >Coefficients:
4
5 (Intercept)           42.80833   0.26831
6 rad_seq_speciesfloridana -0.03981   0.37945
7 rad_seq_speciesfulva   -1.79611   0.37945
8 rad_seq_specieslamellidens -0.71458   0.48371
9 rad_seq_speciesmiamiana -1.85706   0.34904
10 rad_seq_speciespicea   -2.30737   0.31131
11 rad_seq_speciesrudis   -1.47025   0.30843
12 rad_seq_speciestennesseensis -2.05833   0.62925
13 ---
14 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
```

```

15
16 Residual standard error: 0.8049 on 92 degrees of
   freedom
17 Multiple R-squared:  0.5096,    Adjusted R-squared:
   0.4723
18 F-statistic: 13.66 on 7 and 92 DF,  p-value: 5.403e-12

```

Results: It looks like ashmeadi is the reference and all species differ except lamellidens and floridana. I should just do an ANOVA and do a post hoc test to see which ones are diff.

```

1 summary(aov(KO_temp_worker~rad_seq_species,data=Aph.dat))
)
2
3             Df Sum Sq Mean Sq F value Pr(>F)
4 rad_seq_species  7 61.94   8.848  13.66 5.4e-12 ***
5 Residuals       92 59.61   0.648
6 ---
7 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
8   ' '
9   1

```

OK, let's see the pairwise comparisons:

```

1 TukeyHSD(aov(KO_temp_worker~rad_seq_species,data=Aph.dat))
2
3 Tukey multiple comparisons of means
4
5 Fit: aov(formula = KO_temp_worker ~ rad_seq_species,
6 data = Aph.dat)
7 $rad_seq_species

```

			diff	lwr
8	upr p adj			
9	<i>floridana</i> - <i>ashmeadi</i>		-0.03981481	-1.2168021
	1.13717250 1.0000000			
10	<i>picea</i> - <i>ashmeadi</i>		-2.30737179	-3.2729871
	-1.34175651 0.0000000			
11	<i>rudis</i> - <i>ashmeadi</i>		-1.47025119	-2.4269575
	-0.51354490 0.0001832			
12	<i>miamiana</i> - <i>ashmeadi</i>		-1.85705641	-2.9397273
	-0.77438555 0.0000196			
13	<i>lamellidens</i> - <i>ashmeadi</i>		-0.71458333	-2.2149536
	0.78578698 0.8174237			
14	<i>fulva</i> - <i>ashmeadi</i>		-1.79611111	-2.9730984
	-0.61912380 0.0002086			
15	<i>tennesseensis</i> - <i>ashmeadi</i>		-2.05833333	-4.0101460
	-0.10652069 0.0312332			
16	<i>picea</i> - <i>floridana</i>		-2.26755698	-3.2331723
	-1.30194170 0.0000000			
17	<i>rudis</i> - <i>floridana</i>		-1.43043638	-2.3871427
	-0.47373009 0.0003020			
18	<i>miamiana</i> - <i>floridana</i>		-1.81724160	-2.8999125
	-0.73457074 0.0000314			
19	<i>lamellidens</i> - <i>floridana</i>		-0.67476852	-2.1751388
	0.82560180 0.8573138			
20	<i>fulva</i> - <i>floridana</i>		-1.75629630	-2.9332836
	-0.57930899 0.0003129			
21	<i>tennesseensis</i> - <i>floridana</i>		-2.01851852	-3.9703312
	-0.06670587 0.0373589			
22	<i>rudis</i> - <i>picea</i>		0.83712060	0.1571201
	1.51712110 0.0057326			
23	<i>miamiana</i> - <i>picea</i>		0.45031538	-0.3977942
	1.29842501 0.7206279			
24	<i>lamellidens</i> - <i>picea</i>		1.59278846	0.2518094
	2.93376752 0.0088873			

25	fulva-picea	0.51126068	-0.4543546
	1.47687597 0.7234944		
26	tennesseensis-picea	0.24903846	-1.5830877
	2.08116465 0.9998835		
27	miamiana-rudis	-0.38680522	-1.2247575
	0.45114704 0.8401083		
28	lamellidens-rudis	0.75566786	-0.5789103
	2.09024602 0.6506387		
29	fulva-rudis	-0.32585992	-1.2825662
	0.63084637 0.9639642		
30	tennesseensis-rudis	-0.58808214	-2.4155286
	1.23936427 0.9736163		
31	lamellidens-miamiana	1.14247308	-0.2851081
	2.57005430 0.2161909		
32	fulva-miamiana	0.06094530	-1.0217256
	1.14361616 0.9999997		
33	tennesseensis-miamiana	-0.20127692	-2.0977077
	1.69515385 0.9999782		
34	fulva-lamellidens	-1.08152778	-2.5818981
	0.41884254 0.3406000		
35	tennesseensis-lamellidens	-1.34375000	-3.5060138
	0.81851376 0.5358888		
36	tennesseensis-fulva	-0.26222222	-2.2140349
	1.68959042 0.9998923		

Figures:

Points and linear model fits are colored by species

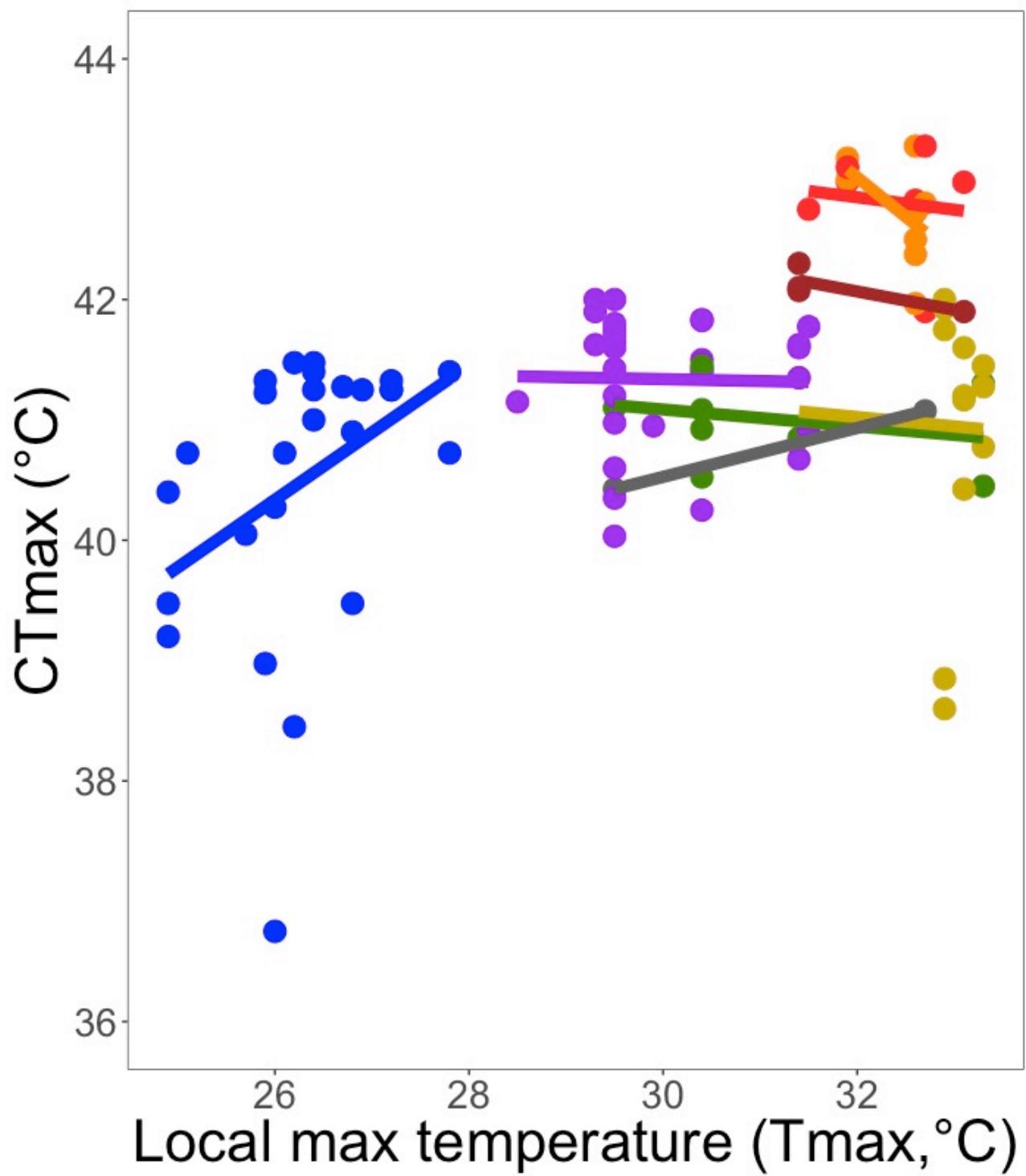
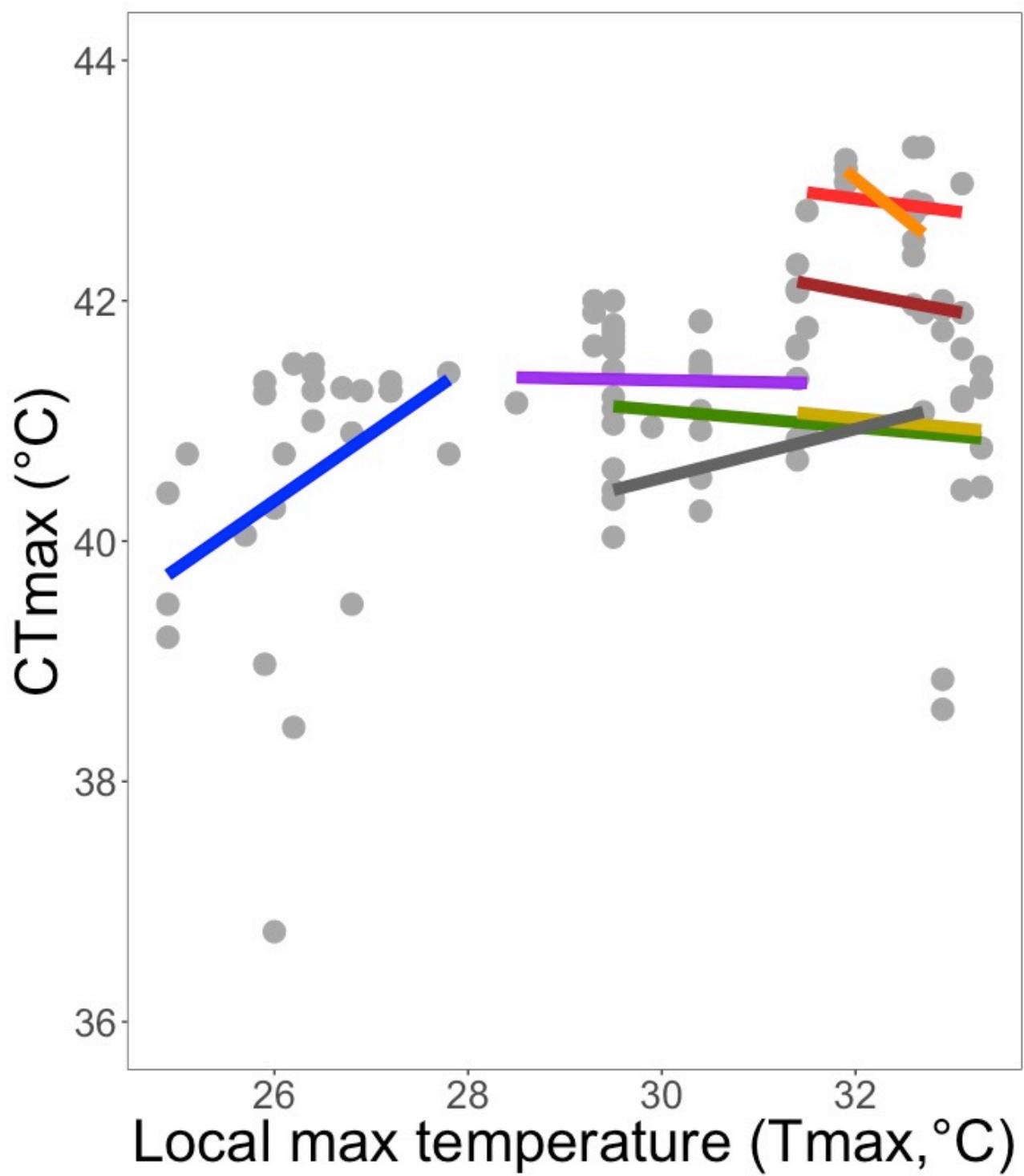
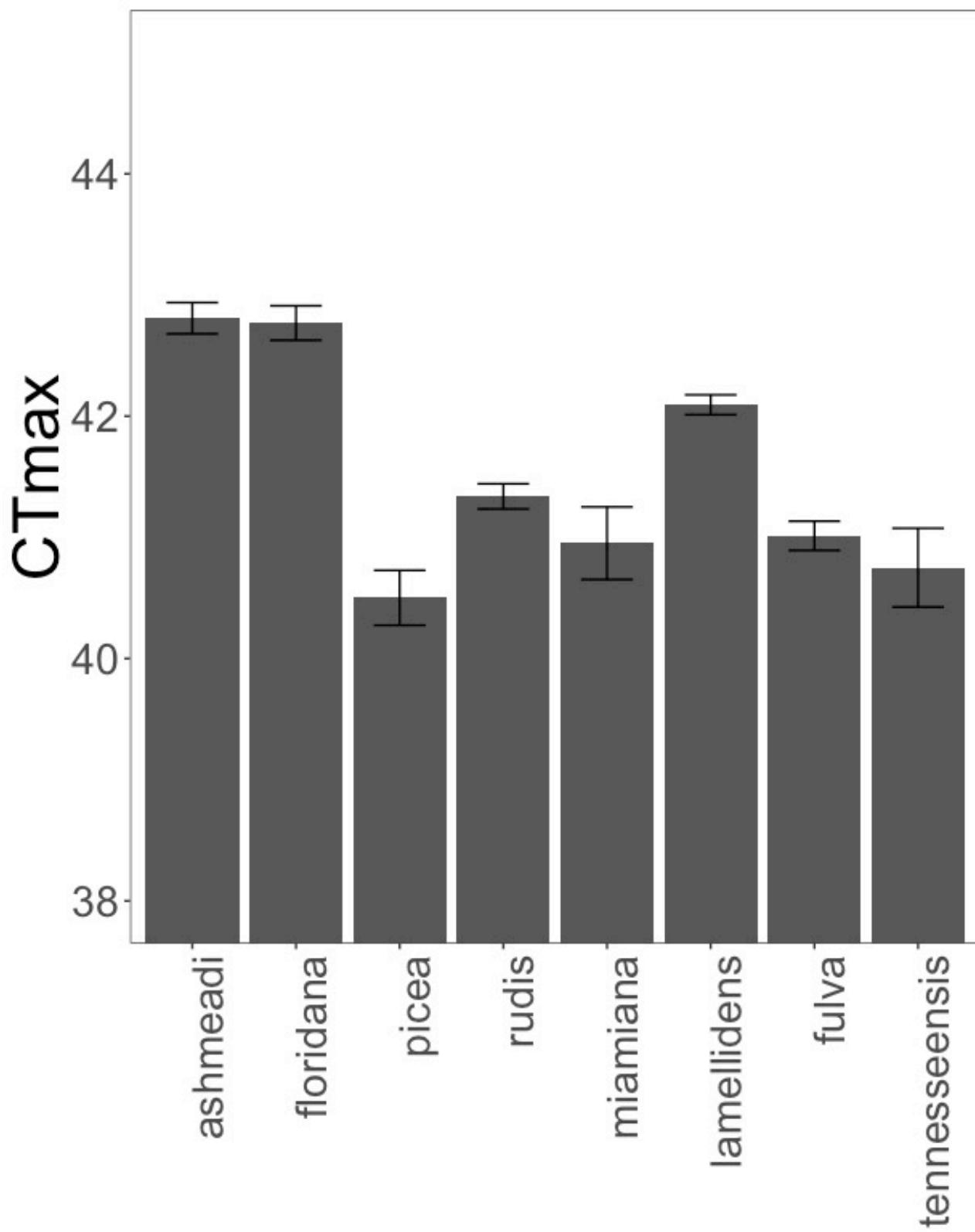


Figure with points colored gray but lines are by species



Barplot option:



SDiamond suggested I do include species as a random effect:

umod3 = additive model between tmax and habitat

umod4= interaction between tmax and habitat

Used anova to compare models

```
1 library(lmerTest) # gives p-values for lmer model
2 umod3<-lmer(KO_temp_worker~bio5+habitat_v2+
(1|rad_seq_species),data=Aph.dat)
3 umod4<-lmer(KO_temp_worker~bio5*habitat_v2+
(1|rad_seq_species),data=Aph.dat)
4
5 anova(umod3,umod4)
>refitting model(s) with ML (instead of REML)
7 Data: Aph.dat
8 Models:
9 object: KO_temp_worker ~ bio5 + habitat_v2 + (1 |
rad_seq_species)
10 ..1: KO_temp_worker ~ bio5 * habitat_v2 + (1 |
rad_seq_species)
11          Df      AIC      BIC logLik deviance Chisq Chi Df
Pr(>Chisq)
12 object  5 259.29 272.31 -124.64    249.28
13 ..1     6 259.85 275.48 -123.92    247.85 1.4396      1
               0.2302
```

Results: models dont differ so go with simpler one (umod3)

```
1 summary(umod3)
2 >Linear mixed model fit by REML t-tests use
Satterthwaite approximations to degrees
3 of freedom [lmerMod]
```

```

4 Formula: KO_temp_worker ~ bio5 + habitat_v2 + (1 |
5   rad_seq_species)
6
7 REML criterion at convergence: 260.2
8
9 Scaled residuals:
10    Min     1Q Median     3Q    Max
11 -4.5651 -0.3550  0.1815  0.5893  1.2790
12
13 Random effects:
14 Groups           Name        Variance Std.Dev.
15 rad_seq_species (Intercept) 0.1556    0.3944
16 Residual          0.6672    0.8168
17 Number of obs: 100, groups: rad_seq_species, 8
18
19 Fixed effects:
20             Estimate Std. Error      df t
21 value Pr(>|t| )
22 (Intercept) 36.942002  1.979357 7.224000
23      18.664 2.23e-07 ***
24 bio5          0.014058  0.006477 7.811000
25      2.171  0.0626 .
26 habitat_v2flat woods  1.050133  0.379272 9.094000
27      2.769  0.0216 *
28
29 ---
30 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
31      0.1 ' ' 1
32
33 Correlation of Fixed Effects:
34            (Intr) bio5
35 bio5       -0.995
36 hbtt_v2fltw  0.243 -0.284

```

Results: We find a significant effect of habitat type and Tmax is marginally significant. I also included a quadratic term which made no difference.

Page 40: 2017-03-04. Visit at UF; feedback on dissertation talk.

Reflecting on visit at UF.

Dissertation talk:

Clancy- part of habitat shifts I see may be due to fire disturbance in flat woods.

Leigh- "I don't believe that the CTmax shifts are related to temperature differences"

In the same vain...some PI named Rob that I met at "zoocial" suggested that I just use land cover. I think I'll use this:

<https://www.mrlc.gov/nlcd2011.php> (this is land cover)

[tree cover data](#)

I got a ton of questions on migration...what the...I didn't focus on this or research this at all. Hmm maybe it would be nice to include pop gen data? Is there viability selection on colony founding among different habitats?

Individual meetings:

Brett: Thought my data were interesting. Suggested to follow up with a biology letters paper by including range size for each species.

Dan Holt: Consider the effects of both abiotic and biotic interactions in setting range limits. You should read one of his old papers. Also, comparing regression tree with maxent models would be a small paper. He tries to publish in American Naturalist every year. It is his favorite journal too.

Bryony: Being more computational will help you.

Michael: BAMM still up in the air.

Christine Miller: Quant gen of sexual selection. males have fancy legs and compete for mates. Doing a phylogenetic approach- reconstructing ancestral states of host plants and morphological innovations.

Baer: mutation accumulation in *C. elegans*. Has done some networks on metabolites. Odd way to plot things, should try fitness on y and metabolite on x.

Getting tree canopy cover data

NLCD 2011 Percent Tree Canopy (Cartographic) [data download](#)

The National Land Cover Database 2011 (NLCD2011) USFS percent tree canopy product was produced through a cooperative project conducted by the Multi-Resolution Land Characteristics (MRLC) Consortium (www.mrlc.gov). The MRLC Consortium is a partnership of federal agencies, consisting of the U.S. Geological Survey, the National Oceanic and Atmospheric Administration, the U.S. Environmental Protection Agency, the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service, the U.S. Forest Service, the National Park Service, the U.S. Fish and Wildlife Service, the Bureau of Land Management, NASA, and the U.S. Army Corps of Engineers. One of the primary goals of the project was to generate current, consistent, and seamless national land cover, percent tree canopy, and percent impervious cover at medium spatial resolution. This product is the cartographic version of the NLCD2011 percent tree canopy cover dataset for CONUS at medium spatial resolution (30 m). It was produced by the USDA Forest Service Remote Sensing Applications Center (RSAC). Tree canopy values range from 0 to 100 percent. The analytic tree canopy layer was produced using a Random Forests? regression algorithm. The cartographic product is a filtered version of the regression algorithm output.

Sessioninfo

```
1 attached base packages:
2 [1] stats      graphics   grDevices utils      datasets
3 [6] methods    base
4
5 other attached packages:
6 [1] dismo_1.1-4  raster_2.5-8  sp_1.2-4
7
8 loaded via a namespace (and not attached):
9 [1] Rcpp_0.12.9    lattice_0.20-34  assertthat_0.1
```

```
10 [ 4] grid_3.3.1          plyr_1.8.4       gtable_0.2.0
11 [ 7] scales_0.4.1        ggplot2_2.2.1    lazyeval_0.2.0
12 [10] rpart_4.1-10        rgdal_1.2-5      tools_3.3.1
13 [13] munsell_0.4.3       yaml_2.1.14     colorspace_1.3-2
14 [16] knitr_1.15.1        tibble_1.2
```

```
1 x<-
  raster("nlcd_2011_USFS_tree_canopy_2011_edition_2014_03_
  31/cartographic_product/nlcd2011_usfs_treecanopy_cartogr
  aphic_3-31-2014.img")
2
3 plot(x)
4 b<-read.csv("20160517_ANBE_ant_sampling.csv",skip=6) #
  original dataset
5 b<-b[-87,] #getting rid of only NA
```



Linking coords with image projections

```
1  
2 coords <- b[, c("lon", "lat")]  
3  
4 names(coords) <- c("x", "y")  
5 coordinates(coords) <- ~x + y  
6 proj4string(coords) <- CRS("+proj=longlat +ellps=WGS84  
+datum=WGS84")  
7 crs_args <- x@crs@projargs  
8 sites_transformed <- spTransform(coords, CRS(crs_args))  
9
```

```
10
11  #extract land cover data for each point, given buffer
12  size
13 #Landcover <- extract(x, sites_transformed, buffer=30)
14 Landcover <- extract(x, sites_transformed, buffer=20)
15
16 Landcover
17 b$tree_canopy_20<-unlist(lapply(Landcover, `[[[` , 1)))
18
```

Page 41: 2017-03-07. To do.

1. Hsp rxn norm analyses: Is there a way to include intraspecific variation?
 - I should include species as a random effect in my regression models with CTmax and Hsp parameters
 - Try the jive model: [dryad](#) and [paper](#)
 - I could try [phylocurve](#) for my analyses of CTmax and Hsp rxn norms.

KGribble itinerary march 20th

format check for thesis

Thermal niche paper edits

Formatting dissertation:

Word did not find any entries for your table of contents.

In your document, select the words to include in the table of contents, and then on the Home tab, under Styles, click a heading style. Repeat for each heading that you want to include, and then insert the table of contents in your document. To manually create a table of contents, on the Document Elements tab, under Table of Contents, point to a style and then click the down arrow button. Click one of the styles under Manual Table of Contents, and then type the entries manually. The caption here should match that in the body of the document exactly.

Questions to ask and things to go over with Sean for thesis format check.

1. What do I need? Print out of thesis?
 2. Double check format for papers in review and already published one
 3. Go over defense notice
 4. Go over formatting for 3rd chapter, hsp rxn norms
 - 5.
-

Page 42: 2017-03-09. General reading notes:

ref: Huey RB, Kearney MR, Krockenberger A *et al.* (2012) Predicting organismal vulnerability to climate warming: roles of behaviour, physiology and adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 1665–1679.

Williams et al. [5] developed an integrative framework for assessing traits that promote vulnerability. They proposed that the vulnerability of a species depends on its sensitivity to environmental change, its exposure to that change, its resilience or ability to recover from perturbations and its potential to adapt to change.

A general framework to think about organismal response to climate warming:

1. thermal sensitivity
2. exposure
3. recovery or resiliency
4. adaptive potential

How do you measure these things?

1. thermal performance curves
2. location, properties of TPCs, behavioral reg, look at life stages
3. ?
4. adaptive potential, additive genetic variance

Looking for good ref for responding to climate change by acclimation, adaptation, and/or migration.

refs:

Moritz C, Agudo R (2013) The Future of Species Under Climate Change: Resilience or Decline? *Science*, **341**, 504–508.

Wiens JJ (2016) Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. *PLOS Biology*, **14**, e2001104.

Chen I-C, Hill JK, Ohlemüller R, Roy DB, Thomas CD (2011) Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science*, **333**, 1024–1026. (range shifts)

Papers on action potentials and CTmax

refs:

Miller NA, Stillman JH (2012) Neural Thermal Performance in Porcelain Crabs, Genus *Petrolisthes*. *Physiological and Biochemical Zoology*, **85**, 29–39.

What is K⁺ conductance doing?

For insects, it has been demonstrated that increasing temperature in *Locusta migratoria* increases the K conductance of neurons, resulting in nerve transmission failures; that similar failures can be induced by adding K to the extracellular fluid (simulating increased K conductance; Money et al. 2009); and that such failures can be prevented by pharmacologically blocking K channels (Wu et al. 2001).

Higher survival to heat stress is associated with decrease nerve membrane K⁺ conductance

Additionally, increased survival after sublethal heat stress has also been associated with a decrease in nerve membrane K conductance (Ramirez et al. 1999; Money et al. 2009).

Review of action potentials:

High permeability indicates that particle *mass* moves easily through a membrane. High conductance indicates that electrical *charge* moves easily through a membrane.

Conductance is the inverse of electrical resistance. If conductance is low, then the resistance of that ion to move across the membrane is high.

K⁺ more concentrated inside cells.

Page 43: 2017-03-09. Dissertation format check with Sean M.

1. fix citation page

- Citation page needs page number "ii"
- need periods between names

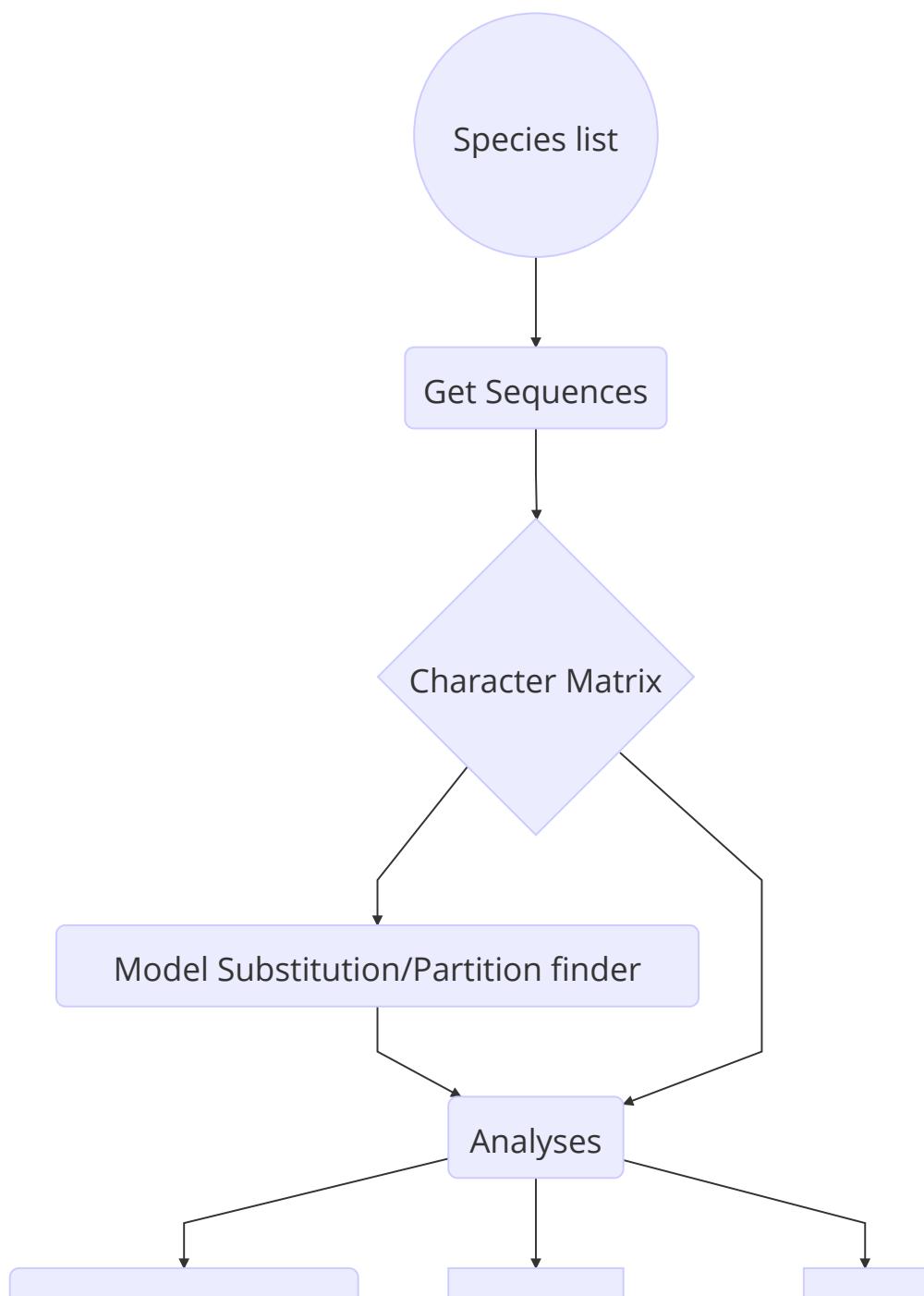
2. some of the pages looked messed up at the bottom, can lower page margin at the bottom of page
 3. Upon defense, sign defense form, scan and upload alongside thesis online.
 - be vigilante for emails because you'll get follow up emails to verify that the formatting is ok
-

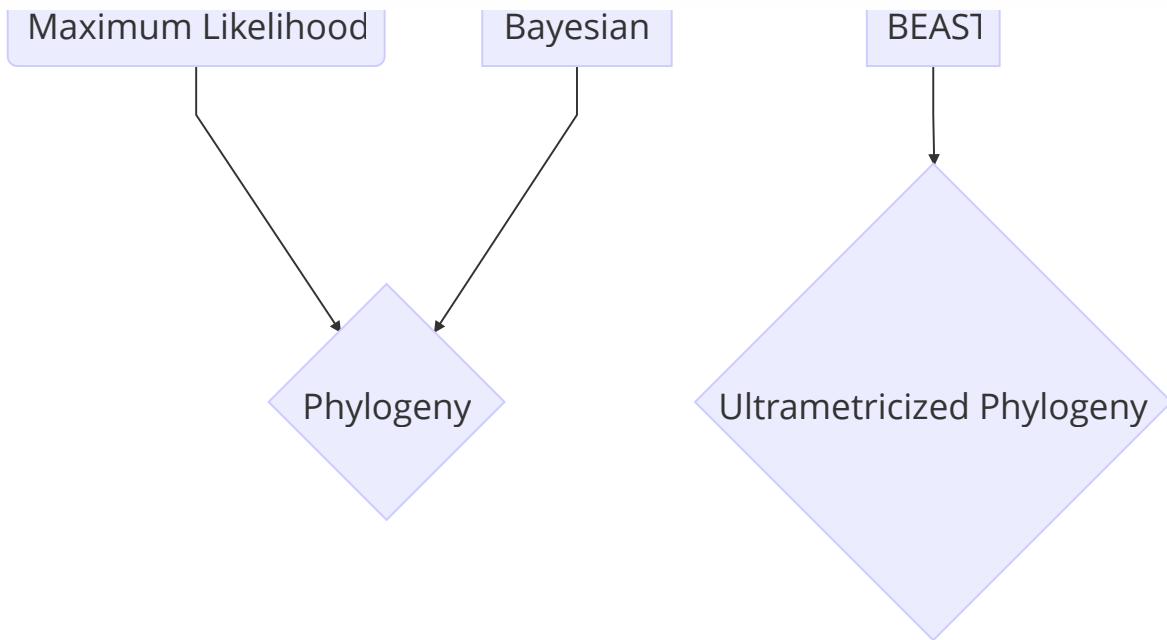
Page 44: 2017-03-09. SHC Lab agenda.

Date	Agenda
Jan. 26th	Organizational meeting
Feb. 2nd	Paper or Data
Feb. 9th	Andrew practice talk (actual talk on Feb. 24th)
Feb. 16th	Katie practice talk (actual talk on Mar. 3rd)
Feb. 23rd	Andrews PNAS feedback
Mar. 2nd	Lab field trip
Mar. 9th	Julias presentation on thesis research
Mar. 16th	Spring Break!!
Mar. 23rd	Caleb Research update
Mar. 29th	Andrews Dissertation Defense _ 3pm Jeffords 234
Mar. 30th	Julia feedback on honors thesis proposal writing review
April 6th	Katies first thesis chapter
April 13th	Paper or Data
April 20th	Student Research Conference practice poster/talks
April 27th	Paper or Data

Page 45: 2017-03-10. General phylogenetics workflow

Drew this up for Emily M. a while ago and just sticking it here for future reference.





Page 46: 2017-03-13: Meeting with SHC for dissertation timeline.

Does SHC need to read over every piece before I send out?

Make sure committee members are ok with me sending stuff out a week out of defense date.

I should send out my thesis 2 weeks ahead of time. Send out March 15th? 17th? 20th?

Layout:

- Abstract: (3/4 page)
 - SHC needs to read
- CH1: Introduction (3-6 pages)
 - 5 pages with references
 - SHC needs to read
- CH2: The evolution of heat shock protein sequences, cis-regulatory elements, and expression profiles in the eusocial Hymenoptera (published BMC Evolutionary Biology)
- Ch3: Molecular adaptations of protection and tolerance predict upper thermal limits in eastern forest ants (aiming to submit to PNAS)
 - go over title,
 - SHC needs to read the whole thing (abstract, intro, results, discussion, methods)
 - I still need to tweak the introduction and discussion
- CH4: Effects of desiccation and starvation on thermal tolerance and the heat shock response in forest ants (In 2nd round of reviews, Journal of Comparative Physiology B)
- CH5: Conclusions and future directions (3 pages)
 - ~ 1 page, have outline, need to fill out

Go over dissertation talk.

Are there any glaring things I need to fix?

Biggest worry: Phylogenetic, PNAS chapter. Conclusions felt short, stilted and pieces didn't stitch together.

Treat phylogenetic part of statistics ; using species as a random effect—not over-ascribing effects of phylogeny. Random effect- there is an independent effect of habitat and temperature. Change in stats might change how I told the story.

In the ms itself, for reference:

Colonies within different lineages differed in CTmax between forest habitats ($\beta=1.37 \pm 0.253$; $t=5.38$, $p <0.001$) and with the local thermal extremes (Tmax, $\beta=0.013 \pm 0.003$; $t=3.52$, $p <0.001$; Full model: $F_{2,97}=33.58$; $R^2=0.40$, $p<0.001$; Fig 2B).

Need to change to random effects model

Changed sentence to:

When including species as a random effect, colonies within different lineages differed in CTmax between forest habitats ($\beta=1.05 \pm 0.379$; $t=2.76$, $p <0.05$) and with the local thermal extremes (Tmax, $\beta=0.014 \pm 0.006$; $t=2.17$, $p <0.001$; Fig 2B).

Amanda suggestion: Broader implications? What types of changes we'd expect to see? turnover? certain populations within a species go extinct. Focus on turnover, different areas where ants are locally adapted to climate. Stick with a consistent flow. Specific—>meaning for ants on a large scale. What else do we need to understand? Expecting to hear a broad statement, but didn't hear it and came to a conclusion on it's

own.

My own notes:

- incorporate ch1 and 2 better... dynamic shifts at the gene copy level, and also how they are regulated.
 - Hsps are absolutely critical for coping with heat stress.
 - hsps can be adaptively tuned over relatively short to broad evolutionary time scales.
 - Evolution of tolerance may make sense because Hsps are costly, and always activating them slows growth, so they should be only turned on when stress is encountered.
 - To respond to climate change, species will need to turn on Hsps all the time?
-

Page 47: 2017-03-13. SHC analysis suggestion hsp rxn norm paper

SHC:

Can you try separating out the deciduous forest species and graph CTmax against PC1, and in a second graph make a box plot comparison of PC1 for deciduous versus pine forest) to show that variation along PC1 significantly explains both patterns of CTmax variation?

Subset, plot, and construct model:

```
1 decj<-subset(jj3,jj3$habitat_v2=="deciduous forest")
#subset
2 ggplot(decj,aes(x=pc1,y=KO_temp_worker))+geom_point(size=5,col="darkgray")+stat_smooth(method="lm",colour="black",size=2.5)+xlab("PC1 scores")+ylab("CTmax (°C)") #
figure
3 summary(lm(KO_temp_worker~pc1,data=decj)) # model
4
5 >Call:
6 lm(formula = KO_temp_worker ~ pc1, data = decj)
7
8 Residuals:
9      Min       1Q   Median       3Q      Max
10 -1.17476 -0.33930  0.08233  0.34954  0.80763
11
12 Coefficients:
13             Estimate Std. Error t value Pr(>|t|)
14 (Intercept) 41.36600    0.09696 426.630 < 2e-16 ***
15 pc1        -0.16944    0.06017  -2.816  0.00852 **
16 ---
17 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
18 0.1 ' ' 1
19 Residual standard error: 0.4823 on 30 degrees of freedom
20 Multiple R-squared:  0.2091,    Adjusted R-squared:
21          0.1827
22 F-statistic: 7.929 on 1 and 30 DF,  p-value: 0.008515
```

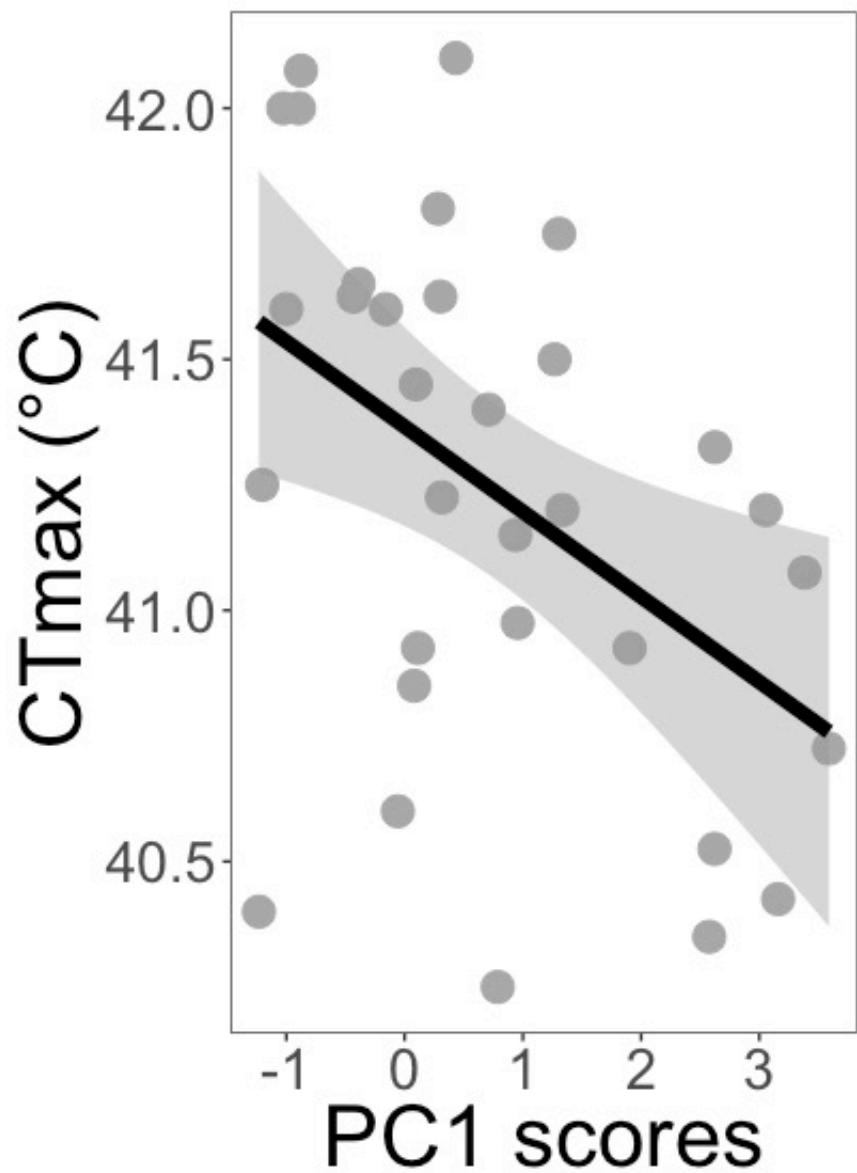


Figure with CT_{max} v pc1 for deciduous species only.

Looks funky, let's double check:

```

1 > decj$rad_seq_species
2 [1] miamiana      miamiana      picea       picea
3 [5] fulva        rудis        rудis       fulva
4 [9] rудis        fulva        rудis       rудis
5 [13] lamellidens fulva        lamellidens picea
6 [17] picea        мiamiana     rудis
tennesseensis
7 [21] picea        rудis        rудis       rудis
8 [25] rудis        rудis        rудis       rудис
9 [29] rудис       rудис       тennesseeensis rудис

```

Ok, no floridana and ashmeadi.

Species as a random effect (deciduous species)

```

1 summary(lmer(KO_temp_worker~pc1+
(1|rad_seq_species),data=decj))
2 Linear mixed model fit by REML t-tests use
Satterthwaite
3 approximations to degrees of freedom [lmerMod]
4 Formula: KO_temp_worker ~ pc1 + (1 | rad_seq_species)
5 Data: decj
6
7 REML criterion at convergence: 48.1
8

```

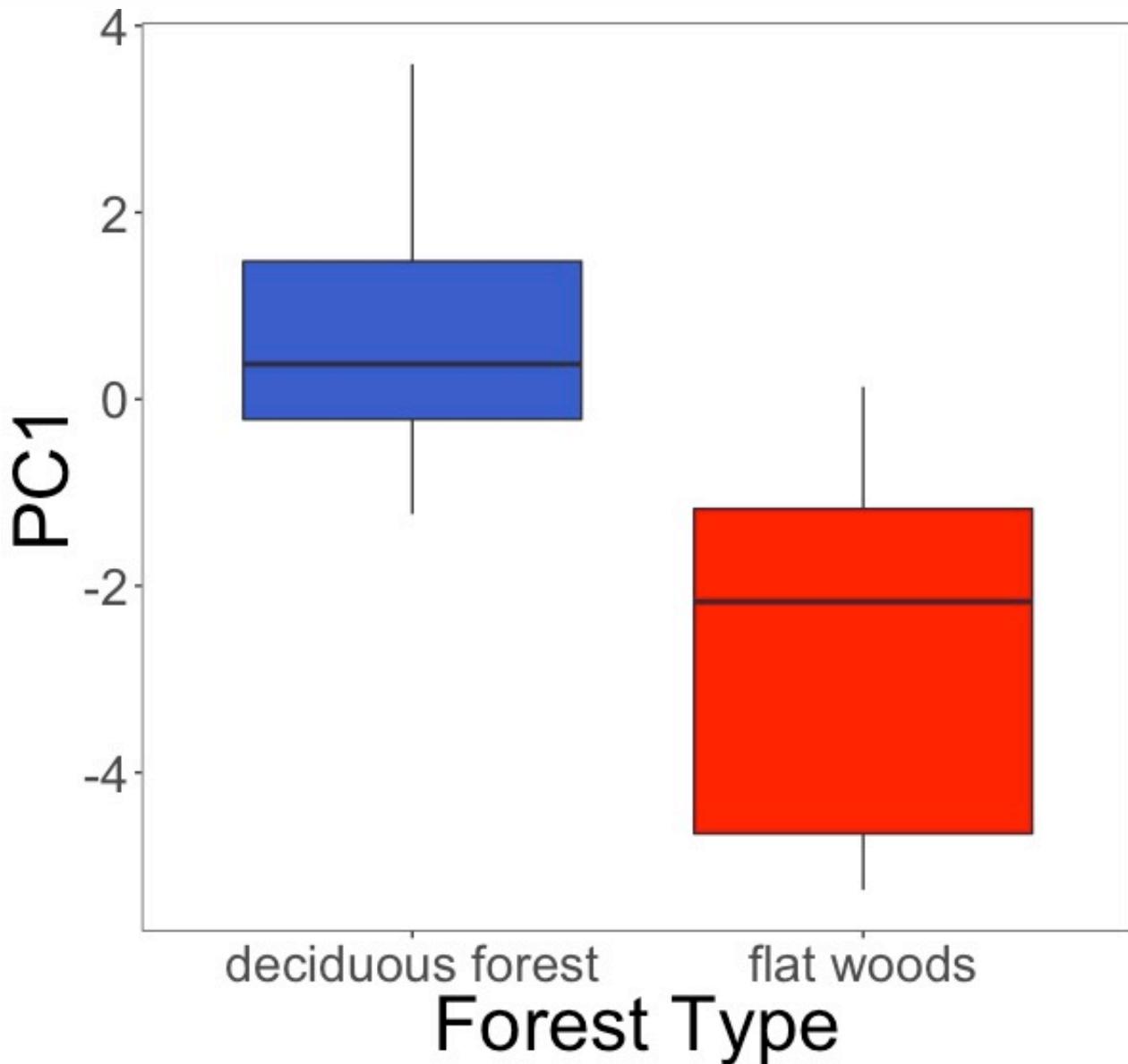
```

9 Scaled residuals:
10      Min       1Q   Median      3Q      Max
11 -2.3133 -0.5270  0.1981  0.5917  1.1821
12
13 Random effects:
14 Groups           Name        Variance Std.Dev.
15 rad_seq_species (Intercept) 0.06714  0.2591
16 Residual          0.19680  0.4436
17 Number of obs: 32, groups: rad_seq_species, 6
18
19 Fixed effects:
20             Estimate Std. Error      df t value
21 Pr(>|t|)
21 (Intercept) 41.41169    0.15293 3.72800 270.791 3.84e-
22 09 ***
22 pc1        -0.16222    0.05958 29.70000 -2.723
22 0.0107 *
23 ---
24 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
24 0.1 ' ' 1
25
26 Correlation of Fixed Effects:
27 (Intr)
28 pc1 -0.385

```

Second part of email: Ok I need to show from the original dataset, the difference in pc1 between forest types:

```
1 ggplot(jj3,aes(x=habitat_v2,y=pc1,fill=habitat_v2))+T+geom_boxplot()+xlab("Forest Type")+ylab("PC1")+scale_fill_manual(name = "", values = c("royalblue3","red"))
2
3 summary(aov(pc1~habitat_v2,data=jj3))
4
5 Df Sum Sq Mean Sq F value    Pr(>F)
6 habitat_v2   1   85.9   85.90   32.96 1.18e-06 ***
7 Residuals   39  101.6    2.61
8 ---
9 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
10   ' ' 1
```



Species as a random effect: testing diff in pc1 between habitats
(spmo1)

```
1 summary(lmer(pc1~habitat_v2+
  (1|rad_seq_species),data=jj3))
2 Linear mixed model fit by REML t-tests use
Satterthwaite
3 approximations to degrees of freedom [lmerMod]
4 Formula: pc1 ~ habitat_v2 + (1 | rad_seq_species)
5 Data: jj3
```

```

6
7 REML criterion at convergence: 153.7
8
9 Scaled residuals:
10      Min       1Q   Median       3Q      Max
11 -1.5600 -0.7065 -0.2086  0.8903  1.7499
12
13 Random effects:
14 Groups           Name        Variance Std.Dev.
15 rad_seq_species (Intercept) 0.06065  0.2463
16 Residual          2.56500  1.6016
17 Number of obs: 41, groups: rad_seq_species, 8
18
19 Fixed effects:
20             Estimate Std. Error      df t
21 value
21 (Intercept) 0.8154     0.3120 2.2230
22 habitat_v2flat woods -3.5297     0.6448 4.3620
23               -5.474
24             Pr(>|t|)
25 (Intercept) 0.1082
25 habitat_v2flat woods 0.0042 **
26 ---
27 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
28 0.1 ' ' 1
29 Correlation of Fixed Effects:
30            (Intr)
31 hbtt_v2fltw -0.484
32

```

Species as a random effect: testing effect of habitats and Tmax on pc1
 (spmo2)

```

1 summary(spmo2)
2 Linear mixed model fit by REML t-tests use
Satterthwaite
3 approximations to degrees of freedom [lmerMod]
4 Formula: pc1 ~ habitat_v2 + bio5 + (1 |
rad_seq_species)
5 Data: jj3
6
7 REML criterion at convergence: 160.3
8
9 Scaled residuals:
10 Min      1Q Median      3Q      Max
11 -1.5469 -0.6572 -0.1900  0.8795  1.6959
12
13 Random effects:
14 Groups           Name        Variance Std.Dev.
15 rad_seq_species (Intercept) 0.1626   0.4032
16 Residual          2.5805   1.6064
17 Number of obs: 41, groups: rad_seq_species, 8
18
19 Fixed effects:
20
21             Estimate Std. Error      df t
22 value
23 (Intercept) 0.9865287 4.3710242 6.3260000
24 0.226
25 habitat_v2flat woods -3.5514575 0.7964546 4.1000000
26 -4.459
27 bio5         -0.0004003 0.0145931 6.6500000
28 -0.027
29
30             Pr(>|t| )
31 (Intercept) 0.8286
32 habitat_v2flat woods 0.0106 *
33 bio5         0.9789
34
35

```

```

29 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'  

   0.1 ' ' 1  

30  

31 Correlation of Fixed Effects:  

32          (Intr) hbt_2w  

33 hbt_v2fltw  0.427  

34 bio5       -0.997 -0.464

```

Ok, let's compare models: 1 with habitat, 1 with habitat and tmax

```

1  

2  

3 anova(spmo1,spmo2)  

4 refitting model(s) with ML (instead of REML)  

5 Data: jj3  

6 Models:  

7 object: pc1 ~ habitat_v2 + (1 | rad_seq_species)  

8 ..1: pc1 ~ habitat_v2 + bio5 + (1 | rad_seq_species)  

9          Df      AIC      BIC    logLik deviance Chisq Chi Df  

10 object  4 161.57 168.42 -76.784     153.57  

11 ..1     5 163.57 172.14 -76.784     153.57      0      1  

12          Pr(>Chisq)  

13 object  

14 ..1      0.9955

```

Temperature range:

```

1 range(Aph.dat$bio5)/10  

2 [1] 24.9 33.3

```

Notes from SHC for the discussion:

- Same pattern of change associated with change in thermal stress within deciduous forest as with shift to more open habitat
 - common regulatory control such that response can be tuned up or down in a coordinated fashion in response to selection?
- Study of extreme thermophiles suggests a shift from induced to constitutive protection, but this may not be representative of thermal diversification in seasonal environments (or less extreme? Or non-desert?). Why not?
- Protection is usually thought of in terms of baseline levels, but important to remember that this can also be achieved through the induction response via graded or early-onset protection. Shift to a more graded response suggests a change in trigger of Hsp gene expression (from direct feedback from denatured proteins to some other temperature sensor – something like HSR1 non-coding RNA?)

Taking out a part of my intro:

Therefore, the patterns of Hsp gene expression may reveal evolutionary adaptations for coping with high temperature (Somero 2011). Evidence for the protection mechanism would be elevated levels of Hsp gene expression at relatively benign temperatures such that organisms are prepared for future lethal temperatures. For example, experimentally elevating Hsps at sublethal levels in fruit flies increased survival to extreme heat shock (Cavicchi et al. 1995). Evidence for the tolerance mechanism would be the increase of Hsp gene expression in concert with temperature increases. For

example, the onset of Hsp induction positively corresponds with thermal limits in marine snails species (Lars and Somero 2002).

This paragrpaah doesn't make that much sense but it has good citations:

Among studies, evidence supporting the protection and tolerance mechanism by is mixed. For example, in controlled heat-shock experiments warm-climate populations or species may express Hsps at lower (Dahlhoff and Rank 2000; Bedulina et al. 2013) or higher levels (Dietz and Somero 1992; Tomanek and Somero 2000). In some cases, warm- and -cool adapted populations or species lack adaptive variation in Hsp expression altogether (Madeira et al. 2015; Lockwood et al. 2010; Jensen et al. 2010), which might imply no role for Hsps in extending upper thermal limits. Moreover, Hsps participate in many other biological processes unrelated to heat (Sørensen et al. 2003). For example, Hsps can induce in response to oxidative, desiccation, and osmotic stress (Morris et al. 2013). However, detailed knockout studies definitively show that thermally responsive Hsps are required for mounting a heat shock response (Bettencourt et al. 2008). Perhaps adaptive tuning of Hsp gene induction is not possible with rapid temperature selection (Sikkink et al. 2014) or in quickly diversifying clades.

2017-03-14: Plopping my discussion here for reference. A lot of it has to be re-written.

The transition from closed canopy forests (deciduous) to open canopy pine/oaksavanna, supported by phylogenetic reconstruction and ancestral trait analyses(Fig), was accompanied by increased upper thermal limits (CTmax; Fig). The extensionof CTmax was associated with differences in the reaction norms of

Hsp expression that reflected both protection and tolerance mechanisms. In support of the protection mechanism, colonies with higher CTmax had higher expression at the low end of the temperature gradient simply by slowing expression rates for *hsp70* and *hsp40*. In support of the tolerance mechanism, colonies with higher CTmax expressed all three Hsps at a higher temperature (Tm) and also to a higher magnitude (maximum expression). Out of the parameters that significantly explained variation in CTmax, six out of twelve parameters diverged between forest type (Fig). The adaptive modification of Hsps altogether highlights complex and non-linear patterns of gene expression.

The structure of habitats is known to alter the thermal experiences of ectotherms (Brett et al. 2014). For temperate forest ants, local environmental temperatures and habitat type have shaped the variation in thermal traits (Fig). Deciduous forest species are significantly buffered from thermal stress than flat woods species due to higher canopy cover (Supplemental), especially over the growing season where foraging activity is highest. These species are known to utilize behavioral thermal regulation such as migrating to new nesting locations when they experience unsuitable environments and workers may maintain optimal body temperatures by avoiding solar radiation in the leaf litter, also known as the Bogert effect (Bogert 1945). Despite these behavioral activities, forest ants still experience selection from the thermal environment because there was still clinal variation in Ctmax (Fig 2). In contrast to deciduous forest species, flat woods species nest near the surface (13-92 cm in depth; Tschinkel 2011) and are not known to migrate whole at all (), which subjects them to thermal stress because surface soil can become superheated (Porter and Gates 1969; Angilletta 2009). If foragers navigate through superheated surface boundary layers, they must rely on physiological mechanisms

to withstand heat stress (). Thermal extremes are known to structure the diversity of upper thermal limits in ant communities even within the same location (Kaspari 2015). For example, in the tropics, ants possess greater CTmax in the canopy and experience more heat bouts than ants in the leaf litter (Kaspari 2015).

The phenotypic divergence of CTmax in forest ants were explained by 68% of the variation in parameters that represent the sigmoidal shape of Hsp expression. The parameters themselves are correlated show large amounts of independent axes of variation for which selection can operate upon and in fact, we find roughly five effective numbers of dimensions from a PCA analysis ($\lambda_{\text{total}}/\lambda_{\text{PC1}}$; supplemental). Perhaps high dimensionality is reflected by the fact that Hsps play other critical roles in the cell (Morris et al. 2013) and are highly pleiotropic. However, only the first PC, which best captured the correlation between slope, Tm, max, but not basal gene expression, related themost to CTmax. Each parameter of the Hsp reaction norm captures the non-linear aspect of Hsp gene expression where detection of adaptive variation or lack thereof depends on the sampled temperature. For example, when sampled at a single temperature slice along the temperature gradient, differences in Hsp reaction norms between open and closed forest habitats showed lower, higher, or no difference at all in Hsp expression. Therefore, sampling dictates where one finds adaptive variation of a trait without a function-valued approach. This is one of the first studies to our knowledge, to empirically capture the non-linear process of Hsp gene induction as a function-valued trait (Stinchcombe).

The net effect of shifts in Hsp expression parameters suggests that colonies with higher CTmax can better maintain protein homeostasis in subtle ways that are highly dynamic. For 2/3 Hsps, we found evidence for protection through graded, but not constitutive means. Constitutive protection may strictly be modulated under different acclimation temperatures (Helms Cahan et al. 2017), which was kept constant (25 °C) in our study, but greater thermal extremes may be needed for species to evolve elevated Hsp expression (Willot et al. 2017). On the other hand, graded protection allows for pre-emptive Hsp expression by elevating expression at the lower end of the temperature gradient when the expression rate is slower. Interestingly, slower expression rate also results in lower expression values before reaching maximum expression levels (Fig 1). However, colonies with higher Ctmax compensate by simultaneously elevating maximum expression to add greater repair at lethal temperatures. Because Hsps act as a sensor for protein denaturation (Richter et al. 2010; Craig and Gross 1991), enhanced tolerance (Tm) suggests that the whole proteome is robust to denaturing for species in open canopy forests. Very few amino acid changes are needed to evolve greater protein stability (Somero 2004; Lockwood and Somero 2012), however, these changes need to be distributed across many proteins across the whole proteome (Field 2001). Another route to greater protein stability involves changing osmolytes and their concentrations to enable higher melting temperatures across all proteins (Stillman and Somero 2001).

Protein homeostasis

Page 48: 2017-03-14. Reading notes- Armstrong et al. 2011; hsp70 and neurophysiology

reference:

Armstrong GAB, Xiao C, Krill JL *et al.* (2011) Glial Hsp70 Protects K+ Homeostasis in the Drosophila Brain during Repetitive Anoxic Depolarization. *PLoS ONE*, **6**, e28994.

Objective: Interesting paper where these authors wanted to understand the role of hsp70 in the ion homeostasis of neurons.

Hypothesis: Hsp70 help maintain K+ ion balance in the brain

Methods: To test for the role of hsp70, they began by manipulating fruit flies by treating them with ouabain, which caused loss of electrical activity in neurons. Ouabain is an inhibitor of Na/K ATPase pump. K gets pumped into cells, and NA gets pumped out. So if it is inhibited, K+ will accumulate outside of cells. Another way they disrupted ion balance was that they put the flies under anoxic conditions (by [Math Processing Error] gas). Their main read out was K+ outside of the cells. Then they measured hsp70 expression in the brain.

Results: Their oubain treatment worked. In hsp70 null background, there were more surges (K+ spikes). The surges were less when Hsp70 was induced in the brain relative to controls.

Page 49: 2017-02-16. Writing notes

A little tidbit I wrote up that ended up not fitting in with the paragraph:

To diminish the cost of protein denaturation, irreversibly damaged proteins are degraded by the 26S proteasome when marked with ubiquitin (Ronai 2016). Ubiquitination and 26S proteasome activity is higher in cold tolerant compared to warm tolerant species under temperature increases (Hofmann and Somero 1996; Todgham et al. 2016). In cold tolerant species Higher temperature sensitivity to protein damage matches with Hsp expression occurring at lower temperatures and further supports that they have less stable proteomes. Therefore clearance of denatured proteins is not likely a source of stress resistance. Instead, the stability of the proteome itself most likely provides stress resistance because unfolding at higher temperatures does not require any additional repair (Hsp) and breakdown (26S proteasome) responses that are costly (Hoekstra and Montooth 2013).

I couldn't fix it so I took it out, but it has good ideas and refs.

Interesting paper that compares the metabolome under different stressors:

ref: Robert Michaud M, Benoit JB, Lopez-Martinez G *et al.* (2008) Metabolomics reveals unique and shared metabolic changes in response to heat shock, freezing and desiccation in the Antarctic midge, *Belgica antarctica*. *Journal of Insect Physiology*, **54**, 645–655.

They measured the metabolome of antarctic midge under desiccation, heat shock, cold shock, and controls. The clustering looks like (newick format):

(desiccation,(cold,(control,heat)))

last paragraph of dissertation conclusion:

oberg or pelini paper on ctmpmax and water weight figure suggests that there is a trade off in 1 subfamily but not another.

cool idea to measure adaptive potential or trade offs between stressors across a whole clade.

also need to understand how the environment influences thermal experiences

Page 50: 2017-03-17. Writing notes

For the conclusion of my thesis, my thoughts are a little scattered. I took this part out:

Not only can the HSF-HSE cascade participate in responding to heat stress, but also it appears to be important for development (Gonsalves et al. 2011; Li et al. 2016) and may be critical for other biological processes. HSF-HSE binding is associated with transcriptionally active regions that are independent of bound chromatin locations under heat shock (Li et al. 2016), suggesting that the HSF has additional function outside the HSR. In fact, HSF binding can co-localize with the E2F transcription factor, which is known to permit development in insects and is influenced by nutritional stress (Britton and Edgar 1998). It is possible for social insects to utilize the cooperation between HSF and E2F to produce alternative reproductive phenotypes, which are in large part driven by environmental cues (Smith et al. 2008). If stressors besides temperature alter the chromatin structure to enforce a closed (heterochromatin) state to deny HSF-HSE binding. Therefore, even if proteins become unfolded under desiccation or starvation, Hsp may not be able to become induced which may explain the cross-susceptibility to coping with heat stress.

Hard to fit this in.

For my hsp rxn discussion, I need to rewrite it, starting with a re-outline for each paragraph:

First part: Focus on CTmax

1. Summary of findings for CTmax and forest types, evolutionary transitions
2. How forest types alter the thermal experiences and in turn, selective regimes. But I need to tie this back to clinal variation that I found too
3. Flat woods forests—SHC had problems with examples at the end. Hmm tie back to selective regimes?

Second part: Focus on hsp rxn norms

4. The transition into flat woods and increase in CTmax were accompanied by dynamic changes in Hsp expression profiles that were captured with a function-valued approach.
 - function valued approach allowed for novel approach to quantify different strategies for hsp expression that correlate with upper thermal limits
 - This is a novel approach over traditional comparisons of typical constitutive and induced expression because
 - Hsp parameters can evolve separately, but the first pc explained 40% of the variation in CTmax
 - give overall relationship, slower expression rate, higher Tm and higher maximum values
 - hsp parameters explained 40% of the variation
 - Furthermore, these parameters can evolve independently
5. Active part of Hsps
 - induced instead of constitutive protection, simply by altering expression rate
 - may not be the most important mechanism
 - interestingly, slower expression rate causes higher expression at the low end of the gradient, but lower expression at the higher end of the gradient before reaching

max expression values.

- Instead higher expression or enhanced response shows investment into active repair for flat woods forests species

6. Passive parts of Hsps:

- hsps are sensors and the higher onset of induction suggests they can resist protein unfolding
-

7. Conclusion:

Taking out a part, has good ideas, but poorly organized:

Hsps expression responds to the thermal sensitivities of enzyme kinetics (Dahlhoff and Rank 2000) and protein aggregation (Willot et al. 2016). Very few amino acid changes are needed to evolve greater protein stability (Somero 2004; Lockwood and Somero 2012) and are enriched in charged amino acids (Leuenberger et al. 2017). However, amino acid changes need to be distributed across many proteins across the whole proteome (Field 2001), which is extremely difficult because many mutations are deleterious and requires high mutation rates. A more attainable route to greater proteome stability involves increasing osmolytes that which can non-specifically extend melting temperatures across all proteins (Stillman and Somero 2001). For example, trehalose is known to act as a chaperone and can stabilize proteins (Richter et al. 2010).

Page 51: 2017-03-20. Writing notes.

Some more writing notes: basically stuff I'm taking out of my discussion.

Traditional approaches with comparisons between constitutive and induced responses may miss the subtle responses found in our study because of unsampled variation. For example, a single measurement of Hsp expression at the middle of the temperature gradient reveals no evidence for adaptive variation between forest types.

Page 52: 2017-03-21. Writing notes

The part that the reviewer wanted me to take out in my discussion because it was too fruit fly focused in my first paper (Nguyen et al. 2016; *BMC Evo Biol*)

Among different species of *Drosophila*, the primacy of *hsp70*'s role in the heat shock response is achieved through an evolutionary increase in copy number from 2 to 5 [16,20]. We detected minimal copy number variation in either *hsp83* or *hsc70*, although it is important to note that lack of detailed manual annotation of non-*Drosophila* genomes may make it difficult to detect highly similar

copies should they be present. Nevertheless, this suggests that Hymenopteran (and other insect) Hsps may be recruited to the heat shock response primarily through expansions in cis-regulatory HSEs across the entire promoter region, along with transcriptional enhancement associated with introns that are lacking in *D. melanogaster *(Additional files 7-8: Figures S4-5)[23-25,54-55].

Hsp rxn norm paper

Took this out of my discussion, couldn't fit it in...

Broadly in ants, the divergence in upper thermal limits is associated with open to forested habitats and nesting locations (Diamond et al. 2012). However, within our study, we were able to recover divergence in upper thermal limits within forest types that differ in canopy cover.

Part of discussion I'm saving:

Acclimation temperature in natural populations change within a season and Hsps are known to track seasonal fluctuations (Dietz and Somero 2004; Banerjee et al. 2014).

Another part: This is at the end of the proteome stability part

The relative contribution of all of these mechanisms are unknown and typically studied in isolation. Nonetheless, resisting heat stress as inferred from the onset of Hsp induction appears to be a critical feature of Hsp expression profiles that is associated with shifts in upper thermal limits.

Maybe I don't even want to pop off on how proteomes can be stable:

Higher proteome stability can be achieved by expressing more thermally stable isoforms (Mastrangelo et al. 2012), protein-protein interactions (Storey and Wu 2013), accumulation of protective osmolytes (Hottiger et al. 1994; Somero 2003), and evolved differences in amino acid sequence (Lockwood[and Somero 2012; Fields et al. 2015).

SHC comments and outline for discussion:

- Primary mode of action of Hsps is enhanced response to damage – how much more?
- But also evidence that more heat-tolerant species are better able to tolerate thermal stress through other means: protein sequence evolution, osmolyte concentrations, metabolic depression and gene expression downregulation(Stanton-Geddes). Because Hsps are expensive, the more frequent the heat stress, the more costly active response is.
- Protection less important contributor to higher thermal tolerance. Protection is usually thought of in terms of baseline levels, but important to remember that this can also be achieved through the induction response via graded or early-onset protection. Shift to a more graded response suggests a change in trigger of Hsp gene expression (from direct feedback from denatured proteins to some other temperature sensor – something like HSR1 non-coding RNA?)
- Study of extreme thermophiles suggests a shift from induced to constitutive protection, but this may not be representative of evolved thermal strategies in seasonal environments (or less extreme? Or non-desert?). Probably not the case that the elevation is occurring in genes other than the ones sampled

here – positive association between inducibility and baseline elevation in hot—desert vs. temperate ant comparison (other literature on the likelihood that non-inducible forms are involved in heat-associated chaperoning?) . Why not? Energetically costly and retards enzyme activity, slowing growth. Most useful when stress is chronic or predictable – typical of species living in predictably extreme environments such as hot deserts (in ants: *Pogonomyrmex* and *Cataglyphis*, desert flies, other taxa). For temperate species, stress is predictable on timescales that are shorter than the individual's lifetime, making baseline shifts more profitable as a plastic acclimation strategy. Evidence that this is what occurs in *Aphaenogaster* (JCB paper), loads of other literature on seasonal acclimation.

- Same pattern of change associated with change in thermal stress within deciduous forest as with shift to more open habitat – common regulatory control such that response can be tuned up or down in a coordinated fashion in response to selection? Any idea of a model that would explain pattern? HSF binding affinity, etc.?

SHC comment on my conclusion:

Do you have anything to say about this? Is there evidence from your work of local adaptation to environmental conditions *within* species?

In response to this sentence:

Further adaptive modifications in the heat shock response will depend on the adaptive potential of Hsp expression (Huey et al. 2012; Tedeschi et al. 2016; Diamond 2016).

2017-03-22: taking the significance out before I submit thesis.

The types evolutionary changes expected for species to persist in a warming world are poorly understood. We found that surviving upper thermal extremes involves coping with the deleterious effects of heat stress by protecting against and tolerating it.

Page 53: 2017-03-22. Prepping for oral defense, reading notes

Submitted thesis to committee. What types of questions will I get for my defense?

1. What types of mechanisms increase proteome stability? Why does a proteome need to be stable?

ref: Fields PA (2001) Review: Protein function at thermal extremes: balancing stability and flexibility. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **129**, 417–431.

Stability is needed to ensure the appropriate geometry for ligand binding, as well as to avoid denaturation, while flexibility is necessary to allow catalysis at a metabolically appropriate rate

- a. A stable protein is needed to bind a substrate but also not unfold and aggregate.
- b. Flexibilty is needed for catalysis because it involves domains in the protein to move.

Stability is defined as the distribution of microstates considered to be in the native form—not sure I like this definition

2. What facilitates protein folding?

- Hsps
- protein-solvant interactions
- it is a challenge to refold proteins because the number of microstates outnumbers the atoms in the universe
- hydrophobic effect—non-polar residues are usually in the interior of globular proteins

3. How far do Aphaenogaster migrate?

- deciduous forest ones move around a lot, but less than 1 meter
-

4. Across the whole clade, is there evidence for phylogenetic niche conservatism?

- potentially, the western clade is a bit reversed. There is more diversity in open canopy species and 1 or 2 closed canopy species.

5. What is the geographical history of NA Aphaenogaster?

- There is evidence for a diversification event from asia to na. *A. japonica* is located in japan and is sister to all of NA *aphaenogaster* species and this split occurred ~ 16 my, roughly the same timing as the bering strait.

- Part of the Stenamini tribe, and Myrmecine subfamily
- Genus is polyphyletic and needs to include Messor.

6. What about membrane stability?

- Adaptive to shift phospholipid head groups, enrich for PC and lipid saturation to confer stability
- Under heat stress, lipids melt and become more fluid like, so organisms respond by becoming more rigid

2017-3-24 cont'd:

7. What can pass through a lipid bilayer(pure)?

- Things that can pass
 - hydrophobic molecules such as gas, steroids
 - small uncharged polar molecules: urea, water, glycerol, ethanol
- Things that CANT pass
 - Large uncharged polar molecules: sugars
 - Ions cannot pass (Na^+ , K^+ , H^+ , Ca^{2+})
 - In active transport: 3 Na^+ goes out, 2 K^+ goes in and produces a negative membrane potential
 -

1. Dissertation defense talk:
 - write out script for parts where I changed the slides (done, need to practice)
 2. Multiple stressor ms reivisions
 - SHC-do waht the reveiwers say; so they want relative expression to the initial timepoint
 3. Ecological genomics overall hw feedback (sent to MP and SK 2017-03-24)
 - 4.
-

Page 55: 2017-03-24. multiple stressors reviewer comments

On putting stuff online, NJG:

Well, the reviews are certainly confidential at all times (even after publication) when you are a reviewer. But the reviews are not confidential once the journal gives them to you as an author, so I don't see any harm in having the on your webpage. However, if any of the reviews were signed (the authors revealed their identity), I would not post those because the authors might not appreciate it.

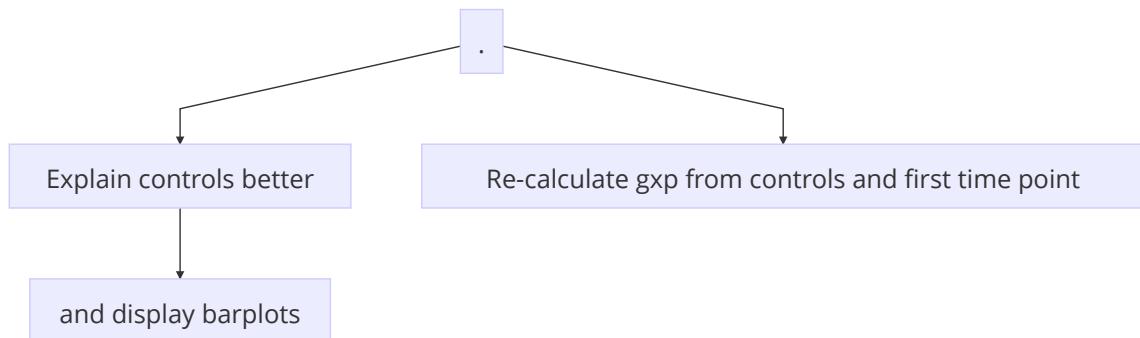
Reviewer comments:

Reviewer #2: The authors have thoroughly addressed the review comments, and in my opinion the manuscript reads much better. The water content results and effects of starvation on dry weight are nice additions. I only have a couple of minor comments to consider before revision:

1. Be careful with tense. When discussing the current results, past tense is typically used, when discussing published results, the present tense is typically used. For example, in line 39, "impairs" should read "impaired," and in line 86 "reduced" should be "reduces." Those were the only two that stuck out to me, but make sure tense is consistent.
2. Line 143: Change "2" to "two"
3. Line 148: Space needed after "used"
4. Line 167: Comma needed after "status"
5. In Figures 2 and 4, it would be helpful to clearly indicate in the caption which group is serving as the reference for the $2\Delta\Delta Ct$ calculations. Also, the authors might want to consider plotting the control groups on the gene expression figures. That is, the control group can be included by setting the mean = 1 scaling the variance accordingly. It is also unclear whether there was a single reference/control group for the induction

experiments, or whether each time point had its own room temperature control that served as a reference. If it was the latter, connecting the points with lines is somewhat misleading, because each point is relative to a different control group. Personally, I think it makes sense to use a single reference group (for example, room temperature controls after the first time point), then express everything relative to that group.

Basically I see two approaches to addressing #5, which was the most important suggestion:



Page 56: 2017-03-28. Definition of phenotypic plasticity and general paper reading

Scheiner 1993; *Annual Review of Ecology and Systematics*

Scheiner SM (1993) Genetics and Evolution of Phenotypic Plasticity. *Annual Review of Ecology and Systematics*, **24**, 35–68.

Scheiner defines phenotypic plasticity as the general effect of the environment on phenotypes, including when the effect is 0.

One way to represent phenotypic plasticity is a *norm of reaction* or *reaction norm*. He refers to the reaction norm as the specific form of that(environmental) effect. These terms can be interchangeable he argues. Plasticity itself is the specific change in a trait with respect to the environment, regardless of the genotype.

Angilletta 2009 defines plasticity as the derivative of a reaction norm. He cites a good reference for how plasticity is defined: Debat and David 2001.

ref: Debat V, David P (2001) Mapping phenotypes: canalization, plasticity and developmental stability. *Trends in Ecology & Evolution*, 16, 555–561.

I like this quote:

The relationship between genotype and phenotype is not one to one.

Hard to understand:

This duality is the basis of 'genetic assimilation', the process by which a response to unusual environments can be converted by selection into a permanent, genetically determined, phenotypic change.

The response can be genetic itself. Poor definition. I'd define genetic assimilation as the evolution of a plastic response into a fixed one.

Wow, this dude named Woltereck made up the term "reaction norm" in 1909.

Definitions of plasticity from this paper:

1. Bradshaw 1965: Plasticity is shown by a genotype when its expression is able to be altered by environmental influences. This may or many not be adaptive.

other definitions are junk. They don't include Scheiner into this paper...what the....

Making sense of JSG's paper:

JSG discusses 3 potential mechanisms of temperature adaptation that is mediated by the transcriptome:

1. Enhanced response hypothesis
 - Inducing a stronger transcriptional response to provide protection.
 - Prediction: cool adapted species strongly upregulate genes at low temperature and down regulate at higher temps
2. Tolerance hypothesis
 - Not really framed as a hypothesis....

proposes that existing inducible stress responses become insufficient or prohibitively costly as environmental stressors increase in frequency, resulting in a shift away from induced response in favor of structural changes

What the...does this mean?

- 1 * I think the hypothesis is this: A shift away from induced response in favor of structural changes. Makes no sense to me...
- 2 * This hypothesis predicts adaptation to stress to be associated with lower transcriptional responsiveness and less sensitivity to temperature perturbations.
- 3 * I'd describe this as resistance.

3. Genetic Assimilation hypothesis

- Selection shifts inducible expression to constitutive expression of stress related genes.

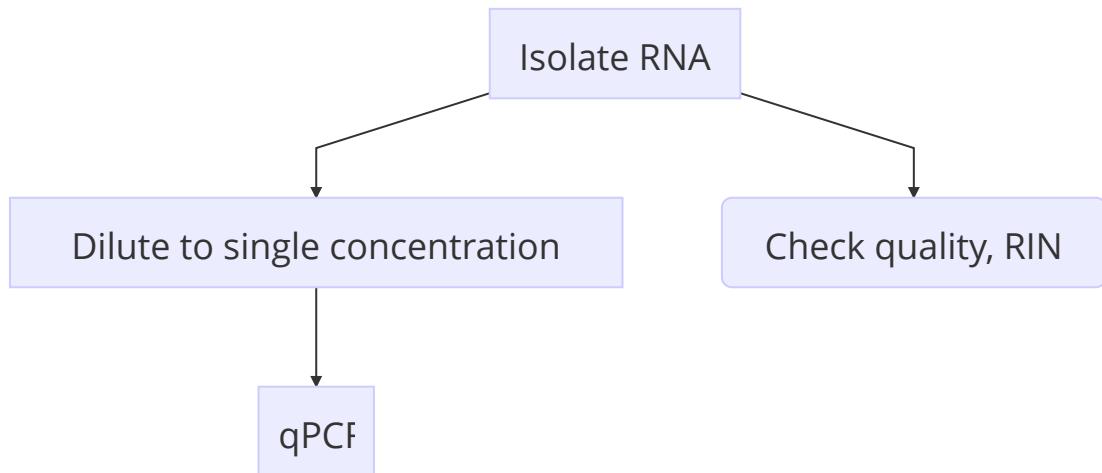
Page 57: 2017-03-28. Meeting with new lab tech, KB.

Curtis' project needs to be finished.

Status:

- **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.**

Workflow



1. **Stressed in nature project**
 - RNA isolations
 - qPCR , measuring gene expression
 - data analysis and reproducible science
 - Github, tracks different versions and facilitates online sharing
 - R, main software for data analysis
2. Maintenance of website
3. Protein related assays
 - protein isolations
 - unfolding curves troubleshooting
 -
4. Proteome stability data analysis
5. **Lab orientation and lab safety.**

Monday, Tuesday, Wednesdays. Wednesday leave at 3:45PM. Maybe Fridays.

Plan for April 3rd-5th:

1. Lab orientation and safety
2. Set up github account
3. Set up R and R studio
4. Organize samples. Find them and place them all in a known location. Prepare for RNA isolations.
 - Let's print out the protocol

April 10th

Begin isolations Monday and Tuesday

Page 58: 2017-03-31. Making thesis fixes, notes

Adding notes from J Preston and Blockwood for Hsp rxn norm chapter.

JP:

In contrast, increases in the inflection point indicates later onset of protein damage (stress resistance; Fig. 1C

Does this mean there is another mechanism that stops protein damage at moderate temperatures OR that Hsps can evolve to be more potent at lower levels?

If thermal tolerance is achieved via greater responsiveness to protein damage, this should be reflected in higher peak expression levels....

**To test this don't you need to know how much protein damage there is? What if protein damage is the same between two species at a moderate temperature, but one of the species (the more tolerant species) takes longer to ramp up Hsps? Doesn't this mean that the tolerant species is less responsive to protein damage at least at moderate temperatures?

I'll add Brent's later.

Blockwood:

For the intro

2nd paragraph in intro

It would be good to include a statement and reference to studies that justify the protective effect of the HSR *in vivo*, i.e. Welte, et al. 1993, Feder et al. 1995, etc. wherein Hsp over expression was engineered with transgenic constructs and increases in whole-organism heat tolerance were measured. This is important because many in the field are confused by thinking of the HSR as a "stress" response and thus see Hsps as bioindicators of stress rather than the protective molecular "heroes" that they are.

3rd paragraph in intro:

You need to cite Brian Bettencourt's work. He tracked alleles of Hsp70 that corresponded to high and low expression in the Australian Drosophila. So, you could probably say something more than just "currently unclear" because we do know something about natural populations. There's also the copepod work out of Ron Burton's lab that has done quite a lot to address this question in

recent years.

In response to:

Hsp gene expression usually follows a non-linear, logistic function
(Kingsolver & Woods 2016);...

Is this really the only reference for this statement? Is this based off of insects only? Or do they cite more broadly into the marine literature? One of the best references for timecourse of hsp expression is Buckley and Somero 2006

For the discussion

Two major considerations: First, I think you need a set of statements about other mechanisms of thermal tolerance. Your hsp expression explains 40% of the variation, but what about the other 60%? It's not all about proteins. This is particularly an important consideration for anyone who appreciates the physiology that underlies CTmax, which likely has a neural (membrane) basis. Further, for aerobic eukaryotes, the membranes of the mitochondria are an important factor. This sort of relates to the OCCLT hypothesis by Portner too. Check out Dahlhoff and Somero 1991 (abalone mitochondria) as well as Stillman and Somero 1999 (porcelain crab heat tolerance). Second, you should briefly discuss the obvious alternative hypothesis that the evolution of CTmax may not have accompanied or been driven by the colonization of different habitats, but rather the species had different CTmax prior to sorting themselves based on their thermal preferences.

Page 59: 2017-03-31. To do list; project update

Project updates:

- **Hsp gene expression + Ctmax project:**
 - Lot of it has to be reviewed and rewritten
 - For intro and discussion: take Brent's comments and rewrite adding refs
 - Methods need to be shrunk for PNAS format
 - Need to do an analysis that can account for intra and interspecific variation
 -
 - Submit to PNAS
- **Multiple stressors ms:**
 - In SHC's hands
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
 - 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:**
 - ?? Lacy's hands
- **Stressed in nature MS: Samples to rerun.**
 - 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.

- I need to find the samples and organize them! Half of the samples are in the chest freezer, half are in our own -80C.
 - qPCR
 - Analyze the data
- **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)
-

Page 60: 2017-04-07. Project updates

Project updates:

- **Hsp gene expression + Ctmax project:**
 - Lot of it has to be reviewed and rewritten
 - For intro and discussion: take Brent's comments and rewrite adding refs
 - Methods need to be shrunk for PNAS format
 - Need to do an analysis that can account for intra and interspecific variation
 - Submit to PNAS
- **Multiple stressors ms:**
 - In SHC's hands
- **Range limits ms:** SHC lab gave verbal edit, still need to incorporate
 - 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:**
 - ?? Lacy's hands
 - **Stressed in nature MS: Samples to rerun.**
 - There are 74 samples: 3 days of RNA isolation + cDNA synthesis.
6 gene targets ran in duplicates is 2 plates per gene = 12 plates
total
 - **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
 - Need gapdh values
-

Page 61: 2017-04-10. To do and project updates

I'm meeting with Wai today at proteomics facility. Ask about:

1. Gapdh values
2. redoing runs: We only found 120-200 proteins, but last time we found 500 proteins.

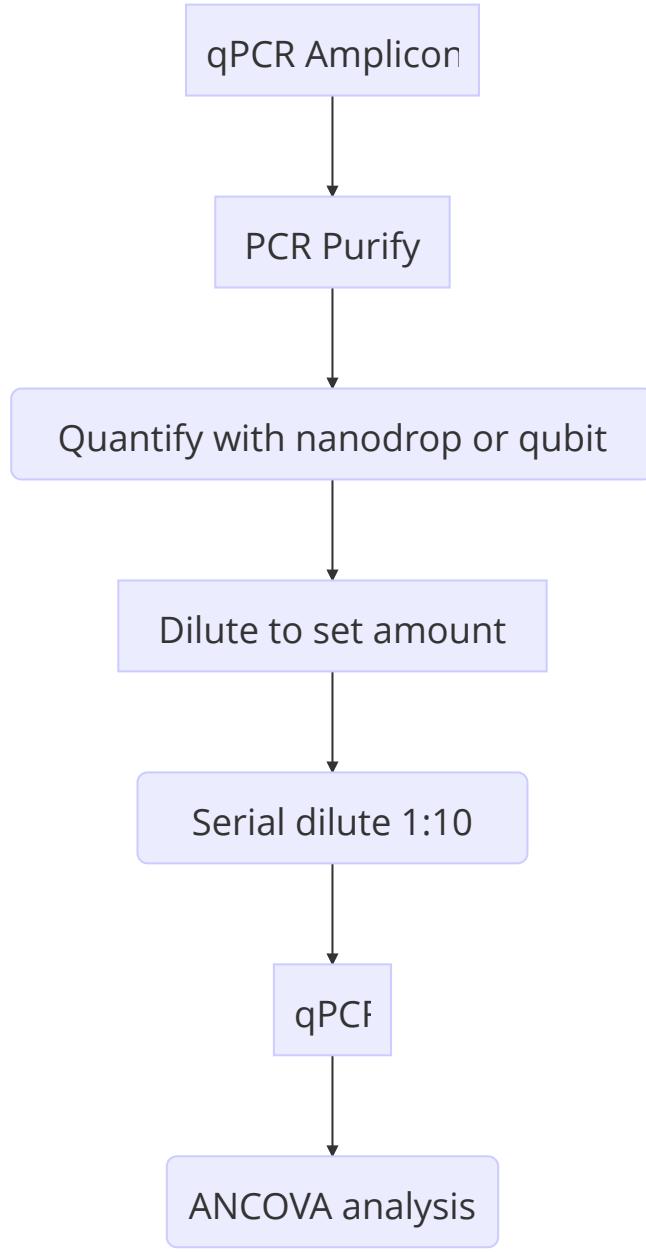
Organize stressed in nature project.

- Input RNA values into spreadsheet.

- Curtis diluted cDNA 1/10 before qpcr; threshold on the machine is 0.5 I think;
 - So I need to dilute cDNA to prep for qpcrs.
-

Page 62: 2017-04-11. Hsp rxn norm project: Species efficiency curves

Workflow



Species list

species	hsp70	hsp40	hsp83_279
ashmeadi	FB2	LPR4	LPR4
barbatus	PB17-10	PB17-10	PB17-10
crematogaster	x	x	X
floridana	FB1	KH7	KH7
fulva	CJ6	CJ6	DUKE7
lamellidens	DUKE8	DUKE9	DUKE9
miamiana	ALA1	ALA4	ALA4
pergandei	EXIT		EXIT
picea	AVON	TED4	PHIL
rudis	POP1	LEX13	POP2
tenn.		SR45	LEX9

PCR purify amplicons

Used Qiagen pcr purify kit according to manufacturer's instructions:

1. took 20 uL of amplicon and added 100 uL (5 vol) of PB buffer
2. Transferred to column, spun top speed 1 min
3. wash with PE buffer 750 uL, spin top speed 1 min
4. Spin again, top speed 1 min
5. Place column in new tubes (with labels) and elute
6. Eluted by adding 30 uL h20 into center of column and waiting 1 min; spun top speed 1 min

ISOLATION_ORDER	HSP70	HSP40	HSP83
1	FB2	LPR4	DUKE9
2	PB17-10	PB17-10	LPR4
3	x	x	PHIL
4	FB1	KH7	KH7
5	CJ6	CJ6	ALA4
6	DUKE8	DUKE9	CREMATOGASTER
7	ALA1	ALA4	LEX9
8	EXIT		DUKE7
9	AVON	TED4	EXIT
10	POP1	LEX13	PB17-10
11		SR45	POP2

Next I should quantify and then load set amount for first dilution step.
THen dilute 1:10 8 times.

Page 63: 2017-4-13. Hsp rxn norm proj: amplicon quant

RNA quant and 1:10 dilution calculations

Species	Colony	Gene	ampliconDNAconc	X10ngcDNA	h20cDNA
ashmeadi	FB2	HSP70	8.34	1.20	98.80
barbatus	PB17-10	HSP70	6.12	1.63	98.37
crematogster	CREMATOGASTER	HSP70	4.65	2.15	97.85
floridana	FB1	HSP70	8.50	1.18	98.82
fulva	CJ6	HSP70	6.71	1.49	98.51
lamellidens	DUKE8	HSP70	2.64	3.79	96.21
miamiana	ALA1	HSP70	6.37	1.57	98.43
pergandei	EXIT	HSP70	6.37	1.57	98.43
picea	AVON	HSP70	5.25	1.90	98.10
rudis	POP1	HSP70	5.13	1.95	98.05
tenn.		HSP70	NA	NA	NA
ashmeadi	LPR4	HSP40	2.06	4.85	95.15
rudis	LEX13	HSP40	2.82	3.55	96.45
miamiana	ALA4	HSP40	2.35	4.26	95.74
tenn.	SR45	HSP40	2.48	4.03	95.97
lamellidens	DUKE9	HSP40	2.45	4.08	95.92
fulva	CJ6	HSP40	0.97	10.33	89.67
barbatus	PB17-10	HSP40	4.77	2.10	97.90
pergandei		HSP40	NA	NA	NA
floridana	KH7	HSP40	2.46	4.07	95.93
picea	TED4	HSP40	2.77	3.61	96.39
crematogster	CREMATOGASTER	HSP40	6.26	1.60	98.40
lamellidens	DUKE9	HSP83	7.36	1.36	98.64
ashmeadi	LPR4	HSP83	12.50	0.80	99.20
picea	PHIL	HSP83	9.50	1.05	98.95
floridana	KH7	HSP83	9.69	1.03	98.97
miamiana	ALA4	HSP83	4.77	2.10	97.90
crematogster	CREMATOGASTER	HSP83	6.12	1.63	98.37
tenn.	LEX9	HSP83	11.10	0.90	99.10
fulva	DUKE7	HSP83	9.25	1.08	98.92
pergandei	EXIT	HSP83	8.14	1.23	98.77
barbatus	PB17-10	HSP83	8.07	1.24	98.76
rudis	POP2	HSP83	7.64	1.31	98.69

Dilute 1: 10 = 0.1ng /uL. Then dilute 1:10 (=0.01ng/uL)for first well for efficiency curve. Dilute by adding 10 uL of sample with 90 uL of water.

Dilute step for efficiency curve (in concentration of ng/uL)

1. 0.01
2. 0.001
3. 0.0001
4. 0.00001
5. 0.000001
6. 0.0000001
7. 0.00000001
8. h20

or....

1. 1 E-2
2. 1 E-3
3. 1 E-4
4. 1 E-5
5. 1 E-6
6. 1 E-7
7. 1 E-8
8. h20

qPCR

Set up hsp70, hsp40, and hsp83 qpcr plates. I made master mix for 100 rxns:

Reagent	Initial conc.	Final conc.	uL for 1 rxn	100 rxn set up
Power sybr green	2x	1x	5	500
Forward Primer	10 uM	250 nM	.25	25
Reverse Primer	10 uM	250 nM	.25	25
nuclease free water	na	na	.5	50
cDNA	0.25 ng/uL	.1 ng/uL	4	na
Total			10	600

Loaded 6 uL master mix into plate and then added 4 uL of template (DNA).

Loaded dilutions for each species H-A. Each species is a row.

- qPCR steps
 1. 95 C, 10 minutes
 2. 40 cycles of: 95 C for 15s, 55 C for 60 seconds and then 70 C for fluorescence acquisition,
- Melt curve analysis
 1. Reactions were heated to 95C 15 s
 2. From 60 C, slowly heat up and measure fluorescence

Results: Hsp70

0.5 fluorescence threshold

It looks like the dilution didn't work for *crematogaster* and *floridana* (Fb1). I didn't mix well for initial 1:10 dilution.

I still need to run hsp40 and hsp83.

Plate layout:

Species are loaded as columns:

1. FB2
2. PB17-10
3. CREMATO
4. FB1
5. CJ6
6. DUKE8
7. ALA1
8. EXIT
9. AVON
10. POP11

Dilution steps are rows:

H. 1 E-2

G. 1 E-3

F. 1 E-4

E. 1 E-5

D. 1 E-6

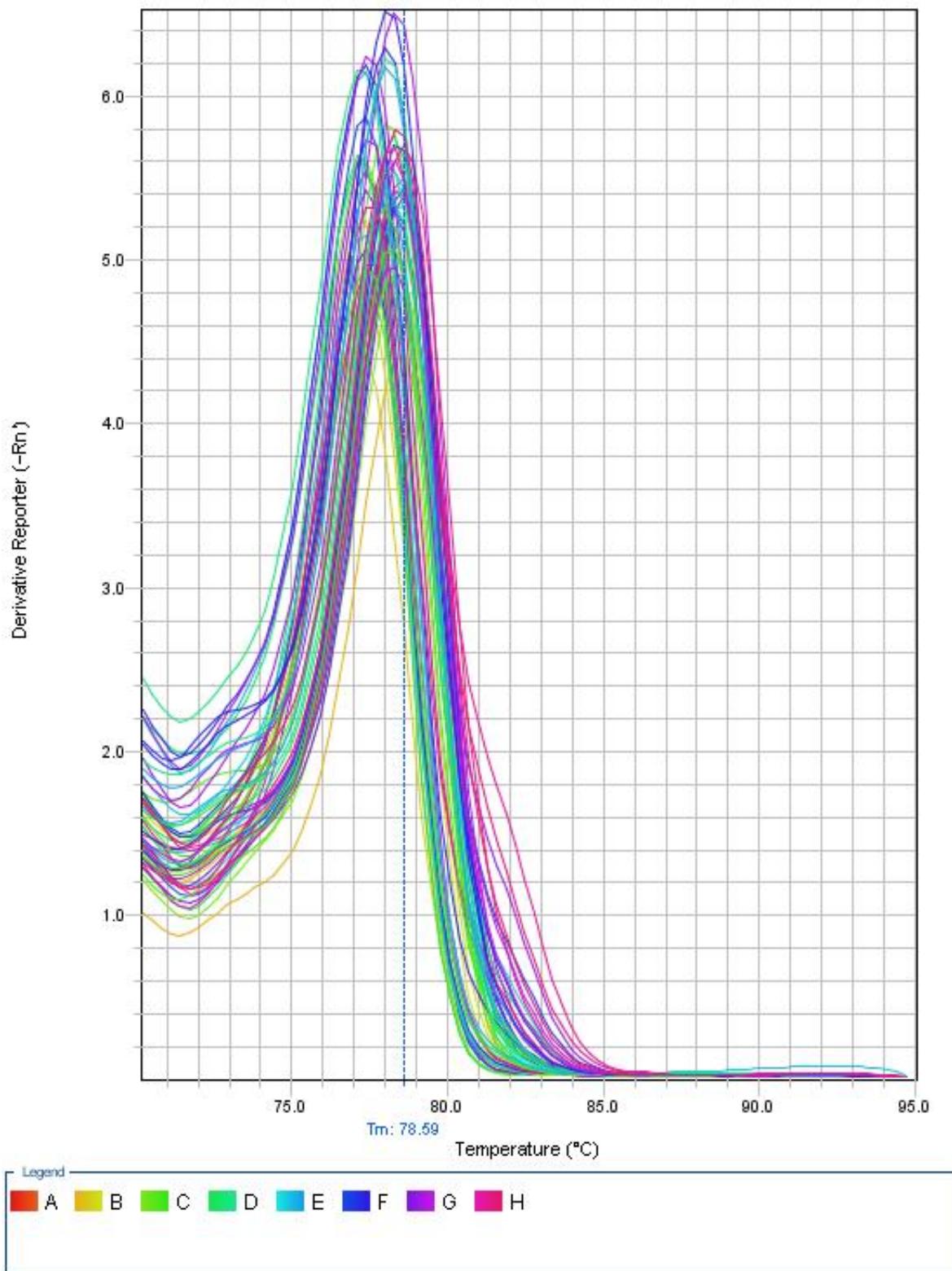
C. 1 E-7

B. 1 E-8

A. h20

hsp70 melt curves

Melt Curve



Looks like a single product.

stats: Fitting ANCOVA model to determine whether species differ in their hsp70 primer efficiency (relationship between cycle threshold and dilution of template amplicons)

```
1 summary(mod1)
2
3 log_dil             Df Sum Sq Mean Sq   F value    Pr(>F)
4 Sample.Name          7   5.2     0.7   20.392 2.36e-
5 log_dil:Sample.Name 7   0.4     0.1   1.684
6 Residuals           40  1.5     0.0
7
8 log_dil              ***
9 Sample.Name            ***
10 log_dil:Sample.Name
11 Residuals
12 ---
13 Signif. codes:
14 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Page 64: 2017-4-14. Hsp rxn norm project: Species efficiency curves cont'd

qPCR: Hsp83 279prim

Species are loaded as columns:

1. DUKE9

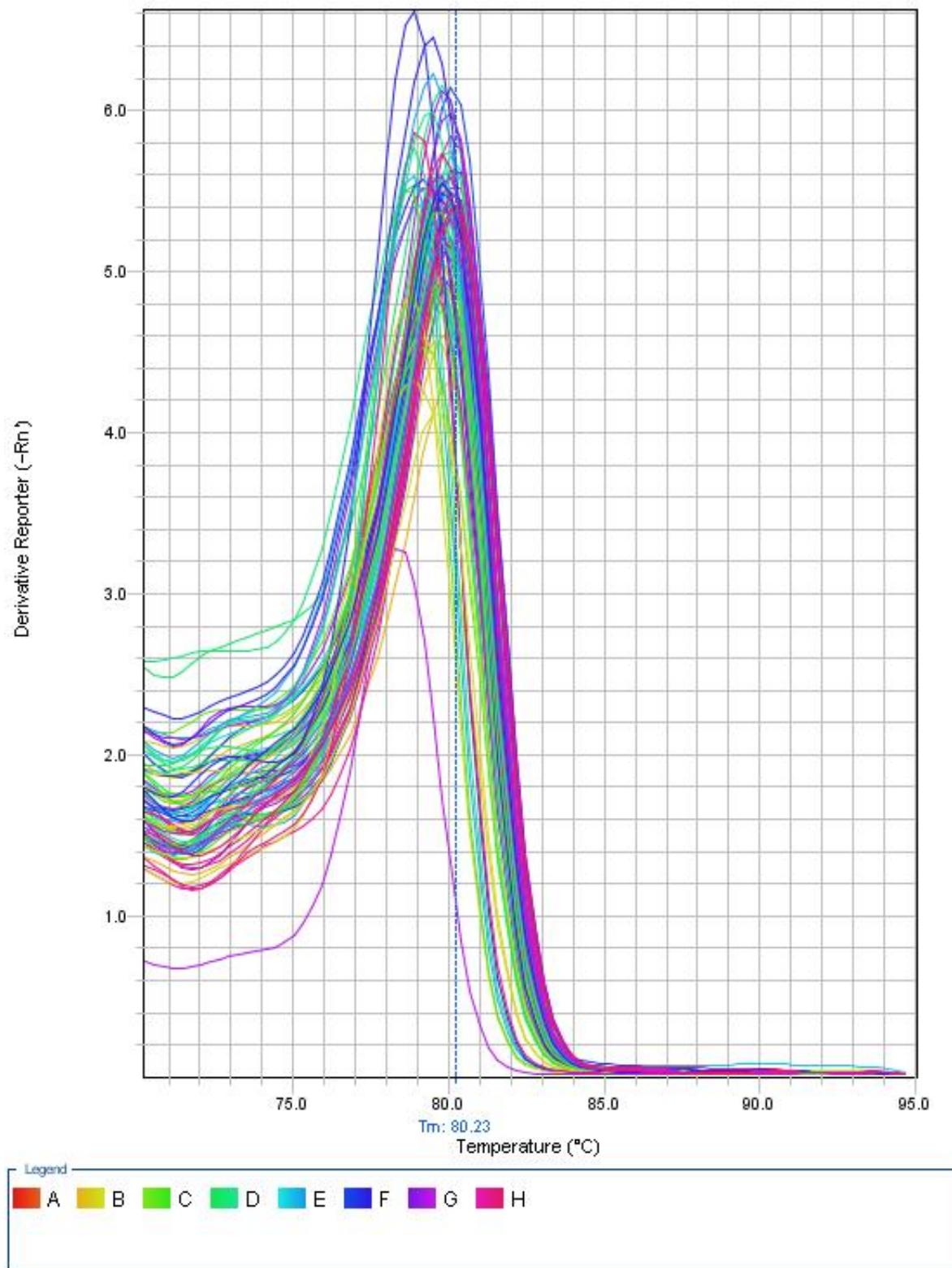
2. LPR4
3. PHIL
4. KH7
5. ALA4
6. CREMATOGASTER
7. LEX9
8. DUKE7
9. EXIT
10. PB17-10
11. POP2

Dilution steps are rows:

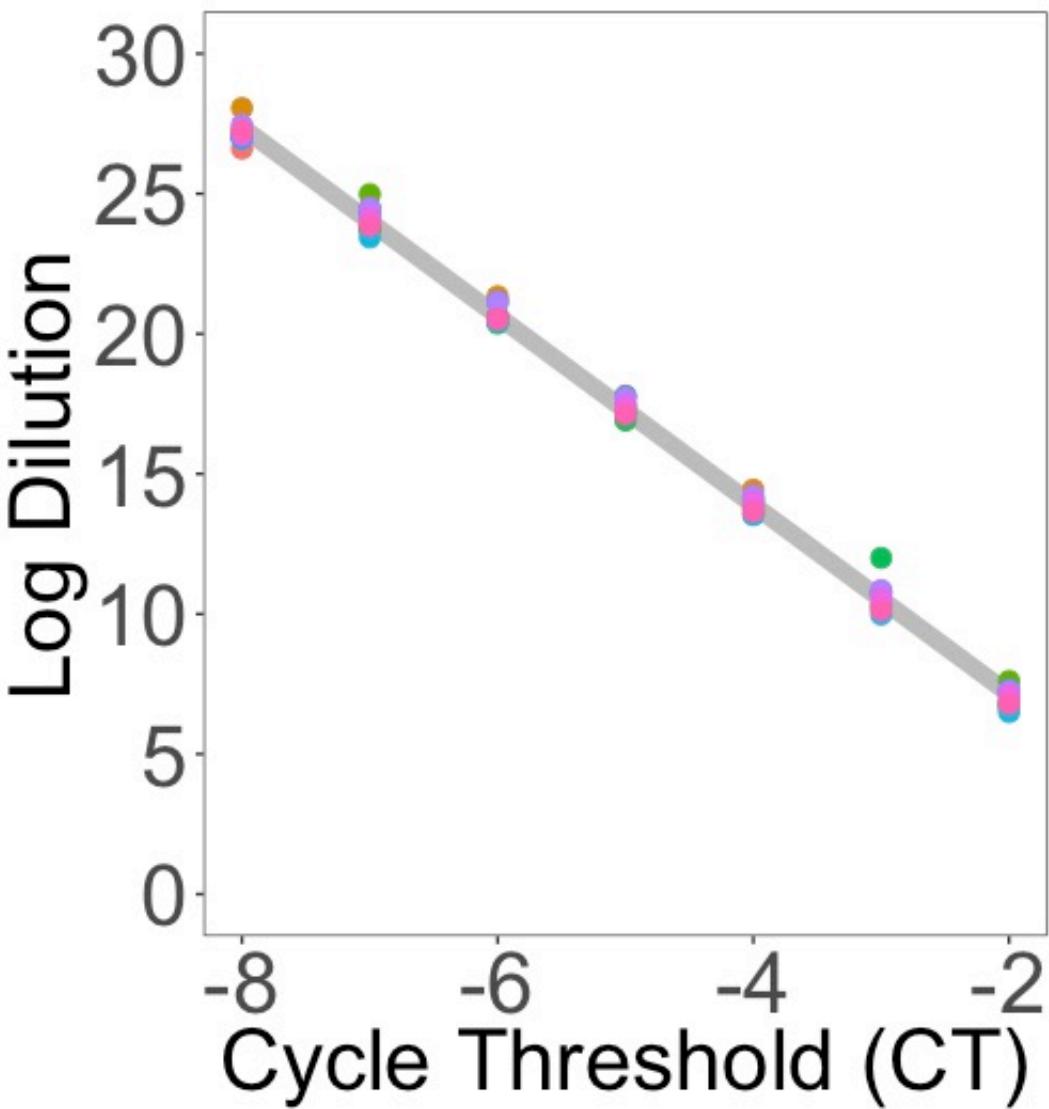
- H. 1 E-2
- G. 1 E-3
- F. 1 E-4
- E. 1 E-5
- D. 1 E-6
- C. 1 E-7
- B. 1 E-8
- A. h20

hsp83 melt curve

Melt Curve



Hsp83 efficiency curve



stats: Fitting ANCOVA model to determine whether species differ in their hsp83 primer efficiency (relationship between cycle threshold and dilution of template amplicons)

```
1 str(hsp83)
2
3 'data.frame': 77 obs. of 6 variables:
4   $ Well_letter: Factor w/ 7 levels "B", "C", "D", "E", ...
5   7 6 5 4 3 2 1 7 6 5 ...
6   $ Well_num   : int 5 5 5 5 5 5 5 6 6 6 ...
7   $ Sample.Name: Factor w/ 11 levels
8   "ALA4", "CREMATOGASTER", ...: 1 1 1 1 1 1 1 2 2 2 ...
9   $ Ct          : num 6.77 10.33 13.8 17.46 20.65 ...
```

```

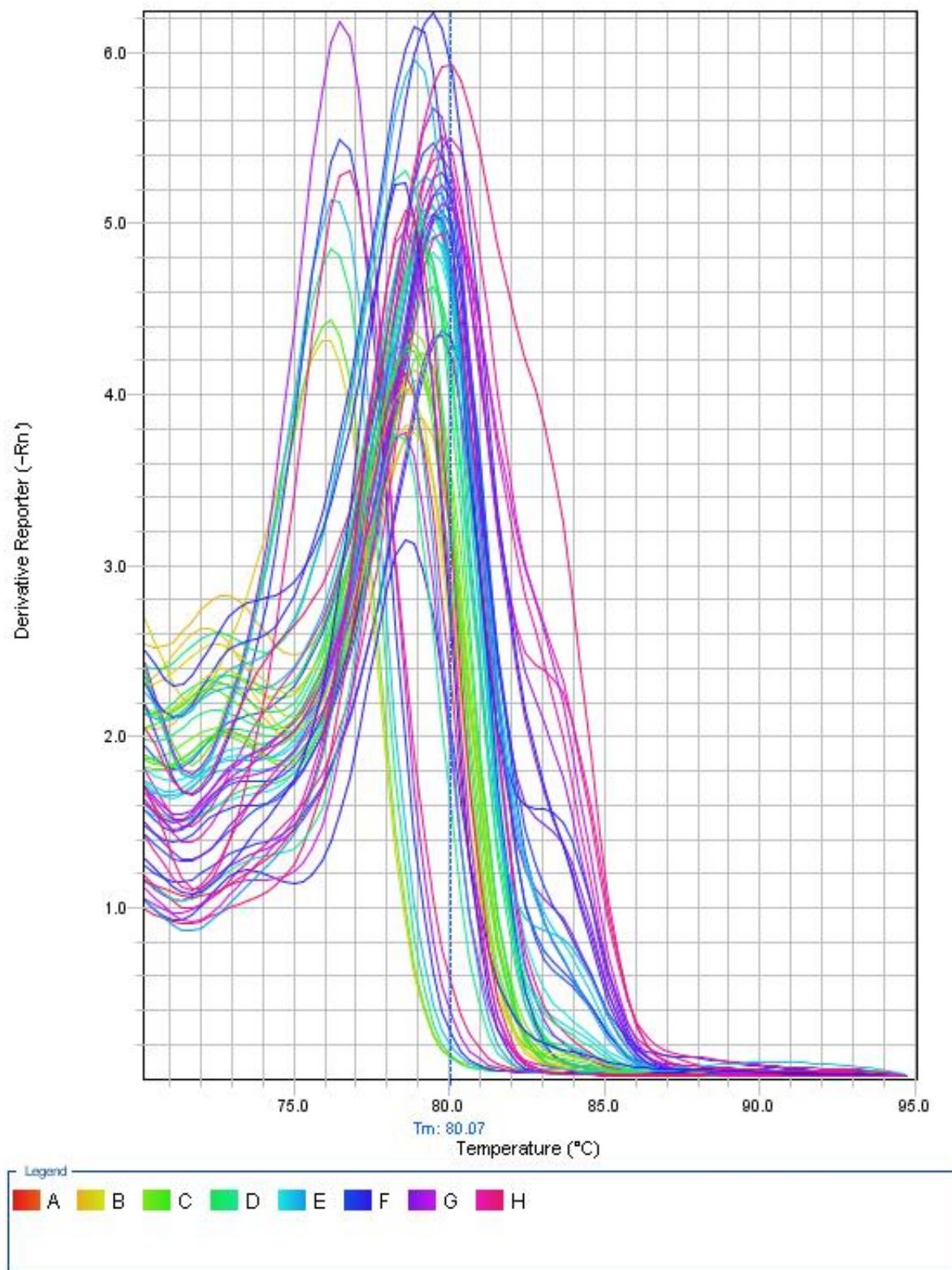
8 $ Dilution    : num  1e-02 1e-03 1e-04 1e-05 1e-06 1e-
9   07 1e-08 1e-02 1e-03 1e-04 ...
10
11 mod2<-aov(Ct~log_dil*Sample.Name,data=hsp83)
12 summary(mod2)
13
14 Df Sum Sq Mean Sq   F value   Pr(>F)
15 log_dil           1   3497   3497 36481.640 < 2e-
16
16 Sample.Name       10      5      1   5.730 7.82e-
17
17 log_dil:Sample.Name 10      1      0   0.728
18   0.695
18 Residuals        55      5      0
19
20 log_dil          ***
21 Sample.Name        ***
22 log_dil:Sample.Name
23 Residuals
24 ---
25 Signif. codes:
26 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
27

```

Hsp40 data

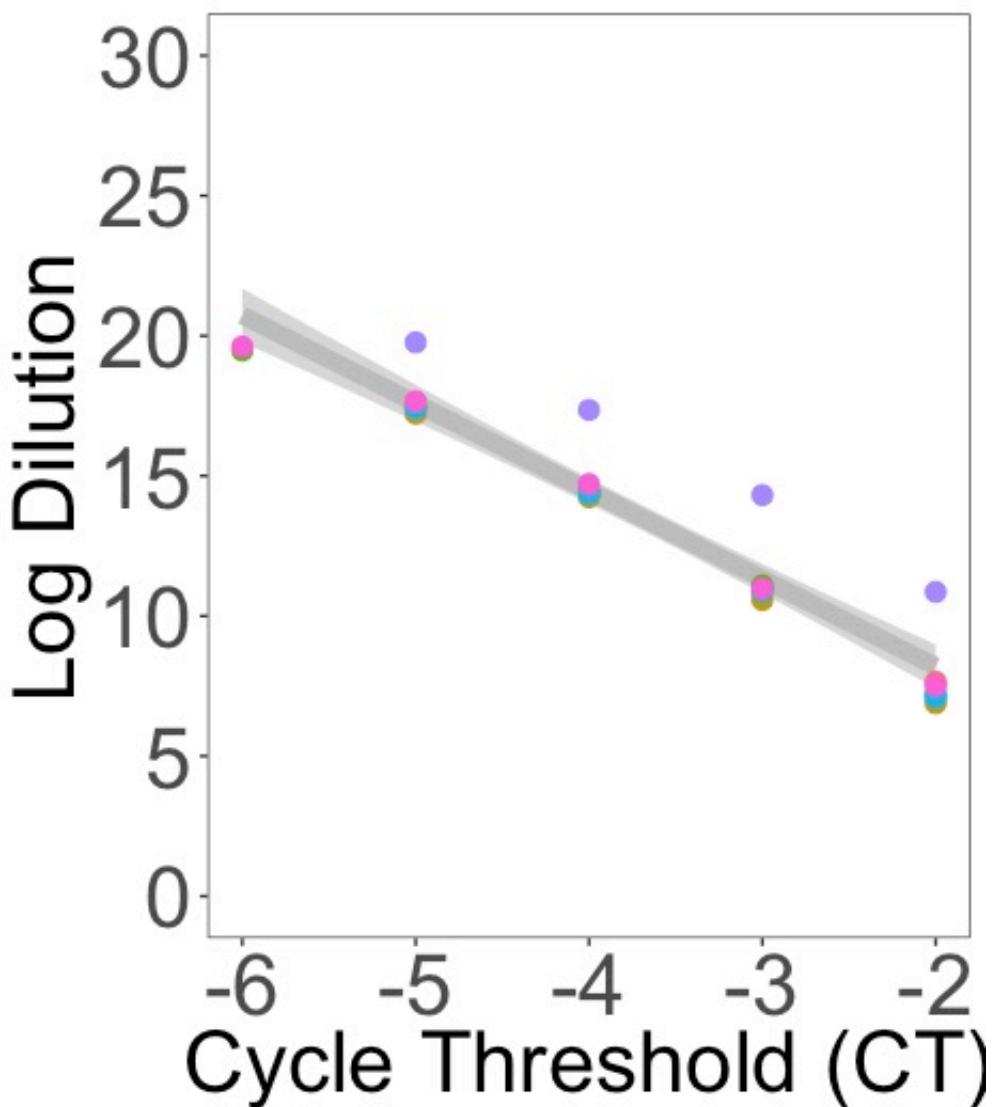
melt curves:

Melt Curve



note: the curves on the left side are hsp70, not hsp40.

sp efficiency curves



stats: Fitting ANCOVA model to determine whether species differ in their hsp40 primer efficiency (relationship between cycle threshold and dilution of template amplicons)

1	Df	Sum Sq	Mean Sq	F value	Pr(>F)
2	log_dil	1	429.8	429.8	2649.521 < 2e-16
3	Sample.Name	6	33.7	5.6	34.614 5.83e-08
4	log_dil:Sample.Name	6	2.0	0.3	2.066 0.119
5	Residuals	15	2.4	0.2	
6					
7					
8	log_dil	***			
9	Sample.Name	***			

```
10 log_dil:Sample.Name
11 Residuals
12 ---
13 Signif. codes:
14 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Page 65: 2017-04-14. Proteome stability project: Meeting with Wai

1. gapdh for loading control
2. which samples to analyze? second run or rerun?
3. proteome search for phosphorylated/ubiquitinated database.

got 1 and 3 down, and searched for both second run and rerun.

next steps:

1. Normalize to gapdh, but I need to find the best peptide that is stable:
 - vVDLIEYVAKA; this peptide looks the most stable. And it does not align with pogo gapdh
 - Blastp sequence against pogo genome:
 2. reanalyze unfolding curves with peptide level data
-

Page 66: 2017-04-19. Proteome stability project: Data analysis

Testing to see if there are differences in the unfolding characteristics of proteins between a warm tolerant (pogos) and cool tolerant species (aphaeno).

Data related: Need to parse out the ones with enough replicates (ones = peptides)

```
1 countrep<-ddply(glob3, .  
  (Protein.Group.Accessions, Sequence, Modifications), summar  
  ize, gene=length(X36))  
2 head(countrep)  
3 [1] 1 5
```

Keep the ones with 5 (2 pogos and 3 aphaeno

```
1  
2 genes6colony<-subset(countrep, countrep$gene==5)  
3 genes6colony<-droplevels(genes6colony)  
4 dim(genes6colony)  
5 [1] 202    4  
6  
7 ##### Combining dataset based on accessions that have  
8 representative colonies for all 5.  
9 Allcombined<-glob3[glob3$Sequence %in%  
10 genes6colony$Sequence,]  
11 dim(Allcombined)  
12 [1] 1052   27
```

Change to wide format:

```

1 comlong<-
2   gather(Allcombined,temperature,unfolding,X30:X65.2)
3 dim(comlong)
4 [1] 10520     19
5 #making species and colony a factor
6 comlong$species<-as.factor(comlong$species)
7 comlong$colony<-as.factor(comlong$colony)

```

Making the temperature actually temperatures:

```

1 comlong$temperature<-
2   as.numeric(sapply(strsplit(comlong$temperature,"X"), '[',
3   2))

```

Fitting unfolding curves for each colony and for each protein(Accession) and modification

```

1
2 paramfit<-ddply(comlong,.
3   (species,colony,Protein.Group.Accessions,Sequence),fail
4   with(f=UFFfun))
5
6 head(paramfit)
7 summary(paramfit$colony)
8 species colony Protein.Group.Accessions
9 1 Aphaenogaster AR13          365266869
10 2 Aphaenogaster AR13         365266869
11 3 Aphaenogaster AR13         365266869
12 4 Aphaenogaster AR13        769831122

```

```

11 5 Aphaenogaster AR13 769831122
12 6 Aphaenogaster AR13 769831122
13 Sequence Estimate Std. Error
14 1 AYQEAFDIAK 0.22414403 0.08297519
15 2 AYQEAFDIAK 41.73987807 1.78323939
16 3 AYQEAFDIAK 0.04075972 0.09419904
17 4 SLPFYTEVLGMTLLQK 0.23564075 0.09353631
18 5 SLPFYTEVLGMTLLQK 42.24986247 1.80218447
19 6 SLPFYTEVLGMTLLQK 0.05660541 0.09785435
20 t value Pr(>|t|) param
21 1 2.7013381 3.057976e-02 slope
22 2 23.4067722 6.595211e-08 Tm
23 3 0.4326978 6.782526e-01 min
24 4 2.5192435 3.985459e-02 slope
25 5 23.4436947 6.523650e-08 Tm
26 6 0.5784660 5.810753e-01 min
27
28 ## Count the colonies
29 summary(paramfit$colony)
30 AR13 ARY D1 P45 P53
31 537 384 561 531 573

```

Only compare the ones with fits for each colony

```

1 num<-ddply(paramfit,.
  (Protein.Group.Accessions,Sequence),summarize,length(co
lony))
2 reppep<-subset(num,num$..1==15)
3
4 dim(reppep)
5 [1] 102    3
6
7 parparamfit<-paramfit[paramfit$Sequence %in%
  reppep$Sequence,]
8 dim(parparamfit)
9 [1] 1530    9
10
11 ##original
12 dim(paramfit)
13 [1] 2586    9

```

Grabbing the parameters:

```

1 slope<-subset(parparamfit,parparamfit$param=="slope")
2 Tm<-subset(parparamfit,parparamfit$param=="Tm")
3 min<-subset(parparamfit,parparamfit$param=="min")

```

Statistics: Fitting anova for every peptide

First, wrote an anova function to grab p values

```

1 aovfit <- function(dat) {
2   aov_fit <- summary(aov(Estimate~species, dat))[[1]]
3   [[ "Pr(>F)"]][1]
}

```

Implement the fit with ddply

```

1 li<-ddply(paramfit, .
  (Sequence), summarize, length=length(colony));dim(li)
2 IDs4mod<-subset(li,li$length==15)
3 dim(li)
4 [1] 200    2
5 > IDs4mod<-subset(li,li$length==15)
6 > dim(IDs4mod)
7 [1] 102    2
8
9 ### actual fit
10 totalmod<-ddply(fitdat, .
  (Sequence,param),failwith(f=aovfit))

```

Significant differences (Note: peptide level analysis)

No FDR

```

1 totalmod[which(totalmod$V1<0.05), ]
          Sequence param      V1
2 126    INVYYNEASGGK      Tm 0.03304835
3 132    IQPVLTSGAGVVTNR   Tm 0.03132238
5 261    TVLIMELINNVAK    Tm 0.04274057

```

With FDR

```

1 newp<-p.adjust(totalmod$V1,method="hochberg")
2 newp[which(newp<0.05)]
3 numeric(0)

```

TM stats and plots

```

1 Tm<-subset(Tm,Tm$Estimate<60) # subsetting out outliers

```

```

2
3 Tmav<-ddply(Tm, .
  (species,Protein.Group.Accessions,Sequence),summarize,E
  stimate=mean(Estimate))
4 head(Tmav)
5
6 species Protein.Group.Accessions
7 1 Aphaenogaster           365266869
8 2 Aphaenogaster           769831122
9 3 Aphaenogaster           769832256
10 4 Aphaenogaster          769833958
11 5 Aphaenogaster          769834659
12 6 Aphaenogaster          769835104
13             Sequence Estimate
14 1           AYQEAFDIAK 45.18603
15 2           SLPFYTEVLGMTLLQK 47.02508
16 3           GIFIVAAK 44.87442
17 4           AKPVVSFIAGLTAPPGR 42.15132
18 5           AADTSLYVK 45.31757
19 6           LAYGTALAK 44.53871

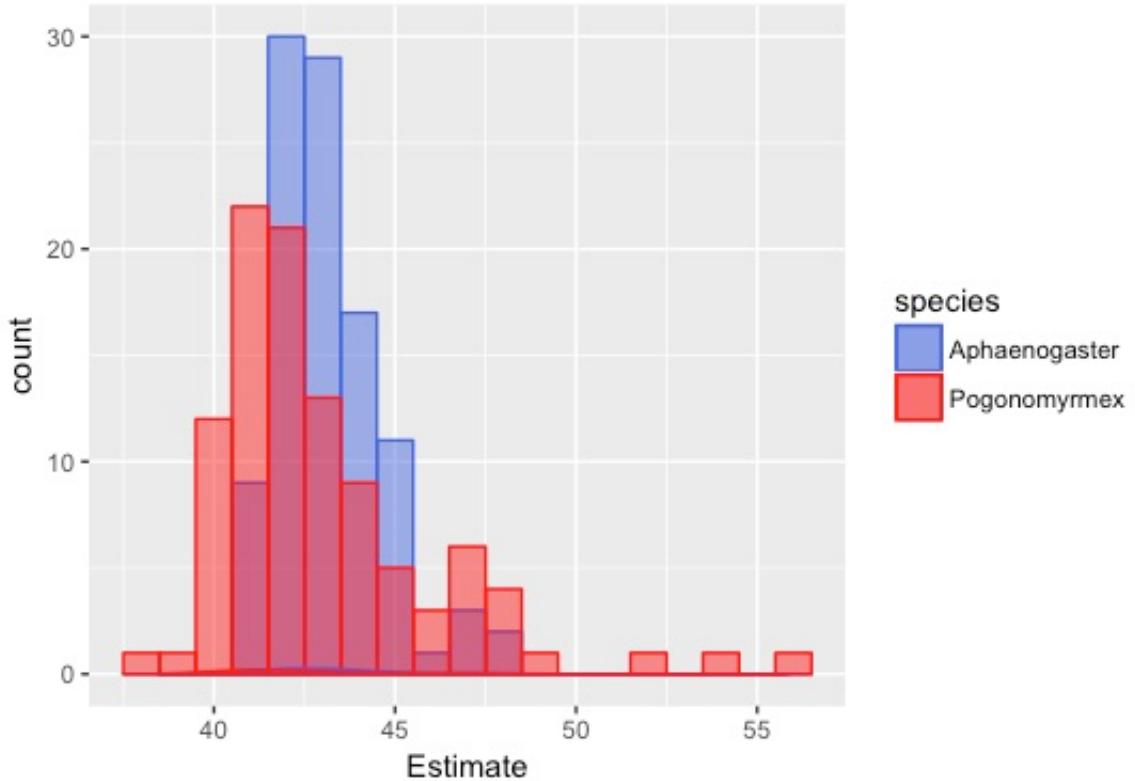
```

Plotting peptide level

```

1 ggplot(Tmav, aes(x=Estimate,colour=species,
  fill=species)) +geom_histogram(binwidth=1, alpha=.5,
  position="identity") +geom_density(alpha=0.25) +ggcol+ggco
12

```



Statistics: Testing differences in overall Tm between aphaeno and pogos.

```

1 summarySE(Tmav, measurevar =
2           "Estimate", groupvars="species")
3
4   species      N Estimate          sd        se
5   1 Aphaenogaster 102 43.13312 1.488930 0.1474261
6   2 Pogonomyrmex 101 42.95960 2.996347 0.2981477
7
8   ci
9   1 0.2924536
10  2 0.5915165
11
12 summary(aov(Estimate~species, data=Tmav))
13
14             Df Sum Sq Mean Sq F value Pr(>F)
15 species         1     1.5    1.528    0.274  0.601
16 Residuals     201 1121.7    5.581

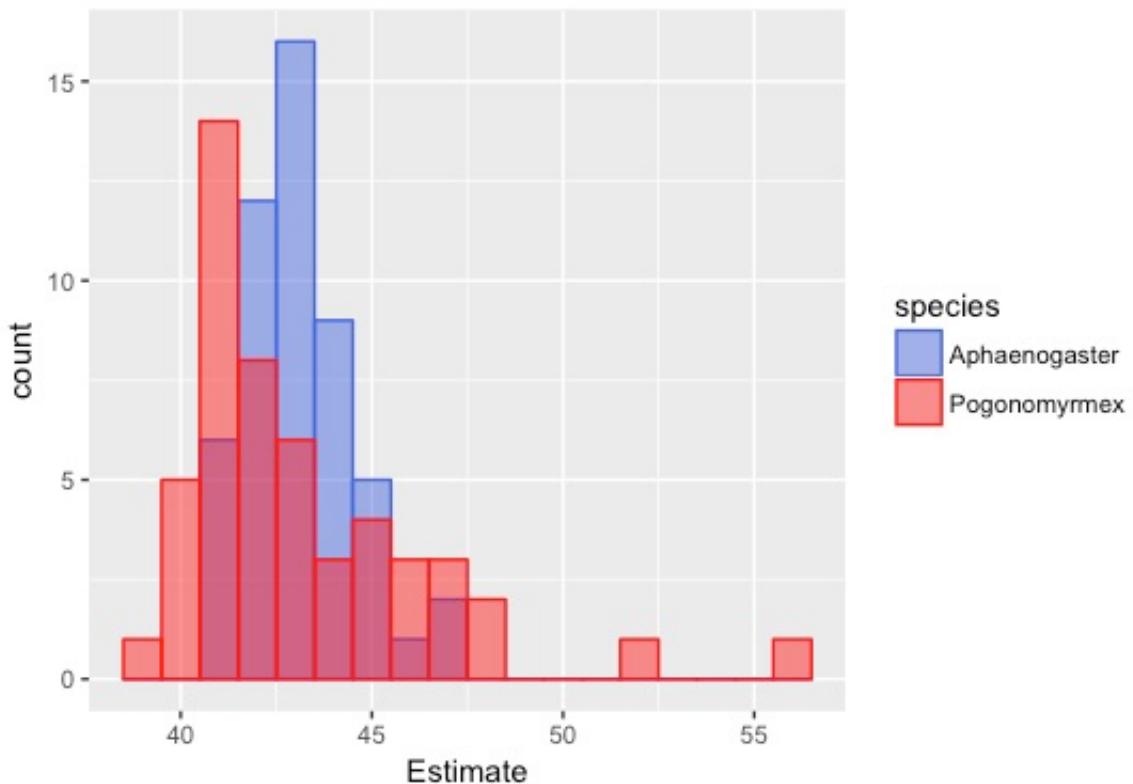
```

Plotting protein level

```

1 ggplot(protlevel, aes(x=Estimate, colour=species,
fill=species)) +geom_histogram(binwidth=1, alpha=.5,
position="identity")+ggcol+ggcol2

```



```

1 protlevel<-summarySE(Tmav,measurevar =
"Estimate",groupvars=c("species","Protein.Group.Accessio
ns"))
2
3 summary(aov(Estimate~species,data=protlevel))
   Df Sum Sq Mean Sq F value Pr(>F)
5 species     1    0.0   0.024   0.004   0.95
6 Residuals  100  596.3   5.963

```

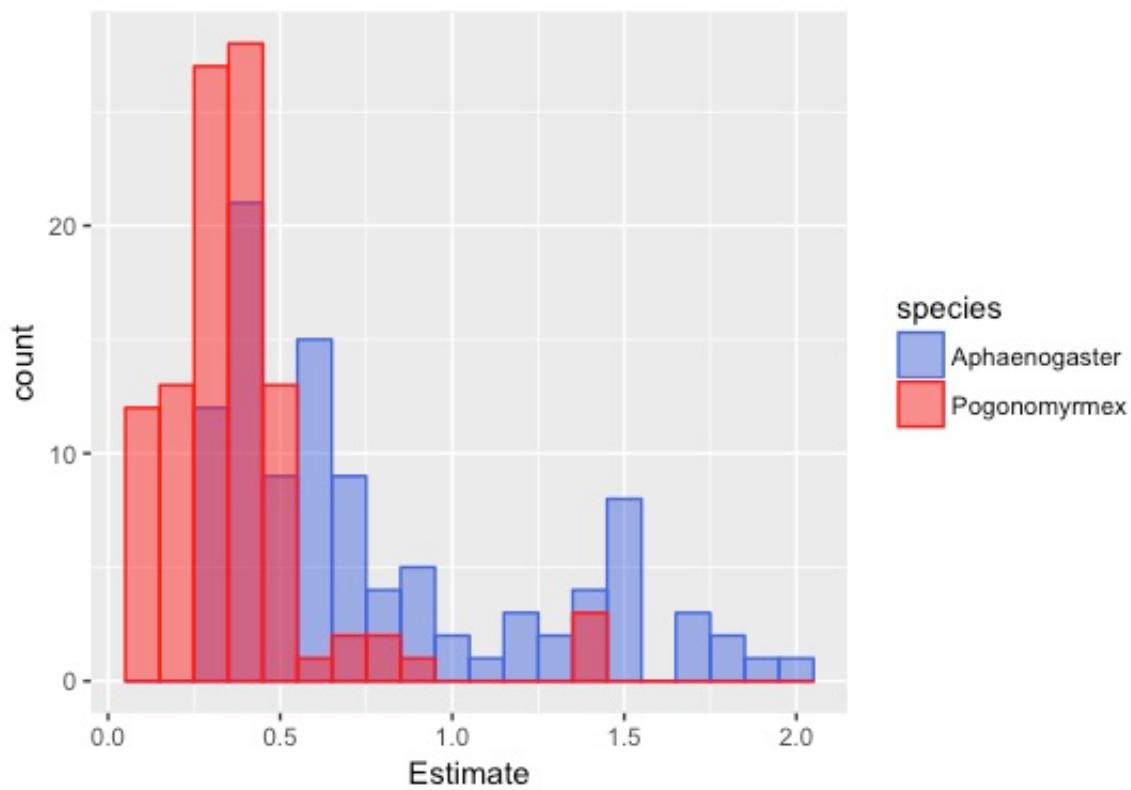
Slope stats and plots

Peptide level

```

1 slav<-ddply(slope,.
  (species,Sequence),summarize,Estimate=mean(Estimate))
2 >head(slav)
3   species           Sequence Estimate
4 1 Aphaenogaster      AADTSLYVK 0.3528855
5 2 Aphaenogaster      ADLVNNLGTIAK 0.3598534
6 3 Aphaenogaster      AGIPLNDNFVK 0.3501090
7 4 Aphaenogaster      AIDVAVK 1.0661180
8 5 Aphaenogaster      AILVDLEPGTMDSVR 0.4839904
9 6 Aphaenogaster      AKPVVSFIAGLTAPPGR 0.7463137
10
11 ggplot(slav, aes(x=Estimate,colour=species,
  fill=species)) +
12   geom_histogram(binwidth=.1, alpha=.5,
  position="identity")+ggcol+ggcol2

```

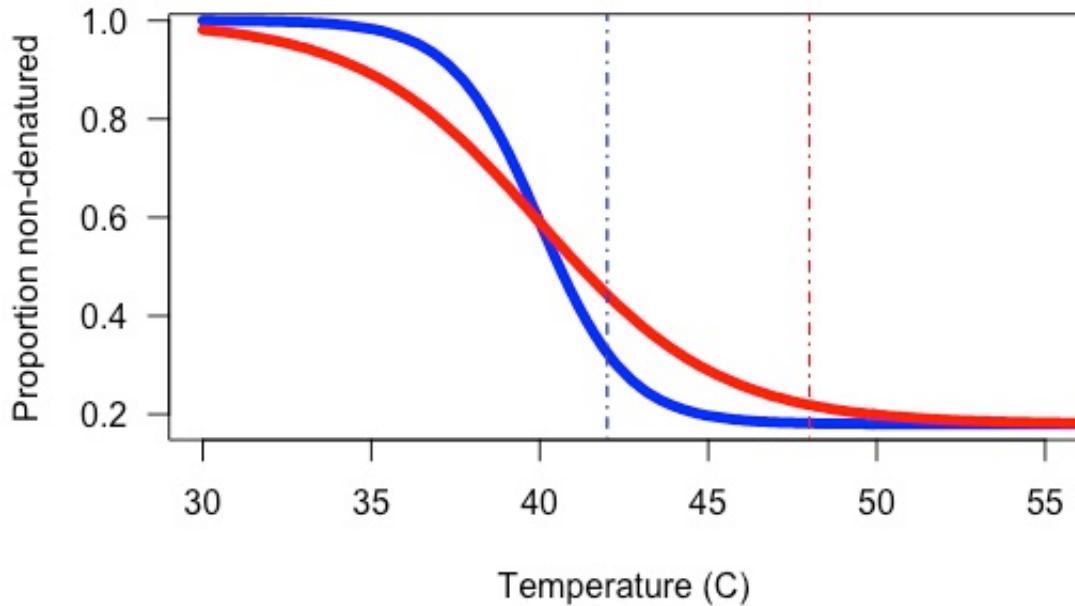


```

1 summarySE(slav,measurevar =
2           "Estimate",groupvars="species")
3             species     N   Estimate        sd
4 1 Aphaenogaster 102 0.7693194 0.4620566
5 2 Pogonomyrmex 102 0.3748694 0.2370752
6             se          ci
7 1 0.04575042 0.09075652
8 2 0.02347395 0.04656600

```

Slope differences between aphaneo (blue) and pogo (red)



Statistics: Testing overall differences in slope between aphaeno and pogo

```

1 summary(aov(Estimate~species,data=slav))
2
3             Df Sum Sq Mean Sq F value    Pr(>F)
4 species       1  7.935   7.935   58.84 7.1e-13
5 Residuals    202 27.240   0.135
6
7 species      ***
8 Residuals
9 ---
10 Signif. codes:
11 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

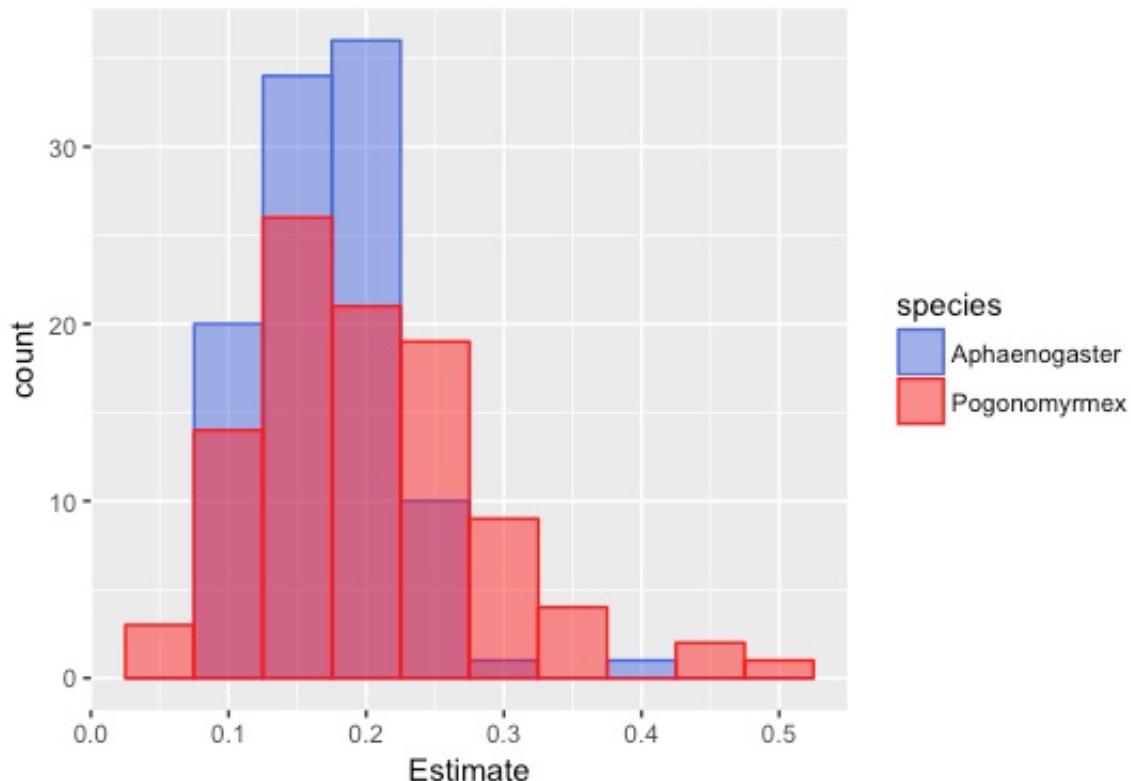
```

Min value, stats and plots

```

1 ggplot(minav, aes(x=Estimate,
2 colour=species,fill=species)) +
  geom_histogram(binwidth=.05, alpha=.5,
position="identity") + ggc1 + ggc2

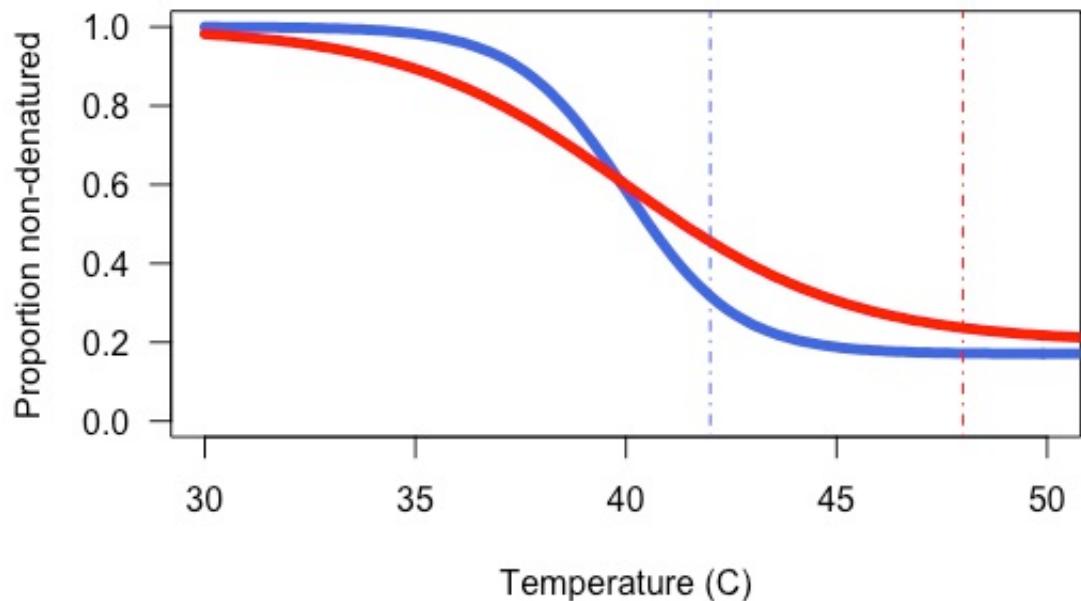
```



Statistics: Testing differences in min (overall) between aphaeno and pogo

```
1 summary(aov(Estimate~species,data=minav))
2
3             Df Sum Sq Mean Sq F value Pr(>F)
4 species       1 0.0392 0.03920   7.949 0.0053 ** 
5 Residuals    199 0.9813 0.00493
6 ---
7 Signif. codes:
8 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
9
10 summarySE(minav,measurevar =
11 "Estimate",groupvars="species")
12
13           species   N   Estimate        sd
14 1 Aphaenogaster 102 0.1704008 0.05211795
15 2 Pogonomyrmex  99 0.1983332 0.08493514
16
17           se          ci
18 1 0.005160446 0.01023694
19 2 0.008536302 0.01694001
```

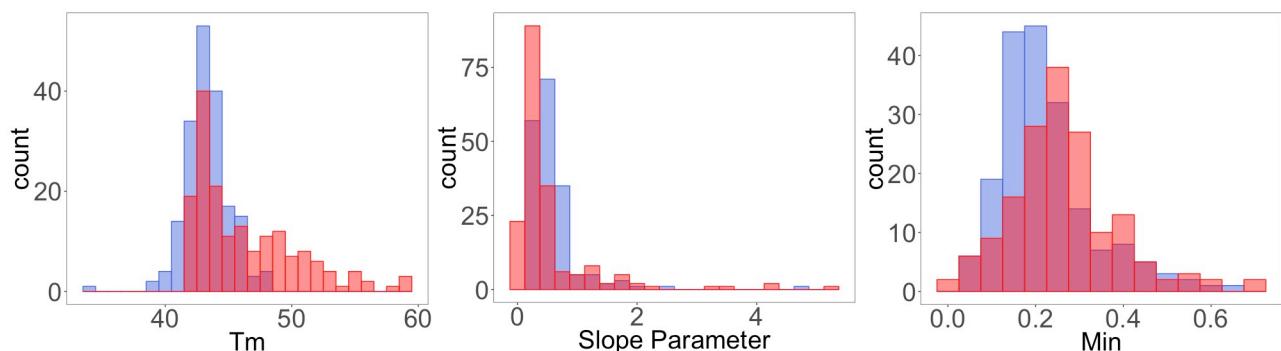
Slope and Min parameter figure both changing:



Page 67: 2017-04-20. Proteome stability project: Re-analysis with normalizing with gapdh

Previous analysis didn't control for loading (gapdh).

updated figs:



Statistics: Difference in whole proteome unfolding properties between aphaeno and pogo

Tm

```
1 summarySE(Tm,measurevar =
2 "Estimate",groupvars="species")
3             species     N Estimate          sd          se
4             ci
5 1 Aphaenogaster 187 43.29463 1.789420 0.1308553
6 0.2581514
7 2 Pogonomyrmex 171 46.40236 4.081091 0.3120889
8 0.6160686
9 > kruskal.test(Estimate~species,data=Tm)
10
11 Kruskal-Wallis rank sum test
12
13 data: Estimate by species
14 Kruskal-Wallis chi-squared = 57.497, df = 1, p-value =
15 3.385e-14
```

Slope

```

1 summarySE(slope,measurevar =
2 "Estimate",groupvars="species")
3             species     N   Estimate        sd       se
4             ci
5 1 Aphaenogaster 181 0.5841720 0.4585716 0.03408536
6 0.06725829
7 2 Pogonomyrmex 181 0.5394447 0.7583165 0.05636522
8 0.11122159
9 > kruskal.test(Estimate~species,data=slope)
10
11 Kruskal-Wallis rank sum test
12
13 data: Estimate by species
14 Kruskal-Wallis chi-squared = 41.402, df = 1, p-value =
15 1.239e-10
16
17

```

Min

```

1 summarySE(min,measurevar =
2 "Estimate",groupvars="species")
3             species     N   Estimate        sd       se
4             ci
5 1 Aphaenogaster 187 0.2240141 0.1080232 0.007899435
6 0.01558401
7 2 Pogonomyrmex 163 0.2633765 0.1235183 0.009674700
8 0.01910478
9 kruskal.test(Estimate~species,data=min)
10
11 Kruskal-Wallis rank sum test
12
13 data: Estimate by species
14 Kruskal-Wallis chi-squared = 15.451, df = 1, p-value =
15 8.469e-05
16

```

Page 68: 2017-04-21. Proteome Stability project: More data analysis; fitting curves separately to each species

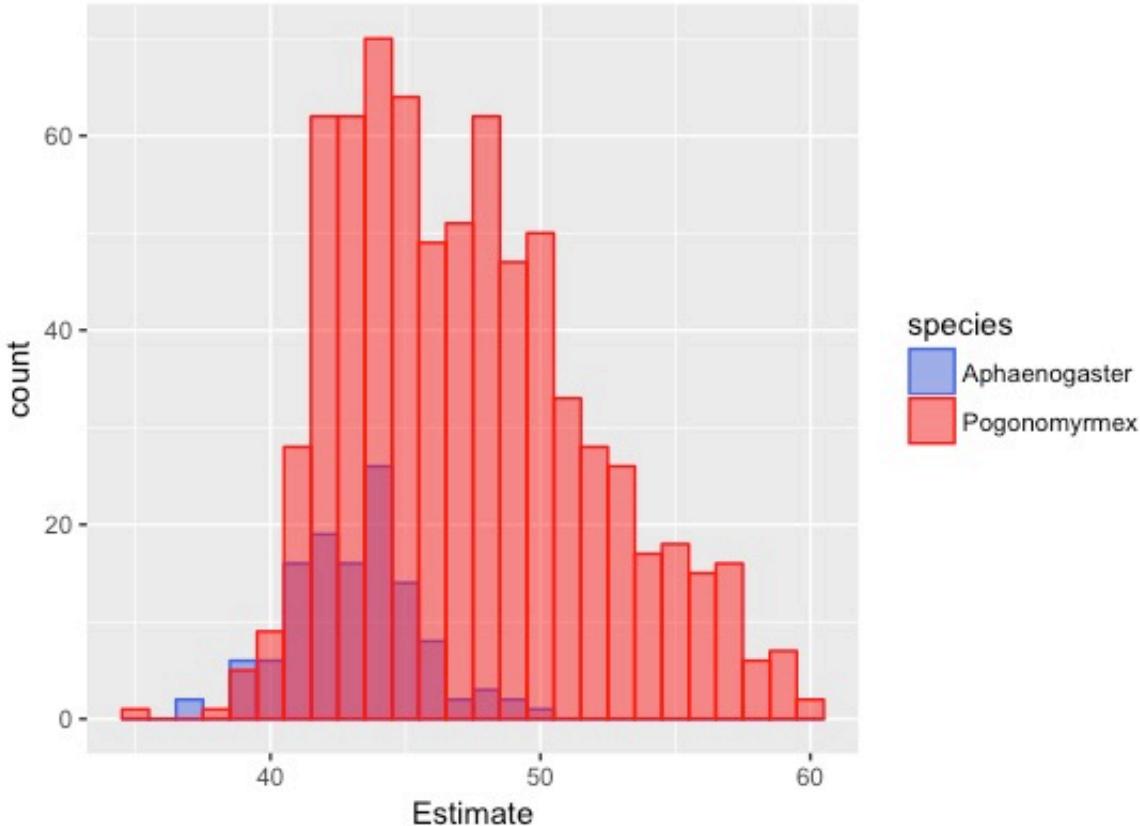
The numbers we get:

Species	CommonPeptides_2	Tm_ave	CI_95	Slope	CI_95.1	Min	CI_95.2
A. rудis	121	43.14	0.421	0.548	0.0958	0.173	0.0166
P. barbatus	729	47.31	0.329	0.455	0.0315	0.254	0.0081

Statistics: differences in overall proteome stability (Tm)

```
1 kruskal.test(Estimate~species,data=combined)
2
3 Kruskal-Wallis rank sum test
4
5 data: Estimate by species
6 Kruskal-Wallis chi-squared = 97.849, df = 1,
7 p-value < 2.2e-16
8
```

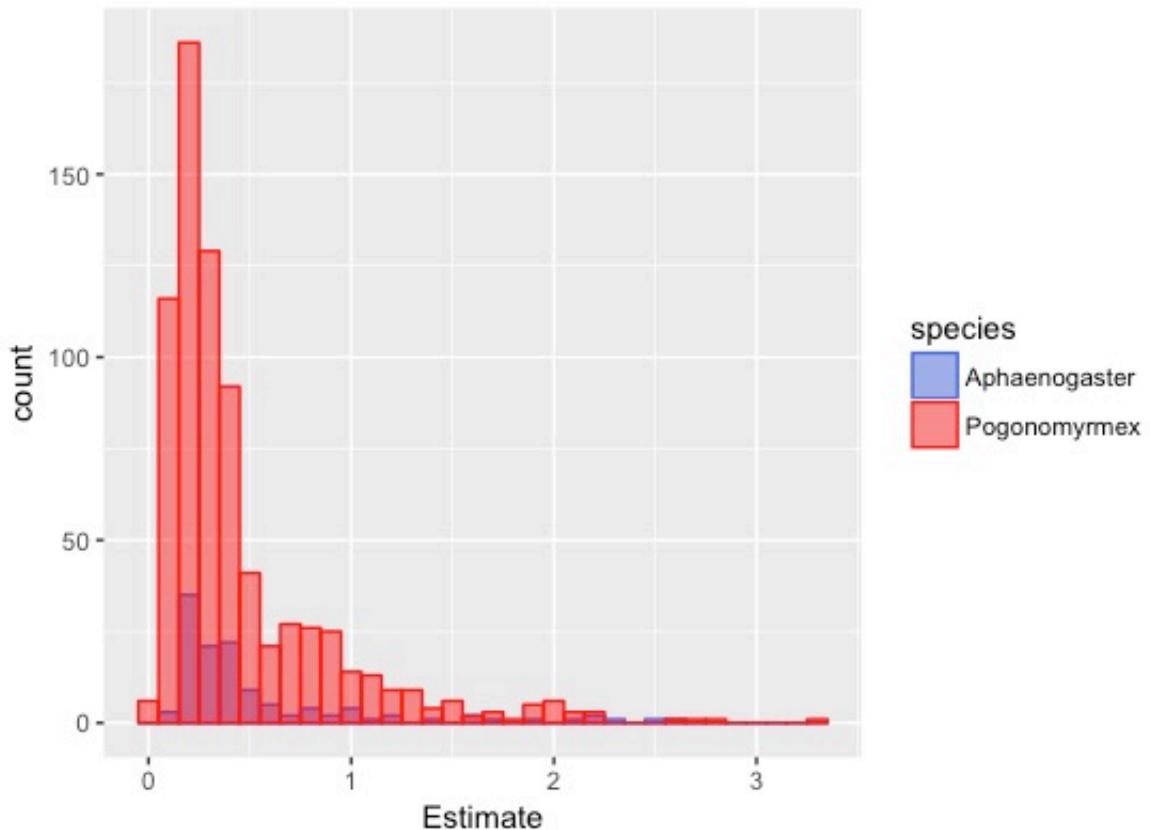
figure:



Statistics: differences in overall unfolding rate (slope)

```
1 kruskal.test(Estimate~species,data=sl.ave.spec)
2
3 Kruskal-Wallis rank sum test
4
5 data: Estimate by species
6 Kruskal-Wallis chi-squared = 1, df = 1,
7 p-value = 0.3173
```

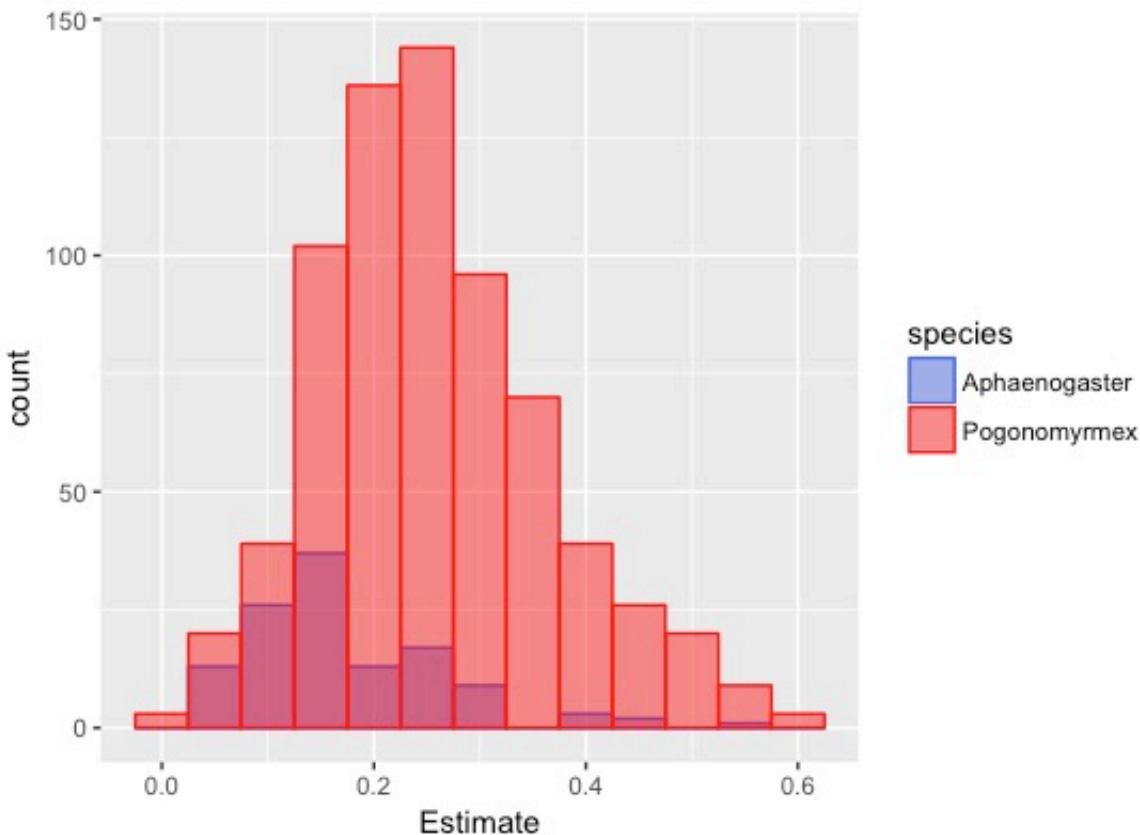
figure:



Statistics: differences in overall min unfolding

```
1 kruskal.test(Estimate~species,data=min.ave)
2
3 Kruskal-Wallis rank sum test
4
5 data: Estimate by species
6 Kruskal-Wallis chi-squared = 65.718, df = 1,
7 p-value = 5.203e-16
8
```

fig:



Summary and future directions

1. There were more peptides identified in pogo than aphaenogaster.
 - Grab fast file with ORFs from Matt Lau.
 - And he needs to send me an annotation table (with gene IDs, GO and Kegg)
 - I need to get an annotation table for pogos
2. One thing to explore is to do an enrichment analysis for proteins that have high Tms
 - Also, if I have aa sequences, I can determine contribution of
 - length
 - aa properties or IDs
3. Searched database for PTMs such as phosphorylation and ubiquitination and found no significant modifications.

Page 69: 2017-04-24. Post doc project ideas: Diapause timing and circadian clocks.

Papers I'm reading, refs:

Levy, R. C., G. M. Kozak, C. B. Wadsworth, B. S. Coates, and E. B. Dopman. 2015. Explaining the sawtooth: latitudinal periodicity in a circadian gene correlates with shifts in generation number. *Journal of Evolutionary Biology* 28:40–53.

Question? What are the molecular targets of selection in temporally diverging reproductively isolated populations?

Hypothesis: Circadian clock/rhythm genes mediate shifts in the timing of life history strategies.

Experimental approach:

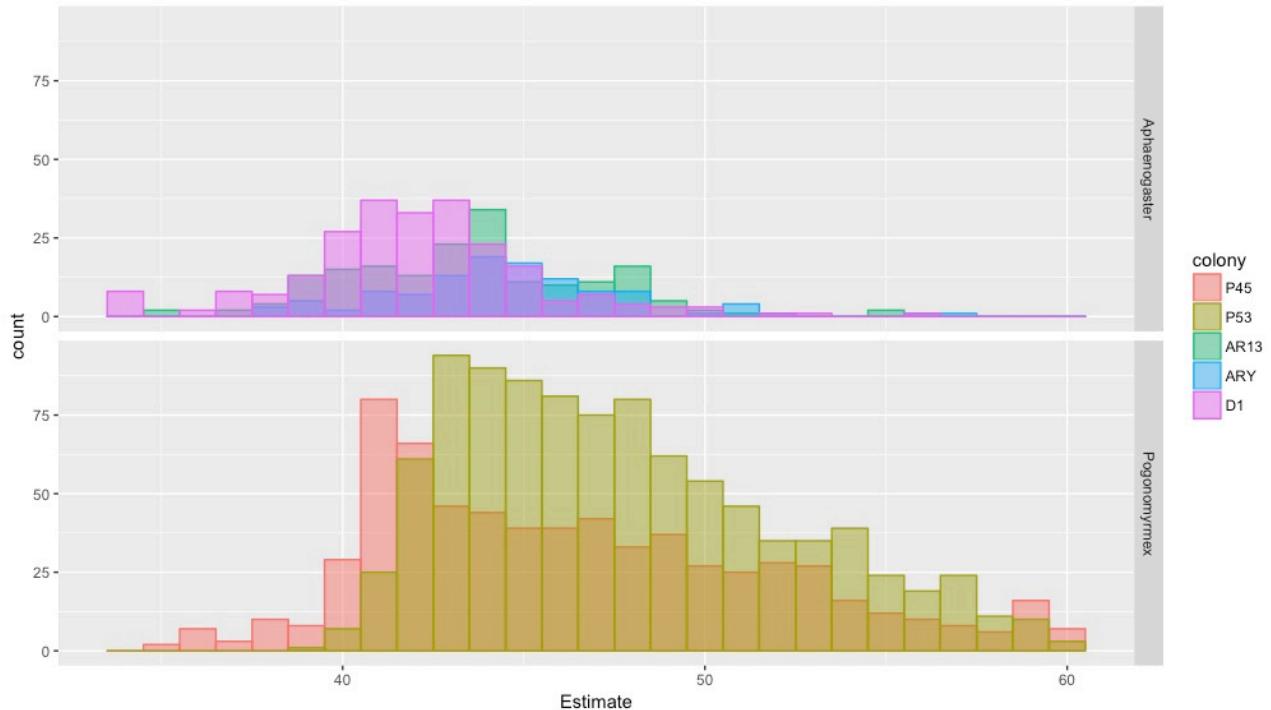
- Knockdown period, timeless, cry1 in Z and E strains with RNAi, inhibitors and compare entry, duration, and exit to diapause.
 - Which inhibitors?
- To determine if inhibitors perturb protein protein interactions, I will measure the proteome correlates of circ clocks with immunoprecipitation.

Predictions:

- Z strain (univoltine) will phenocopy E strain.
 - Both strains will delay?
-

Page 70: 2017-04-25. Proteome stability, individual colony variation and other stats

No filter for common peptides. So this figure has a mix of unique and overlapping peptides among all colonies.



Revisiting Statistics: ANOVAs for each parameter and each gene:

function I used

```
1 aovfit <- function(dat) {  
2   aov_fit <- summary(aov(Estimate~species, dat))[[1]]  
3   [[ "Pr(>F)" ]][[1]]  
4 }
```

peptide level :

```

1 totalmod<-ddply(fitdat,.
  (Sequence,param),failwith(f=aovfit))
2 totalmod[which(totalmod$V1<0.05),]
3
4   Sequence param      V1      adj
5 1       AADTSLYVK    min 0.001916678 0.2970850
6 7       AILVDLEPGTMDSVR min 0.037236005 0.9920444
7 13      ALGLPIERPK   min 0.042182366 0.9920444
8 26      APFDPPGPGTPK slope 0.034160311 0.9920444
9 42      DAGTISGLVVMR   Tm 0.045975707 0.9920444
10 61     INVYYNEASGGK   min 0.012364019 0.9920444
11 76     LLEELEQGQK    min 0.015446495 0.9920444
12 79     LPVVAATIYR    min 0.012437384 0.9920444
13 93     NPVCTDYFSGQK   Tm 0.019241119 0.9920444
14 100    SSLPEHVVK    min 0.001634925 0.2550484
15 108    TIVDITSHK    Tm 0.014540086 0.9920444
16 129    VIDPFTIKPIDAQTIIK Tm 0.006483948 0.9920444
17 142    VVPPILLETGK   min 0.018339626 0.9920444

```

protein level :

```

1 protmod1<-ddply(proteinlev,.
  (Protein.Group.Accessions,param),failwith(f=aovfit))
2 protmod1[protmod1$V1<0.05,]
3
4   Protein.Group.Accessions param      V1
5 7             769834659  min 0.001916678
6 13            769835200  min 0.015446495
7 17            769838592 slope 0.034160311
8 24            769842741   Tm 0.045975707
9 54            769850975   Tm 0.035431063
10 55            769851924  min 0.028747970
11
12 protmod1$adj<-p.adjust(protmod1$V1,method="hochberg")
13 protmod1[protmod1$adj<0.05,]
14 [1] Protein.Group.Accessions param

```

```
15 [3] v1 adj  
16 <0 rows> (or 0-length row.names)
```

It looks like there are whole proteome wide changes in stability(unfolding) that are hard to detect(we have low power maybe bc of sample sizes) at the peptide/protein level.

Page 71: 2017-04-26. Climate cascade meeting: Project updates; to do list

- **Stressed in nature MS: Samples to rerun.**
 - 2 more plates, 18s rRNA; RIN values in
 - Need to analyze data; show results next week.
 - Get the deltas too to get at the past historical temperatures they were experiencing. Do regression with deltas.
- **Proteome stability project:**
 1. There were more peptides identified in pogo than aphaenogaster.
 - Grab fast file with ORFs from Matt Lau.
 - And he needs to send me an annotation table (with gene IDs, GO and Kegg)
 - I need to get an annotation table for pogos
 2. One thing to explore is to do an enrichment analysis for proteins that have high Tms
 - Also, if I have aa sequences, I can determine contribution of

- length
 - aa properties or IDs
3. Searched database for PTMs such as phosphorylation and ubiquitination and found no significant modifications.

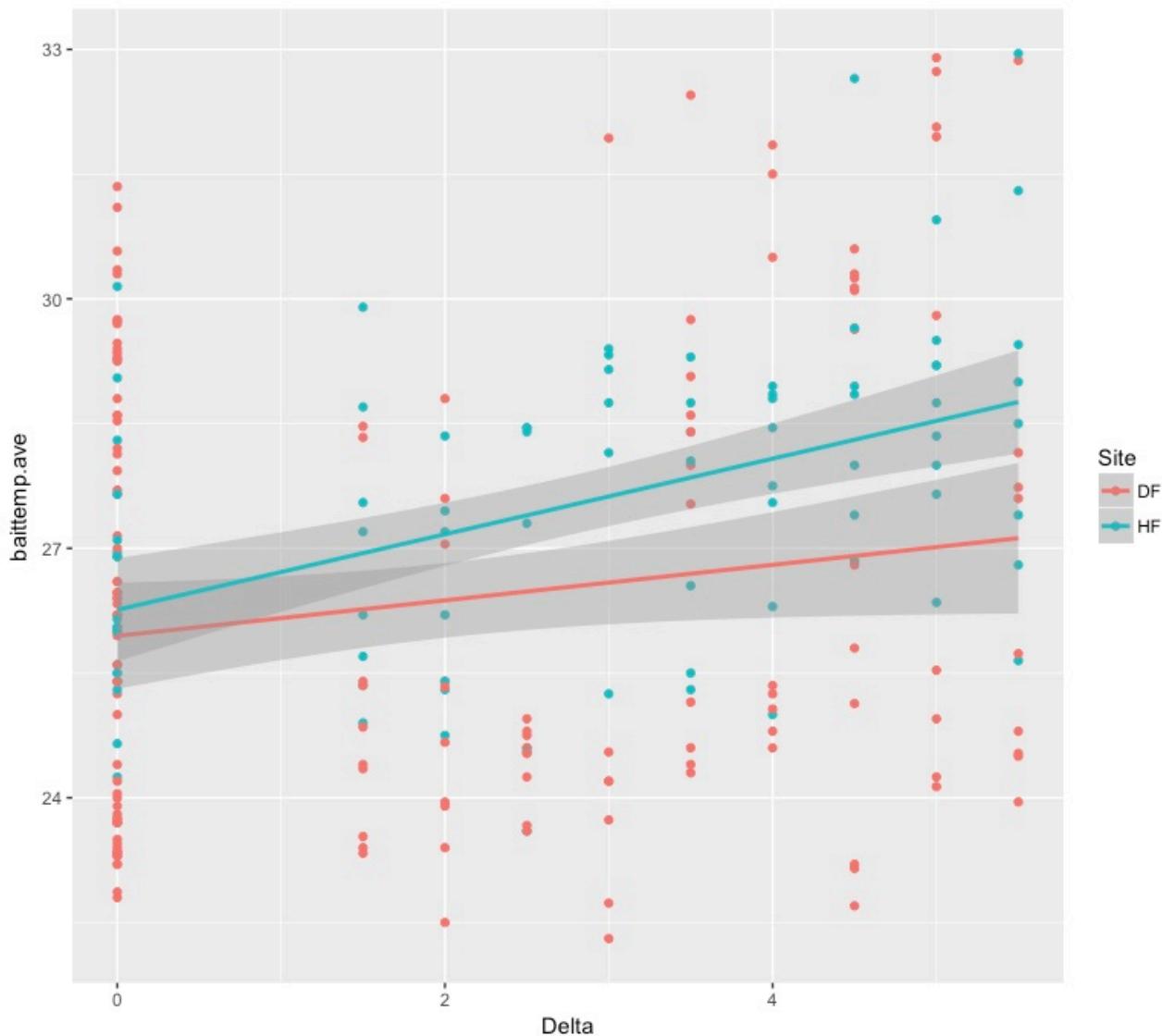
Notes from SHC and NGotelli:

1. Do a randomization test:
 - Shuffle the labels-- go to the colony level
 - Compare different medians
2. Do KS test to determine difference in cumulative density function
3. Plot density curves
4. Plot Aph Tm vs Pogo Tm
5. Create story board for manuscript ; focus on narrative and then construct figures accordingly
6. Try to analyze by block which might account for variation
intraspecific variation
7. NGotelli comment: Aph Tms are within the range of pogo Tms

Page 72: 2017-04-28. Stressed in nature project: data analysis

Quick and dirty: ants were sampled at warming chambers for 2 sites; duke forest and harvard forest.

Looking at bait temperatures against targeted delta chambers



```

1 summary(lm(baittemp.ave~Delta*Site,data=warm))
2
3 Call:
4 lm(formula = baittemp.ave ~ Delta * Site, data = warm)
5
6 Residuals:
7   Min     1Q Median     3Q    Max
8 -4.2807 -2.0306 -0.2349  1.5072  5.8855
9
10 Coefficients:
11                               Estimate Std. Error t value Pr(>|t| )

```

```

12 (Intercept) 25.94661    0.28228   91.919   <2e-16 ***
13 Delta        0.21358    0.09977   2.141    0.0333 *
14 SiteHF       0.31354    0.53701   0.584    0.5599
15 Delta:SiteHF 0.24087    0.16926   1.423    0.1561
16 ---
17 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
18
19 Residual standard error: 2.417 on 230 degrees of freedom
20 Multiple R-squared:  0.107, Adjusted R-squared:
21          0.09532
22 F-statistic: 9.183 on 3 and 230 DF,  p-value: 9.19e-06

```

The lines are not sig diff. Delta and bait temp are weakly correlated too.

Getting gene counts from [MCMC.qpcr](#)

Tutorial [here](#)

getting gene counts

```

1 gene<-c("CT_18s","CT_40","CT_70","CT_83","CT_actin") # 
  list the genes
2 amp<-data.frame(gene,efficiency=rep(2,length(gene))) #
  input efficiencies
3 names(warm)[6]<-"sample" # rename sample
4
5
6 ### getting gene counts
7 dd<-
8 cq2counts(data=warm,genecols=c(14:18),condcols=c(13,5:6
,2:4,20:21),Cq1=37,effic=amp)

```

```

8 dd$baittemp.ave<-
9   as.numeric(as.character(dd$baittemp.ave))
10 dd$RIN_Value<-as.numeric(as.character(dd$RIN_Value))
11 dd$Year_collect<-
12   as.numeric(as.character(dd$Year_collect))
13 str(dd)
14
15 'data.frame': 1135 obs. of 11 variables:
16   $ count : num 713721 1704598 2884720 444863
17   $ 1145861 ...
18   $ gene : Factor w/ 5 levels
19   "CT_18s", "CT_40", ...
20   $ Isolation.Date: Factor w/ 29 levels
21   "20150624", "20150710", ...
22   $ Cham : Factor w/ 15 levels
23   "1", "2", "3", "4", ...
24   $ Sample : Factor w/ 3 levels "1", "2", "3": 2 2
25   2 2 2 3 2 2 3 3 ...
26   $ Year_collect : num 2014 2014 2014 2013 2014 ...
27   $ Site : Factor w/ 2 levels "DF", "HF": 1 1
28   1 1 2 1 1 1 2 ...
29   $ Vial.me : Factor w/ 238 levels "", "DF 1.1", "DF
30   1.2", ...
31   $ RIN_Value : num 7 6 5.5 1 2.9 1 1 1.1 2 1 ...
32   $ baittemp.ave : num 28.6 28.1 28.6 29.8 32.7 ...
33   $ Siteyear : chr "DF 2014" "DF 2014" "DF 2014"
34   "DF 2013" ...

```

Statistics: Fitting an ANOVA to test interaction between gene, year, baittemp, year collected, and delta.

```

1 fullmod<-
2   aov(log(count+1)~RIN_Value+gene*Site*Year_collect*baittemp.ave*Delta,data=dd)
3
4 Step: AIC=1508.51
5 log(count + 1) ~ RIN_Value + gene + Site + Year_collect +
6   + baittemp.ave +
7     Delta + gene:Site + gene:Year_collect +
8     Site:Year_collect +
9       gene:baittemp.ave + Site:baittemp.ave +
10      Year_collect:baittemp.ave +
11        Site:Delta + Year_collect:Delta +
12        baittemp.ave:Delta + gene:Site:Year_collect +
13          gene:Year_collect:baittemp.ave +
14          Site:Year_collect:baittemp.ave +
15            Site:Year_collect:Delta + Site:baittemp.ave:Delta
16
17
18
19
```

	Df	Sum of Sq	RSS
AIC			
<none>		3678.1	
1508.5			
- Site:Year_collect:baittemp.ave	1	8.487	3686.6
1509.1			
- gene:Site:Year_collect	4	31.985	3710.1
1510.3			
- Site:Year_collect:Delta	9	74.585	3752.7
1513.3			
+ Year_collect:baittemp.ave:Delta	9	42.338	3635.8
1513.4			
+ gene:Site:baittemp.ave	4	5.817	3672.3
1514.7			
- gene:Year_collect:baittemp.ave	4	57.795	3735.9
1518.2			

		1	58.863	3737.0
20	- RIN_Value	1	58.863	3737.0
	1524.5			
21	- Site:baittemp.ave:Delta	9	213.129	3891.3
	1554.4			
22	+ gene:Delta	36	53.133	3625.0
	1564.0			
23		Df	Sum Sq	Mean Sq F
	value Pr(>F)			
24	RIN_Value	1	2509	2509
	714.820 < 2e-16 ***			
25	gene	4	20041	5010
	1427.529 < 2e-16 ***			
26	Site	1	18	18
	5.027 0.025162 *			
27	Year_collect	1	663	663
	188.912 < 2e-16 ***			
28	baittemp.ave	1	2	2
	0.544 0.460803			
29	Delta	9	103	11
	3.261 0.000637 ***			
30	gene:Site	4	71	18
	5.031 0.000511 ***			
31	gene:Year_collect	4	221	55
	15.734 1.64e-12 ***			
32	Site:Year_collect	1	64	64
	18.251 2.11e-05 ***			
33	gene:baittemp.ave	4	150	38
	10.706 1.66e-08 ***			
34	Site:baittemp.ave	1	136	136
	38.667 7.25e-10 ***			
35	Year_collect:baittemp.ave	1	973	973
	277.328 < 2e-16 ***			
36	Site:Delta	9	184	20
	5.819 6.19e-08 ***			

```

37 Year_collect:Delta          9    273     30
      8.630 1.46e-12 ***
38 baittemp.ave:Delta          9    146     16
      4.613 5.32e-06 ***
39 gene:Site:Year_collect     4     19      5
      1.342 0.252392
40 gene:Year_collect:baittemp.ave 4     59      15
      4.187 0.002279 **
41 Site:Year_collect:baittemp.ave 1     95      95
      27.031 2.41e-07 ***
42 Site:Year_collect:Delta      9     61      7
      1.946 0.042393 *
43 Site:baittemp.ave:Delta      9    213     24
      6.747 1.89e-09 ***
44 Residuals                   1048   3678     4
45 ---
46 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

model selection with dredge() function in MuMin package

can't do full model:

```

1 fullmod2<-
  aov(log(count+1)~RIN_Value+gene+Site*Year_collect*baittemp.ave,data=dd)
2 a<-dredge(fullmod2)

```

```

1 get.models(a,subset=delta<4)
2
3 summary(aov(formula = log(count + 1) ~ baittemp.ave +
  gene + RIN_Value +

```

```

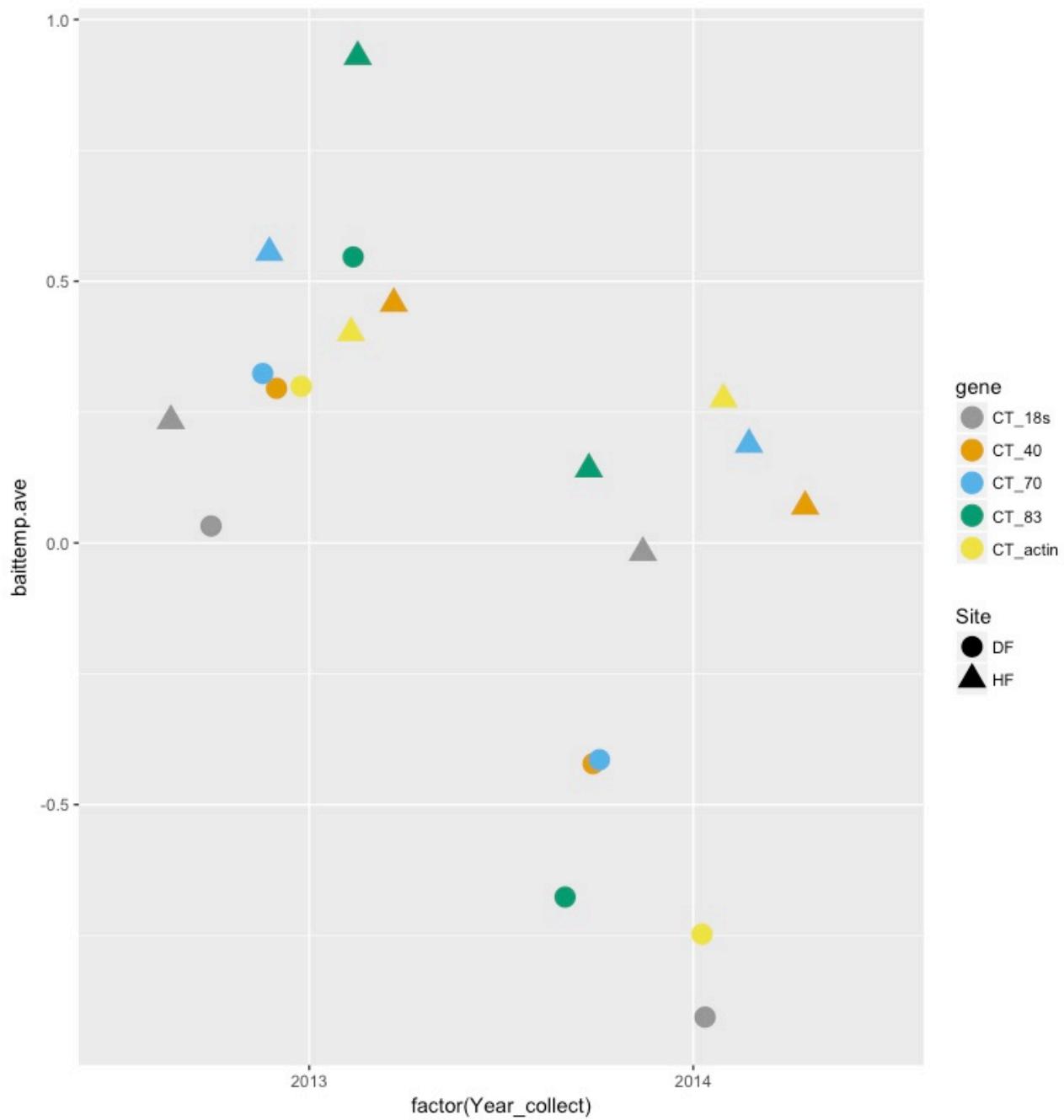
4   Site + Year_collect + baittemp.ave:Site +
baittemp.ave:Year_collect +
      Site:Year_collect + baittemp.ave:Site:Year_collect
+ 1, data = dd))

6
7   Df Sum Sq Mean Sq F value    Pr(>F)
8 baittemp.ave                               1     62      62
13.439 0.000258 ***
9 gene                                      4 19917     4979
10 1072.077 < 2e-16 ***
10 RIN_Value                                1    2579     2579
11 555.191 < 2e-16 ***
11 Site                                     1     24      24
12 5.262 0.021984 *
12 Year_collect                             1     650     650
13 139.901 < 2e-16 ***
13 baittemp.ave:Site                        1     126     126
27.070 2.33e-07 ***
14 baittemp.ave:Year_collect                1     838     838
180.393 < 2e-16 ***
15 Site:Year_collect                       1     171     171
36.839 1.75e-09 ***
16 baittemp.ave:Site:Year_collect          1     100     100
21.439 4.08e-06 ***
17 Residuals                                1122    5211      5

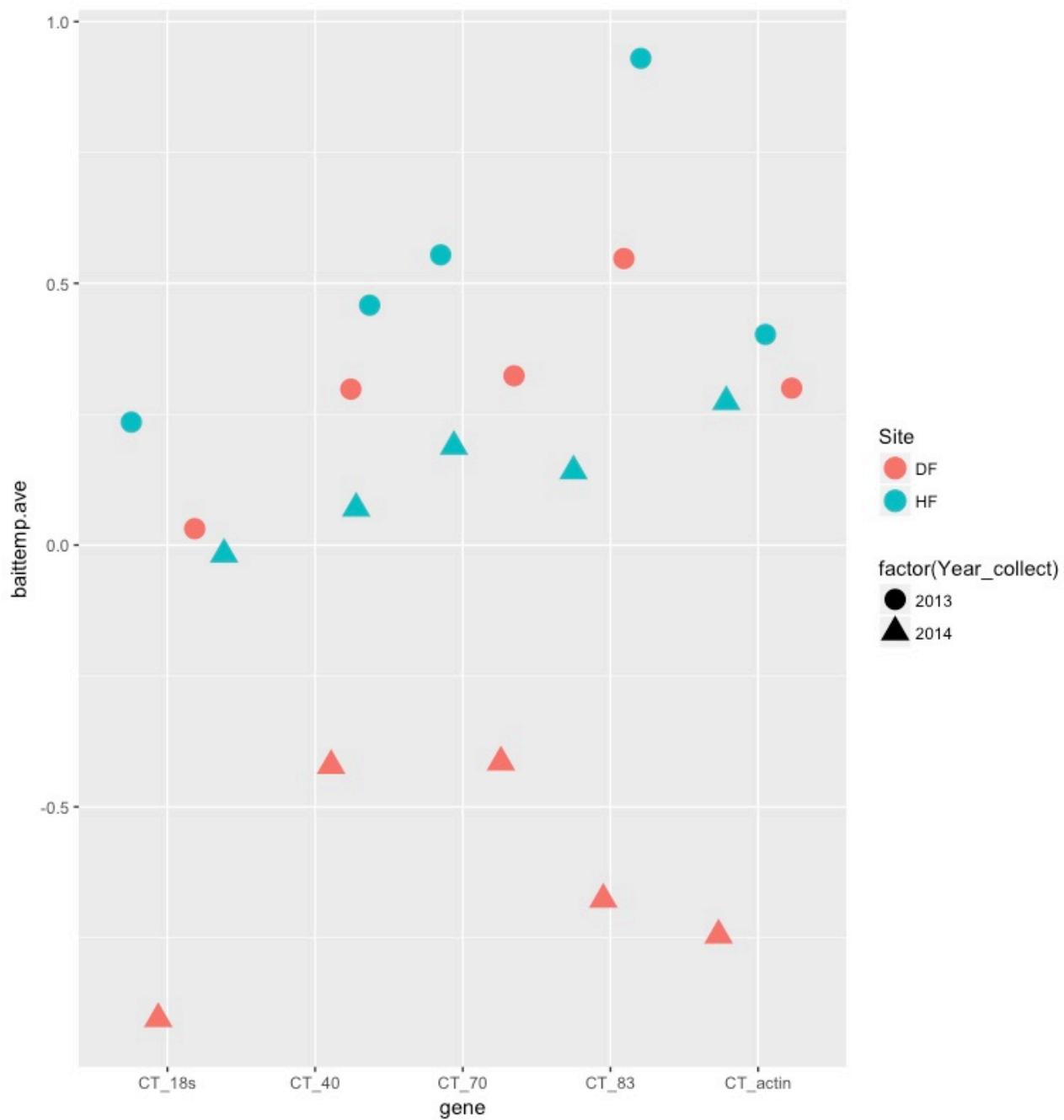
```

Plotting:

Estimated slope between bait temp and gene count, against years;
genes are colored, site are different symbols



Estimated slope between bait temp and gene count, against genes



Statistics: 18s

```

1 mod5<-
  aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.a
  ve*Delta,data=gxp18s)
2
3 Step:  AIC=494.96

```

```

4 log(count + 1) ~ Site + Year_collect + baittemp.ave +
Delta +
Site:Year_collect + Site:baittemp.ave +
Year_collect:baittemp.ave +
Site:Delta + Year_collect:Delta +
baittemp.ave:Delta + Site:Year_collect:baittemp.ave

7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

```

		Df	Sum of Sq	RSS	
AIC					
<none>				1782.5	
494.96					
+ RIN_Value	1	14.778	1767.7		
495.05					
- baittemp.ave:Delta	1	19.913	1802.4		
495.52					
- Site:Year_collect:baittemp.ave	1	20.865	1803.3		
495.64					
+ Year_collect:baittemp.ave:Delta	1	5.637	1776.8		
496.24					
+ Site:Year_collect:Delta	1	3.601	1778.9		
496.50					
- Year_collect:Delta	1	28.524	1811.0		
496.62					
+ Site:baittemp.ave:Delta	1	0.017	1782.4		
496.96					
- Site:Delta	1	64.748	1847.2		
501.17					
		Df	Sum Sq	Mean Sq	F
value	Pr(>F)				
Site		1	111.4	111.4	
13.620	0.000283 ***				
Year_collect		1	906.3	906.3	
110.844	< 2e-16 ***				
baittemp.ave		1	73.4	73.4	
8.982	0.003042 **				

22	Delta		1	0.0	0.0
	0.001 0.975781				
23	Site:Year_collect		1	2.1	2.1
	0.258 0.612198				
24	Site:baittemp.ave		1	58.5	58.5
	7.152 0.008053 **				
25	Year_collect:baittemp.ave		1	212.9	212.9
	26.033 7.29e-07 ***				
26	Site:Delta		1	65.0	65.0
	7.944 0.005269 **				
27	Year_collect:Delta		1	50.0	50.0
	6.116 0.014160 *				
28	baittemp.ave:Delta		1	19.1	19.1
	2.330 0.128354				
29	Site:Year_collect:baittemp.ave		1	20.9	20.9
	2.552 0.111612				
30	Residuals		218	1782.5	8.2
31	---				
32	Signif. codes:	0 *** 0.001 ** 0.01 * 0.05 .			
	0.1 ' ' 1				

Statistics: actin

```

1 mod4<-
2 aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.a
ve*Delta,data=gxpactin)
3 Start: AIC=295.09
4 log(count + 1) ~ RIN_Value + Site * Year_collect *
baittemp.ave *
5 Delta
6
7 Df Sum of Sq
8 RSS      AIC

```

```

7 <none>
    717.50 295.08
8 - Site:Year_collect:baittemp.ave:Delta 1     8.0973
    725.60 295.62
9 - RIN_Value                               1   25.0229
    742.53 300.83

10
    Df Sum Sq Mean Sq
F value   Pr(>F)
11 RIN_Value                               1   429.7   429.7
    125.168 < 2e-16 ***
12 Site                                    1     9.8     9.8
    2.865 0.09203 .
13 Year_collect                           1   110.3   110.3
    32.137 4.73e-08 ***
14 baittemp.ave                            1     3.2     3.2
    0.936 0.33439
15 Delta                                   1     0.0     0.0
    0.008 0.92832
16 Site:Year_collect                      1   23.4   23.4
    6.824 0.00965 **
17 Site:baittemp.ave                      1   19.5   19.5
    5.692 0.01793 *
18 Year_collect:baittemp.ave              1   224.0   224.0
    65.243 5.21e-14 ***
19 Site:Delta                             1     2.2     2.2
    0.649 0.42141
20 Year_collect:Delta                     1   27.2   27.2
    7.916 0.00537 **
21 baittemp.ave:Delta                    1     1.0     1.0
    0.299 0.58478
22 Site:Year_collect:baittemp.ave       1   37.2   37.2
    10.835 0.00117 **
23 Site:Year_collect:Delta               1     1.6     1.6
    0.479 0.48975

```

24	Site:baittemp.ave:Delta	1	0.0	0.0
	0.003 0.95749			
25	Year_collect:baittemp.ave:Delta	1	0.0	0.0
	0.011 0.91665			
26	Site:Year_collect:baittemp.ave:Delta	1	8.1	8.1
	2.359 0.12610			
27	Residuals	209	717.5	3.4
28	---			

Statistics: hsp70

1	mod1<-			
	aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.ave*Delta,data=gxp70)			
2				
3	Step: AIC=200.41			
4	log(count + 1) ~ RIN_Value + Site + Year_collect +			
	baittemp.ave +			
	Delta + Site:Year_collect + Site:baittemp.ave +			
	Year_collect:baittemp.ave +			
	Site:Delta + Year_collect:Delta +			
	Site:Year_collect:Delta			
7				
8		Df	Sum of Sq	RSS
	AIC			
9	<none>			494.75
	200.41			
10	- Site:baittemp.ave	1	4.583	499.33
	200.52			
11	+ baittemp.ave:Delta	1	0.689	494.06
	202.09			
12	+ Site:Year_collect:baittemp.ave	1	0.002	494.75
	202.40			

```

13 - Site:Year_collect:Delta           1    15.443 510.19
205.44
14 - RIN_Value                      1    32.820 527.57
213.11
15 - Year_collect:baittemp.ave     1    152.325 647.07
259.87
16
Df Sum Sq Mean Sq F value
Pr(>F)
17 RIN_Value                         1    353.3 353.3 154.963 <
2e-16 ***
18 Site                            1    12.4   12.4   5.454
0.020431 *
19 Year_collect                     1    27.1   27.1   11.898
0.000675 ***
20 baittemp.ave                     1    7.9    7.9    3.449
0.064641 .
21 Delta                           1    0.1    0.1    0.053
0.818551
22 Site:Year_collect               1    0.6    0.6    0.263
0.608702
23 Site:baittemp.ave               1    14.8   14.8   6.502
0.011462 *
24 Year_collect:baittemp.ave      1    120.5 120.5 52.858
6.41e-12 ***
25 Site:Delta                        1    1.2    1.2    0.541
0.463016
26 Year_collect:Delta              1    26.7   26.7   11.696
0.000748 ***
27 Site:Year_collect:Delta         1    15.4   15.4   6.773
0.009890 **
28 Residuals                       217  494.7  2.3
29 ---

```

dredging:

```

1 mod1<-
2   aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.a
3   ve*Delta,data=gxp70)
4 h70d<-dredge(mod1)
5 summary(model.avg(h70d,delta<4))
6 (conditional average)
7
8
9
10
11
12
13
14
15
16
17
18
```

		Estimate	Std. Error
Adjusted SE z value			
(Intercept)		-4.151e+04	5.773e+03
5.800e+03	7.157		
baittemp.ave		1.506e+03	2.155e+02
2.166e+02	6.954		
Delta		-2.422e+02	1.055e+03
1.058e+03	0.229		
RIN_Value		2.611e-01	6.662e-02
6.697e-02	3.899		
SiteHF		2.422e+03	6.509e+03
6.533e+03	0.371		
Year_collect		2.062e+01	2.867e+00
2.881e+00	7.158		
baittemp.ave:Year_collect		-7.480e-01	1.071e-01
1.076e-01	6.954		
Delta:SiteHF		-1.244e+03	4.719e+02
4.745e+02	2.622		
Delta:Year_collect		1.202e-01	5.238e-01
5.254e-01	0.229		
SiteHF:Year_collect		-1.305e+00	3.347e+00
3.359e+00	0.389		
Delta:SiteHF:Year_collect		6.179e-01	2.344e-01
2.357e-01	2.622		
baittemp.ave:SiteHF		-6.104e+01	3.187e+02
3.198e+02	0.191		
baittemp.ave:Delta		-2.946e+01	6.526e+01
6.549e+01	0.450		

```

19 baittemp.ave:SiteHF:Year_collect 1.456e-01 3.212e-01
      3.225e-01 0.451
20 baittemp.ave:Delta:Year_collect 4.222e-02 4.320e-02
      4.344e-02 0.972
21 Pr(>|z|)
22 (Intercept) < 2e-16 ***
23 baittemp.ave < 2e-16 ***
24 Delta 0.81891
25 RIN_Value 9.66e-05 ***
26 SiteHF 0.71089
27 Year_collect < 2e-16 ***
28 baittemp.ave:Year_collect < 2e-16 ***
29 Delta:SiteHF 0.00875 **
30 Delta:Year_collect 0.81900
31 SiteHF:Year_collect 0.69760
32 Delta:SiteHF:Year_collect 0.00875 **
33 baittemp.ave:SiteHF 0.84862
34 baittemp.ave:Delta 0.65281
35 baittemp.ave:SiteHF:Year_collect 0.65174
36 baittemp.ave:Delta:Year_collect 0.33112
37 ---
38 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
      0.1 ' ' 1
39
40 Relative variable importance:
41                               baittemp.ave Delta RIN_Value Site
42 Year_collect
43 Importance: 1.00 1.00 1.00 1.00
44 N containing models: 10 10 10 10
45 baittemp.ave:Year_collect
        Delta:Year_collect
46 Importance: 1.00 1.00

```

```
46 N containing models: 10 10

47 Site:Year_collect Delta:Site
Delta:Site:Year_collect
48 Importance: 0.92 0.81 0.81

49 N containing models: 9 7 7

50 baittemp.ave:Site
baittemp.ave:Delta
51 Importance: 0.61 0.28

52 N containing models: 7 4

53 baittemp.ave:Site:Year_collect
54 Importance: 0.13
55 N containing models: 2
56 baittemp.ave:Delta:Year_collect
57 Importance: 0.10
58 N containing models: 2
```

fig showing the bait temp by year interaction

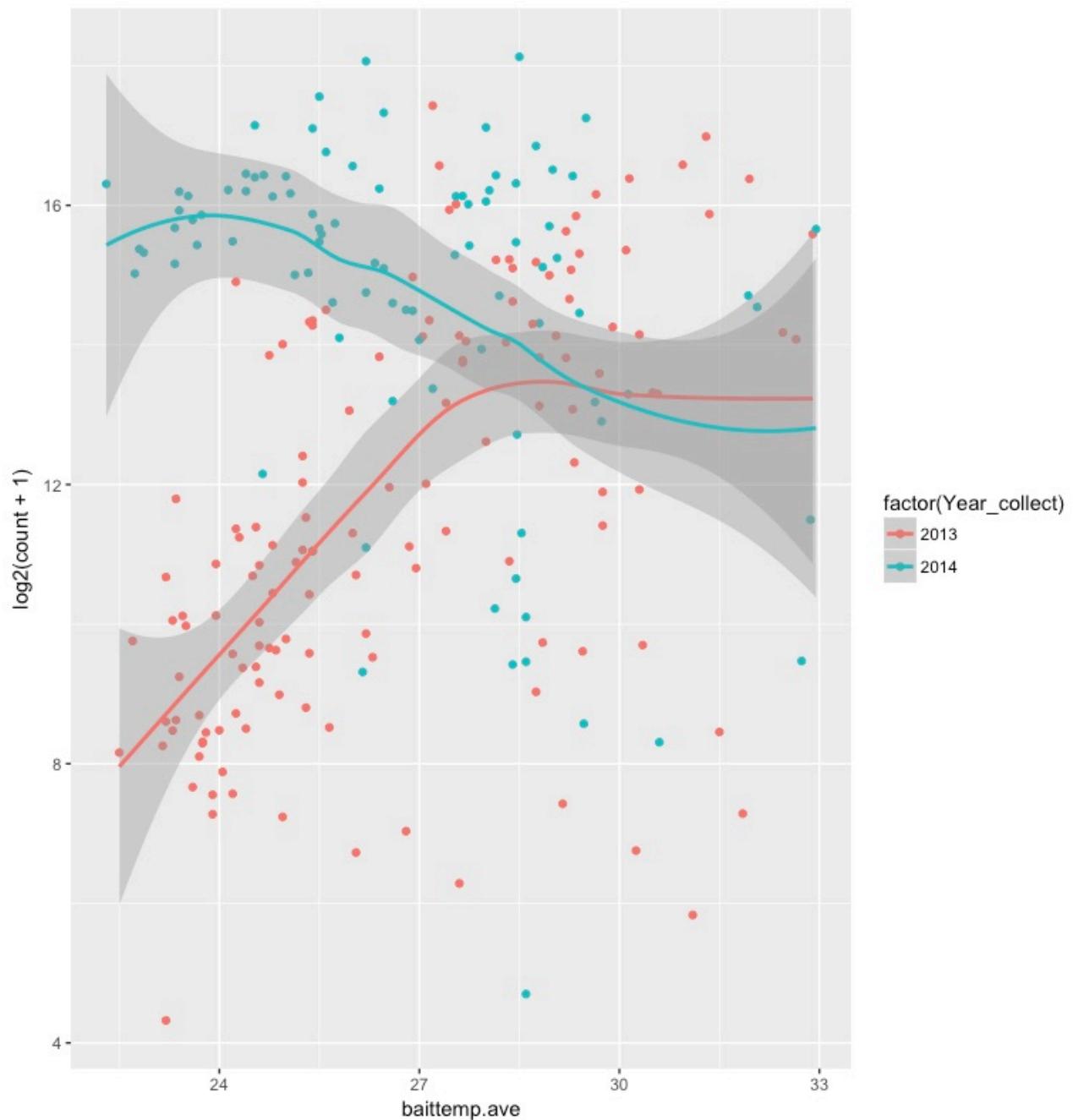
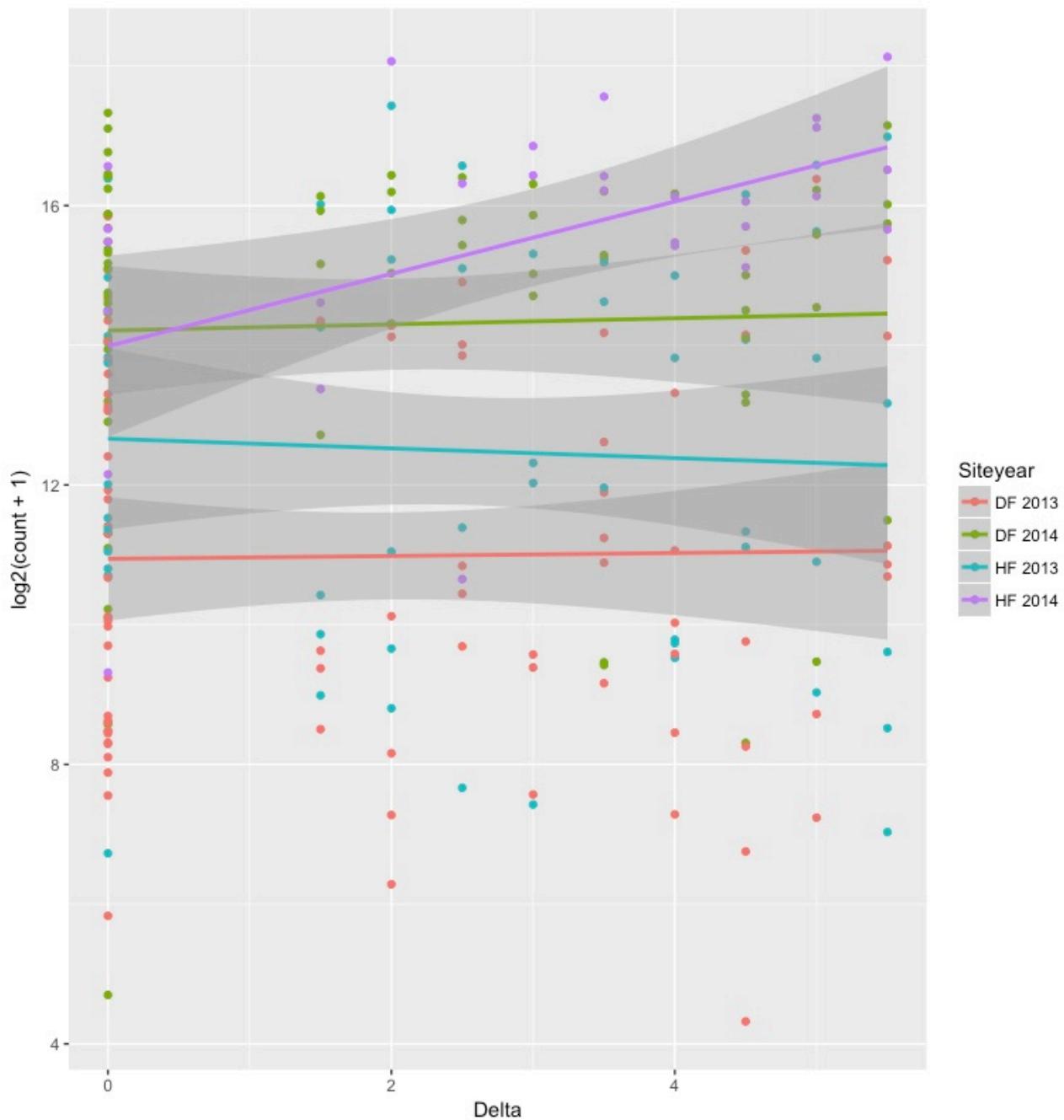


fig showing delta by site by year interaction



Statistics: hsp40

```

1 mod2<-
  aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.a
  ve*Delta,data=gxp40)
2
3 Step:  AIC=117.76

```

```

4 log(count + 1) ~ RIN_Value + Site + Year_collect +
baittemp.ave +
      Delta + Site:Year_collect + Site:baittemp.ave +
Year_collect:baittemp.ave +
      Site:Delta + Year_collect:Delta +
baittemp.ave:Delta + Site:Year_collect:Delta +
      Year_collect:baittemp.ave:Delta

8

9                               Df Sum of Sq      RSS
AIC

10 <none>                           332.63
117.76

11 - Site:Year_collect:Delta        1   3.1659 335.79
117.87

12 + Site:baittemp.ave:Delta       1   1.0630 331.56
119.05

13 - RIN_Value                     1   5.0316 337.66
119.10

14 - Year_collect:baittemp.ave:Delta 1   5.3788 338.00
119.33

15 + Site:Year_collect:baittemp.ave 1   0.3536 332.27
119.53

16 - Site:baittemp.ave             1   7.4288 340.05
120.67

17                               Df Sum Sq Mean Sq F
value    Pr(>F)
18 RIN_Value                      1  262.9  262.92
164.412 < 2e-16 ***
19 Site                          1   0.6   0.59
0.368 0.544494
20 Year_collect                  1   79.0  78.96
49.376 2.97e-11 ***
21 baittemp.ave                   1   1.1   1.12
0.699 0.404232

```

```

22 Delta                               1     0.1      0.07
    0.046 0.829526
23 Site:Year_collect                  1     7.9      7.94
    4.962 0.026977 *
24 Site:baittemp.ave                 1    17.3     17.28
    10.807 0.001187 **
25 Year_collect:baittemp.ave        1   114.3    114.34
    71.500 4.87e-15 ***
26 Site:Delta                          1     9.8      9.81
    6.133 0.014068 *
27 Year_collect:Delta                1    22.1     22.13
    13.841 0.000256 ***
28 baittemp.ave:Delta                 1     5.2      5.16
    3.228 0.073855 .
29 Site:Year_collect:Delta          1     4.3      4.26
    2.664 0.104143
30 Year_collect:baittemp.ave:Delta  1     5.4      5.38
    3.363 0.068086 .
31 Residuals                         208   332.6    1.60

32 ---
33 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
    0.1 ' ' 1

```

Statistics: hsp83

```

1 mod3<-
  aov(log(count+1)~RIN_Value+Site*Year_collect*baittemp.a
  ve*Delta,data=gxp83)
2 Step: AIC=368.17
3 log(count + 1) ~ RIN_Value + Site + Year_collect +
  baittemp.ave +
4     Delta + Site:Year_collect + Site:baittemp.ave +
  Year_collect:baittemp.ave +

```

```

5   Year_collect:Delta + baittemp.ave:Delta +
Year_collect:baittemp.ave:Delta
6
7                               Df Sum of Sq      RSS
AIC
8 <none>                                1031.6
368.17
9 + Site:Delta                         1     8.143 1023.5
368.37
10 + Site:Year_collect:baittemp.ave    1     2.224 1029.4
369.68
11 - Site:baittemp.ave                 1    20.209 1051.8
370.60
12 - Year_collect:baittemp.ave:Delta  1    31.913 1063.5
373.12
13 - RIN_Value                         1    58.793 1090.4
378.81
14 - Site:Year_collect                1    71.242 1102.8
381.40
15                               Df Sum Sq Mean Sq F
value   Pr(>F)
16 RIN_Value                           1 1073.4 1073.4
224.743 < 2e-16 ***
17 Site                                1   0.0   0.0
0.001 0.97546
18 Year_collect                         1  227.3  227.3
47.590 5.73e-11 ***
19 baittemp.ave                         1   26.7   26.7
5.595 0.01890 *
20 Delta                               1   0.2   0.2
0.038 0.84575
21 Site:Year_collect                   1   46.4   46.4
9.723 0.00207 **
22 Site:baittemp.ave                  1   38.9   38.9
8.154 0.00472 **

```

```

23 Year_collect:baittemp.ave      1   361.6   361.6
    75.718 8.44e-16 ***
24 Year_collect:Delta           1   55.4    55.4
    11.591  0.00079 ***
25 baittemp.ave:Delta          1   0.3     0.3
    0.060  0.80745
26 Year_collect:baittemp.ave:Delta 1   31.9    31.9
    6.682  0.01040 *
27 Residuals                   216 1031.6   4.8
28 ---
29 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'.
    0.1 ' ' 1

```

Page 73: 2017-04-30. Stressed in nature project: Re-analysis

To analyze all genes into 1 model; using the MCMC.qpcr package:
some data parsing first:

```

1 > warm<-
2   read.csv("../Data/20170427_final_dataset_2013_2014_year
3   s_HF_DF.csv",skip=10)
4 > warm$baittemp.ave<-
5   apply(warm[,8:11],1,mean,na.rm=TRUE)
6 > warm<-warm[!is.na(warm$RIN_Value),]
7 > warm$Cham<-as.factor(warm$Cham)
8 > warm$Sample<-as.factor(warm$Sample)
9 > warm$Collection.Date<-
10  as.factor(as.character(warm$Collection.Date))
11 > chd<-read.csv("../Data/2017_chamber_dat.csv")

```

```

8 > chd$Cham<-as.factor(chd$Cham)
9 > str(chd)
10 'data.frame':   30 obs. of  3 variables:
11   $ Cham : Factor w/ 15 levels "1","2","3","4",...: 1 2 3
12     4 5 6 7 8 9 10 ...
13   $ Site : Factor w/ 2 levels "DF","HF": 2 2 2 2 2 2 2 2
14     2 2 ...
13   $ Delta: num  5.5 4 2 0 5 0 3 1.5 2.5 4.5 ...
14 > warm<-inner_join(warm,chd,by=c("Cham","Site"))

```

getting gene counts:

```

1 gene<-c("CT_18s","CT_40","CT_70","CT_83","CT_actin") # 
list the genes
2 amp<-data.frame(gene,efficiency=rep(2,length(gene))) #
input efficiencies
3 names(warm)[6]<- "sample"
4
5 dd<-
6 cq2counts(data=warm,genecols=c(14:18),condcols=c(13,5:6
,1:4,20:22),Cq1=34,effic=amp)
7
8 #### more data manipulations; changing between
factor/class
9 dd$baittemp.ave<-
10 as.numeric(as.character(dd$baittemp.ave))
11 dd$RIN_Value<-as.numeric(as.character(dd$RIN_Value))
12 dd$Year_collect<-
13 as.numeric(as.character(dd$Year_collect))
14 dd$Siteyear<-paste(dd$Site,dd$Year_collect) # for
plotting purposes
15 dd$Delta<-as.numeric(as.character(dd$Delta))
16

```

Getting Julian days:

```

1  ### messing with dates
2  dd$Collection.Date<-as.character(dd$Collection.Date)
3  v<-paste(substr(dd$Collection.Date,1,4),"-"
4    ,substr(dd$Collection.Date,5,6),"-"
5    ,substr(dd$Collection.Date,7,8))
6
7  dd$Date<-gsub(" ", "", v, fixed = TRUE)
8
9  dd$JulianDay<-as.numeric(format(as.Date(dd$Date), "%j"))
10 # finding the latest julian day
11
12 dd$Jdaycont<-
13 ifelse(dd$Year_collect==2013,dd$JulianDay,dd$JulianDay+
14 253) ## adding
15 #latest julian day to make it continuous

```

Statistics: specifying bayesian model to include all genes, no control genes

```

1 mm=mcmc.qpcr(
2   fixed="RIN_Value+Jdaycont*Site*baittemp.ave*Delta",
3   random="Vial.me",
4   data=dd)

```

output

```

1 Iterations = 3001:12991
2 Thinning interval = 10
3 Sample size = 1000
4
5 DIC: 13001.7
6
7 G-structure: ~sample

```

```

8
9      post.mean  l-95% CI u-95% CI eff.samp
10 sample     0.0292 1.741e-43  0.07663    608.2
11
12      ~idh(gene):Vial.me
13
14      post.mean  l-95% CI u-95% CI
15      eff.samp
16 geneCT_18s.Vial.me    0.001257 3.491e-16 0.004498
17      40.48
18 geneCT_40.Vial.me    0.027512 2.233e-07 0.126707
19      41.07
20 geneCT_70.Vial.me    0.001494 1.029e-13 0.009846
21      36.60
22 geneCT_83.Vial.me    0.005245 1.123e-09 0.021233
23      40.93
24 geneCT_actin.Vial.me 0.003338 4.250e-08 0.020209
25      25.33
26
27      R-structure: ~idh(gene):units
28
29      post.mean  l-95% CI u-95% CI
30      eff.samp
31 geneCT_18s.units     8.187   6.7688   9.743
32      1000.0
33 geneCT_40.units      1.162   0.9151   1.435
34      271.7
35 geneCT_70.units      2.078   1.7054   2.483
36      1000.0
37 geneCT_83.units      4.406   3.6245   5.385
38      208.8
39 geneCT_actin.units   3.114   2.5009   3.740
40      809.4

```

```

30 Location effects: count ~ 0 + gene + gene:RIN_Value
+ gene:Jdaycont * Site * baittemp.ave * Delta
31
32
33 post.mean    l-95% CI    u-95% CI eff.samp
34 geneCT_18s
35     -2.133e+01 -2.668e+01 -1.618e+01    1121.3
36 geneCT_40
37     -3.130e+01 -3.629e+01 -2.598e+01    1150.3
38 geneCT_70
39     -2.838e+01 -3.349e+01 -2.302e+01    1139.9
40 geneCT_83
41     -3.257e+01 -3.746e+01 -2.696e+01    1155.2
42 geneCT_actin
43     -3.005e+01 -3.514e+01 -2.487e+01    1159.6
44 SiteHF
45     8.902e+00 -8.067e+00  2.217e+01    1000.0
46 baittemp.ave
47     1.260e+00  1.054e+00  1.453e+00    1147.9
48 Delta
49     1.813e+00  2.477e-01  3.300e+00    1000.0
50 geneCT_18s:RIN_Value
51     1.818e-01 -5.029e-02  4.004e-01    1000.0
52 geneCT_40:RIN_Value
53     1.044e-01  1.507e-02  2.070e-01    458.6
54 geneCT_70:RIN_Value
55     2.444e-01  1.194e-01  3.594e-01    1000.0
56 geneCT_83:RIN_Value
57     2.985e-01  1.199e-01  4.684e-01    1000.0
58 geneCT_actin:RIN_Value
59     1.592e-01  1.092e-02  3.006e-01    1000.0
60 geneCT_18s:Jdaycont
61     1.614e-01  1.377e-01  1.902e-01    1000.0
62 geneCT_40:Jdaycont
63     1.152e-01  9.677e-02  1.310e-01    1000.0

```

48	geneCT_70:Jdaycont				
	1.088e-01	9.180e-02	1.295e-01	1000.0	
49	geneCT_83:Jdaycont				
	1.274e-01	1.028e-01	1.507e-01	1000.0	
50	geneCT_actin:Jdaycont				
	1.258e-01	1.065e-01	1.466e-01	1000.0	
51	SiteHF:baittemp.ave				
	-3.373e-01	-8.651e-01	2.879e-01	1000.0	
52	SiteHF:Delta				
	-1.691e+00	-5.332e+00	2.575e+00	912.4	
53	baittemp.ave:Delta				
	-7.949e-02	-1.408e-01	-2.348e-02	1000.0	
54	geneCT_18s:Jdaycont:SiteHF				
	-6.293e-02	-1.508e-01	2.361e-02	1000.0	
55	geneCT_40:Jdaycont:SiteHF				
	-3.240e-02	-8.934e-02	2.572e-02	1000.0	
56	geneCT_70:Jdaycont:SiteHF				
	-4.520e-02	-1.026e-01	1.781e-02	1000.0	
57	geneCT_83:Jdaycont:SiteHF				
	-6.173e-02	-1.316e-01	1.224e-02	1000.0	
58	geneCT_actin:Jdaycont:SiteHF				
	-5.739e-02	-1.235e-01	1.120e-02	1000.0	
59	geneCT_18s:Jdaycont:baittemp.ave				
	-5.828e-03	-6.852e-03	-4.893e-03	1000.0	
60	geneCT_40:Jdaycont:baittemp.ave				
	-4.200e-03	-4.832e-03	-3.531e-03	1000.0	
61	geneCT_70:Jdaycont:baittemp.ave				
	-4.032e-03	-4.785e-03	-3.365e-03	1104.7	
62	geneCT_83:Jdaycont:baittemp.ave				
	-4.441e-03	-5.307e-03	-3.522e-03	1000.0	
63	geneCT_actin:Jdaycont:baittemp.ave				
	-4.509e-03	-5.234e-03	-3.713e-03	1000.0	
64	geneCT_18s:Jdaycont:Delta				
	-1.136e-02	-1.954e-02	-3.116e-03	1000.0	

```
65 geneCT_40:Jdaycont:Delta  
    -7.652e-03 -1.244e-02 -2.561e-03 1000.0  
66 geneCT_70:Jdaycont:Delta  
    -6.787e-03 -1.287e-02 -1.802e-03 1000.0  
67 geneCT_83:Jdaycont:Delta  
    -6.808e-03 -1.368e-02 -3.982e-04 890.1  
68 geneCT_actin:Jdaycont:Delta  
    -4.889e-03 -1.087e-02 1.344e-03 747.7  
69 SiteHF:baittemp.ave:Delta  
    4.847e-02 -9.596e-02 1.992e-01 1000.0  
70 geneCT_18s:Jdaycont:SiteHF:baittemp.ave  
    2.659e-03 -6.000e-04 5.992e-03 1000.0  
71 geneCT_40:Jdaycont:SiteHF:baittemp.ave  
    1.274e-03 -8.384e-04 3.507e-03 1000.0  
72 geneCT_70:Jdaycont:SiteHF:baittemp.ave  
    1.723e-03 -6.570e-04 3.918e-03 1000.0  
73 geneCT_83:Jdaycont:SiteHF:baittemp.ave  
    2.440e-03 -1.164e-04 5.299e-03 1000.0  
74 geneCT_actin:Jdaycont:SiteHF:baittemp.ave  
    2.146e-03 -4.782e-04 4.648e-03 1000.0  
75 geneCT_18s:Jdaycont:SiteHF:Delta  
    4.693e-03 -1.550e-02 2.612e-02 1018.7  
76 geneCT_40:Jdaycont:SiteHF:Delta  
    1.524e-03 -1.189e-02 1.431e-02 882.5  
77 geneCT_70:Jdaycont:SiteHF:Delta  
    7.593e-03 -8.370e-03 2.111e-02 1000.0  
78 geneCT_83:Jdaycont:SiteHF:Delta  
    2.915e-03 -1.576e-02 1.984e-02 1000.0  
79 geneCT_actin:Jdaycont:SiteHF:Delta  
    5.448e-03 -1.136e-02 2.111e-02 1000.0  
80 geneCT_18s:Jdaycont:baittemp.ave:Delta  
    4.781e-04 1.641e-04 7.779e-04 1000.0  
81 geneCT_40:Jdaycont:baittemp.ave:Delta  
    3.269e-04 1.358e-04 5.056e-04 1000.0
```

82	geneCT_70:Jdaycont:baittemp.ave:Delta				
	2.900e-04	1.029e-04	5.131e-04	1000.0	
83	geneCT_83:Jdaycont:baittemp.ave:Delta				
	3.028e-04	5.222e-05	5.459e-04	904.2	
84	geneCT_actin:Jdaycont:baittemp.ave:Delta				
	2.241e-04	9.050e-06	4.616e-04	749.3	
85	geneCT_18s:Jdaycont:SiteHF:baittemp.ave:Delta				
	-1.847e-04	-9.327e-04	6.063e-04	1000.0	
86	geneCT_40:Jdaycont:SiteHF:baittemp.ave:Delta				
	-2.955e-05	-5.167e-04	4.537e-04	881.9	
87	geneCT_70:Jdaycont:SiteHF:baittemp.ave:Delta				
	-2.312e-04	-7.787e-04	3.257e-04	1000.0	
88	geneCT_83:Jdaycont:SiteHF:baittemp.ave:Delta				
	-9.858e-05	-6.929e-04	6.199e-04	1000.0	
89	geneCT_actin:Jdaycont:SiteHF:baittemp.ave:Delta				
	-1.592e-04	-7.268e-04	4.714e-04	1000.0	
90					pMCMC
91	geneCT_18s				<0.001

92	geneCT_40				<0.001

93	geneCT_70				<0.001

94	geneCT_83				<0.001

95	geneCT_actin				<0.001

96	SiteHF				0.232
97	baittemp.ave				<0.001

98	Delta				0.024
	*				

99	geneCT_18s:RIN_Value	0.122
100	geneCT_40:RIN_Value	0.040
	*	
101	geneCT_70:RIN_Value	<0.001

102	geneCT_83:RIN_Value	0.002
	**	
103	geneCT_actin:RIN_Value	0.032
	*	
104	geneCT_18s:Jdaycont	<0.001

105	geneCT_40:Jdaycont	<0.001

106	geneCT_70:Jdaycont	<0.001

107	geneCT_83:Jdaycont	<0.001

108	geneCT_actin:Jdaycont	<0.001

109	SiteHF:baittemp.ave	0.236
110	SiteHF:Delta	0.420
111	baittemp.ave:Delta	0.014
	*	
112	geneCT_18s:Jdaycont:SiteHF	0.142
113	geneCT_40:Jdaycont:SiteHF	0.282
114	geneCT_70:Jdaycont:SiteHF	0.134
115	geneCT_83:Jdaycont:SiteHF	0.100
	.	

116	geneCT_actin:Jdaycont:SiteHF	0.110
117	geneCT_18s:Jdaycont:baittemp.ave	<0.001

118	geneCT_40:Jdaycont:baittemp.ave	<0.001

119	geneCT_70:Jdaycont:baittemp.ave	<0.001

120	geneCT_83:Jdaycont:baittemp.ave	<0.001

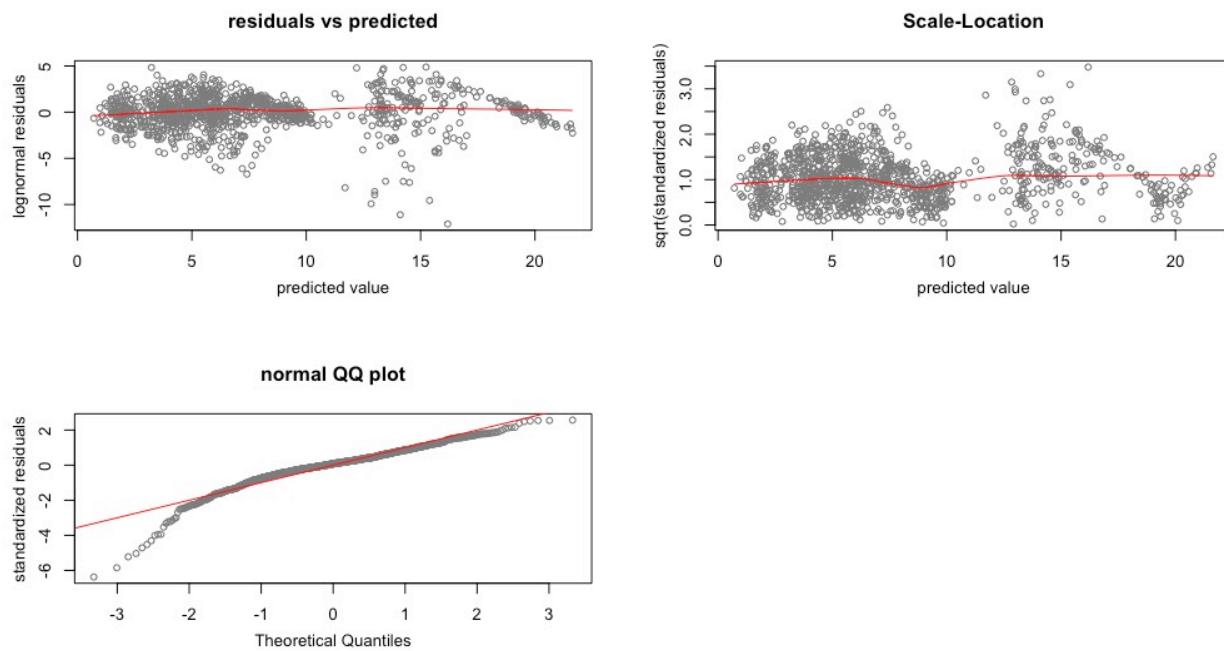
121	geneCT_actin:Jdaycont:baittemp.ave	<0.001

122	geneCT_18s:Jdaycont:Delta	0.004
	**	
123	geneCT_40:Jdaycont:Delta	0.002
	**	
124	geneCT_70:Jdaycont:Delta	0.020
	*	
125	geneCT_83:Jdaycont:Delta	0.044
	*	
126	geneCT_actin:Jdaycont:Delta	0.146
127	SiteHF:baittemp.ave:Delta	0.540
128	geneCT_18s:Jdaycont:SiteHF:baittemp.ave	0.098
	.	
129	geneCT_40:Jdaycont:SiteHF:baittemp.ave	0.266
130	geneCT_70:Jdaycont:SiteHF:baittemp.ave	0.130
131	geneCT_83:Jdaycont:SiteHF:baittemp.ave	0.080
	.	
132	geneCT_actin:Jdaycont:SiteHF:baittemp.ave	0.114

133	geneCT_18s:Jdaycont:SiteHF:Delta	0.688
134	geneCT_40:Jdaycont:SiteHF:Delta	0.840
135	geneCT_70:Jdaycont:SiteHF:Delta	0.330
136	geneCT_83:Jdaycont:SiteHF:Delta	0.728
137	geneCT_actin:Jdaycont:SiteHF:Delta	0.520
138	geneCT_18s:Jdaycont:baittemp.ave:Delta ***	<0.001
139	geneCT_40:Jdaycont:baittemp.ave:Delta ***	<0.001
140	geneCT_70:Jdaycont:baittemp.ave:Delta *	0.010
141	geneCT_83:Jdaycont:baittemp.ave:Delta *	0.016
142	geneCT_actin:Jdaycont:baittemp.ave:Delta .	0.068
143	geneCT_18s:Jdaycont:SiteHF:baittemp.ave:Delta	0.654
144	geneCT_40:Jdaycont:SiteHF:baittemp.ave:Delta	0.902
145	geneCT_70:Jdaycont:SiteHF:baittemp.ave:Delta	0.430
146	geneCT_83:Jdaycont:SiteHF:baittemp.ave:Delta	0.756
147	geneCT_actin:Jdaycont:SiteHF:baittemp.ave:Delta	0.620
148	---	
149	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	

"control genes" are not stable...or constant across treatments

model diagnostics



Statistics: fitting with control genes

```
1 mm2=mcmc.qpcr(  
2   fixed="RIN_Value+Jdaycont*Site*baittemp.ave*Delta",  
3   random="Vial.me", controls=c("CT_18s", "CTactin"),  
4   data=dd)  
5 summary(mm2)
```

Output:

```
1 Iterations = 3001:12991  
2 Thinning interval = 10  
3 Sample size = 1000  
4  
5 DIC: 12982.46  
6  
7 G-structure: ~sample  
8  
9      post.mean l-95% CI u-95% CI eff.samp
```

```

10 sample 0.06628 2.181e-54 6.863e-05      1000
11
12 ~idh(gene):Vial.me
13
14 post.mean l-95% CI u-95% CI
15 eff.samp
16 geneCT_40.Vial.me 0.01051 5.492e-08 0.06105
17 13.28
18 geneCT_70.Vial.me 0.22941 1.663e-04 0.78709
19 37.78
20 geneCT_83.Vial.me 0.59330 4.545e-07 2.71722
21 11.88
22 geneCT_actin.Vial.me 0.27245 1.185e-06 1.28528
23 13.71
24 geneCT_18s.Vial.me 0.03488 4.961e-20 0.06262
25 28.64
26
27 R-structure: ~idh(gene):units
28
29
30 post.mean l-95% CI u-95% CI eff.samp
31 geneCT_40.units 1.190 0.9469 1.459 587.04
32 geneCT_70.units 1.871 1.2816 2.486 79.41
33 geneCT_83.units 3.884 1.7517 5.411 21.65
34 geneCT_actin.units 2.852 1.7793 3.839 23.67
35 geneCT_18s.units 8.109 6.6364 9.655 908.67
36
37 Location effects: count ~ 0 + gene + +gene:RIN_Value +
38 gene:Jdaycont * Site * baittemp.ave * Delta
39
40
41 post.mean l-95% CI u-95% CI eff.samp pMCMC
42 geneCT_40
43 -3.107e+01 -3.559e+01 -2.616e+01 1000.0 <0.001 ***
44 geneCT_70
45 -2.817e+01 -3.312e+01 -2.354e+01 1022.4 <0.001 ***

```

35	geneCT_83	-3.231e+01	-3.692e+01	-2.722e+01	1000.0	<0.001	***
36	geneCT_actin	-2.982e+01	-3.456e+01	-2.496e+01	1057.2	<0.001	***
37	geneCT_18s	-2.112e+01	-2.561e+01	-1.617e+01	1000.0	<0.001	***
38	SiteHF	8.508e+00	-6.382e+00	2.209e+01	1000.0	0.250	
39	baittemp.ave	1.251e+00	1.050e+00	1.419e+00	1000.0	<0.001	***
40	Delta	1.712e+00	2.057e-01	3.224e+00	1000.0	0.034	*
41	geneCT_40:RIN_Value	9.921e-02	-4.455e-03	1.961e-01	1000.0	0.062	.
42	geneCT_70:RIN_Value	2.465e-01	1.478e-01	3.764e-01	1000.0	<0.001	***
43	geneCT_83:RIN_Value	2.911e-01	1.060e-01	4.576e-01	1000.0	0.002	**
44	geneCT_actin:RIN_Value	1.520e-01	1.671e-03	2.945e-01	1000.0	0.050	.
45	geneCT_18s:RIN_Value	1.747e-01	-4.381e-02	4.105e-01	1000.0	0.134	
46	geneCT_40:Jdaycont	1.145e-01	9.705e-02	1.315e-01	979.5	<0.001	***
47	geneCT_70:Jdaycont	1.083e-01	8.998e-02	1.263e-01	1000.0	<0.001	***
48	geneCT_83:Jdaycont	1.272e-01	1.059e-01	1.501e-01	843.5	<0.001	***
49	geneCT_actin:Jdaycont	1.255e-01	1.066e-01	1.468e-01	1103.4	<0.001	***
50	geneCT_18s:Jdaycont	1.609e-01	1.345e-01	1.875e-01	912.4	<0.001	***
51	SiteHF:baittemp.ave	-3.221e-01	-8.854e-01	2.107e-01	1000.0	0.248	

52	SiteHF:Delta					
	-1.603e+00 -5.566e+00 1.967e+00		890.2	0.404		
53	baittemp.ave:Delta					
	-7.549e-02 -1.310e-01 -1.567e-02		1000.0	0.010 **		
54	geneCT_40:Jdaycont:SiteHF					
	-2.990e-02 -8.217e-02 2.259e-02		967.4	0.298		
55	geneCT_70:Jdaycont:SiteHF					
	-4.375e-02 -1.042e-01 1.393e-02		1000.0	0.134		
56	geneCT_83:Jdaycont:SiteHF					
	-6.178e-02 -1.275e-01 1.679e-02		1000.0	0.096 .		
57	geneCT_actin:Jdaycont:SiteHF					
	-5.675e-02 -1.209e-01 6.528e-03		1000.0	0.076 .		
58	geneCT_18s:Jdaycont:SiteHF					
	-5.912e-02 -1.412e-01 2.562e-02		1000.0	0.172		
59	geneCT_40:Jdaycont:baittemp.ave					
	-4.167e-03 -4.829e-03 -3.517e-03		1000.0	<0.001 ***		
60	geneCT_70:Jdaycont:baittemp.ave					
	-4.012e-03 -4.651e-03 -3.302e-03		1000.0	<0.001 ***		
61	geneCT_83:Jdaycont:baittemp.ave					
	-4.431e-03 -5.231e-03 -3.613e-03		1000.0	<0.001 ***		
62	geneCT_actin:Jdaycont:baittemp.ave					
	-4.495e-03 -5.236e-03 -3.742e-03		1000.0	<0.001 ***		
63	geneCT_18s:Jdaycont:baittemp.ave					
	-5.801e-03 -6.805e-03 -4.873e-03		874.7	<0.001 ***		
64	geneCT_40:Jdaycont:Delta					
	-7.386e-03 -1.297e-02 -2.928e-03		1000.0	0.004 **		
65	geneCT_70:Jdaycont:Delta					
	-6.513e-03 -1.160e-02 -5.976e-04		1000.0	0.016 *		
66	geneCT_83:Jdaycont:Delta					
	-6.738e-03 -1.344e-02 -1.580e-04		1000.0	0.050 *		
67	geneCT_actin:Jdaycont:Delta					
	-4.677e-03 -1.078e-02 9.420e-04		1000.0	0.126		
68	geneCT_18s:Jdaycont:Delta					
	-1.088e-02 -1.775e-02 -2.091e-03		913.1	0.006 **		

69	SiteHF:baittemp.ave:Delta						
	4.473e-02 -9.274e-02 1.865e-01	883.6	0.540				
70	geneCT_40:Jdaycont:SiteHF:baittemp.ave						
	1.180e-03 -8.438e-04 3.177e-03	953.9	0.268				
71	geneCT_70:Jdaycont:SiteHF:baittemp.ave						
	1.667e-03 -4.613e-04 4.046e-03	1000.0	0.146				
72	geneCT_83:Jdaycont:SiteHF:baittemp.ave						
	2.443e-03 -5.353e-04 4.971e-03	1000.0	0.080	.			
73	geneCT_actin:Jdaycont:SiteHF:baittemp.ave						
	2.120e-03 -2.446e-04 4.658e-03	1000.0	0.080	.			
74	geneCT_18s:Jdaycont:SiteHF:baittemp.ave						
	2.511e-03 -7.564e-04 5.644e-03	1000.0	0.138				
75	geneCT_40:Jdaycont:SiteHF:Delta						
	1.144e-03 -1.164e-02 1.390e-02	911.9	0.846				
76	geneCT_70:Jdaycont:SiteHF:Delta						
	7.344e-03 -6.505e-03 2.243e-02	780.1	0.320				
77	geneCT_83:Jdaycont:SiteHF:Delta						
	2.850e-03 -1.492e-02 1.992e-02	706.7	0.758				
78	geneCT_actin:Jdaycont:SiteHF:Delta						
	5.225e-03 -1.075e-02 2.028e-02	904.6	0.516				
79	geneCT_18s:Jdaycont:SiteHF:Delta						
	3.949e-03 -1.684e-02 2.472e-02	1118.7	0.712				
80	geneCT_40:Jdaycont:baittemp.ave:Delta						
	3.160e-04 1.453e-04 5.166e-04	1000.0	0.002	**			
81	geneCT_70:Jdaycont:baittemp.ave:Delta						
	2.795e-04 7.007e-05 4.703e-04	1000.0	0.006	**			
82	geneCT_83:Jdaycont:baittemp.ave:Delta						
	2.992e-04 6.755e-05 5.555e-04	1000.0	0.020	*			
83	geneCT_actin:Jdaycont:baittemp.ave:Delta						
	2.154e-04 -1.270e-05 4.238e-04	1000.0	0.048	*			
84	geneCT_18s:Jdaycont:baittemp.ave:Delta						
	4.596e-04 1.451e-04 7.155e-04	905.2	0.004	**			
85	geneCT_40:Jdaycont:SiteHF:baittemp.ave:Delta						
	-1.369e-05 -4.662e-04 4.803e-04	896.1	0.958				

```

86 geneCT_70:Jdaycont:SiteHF:baittemp.ave:Delta
     -2.208e-04 -7.666e-04  2.911e-04      807.4  0.426
87 geneCT_83:Jdaycont:SiteHF:baittemp.ave:Delta
     -9.532e-05 -6.881e-04  5.898e-04      696.0  0.790
88 geneCT_actin:Jdaycont:SiteHF:baittemp.ave:Delta
     -1.489e-04 -7.156e-04  4.260e-04      899.7  0.634
89 geneCT_18s:Jdaycont:SiteHF:baittemp.ave:Delta
     -1.549e-04 -9.061e-04  6.138e-04     1116.7  0.702
90 ---
91 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
   0.1 ' ' 1

```

Statistics: "Classic" model (delta delta ct) with control genes

```

1 summary(classicmod)
2
3 Iterations = 3001:12991
4 Thinning interval = 10
5 Sample size = 1000
6
7 DIC: 4920.499
8
9 G-structure: ~idh(gene):Vial.me
10
11                         post.mean  l-95% CI u-95% CI
12 eff.samp
13 geneCT_40.Vial.me      0.005448 2.373e-14  0.02023
   39.20
14 geneCT_70.Vial.me      0.009244 1.900e-09  0.06938
   15.33
15 geneCT_83.Vial.me      3.759275 2.403e+00  5.27956
   795.80

```

```

15 geneCT_actin.Vial.me 1.428124 4.243e-01 2.41822
297.78
16 geneCT_18s.Vial.me 6.480444 4.572e+00 8.27869
882.87
17
18 ~idh(gene):sample
19
20 post.mean l-95% CI u-95% CI
eff.samp
21 geneCT_40.sample 7.525e-04 5.851e-59 6.254e-08
1000.0
22 geneCT_70.sample 1.611e-02 2.763e-59 2.056e-02
808.5
23 geneCT_83.sample 2.277e-40 2.463e-138 1.892e-46
0.0
24 geneCT_actin.sample 8.960e-04 9.595e-62 3.760e-07
488.2
25 geneCT_18s.sample 2.551e-02 5.330e-65 1.510e-02
812.1
26
27 R-structure: ~units
28
29 post.mean l-95% CI u-95% CI eff.samp
30 units 3.251 2.838 3.664 662.3
31
32 Location effects: count ~ 0 + gene + +gene:RIN_Value +
gene:Site * Jdaycont * Delta
33
34 post.mean l-95%
35 CI u-95% CI eff.samp pMCMC
36 geneCT_40 -2.6284204
-3.7176141 -1.4323167 879.9 <0.001 ***
geneCT_70 -3.2528194
-4.3525418 -2.0888514 1000.0 <0.001 ***

```

37	geneCT_83		-5.3863914
	-7.0044499 -3.6856766	1000.0 <0.001 ***	
38	geneCT_actin		-3.5034426
	-4.7385966 -2.0511151	1150.8 <0.001 ***	
39	geneCT_18s		-4.1354609
	-6.0388715 -2.1572579	1000.0 <0.001 ***	
40	Jdaycont		0.0069301
	-0.0022933 0.0161707	1000.0 0.156	
41	Delta		-1.0012178
	-1.7488433 0.0055000	1000.0 0.028 *	
42	geneCT_40:RIN_Value		0.2189257
	0.0713273 0.3608404	1000.0 0.004 **	
43	geneCT_70:RIN_Value		0.3380169
	0.1844881 0.4760341	1000.0 <0.001 ***	
44	geneCT_83:RIN_Value		0.4467853
	0.2559811 0.6592500	1000.0 <0.001 ***	
45	geneCT_actin:RIN_Value		0.3079191
	0.1534536 0.4931644	1000.0 <0.001 ***	
46	geneCT_18s:RIN_Value		0.3417462
	0.1006646 0.5934699	1000.0 0.008 **	
47	geneCT_40:SiteHF		-0.1231226
	-1.9166395 2.0395052	1000.0 0.918	
48	geneCT_70:SiteHF		1.5101947
	-0.4924095 3.4462324	892.6 0.126	
49	geneCT_83:SiteHF		-1.1804353
	-4.2953270 1.6625582	1000.0 0.496	
50	geneCT_actin:SiteHF		-0.1028772
	-2.7037113 2.3326229	870.1 0.922	
51	geneCT_18s:SiteHF		2.1563027
	-1.3683731 5.5980660	1000.0 0.244	
52	Jdaycont:Delta		0.0020503
	-0.0008895 0.0044592	1000.0 0.132	
53	geneCT_40:SiteDF:Jdaycont		-0.0028295
	-0.0141627 0.0061651	1000.0 0.592	

54	geneCT_70:SiteDF:Jdaycont		-0.0029284	
	-0.0130408	0.0068426	1000.0	0.584
55	geneCT_83:SiteDF:Jdaycont		0.0019250	
	-0.0073718	0.0139120	1000.0	0.722
56	geneCT_actin:SiteDF:Jdaycont		-0.0004737	
	-0.0103707	0.0100169	1000.0	0.944
57	geneCT_18s:SiteDF:Jdaycont		-0.0012918	
	-0.0123587	0.0089910	1000.0	0.822
58	geneCT_40:SiteHF:Jdaycont		-0.0006130	
	-0.0113985	0.0103346	1000.0	0.894
59	geneCT_70:SiteHF:Jdaycont		-0.0062379	
	-0.0176881	0.0048022	1000.0	0.280
60	geneCT_83:SiteHF:Jdaycont		0.0080759	
	-0.0042720	0.0206776	1000.0	0.214
61	geneCT_actin:SiteHF:Jdaycont		-0.0006060	
	-0.0120543	0.0114044	1000.0	0.964
62	geneCT_40:SiteDF:Delta		0.8353265	
	-0.0849807	1.7891522	1000.0	0.090 .
63	geneCT_70:SiteDF:Delta		0.9272206	
	-0.1068245	1.8419725	1000.0	0.066 .
64	geneCT_83:SiteDF:Delta		0.8385269	
	-0.2175849	1.8485020	1000.0	0.118
65	geneCT_actin:SiteDF:Delta		1.0904933	
	-0.0078081	1.9762069	1000.0	0.032 *
66	geneCT_18s:SiteDF:Delta		0.7533491	
	-0.3674505	1.8339046	1000.0	0.188
67	geneCT_40:SiteHF:Delta		0.6427846	
	-0.3309434	1.6253035	1000.0	0.222
68	geneCT_70:SiteHF:Delta		0.5351106	
	-0.4302433	1.5825929	1000.0	0.298
69	geneCT_83:SiteHF:Delta		0.7188371	
	-0.3859249	1.9097256	1000.0	0.230
70	geneCT_actin:SiteHF:Delta		0.5040863	
	-0.6524067	1.4469885	1000.0	0.362

```

71 geneCT_40:SiteDF:Jdaycont:Delta      -0.0014557
    -0.0041871  0.0013708   1000.0  0.284
72 geneCT_70:SiteDF:Jdaycont:Delta      -0.0018350
    -0.0043511  0.0014005   1000.0  0.214
73 geneCT_83:SiteDF:Jdaycont:Delta      -0.0013770
    -0.0046490  0.0014063   1000.0  0.378
74 geneCT_actin:SiteDF:Jdaycont:Delta  -0.0023920
    -0.0055308  0.0004937   1000.0  0.126
75 geneCT_18s:SiteDF:Jdaycont:Delta    -0.0010735
    -0.0043216  0.0021393   1000.0  0.494
76 geneCT_40:SiteHF:Jdaycont:Delta     -0.0010914
    -0.0040039  0.0019467   1000.0  0.496
77 geneCT_70:SiteHF:Jdaycont:Delta     -0.0003770
    -0.0032484  0.0029631   1000.0  0.820
78 geneCT_83:SiteHF:Jdaycont:Delta     -0.0012333
    -0.0048577  0.0022097   1000.0  0.498
79 geneCT_actin:SiteHF:Jdaycont:Delta -0.0004587
    -0.0036330  0.0028876   1000.0  0.764
80 ---
81 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.''
                  0.1 ' ' 1
82

```

Yeah non of this is right...vial.me should not be a random effect. Vial.me should be labeled as the "sample".

Also, how do I deal with the fact that the chambers are repeatedly sampled? I have to include a repeated measures aspect into model.

Also the model looks so complicated. What am I testing?

Main effects: Rin_VAlue - to control for quality

Interactions:

3 way between, Site/Julianday Delta

3 way between, Site/*Julian day* bait temperature

Other variables:

1. collection date
 2. Year collected
-

Page 74: 2017-05-01 & 2017-05-02. Stressed in nature project: re-analysis

I dont have enough samples per chamber to include it as a random effect.

```
1 sumda<-ddply(dd, .  
  (Site,Collection.Date,Delta,Cham,gene),summarize,num=len  
  gth(Date))  
2 > range(sumda$num)  
3 [1] 1 3
```

I'll need to average the genes.

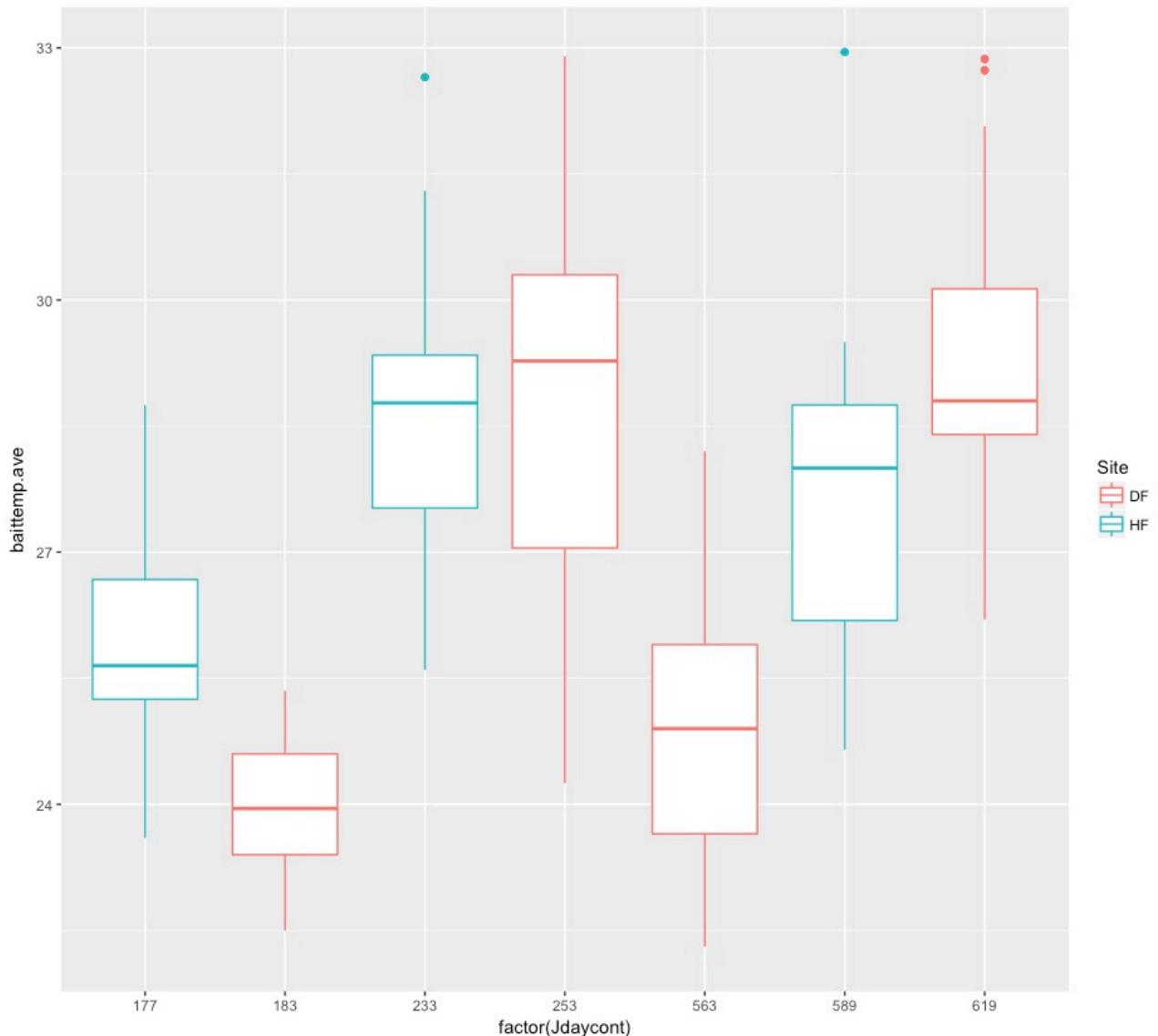
```

1 ddave<-ddply(dd,.
2   (Site,Year_collect,Collection.Date,Delta,gene,Cham),sum
3   marize,count=mean(count),RIN_Value=mean(RIN_Value),bait
4   temp.ave=mean(baittemp.ave))
5
6 'data.frame': 454 obs. of 9 variables:
7 $ Site           : Factor w/ 2 levels "DF","HF": 1 1 1
8   1 1 1 1 1 1 1 ...
9 $ Year_collect  : num  2013 2013 2013 2013 2013 ...
10 $ Collection.Date: Factor w/ 7 levels
11   "20130626","20130702",...
12 $ Delta          : num  0 0 0 0 0 0 0 0 0 ...
13 $ gene           : Factor w/ 5 levels
14   "CT_18s","CT_40",...
15 $ Cham           : Factor w/ 15 levels
16   "1","2","3","4",...
17 $ count          : num  1172618 579994 1640821
18   11910265 737550 ...
19 $ RIN_Value      : num  4.9 2.87 2.7 2.67 1.87 ...
20 $ baittemp.ave   : num  23.2 23.8 23.3 23.4 23.8 ...

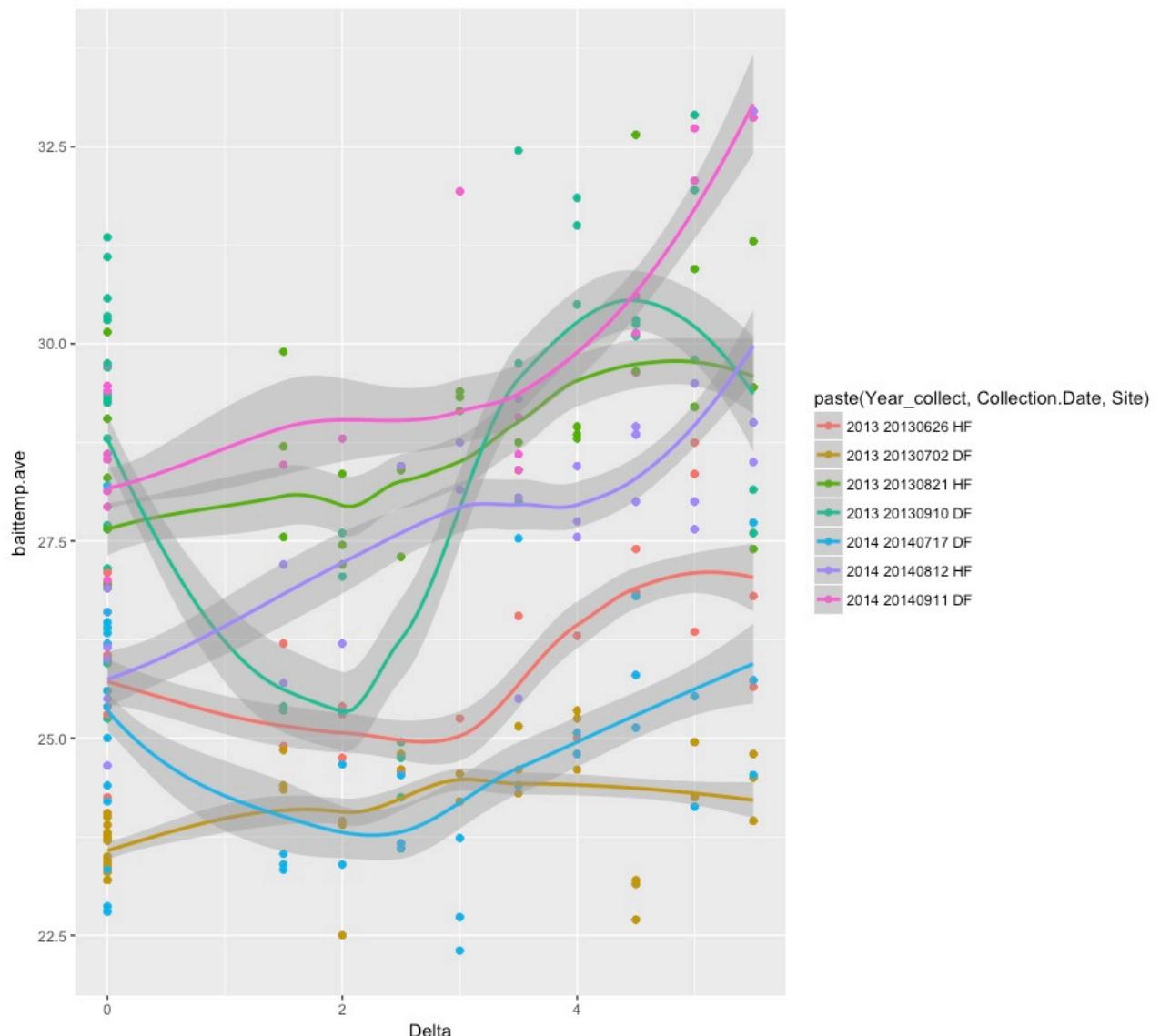
```

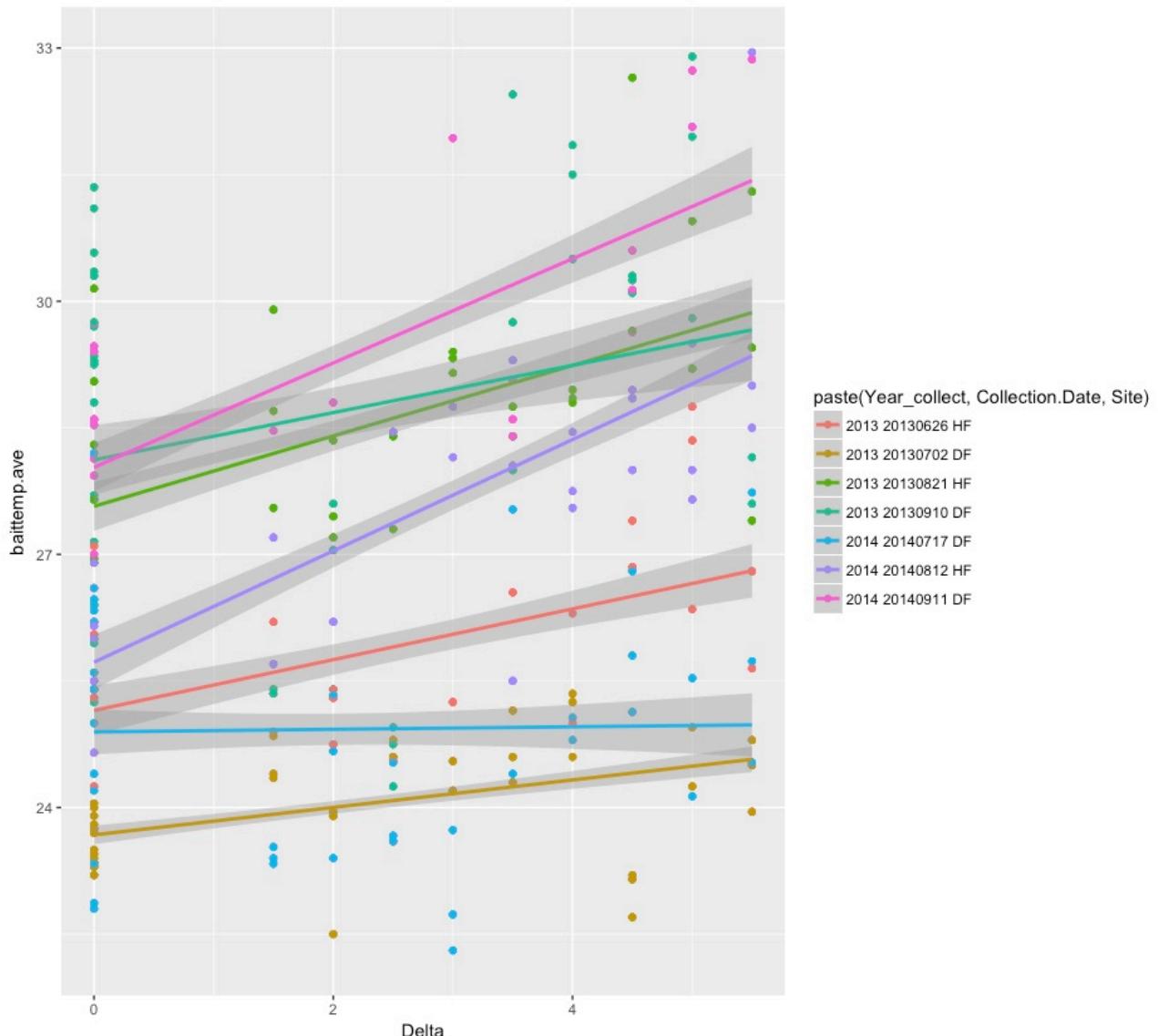
Figures for environmental conditions

bait temp vs julian day



plot with bait temp vs delta; colored by year, collection date, site



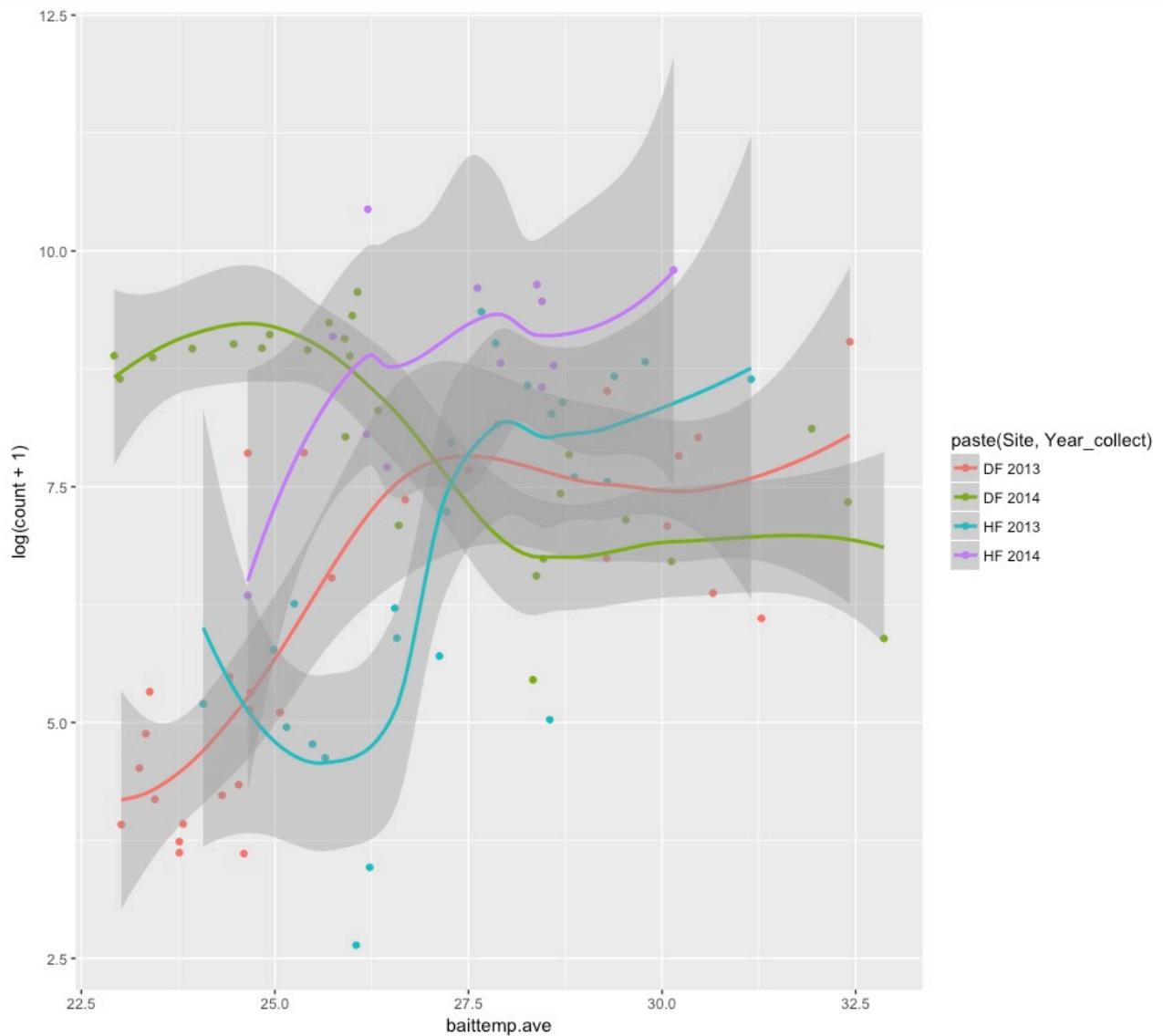


Models don't converge when testing for ...

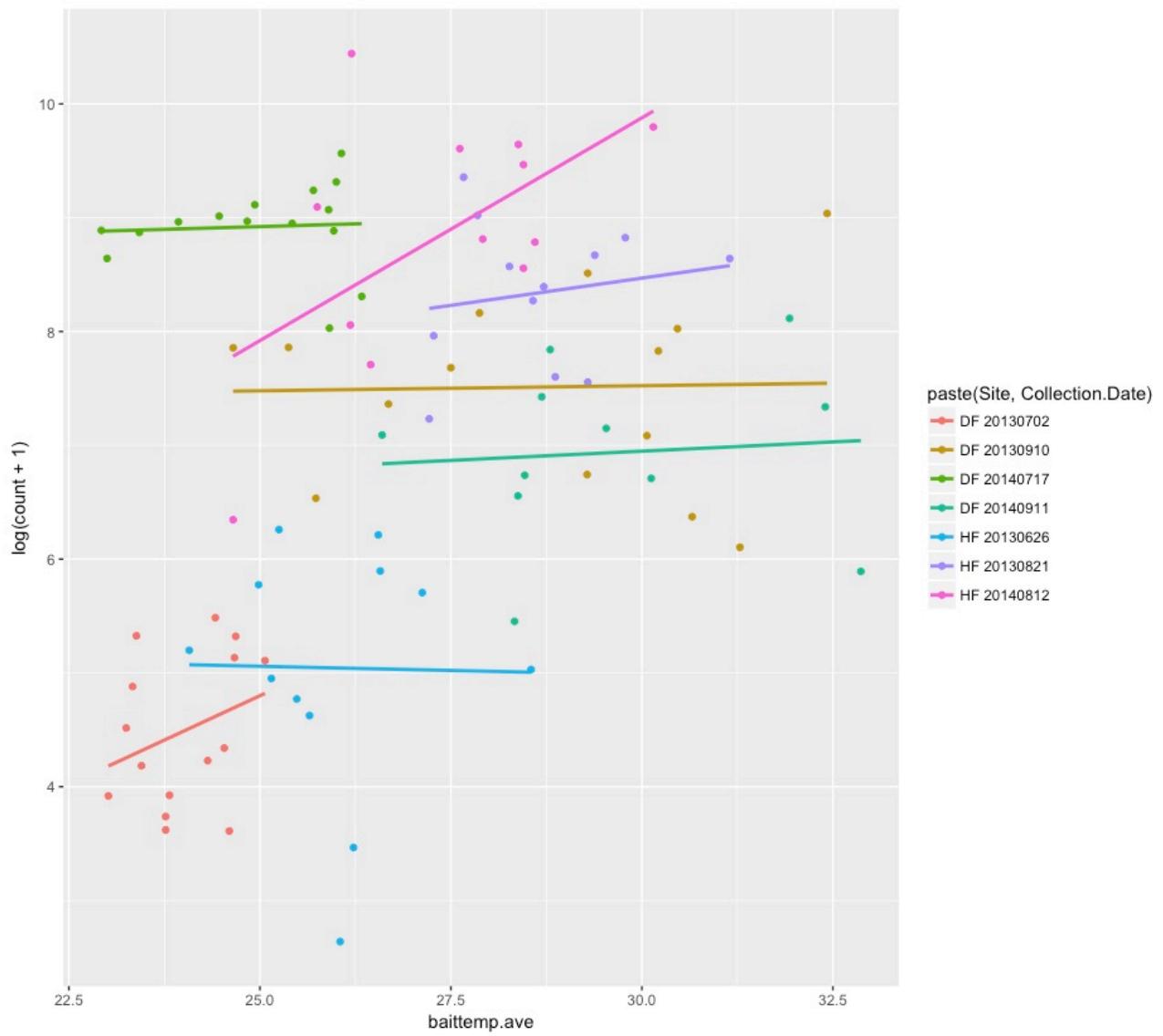
```
1 mod22<-
  glm(log(count+1)~RIN_Value+Site*baittemp.ave*Year_collec
  t*Delta,family=poisson,data=gxp70)
```

Parsing out hsp70 to visualize:

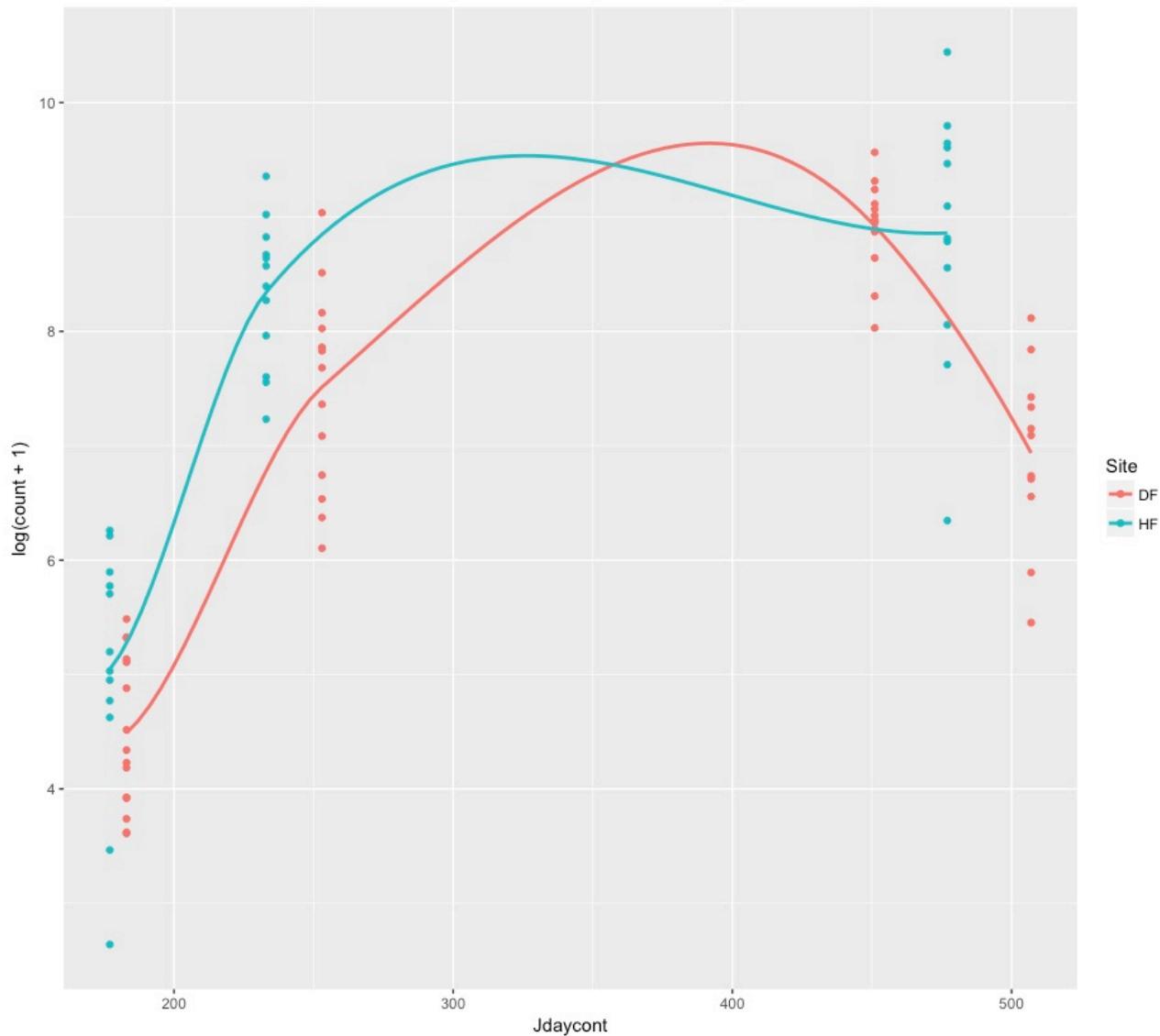
counts vs bait temp colored by year



counts vs bait temp colored by site and collection



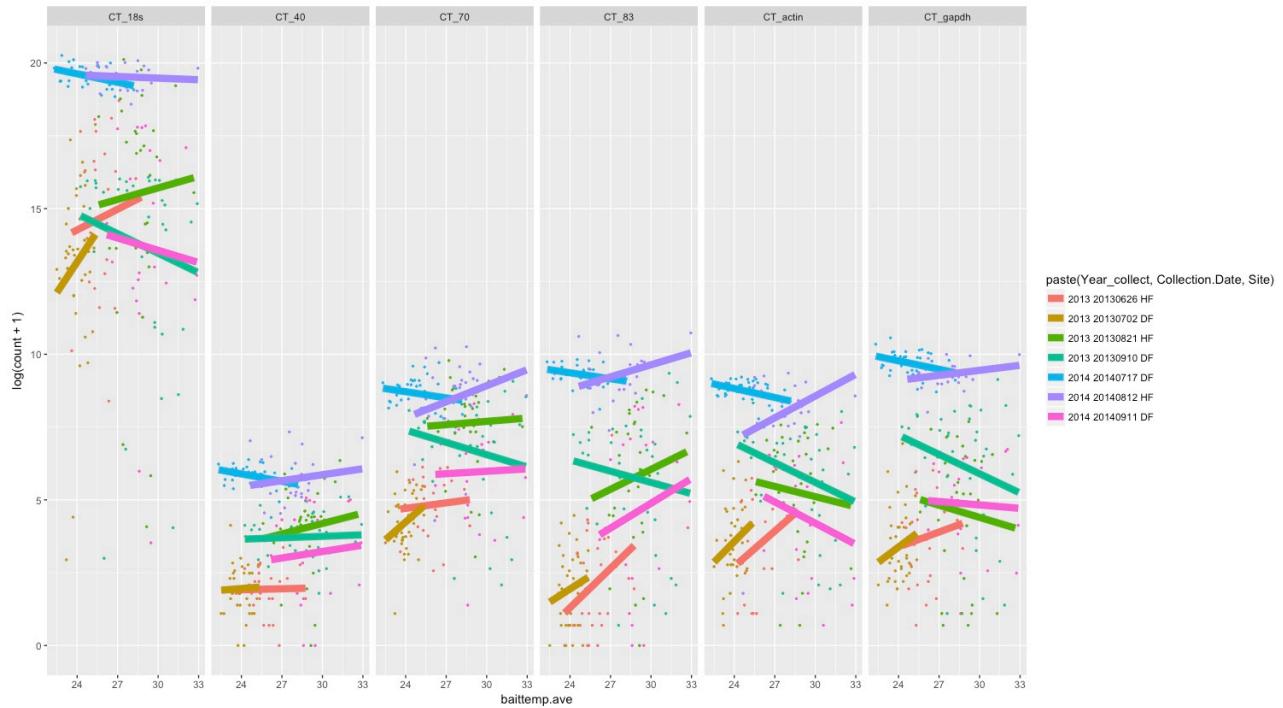
counts vs julian day colored by site



OK SHOWING ALL GENES

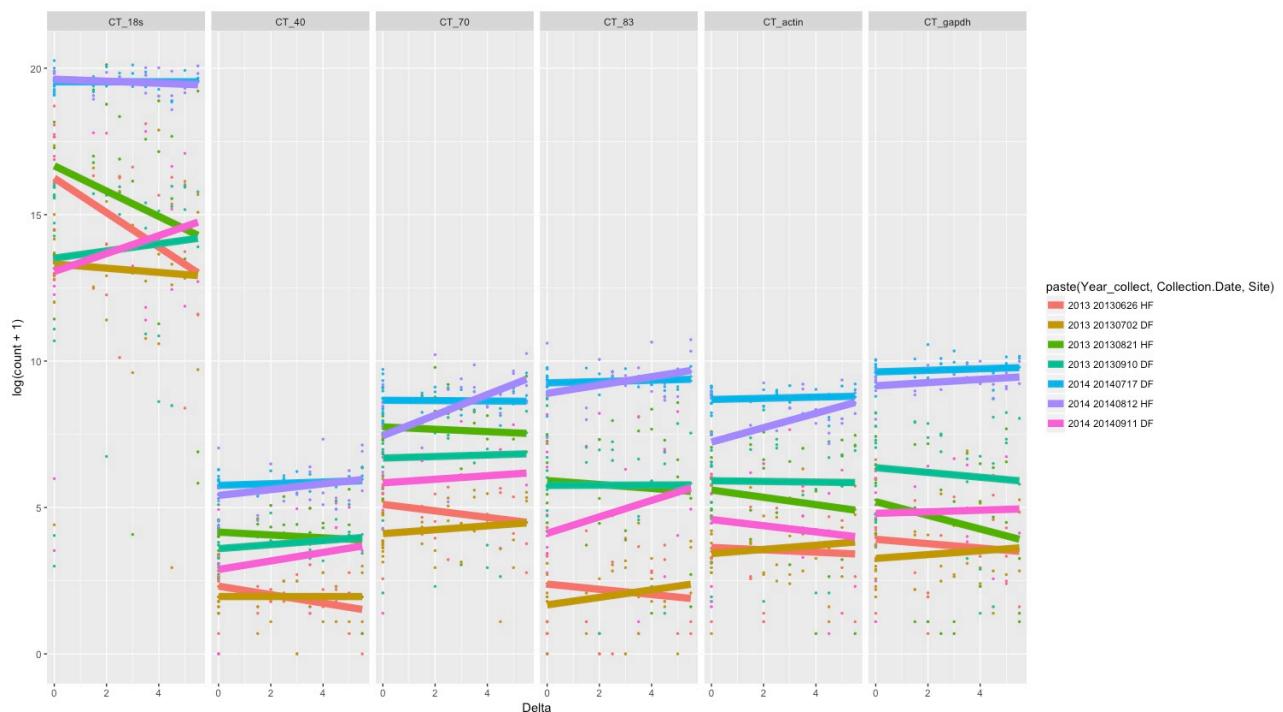
counts vs bait

```
1 ggplot(dd,aes(x=baittemp.ave,y=log(count+1),colour=paste
(Year_collect,Collection.Date,Site)))+geom_point(size=0.
5)+stat_smooth(size=3,method="lm",se=FALSE)+facet_grid(.~gene)
```



counts vs delta

```
1 ggplot(dd,aes(x=Delta,y=log(count+1),colour= paste(Year_collect,Collection.Date,Site)))+geom_point(size=0.5)+stat_smooth(size=3,method="lm",se=FALSE)+facet_grid(.~gene)
```



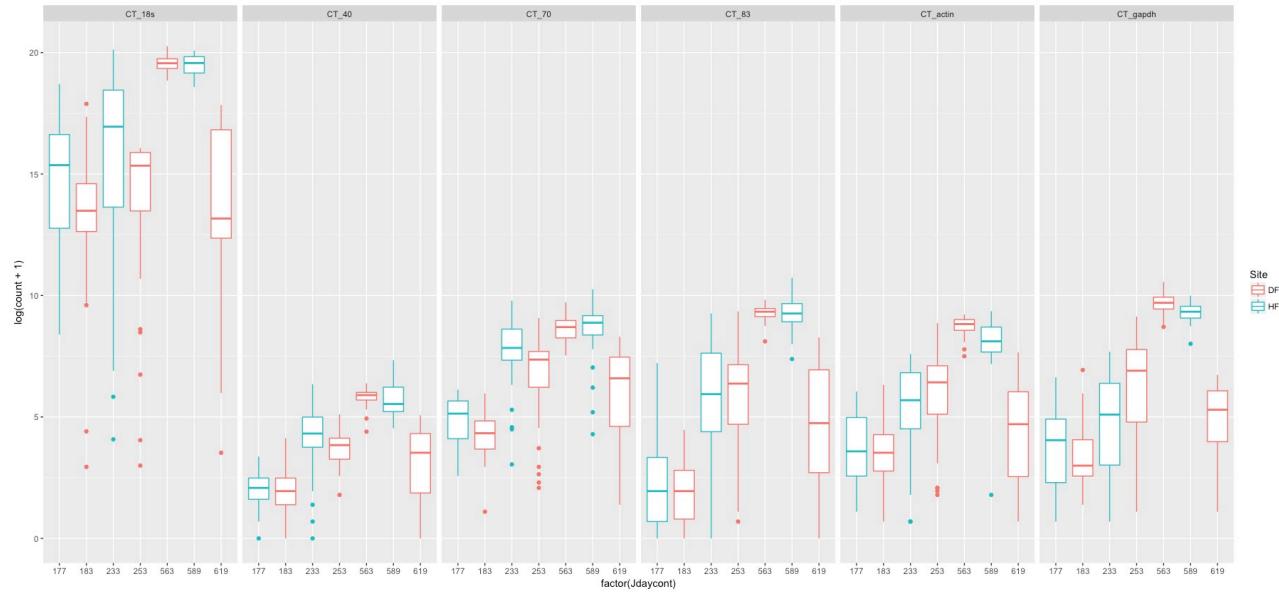
counts vs julian day

numbers for each collection day

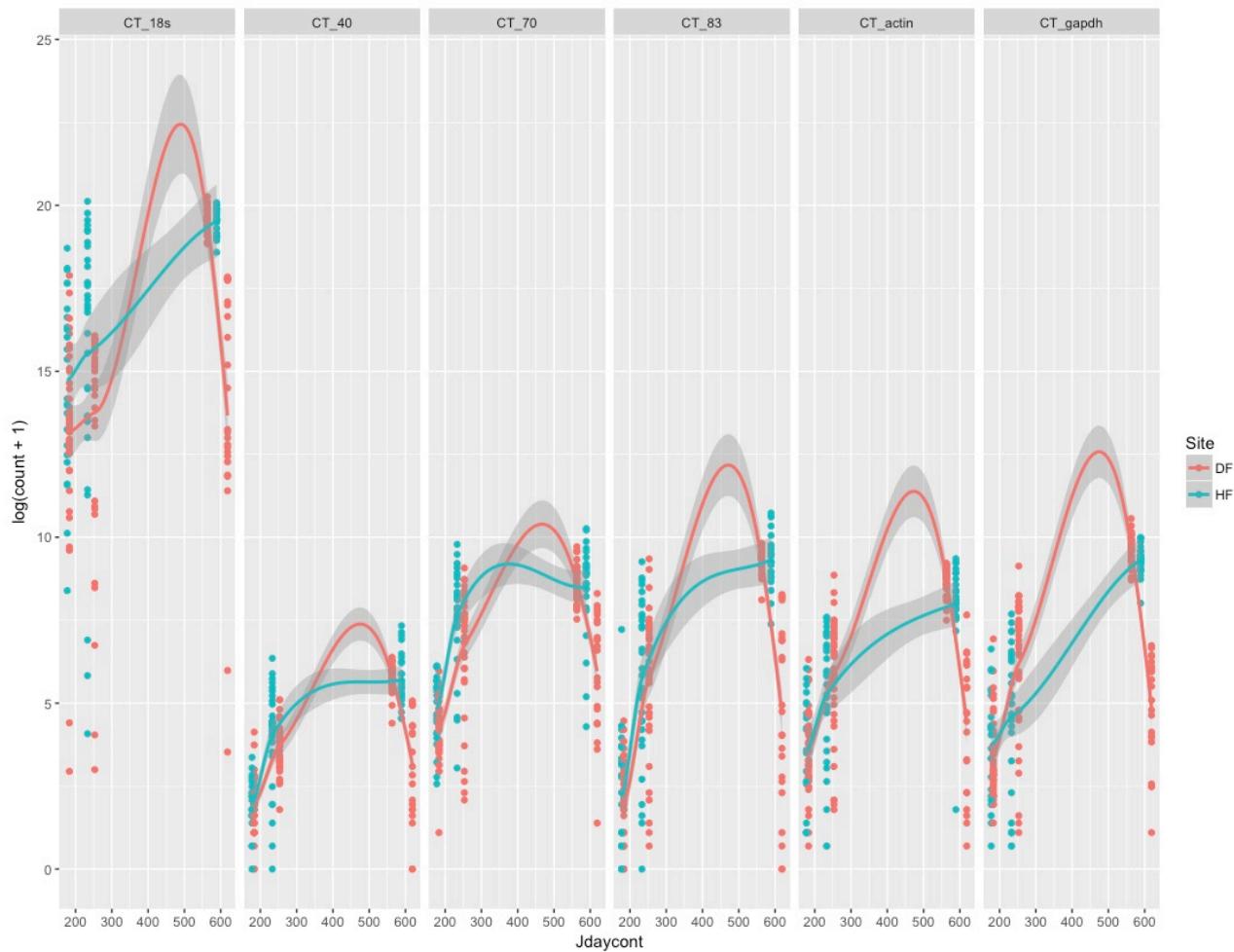
```
1 summary(factor(dd$Jdaycont))
2 177 183 233 253 563 589 619
3 143 251 192 230 240 164 136
4
5 ddply(dd,.(Site,JulianDay),summarize,length(Delta))
6   Site JulianDay length(Delta)
7 1   DF      183      251
8 2   DF      198      240
9 3   DF      253      230
10 4   DF      254      136
11 5   HF      177      143
12 6   HF      224      164
13 7   HF      233      192
14
15 ddply(dd,.
16   (Site,Year_collect,JulianDay),summarize,length(Delta))
17   Site Year_collect JulianDay length(Delta)
18 1   DF      2013      183      251
19 2   DF      2013      253      230
20 3   DF      2014      198      240
21 4   DF      2014      254      136
22 5   HF      2013      177      143
23 6   HF      2013      233      192
24 7   HF      2014      224      164
```

plots by continuous days

```
1 ggplot(dd,aes(x=factor(Jdaycont),y=log(count+1),color=Site))+geom_boxplot() + facet_grid(.~gene)
```

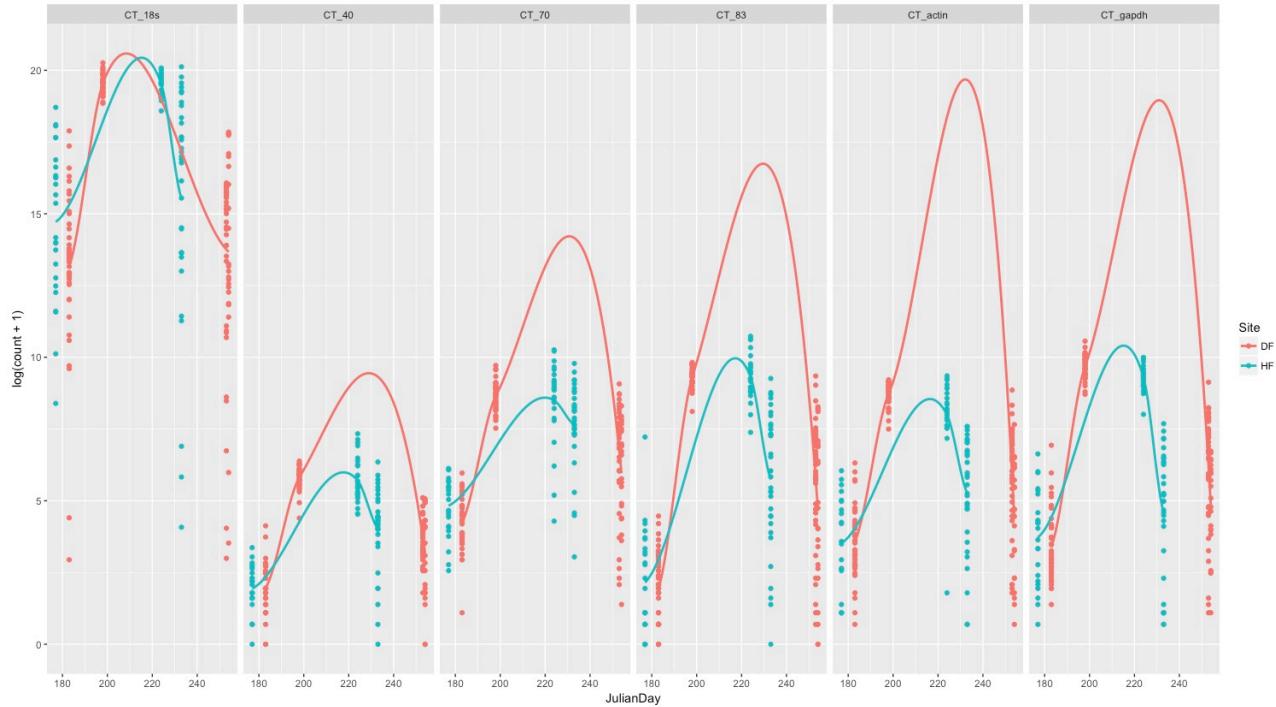


```
1 ggplot(dd,aes(x=Jdaycont,y=log(count+1),color=Site))+geom_point()+stat_smooth() +facet_grid(.~gene)
```



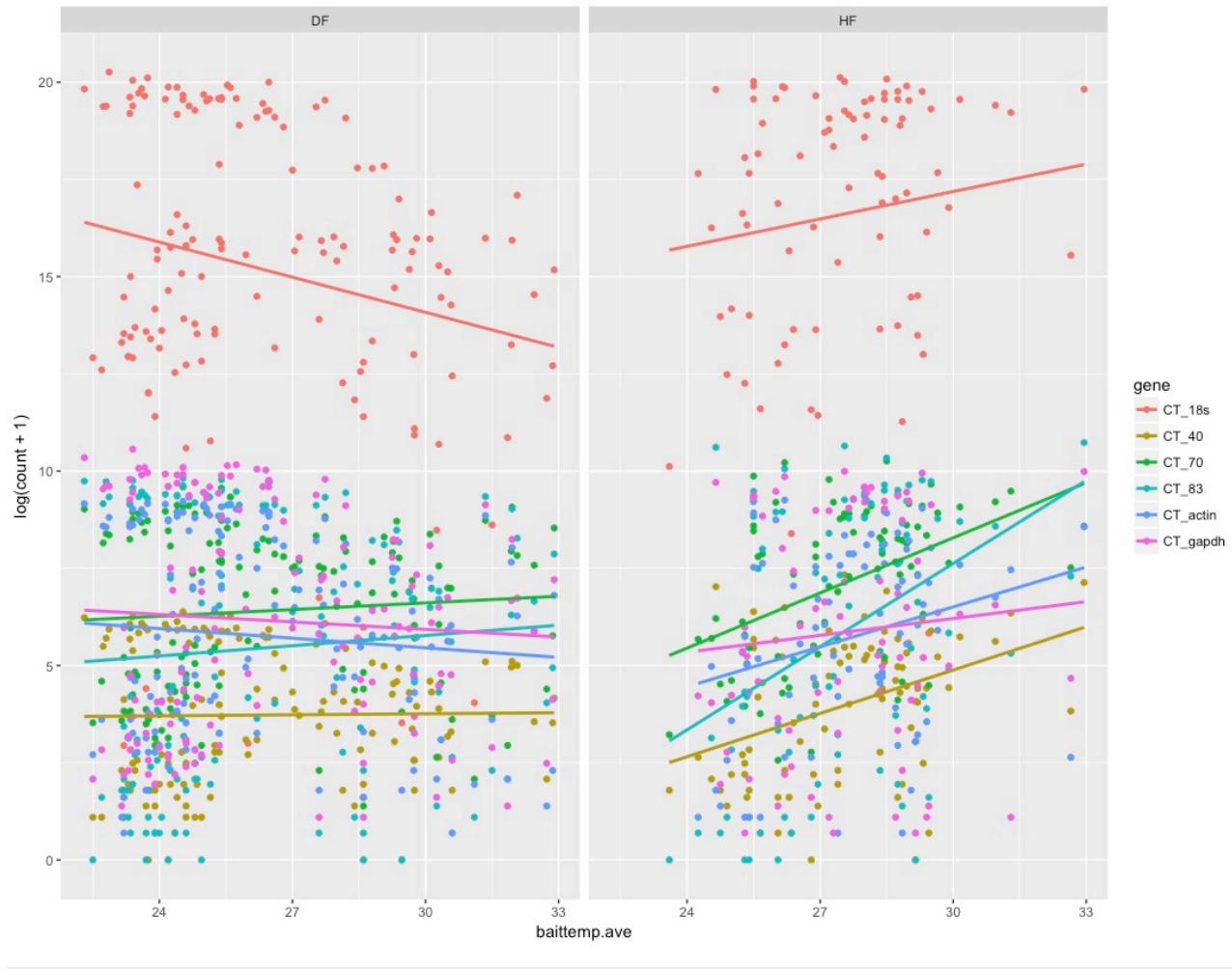
plot just by day

```
1 ggplot(dd,aes(x=JulianDay,y=log(count+1),color=Site))+geom_point()+geom_smooth(se=FALSE)+facet_grid(.~gene)
```



Site level differences

```
1 ggplot(dd,aes(x=baittemp.ave,y=log(count+1),color=Site)) +geom_point()+geom_smooth(method="lm",se=FALSE)+facet_grid(.~Site)
```



Page 75: 2017-05-03. Stressed in nature project: Analyses...again

Using mcmc.qpcr package.

Model 1: Naive - Testing for site by temperature interaction while controlling for season

Model testing for the effect of Site by bait temp and site by delta, controlling for RIN and Julian Day (globla fixed effects); including chamber as a random effect.

- Fixed effects: Site * bait + Site * Delta
- Global fixed effects: RIN value (rna quality) and Julian Day continuous
- Random factor: Chamber

```
1 mm=mcmc.qpcr(  
2   fixed="Site*baittemp.ave+Site*Delta",  
3   random=c("Cham"),globalFixed=c("RIN_Value","Jdaycont"),  
4   data=dd,pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =  
5   TRUE,pl=TRUE)  
6 summary(mm)
```

Model 1 output: Naive - Testing for site by temperature interaction while controlling for season

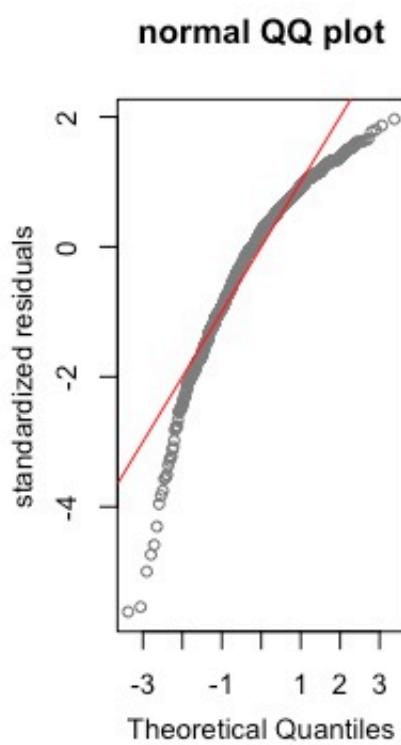
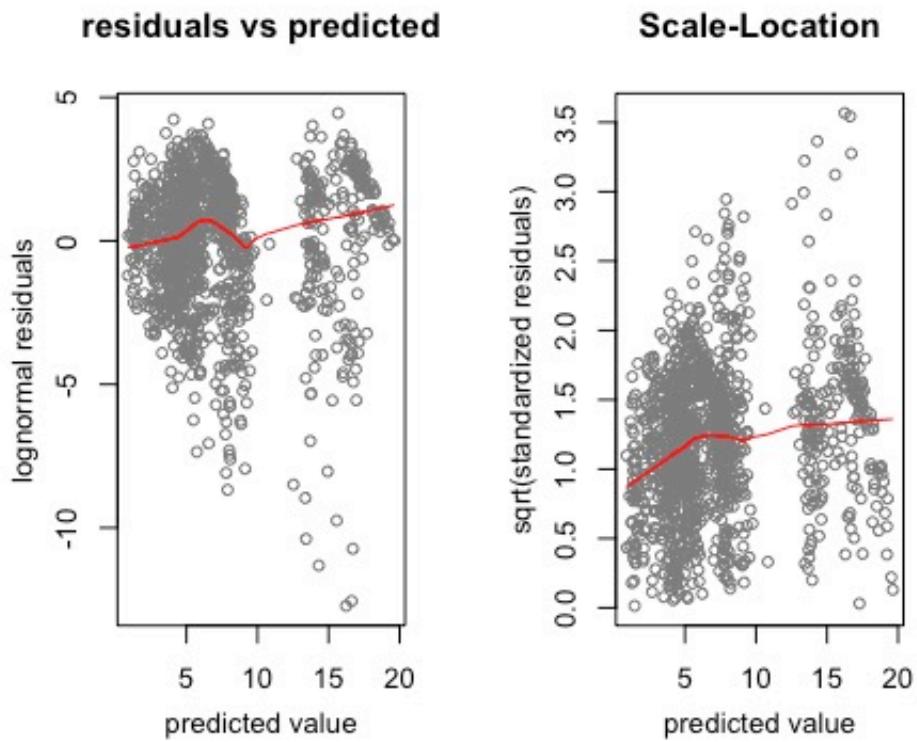
```
1 Iterations = 3001:49991  
2 Thinning interval = 10  
3 Sample size = 4700  
4  
5 DIC: 15013.3  
6  
7 G-structure: ~sample  
8  
9          post.mean l-95% CI u-95% CI eff.samp  
10 sample     2.971    2.338    3.624    3132  
11  
12          ~idh(gene):Cham  
13
```

14		post.mean	l-95% CI	u-95% CI	eff.samp
15	geneCT_18s.Cham	0.5831	0.1978	1.0303	4700
16	geneCT_40.Cham	0.3985	0.1705	0.7036	4811
17	geneCT_70.Cham	0.3942	0.1666	0.6862	4700
18	geneCT_83.Cham	0.4886	0.2011	0.8666	4700
19	geneCT_actin.Cham	0.4596	0.1884	0.8179	4996
20	geneCT_gapdh.Cham	0.5304	0.1953	0.9395	4700
21					
22	R-structure: ~idh(gene):units				
23					
24		post.mean	l-95% CI	u-95% CI	eff.samp
25	geneCT_18s.units	4.7506	3.7644	5.6788	4700
26	geneCT_40.units	0.4917	0.3377	0.6555	2027
27	geneCT_70.units	0.7914	0.6072	0.9920	3157
28	geneCT_83.units	2.1489	1.6681	2.6629	1060
29	geneCT_actin.units	1.2714	0.9745	1.5733	2346
30	geneCT_gapdh.units	1.7054	1.3214	2.1022	2197
31					
32	Location effects: count ~ 0 + gene + RIN_Value + Jdaycont + +gene:Site * baittemp.ave + gene:Site * Delta				
33					
34		post.mean	l-95% CI		
	u-95% CI	eff.samp	pMCMC		
35	geneCT_18s			16.110765	12.281529
	19.996436	4700 < 2e-04	***		
36	geneCT_40			-2.531363	-4.724173
	-0.348647	3418 0.02170	*		
37	geneCT_70			-1.967982	-4.259055
	0.126263	3489 0.07404	.		
38	geneCT_83			-4.654332	-7.732704
	-1.898907	3394 0.00213	**		
39	geneCT_actin			1.126526	-1.291423
	3.750314	3925 0.36681			

40	geneCT_gapdh		0.552040	-2.374267
	3.209835	4382 0.68936		
41	RIN_Value		0.199144	0.079509
	0.314567	4253 < 2e-04 ***		
42	Jdaycont		0.006478	0.004881
	0.008213	4700 < 2e-04 ***		
43	baittemp.ave		0.040939	-0.247876
	0.344168	4700 0.80255		
44	Delta		-0.112723	-0.400060
	0.198869	4700 0.47489		
45	geneCT_18s:SiteHF		-8.646499	-19.103973
	2.627596	4932 0.11277		
46	geneCT_40:SiteHF		-5.705043	-13.030149
	1.312838	4172 0.12511		
47	geneCT_70:SiteHF		-3.986178	-11.408322
	3.139362	4700 0.28638		
48	geneCT_83:SiteHF		-11.059021	-19.950986
	-2.446486	2618 0.01447 *		
49	geneCT_actin:SiteHF		-5.294406	-13.214177
	2.828587	4700 0.18426		
50	geneCT_gapdh:SiteHF		1.198255	-6.731170
	9.986804	4021 0.78298		
51	geneCT_18s:SiteDF:baittemp.ave		-0.215198	-0.549081
	0.125987	4700 0.20511		
52	geneCT_40:SiteDF:baittemp.ave		0.052836	-0.250665
	0.373404	4700 0.74894		
53	geneCT_70:SiteDF:baittemp.ave		0.151878	-0.162515
	0.459935	4700 0.35064		
54	geneCT_83:SiteDF:baittemp.ave		0.202728	-0.118489
	0.515906	4067 0.20809		
55	geneCT_actin:SiteDF:baittemp.ave		0.007118	-0.314216
	0.317920	4700 0.95447		
56	geneCT_gapdh:SiteDF:baittemp.ave		0.027289	-0.276939
	0.361853	4102 0.87106		

57	geneCT_18s:SiteHF:baittemp.ave	0.213395	-0.135902
	0.543047 4418 0.21830		
58	geneCT_40:SiteHF:baittemp.ave	0.302935	0.085426
	0.506891 3002 0.01021 *		
59	geneCT_70:SiteHF:baittemp.ave	0.330920	0.091069
	0.539425 3979 0.00511 **		
60	geneCT_83:SiteHF:baittemp.ave	0.648529	0.382448
	0.926608 2402 < 2e-04 ***		
61	geneCT_actin:SiteHF:baittemp.ave	0.201615	-0.048443
	0.451531 3650 0.11660		
62	geneCT_18s:SiteDF:Delta	0.232695	-0.186446
	0.635062 4700 0.26596		
63	geneCT_40:SiteDF:Delta	0.116132	-0.245279
	0.480588 4700 0.53617		
64	geneCT_70:SiteDF:Delta	0.047066	-0.308663
	0.421765 4700 0.79745		
65	geneCT_83:SiteDF:Delta	0.204531	-0.174726
	0.591889 4700 0.28723		
66	geneCT_actin:SiteDF:Delta	0.070546	-0.317046
	0.432616 4700 0.71787		
67	geneCT_gapdh:SiteDF:Delta	0.248740	-0.106857
	0.664952 4700 0.20851		
68	geneCT_18s:SiteHF:Delta	-0.298719	-0.669543
	0.080339 4700 0.11872		
69	geneCT_40:SiteHF:Delta	-0.140106	-0.379562
	0.099732 3397 0.26298		
70	geneCT_70:SiteHF:Delta	-0.062122	-0.301972
	0.191591 4311 0.63106		
71	geneCT_83:SiteHF:Delta	-0.155097	-0.457374
	0.136096 4052 0.31745		
72	geneCT_actin:SiteHF:Delta	0.031560	-0.229501
	0.318134 4183 0.80255		
73	---		
74	Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1		

diagnostics:



Model 2: Naive-Testing for site by temperature by season (julian day) interaction

Model testing for effect of Site by bait temp by julian day continuous; controlling for RIN and including chamber as a random effect

- Fixed effects: Site * bait + Site * Delta
- Global fixed effects: RIN value (rna quality) and Julian Day continuous
- Random factor: Chamber

```
1 mm2=mcmc.qpcr(  
2   fixed="Site*baittemp.ave*Jdaycont",  
3   random=c("Cham"),globalFixed=c("RIN_Value"),  
4   data=dd,pr=TRUE,vprior="iw",nitt=50000,include=0,geneSpe  
cRes = TRUE,pl=TRUE)  
5 summary(mm2)  
6
```

MOdel 2 output: Naive-Testing for site by temperature by season (julian day) interaction

```
1  
2 Iterations = 3001:49991  
3 Thinning interval = 10  
4 Sample size = 4700  
5  
6 DIC: 15023.12  
7  
8 G-structure: ~sample
```

```

9
10      post.mean l-95% CI u-95% CI eff.samp
11 sample     1.917    1.514    2.347    3390
12
13      ~idh(gene):Cham
14
15      post.mean l-95% CI u-95% CI eff.samp
16 geneCT_18s.Cham     0.6301   0.2387   1.1668   4700
17 geneCT_40.Cham      0.3843   0.1727   0.6856   4700
18 geneCT_70.Cham      0.3639   0.1597   0.6361   4700
19 geneCT_83.Cham      0.4681   0.1787   0.8216   4700
20 geneCT_actin.Cham   0.4103   0.1743   0.7152   5230
21 geneCT_gapdh.Cham   0.4443   0.1841   0.7888   5332
22
23 R-structure: ~idh(gene):units
24
25      post.mean l-95% CI u-95% CI eff.samp
26 geneCT_18s.units     4.4459   3.6026   5.2871   4700.0
27 geneCT_40.units       0.3735   0.2619   0.4933   2511.4
28 geneCT_70.units       0.6003   0.4567   0.7375   3188.9
29 geneCT_83.units       1.1112   0.8506   1.3815   905.5
30 geneCT_actin.units   1.1518   0.9038   1.4147   2207.6
31 geneCT_gapdh.units   1.0840   0.8476   1.3569   1719.8
32
33 Location effects: count ~ 0 + gene + RIN_Value +
34 + gene:Site * baittemp.ave * Jdaycont
35
36      post.mean
37      l-95% CI   u-95% CI eff.samp      pMCMC
38 geneCT_18s           -2.697e+00
39      -1.060e+01  5.214e+00      4700  0.518723
40 geneCT_40             -1.215e+01
41      -1.585e+01 -7.936e+00      2962  < 2e-04 ***
```

39	geneCT_83				-3.134e+01
	-3.626e+01 -2.646e+01	2357	< 2e-04	***	
40	geneCT_actin				-1.958e+01
	-2.451e+01 -1.468e+01	4107	< 2e-04	***	
41	geneCT_gapdh				-2.243e+01
	-2.731e+01 -1.775e+01	3686	< 2e-04	***	
42	RIN_Value				2.047e-01
	1.102e-01 3.090e-01	4147	0.000426	***	
43	baittemp.ave				-1.958e-01
	-6.238e-01 2.510e-01	4700	0.382979		
44	Jdaycont				-5.680e-04
	-3.427e-02 3.148e-02	4700	0.972340		
45	geneCT_18s:SiteHF				1.539e+01
	-3.751e+00 3.350e+01	4700	0.108936		
46	geneCT_40:SiteHF				1.293e+00
	-1.058e+01 1.215e+01	3853	0.809362		
47	geneCT_70:SiteHF				6.747e+00
	-5.146e+00 1.809e+01	4700	0.250638		
48	geneCT_83:SiteHF				7.440e+00
	-6.279e+00 2.034e+01	2260	0.265957		
49	geneCT_actin:SiteHF				1.731e+01
	4.324e+00 3.092e+01	3714	0.011064	*	
50	geneCT_gapdh:SiteHF				2.841e+01
	1.519e+01 4.106e+01	4490	< 2e-04	***	
51	baittemp.ave:Jdaycont				4.655e-04
	-7.468e-04 1.643e-03	4700	0.425532		
52	geneCT_18s:SiteDF:baittemp.ave				7.159e-01
	1.879e-01 1.248e+00	4700	0.007234	**	
53	geneCT_40:SiteDF:baittemp.ave				6.869e-01
	1.890e-01 1.132e+00	4101	0.002979	**	
54	geneCT_70:SiteDF:baittemp.ave				9.480e-01
	4.694e-01 1.406e+00	4497	< 2e-04	***	
55	geneCT_83:SiteDF:baittemp.ave				1.435e+00
	9.523e-01 1.918e+00	4700	< 2e-04	***	

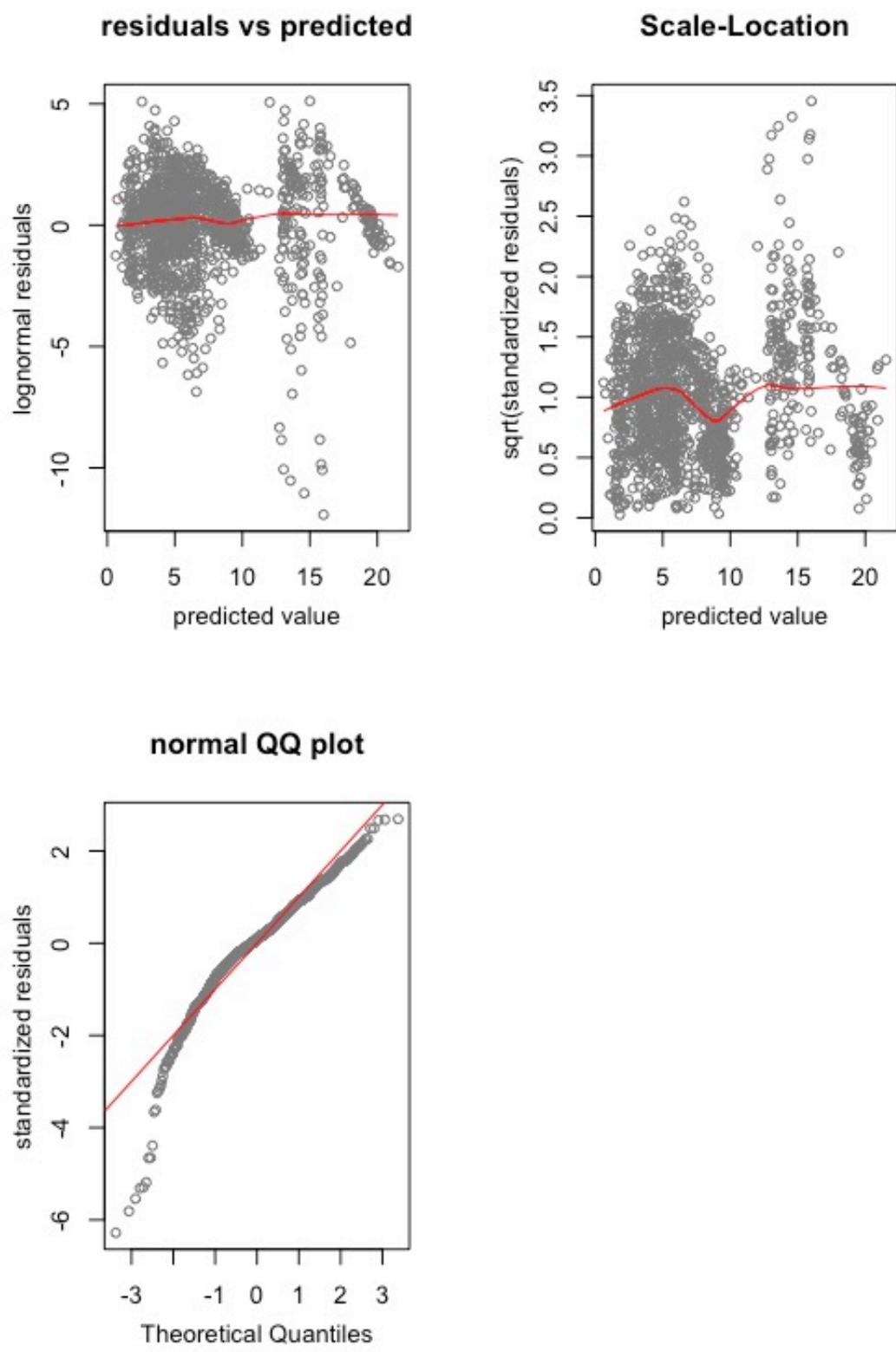
56	geneCT_actin:SiteDF:baittemp.ave		1.057e+00
	5.651e-01 1.514e+00	4700 < 2e-04	***
57	geneCT_gapdh:SiteDF:baittemp.ave		1.147e+00
	6.529e-01 1.612e+00	4506 < 2e-04	***
58	geneCT_18s:SiteHF:baittemp.ave		1.760e-01
	-4.193e-01 7.441e-01	4700 0.566383	
59	geneCT_40:SiteHF:baittemp.ave		6.176e-01
	3.030e-01 9.288e-01	3589 < 2e-04	***
60	geneCT_70:SiteHF:baittemp.ave		6.932e-01
	3.741e-01 1.028e+00	4700 < 2e-04	***
61	geneCT_83:SiteHF:baittemp.ave		1.085e+00
	7.130e-01 1.493e+00	2729 < 2e-04	***
62	geneCT_actin:SiteHF:baittemp.ave		3.454e-01
	-6.321e-02 7.128e-01	3574 0.085532	.
63	geneCT_18s:SiteDF:Jdaycont		7.762e-02
	3.948e-02 1.162e-01	4700 < 2e-04	***
64	geneCT_40:SiteDF:Jdaycont		5.000e-02
	1.422e-02 8.342e-02	4497 0.002553	**
65	geneCT_70:SiteDF:Jdaycont		6.005e-02
	2.730e-02 9.631e-02	4472 0.000851	***
66	geneCT_83:SiteDF:Jdaycont		9.879e-02
	6.440e-02 1.345e-01	4700 < 2e-04	***
67	geneCT_actin:SiteDF:Jdaycont		7.954e-02
	4.401e-02 1.142e-01	4700 < 2e-04	***
68	geneCT_gapdh:SiteDF:Jdaycont		8.735e-02
	5.198e-02 1.232e-01	4700 < 2e-04	***
69	geneCT_18s:SiteHF:Jdaycont		4.785e-03
	-3.861e-02 5.004e-02	4700 0.826383	
70	geneCT_40:SiteHF:Jdaycont		2.291e-02
	1.350e-04 4.613e-02	4243 0.051064	.
71	geneCT_70:SiteHF:Jdaycont		2.117e-02
	-3.446e-03 4.640e-02	4700 0.092766	.
72	geneCT_83:SiteHF:Jdaycont		4.531e-02
	1.793e-02 7.605e-02	3481 0.002128	**

```

73 geneCT_actin:SiteHF:Jdaycont           5.129e-03
    -2.447e-02  3.241e-02      4202  0.741702
74 geneCT_18s:SiteDF:baittemp.ave:Jdaycont -3.044e-03
    -4.453e-03 -1.623e-03      4700  < 2e-04 ***
75 geneCT_40:SiteDF:baittemp.ave:Jdaycont -2.162e-03
    -3.423e-03 -8.991e-04      4509  0.000851 ***
76 geneCT_70:SiteDF:baittemp.ave:Jdaycont -2.556e-03
    -3.897e-03 -1.369e-03      4493  < 2e-04 ***
77 geneCT_83:SiteDF:baittemp.ave:Jdaycont -3.850e-03
    -5.103e-03 -2.532e-03      4700  < 2e-04 ***
78 geneCT_actin:SiteDF:baittemp.ave:Jdaycont -3.274e-03
    -4.538e-03 -1.981e-03      4700  < 2e-04 ***
79 geneCT_gapdh:SiteDF:baittemp.ave:Jdaycont -3.486e-03
    -4.756e-03 -2.131e-03      4700  < 2e-04 ***
80 geneCT_18s:SiteHF:baittemp.ave:Jdaycont -2.595e-04
    -1.828e-03  1.393e-03      4700  0.745106
81 geneCT_40:SiteHF:baittemp.ave:Jdaycont -1.062e-03
    -1.879e-03 -2.114e-04      4265  0.014894 *
82 geneCT_70:SiteHF:baittemp.ave:Jdaycont -1.059e-03
    -1.937e-03 -1.224e-04      4700  0.018723 *
83 geneCT_83:SiteHF:baittemp.ave:Jdaycont -1.640e-03
    -2.709e-03 -6.076e-04      3510  0.003404 **
84 geneCT_actin:SiteHF:baittemp.ave:Jdaycont -3.494e-04
    -1.347e-03  7.197e-04      4217  0.517872
85 ---
86 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

diagnostics



Model 3: Soft normalization- Including 18s, actin, and gapdh to normalize

Model testing for effect of Site by bait temp; controlling for RIN and julian day continuous and including chamber as a random effect

- Fixed effects: Site * bait + Site * Delta
- Global fixed effects: RIN value (rna quality) and Julian Day continuous
- Random factor: Chamber

```
1 softnorm=mcmc.qpcr(fixed="Site*baittemp.ave+Site*Delta",
2   random="Cham",data=dd,pr=TRUE,vprior="iw",nitt=10000,gen
3   eSpecRes =
4   TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_actin","CT_ga
5   pdh","CT_18s"),globalFixed=c("RIN_Value","Jdaycont"))
6 summary(softnorm)
7
```

Model 3 output: Soft normalization- Including 18s, actin, and gapdh to normalize

```
1
2 Iterations = 3001:9991
3 Thinning interval = 10
4 Sample size = 700
5
6 DIC: 9005.386
7
8 G-structure: ~sample
9
10 post.mean l-95% CI u-95% CI eff.samp
11 sample     2.517    1.983    3.093      700
12
13 ~idh(gene):Cham
14
15 post.mean l-95% CI u-95% CI eff.samp
```

```

16 geneNORM.Cham      0.4293   0.1489   0.7762   518.2
17 geneCT_40.Cham    0.3077   0.1190   0.5434   700.0
18 geneCT_70.Cham    0.3022   0.1199   0.5436   700.0
19 geneCT_83.Cham    0.4124   0.1460   0.7656   700.0
20
21 R-structure: ~idh(gene):units
22
23             post.mean l-95% CI u-95% CI eff.samp
24 geneNORM.units     1.1929   0.9095   1.4661   474.6
25 geneCT_40.units    0.3536   0.2313   0.4828   304.2
26 geneCT_70.units    0.7093   0.5239   0.8606   485.2
27 geneCT_83.units    2.5393   1.9813   3.1185   115.3
28
29 Location effects: count ~ gene + RIN_Value + Jdaycont
+ Site * baittemp.ave + Site * Delta + gene:Site * baittemp.ave + gene:Site * Delta
30
31             post.mean   l-95% CI
32             u-95% CI eff.samp   pMCMC
33 (Intercept)          8.684991  6.169046
34             11.086840  577.4 < 0.001 ** 
35 geneCT_40            -9.671732 -12.106692
36             -7.552922  468.7 < 0.001 ** 
37 geneCT_70            -9.182214 -11.530024
38             -7.134289  514.8 < 0.001 ** 
39 geneCT_83            -11.866716 -14.977469
40             -8.537719  535.9 < 0.001 ** 
41 RIN_Value            0.211073  0.107941
42             0.331795  700.0 < 0.001 ** 
43 Jdaycont              0.005798  0.004379
44             0.007320  700.0 < 0.001 ** 
45 SiteHF                -4.721952 -11.937967
46             3.212892  411.7 0.20286
47 baittemp.ave           -0.113720 -0.214446
48             -0.025546  438.3 0.01429 *

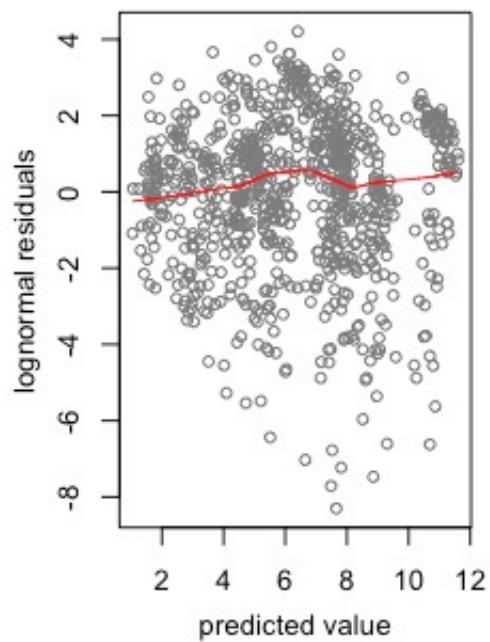
```

40	Delta		0.084153	-0.108549
	0.293981	700.0 0.41143		
41	SiteHF:baittemp.ave		0.209194	-0.101481
	0.482602	700.0 0.14857		
42	SiteHF:Delta		-0.203374	-0.514559
	0.153917	700.0 0.23429		
43	geneCT_40:SiteHF		-1.897398	-6.878381
	3.333412	700.0 0.48000		
44	geneCT_70:SiteHF		-0.323395	-6.073259
	5.043501	610.0 0.92000		
45	geneCT_83:SiteHF		-7.705725	-14.989465
	0.375533	601.6 0.06857 .		
46	geneCT_40:SiteDF:baittemp.ave		0.157087	0.069875
	0.242143	505.6 < 0.001 **		
47	geneCT_70:SiteDF:baittemp.ave		0.256887	0.166474
	0.339240	515.3 < 0.001 **		
48	geneCT_83:SiteDF:baittemp.ave		0.308844	0.178066
	0.427849	599.9 < 0.001 **		
49	geneCT_40:SiteHF:baittemp.ave		0.230984	0.060239
	0.419166	700.0 0.01429 *		
50	geneCT_70:SiteHF:baittemp.ave		0.266757	0.080296
	0.462425	542.0 0.00857 **		
51	geneCT_83:SiteHF:baittemp.ave		0.597438	0.342265
	0.859273	494.2 < 0.001 **		
52	geneCT_40:SiteDF:Delta		-0.065660	-0.240177
	0.106559	700.0 0.42286		
53	geneCT_70:SiteDF:Delta		-0.127074	-0.322574
	0.049327	882.0 0.15429		
54	geneCT_83:SiteDF:Delta		0.015443	-0.224270
	0.231865	700.0 0.90857		
55	geneCT_40:SiteHF:Delta		-0.124124	-0.320080
	0.061610	700.0 0.22571		
56	geneCT_70:SiteHF:Delta		-0.050784	-0.270149
	0.156429	700.0 0.66571		

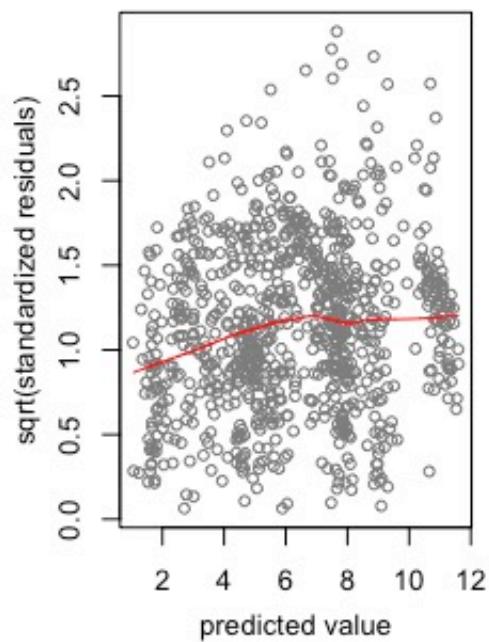
```
57 geneCT_83:SiteHF:Delta      -0.166027  -0.435081  
58 0.113876    607.6 0.26000  
58 ---  
59 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'  
59 0.1 ' ' 1
```

diagnostics

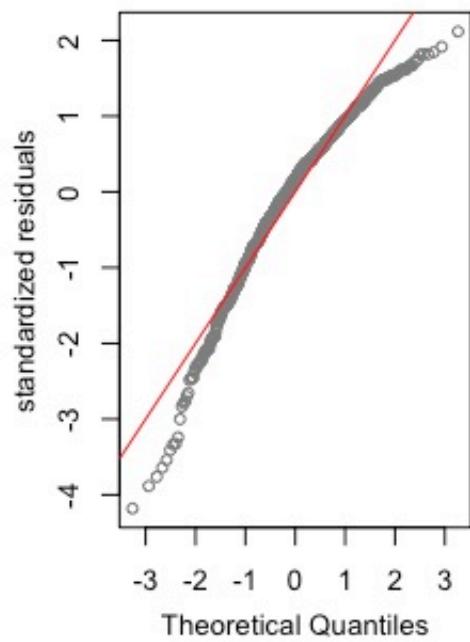
residuals vs predicted



Scale-Location



normal QQ plot



Model 4: Soft normalization- Including 18s, actin, and gapdh to normalize; include

interaction with season

Model testing for effect of Site by bait temp by julian day continuous; controlling for RIN and including chamber as a random effect

- Fixed effects: Site * bait*Jdaycont
- Global fixed effects: RIN value (rna quality)
- Random factor: Chamber

```
1 softnorm2=mcmc.qpcr(fixed="Site*baittemp.ave*Jdaycont",r  
andom="Cham",data=dd,pr=TRUE,vprior="iw",nitt=10000,gene  
SpecRes =  
TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_actin","CT_ga  
pdh","CT_18s"),globalFixed=c("RIN_Value"))  
2 summary(softnorm2)
```

Model 4 output: Soft normalization- Including 18s, actin, and gapdh to normalize; include interaction with season

```
1 Iterations = 3001:9991  
2 Thinning interval = 10  
3 Sample size = 700  
4  
5 DIC: 9005.803  
6  
7 G-structure: ~sample  
8  
9          post.mean l-95% CI u-95% CI eff.samp  
10 sample     1.818    1.433    2.249    610.9  
11  
12          ~idh(gene):Cham  
13
```

```

14          post.mean l-95% CI u-95% CI eff.samp
15 geneNORM.Cham      0.3805    0.1372    0.6941    790.7
16 geneCT_40.Cham    0.2881    0.1069    0.5029    700.0
17 geneCT_70.Cham    0.2813    0.1253    0.4923    700.0
18 geneCT_83.Cham    0.3909    0.1561    0.6891    700.0
19
20 R-structure: ~idh(gene):units
21
22          post.mean l-95% CI u-95% CI eff.samp
23 geneNORM.units     0.8994    0.7000    1.0903    700.0
24 geneCT_40.units    0.3020    0.2036    0.3963    344.5
25 geneCT_70.units    0.5810    0.4347    0.7147    576.4
26 geneCT_83.units    1.1998    0.9383    1.4820    237.8
27
28 Location effects: count ~ gene + RIN_Value + Site *
baittemp.ave * Jdaycont + gene:Site * baittemp.ave *
Jdaycont
29
30          post.mean l-
31 95% CI   u-95% CI eff.samp pMCMC
32 (Intercept)           -8.661e+00
-1.331e+01 -4.130e+00    700.0 < 0.001 ** 
33 geneCT_40              -1.903e+00
-6.363e+00  2.411e+00    700.0  0.39143
34 geneCT_70              -5.688e+00
-9.774e+00 -1.181e+00    700.0  0.00857 ** 
35 geneCT_83              -2.090e+01
-2.622e+01 -1.549e+01    700.0 < 0.001 ** 
36 RIN_Value                1.981e-01
9.955e-02  3.036e-01    575.0 < 0.001 ** 
37 SiteHF                  1.648e+01
4.298e+00  2.825e+01    700.0  0.00286 ** 
38 baittemp.ave               5.477e-01
3.813e-01  7.313e-01    700.0 < 0.001 ** 

```

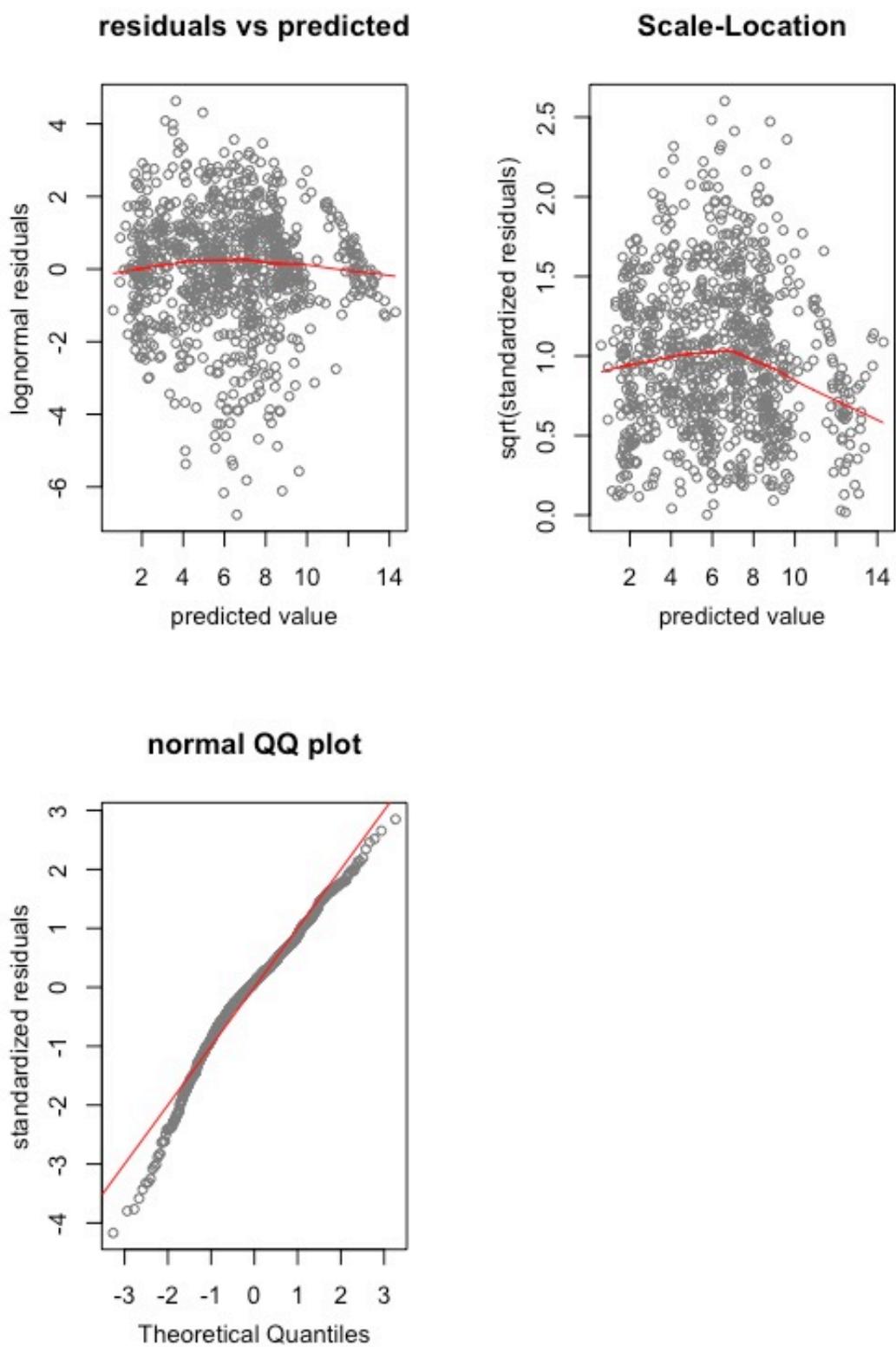
38	Jdaycont		6.788e-02
	5.544e-02	8.040e-02	700.0 < 0.001 **
39	SiteHF:baittemp.ave		-6.478e-01
	-1.100e+00	-1.991e-01	700.0 < 0.001 **
40	SiteHF:Jdaycont		-7.108e-02
	-1.052e-01	-4.008e-02	700.0 < 0.001 **
41	baittemp.ave:Jdaycont		-2.314e-03
	-2.771e-03	-1.835e-03	700.0 < 0.001 **
42	geneCT_40:SiteHF		-1.646e+01
	-2.491e+01	-6.981e+00	485.8 < 0.001 **
43	geneCT_70:SiteHF		-1.118e+01
	-2.052e+01	-2.765e+00	700.0 0.02571 *
44	geneCT_83:SiteHF		-1.103e+01
	-2.284e+01	-5.886e-01	700.0 0.04571 *
45	SiteHF:baittemp.ave:Jdaycont		2.783e-03
	1.619e-03	4.034e-03	700.0 < 0.001 **
46	geneCT_40:SiteDF:baittemp.ave		-1.120e-01
	-2.825e-01	5.198e-02	700.0 0.19429
47	geneCT_70:SiteDF:baittemp.ave		1.504e-01
	-2.890e-02	3.083e-01	700.0 0.08571 .
48	geneCT_83:SiteDF:baittemp.ave		6.275e-01
	4.305e-01	8.497e-01	700.0 < 0.001 **
49	geneCT_40:SiteHF:baittemp.ave		5.121e-01
	2.556e-01	7.882e-01	373.2 < 0.001 **
50	geneCT_70:SiteHF:baittemp.ave		5.965e-01
	3.218e-01	9.269e-01	700.0 < 0.001 **
51	geneCT_83:SiteHF:baittemp.ave		9.978e-01
	6.902e-01	1.447e+00	618.8 < 0.001 **
52	geneCT_40:SiteDF:Jdaycont		-2.229e-02
	-3.189e-02	-1.169e-02	700.0 < 0.001 **
53	geneCT_70:SiteDF:Jdaycont		-1.194e-02
	-2.321e-02	-1.849e-03	700.0 0.02571 *
54	geneCT_83:SiteDF:Jdaycont		2.635e-02
	1.412e-02	4.098e-02	700.0 < 0.001 **

```

55 geneCT_40:SiteHF:Jdaycont           2.447e-02
      5.219e-03  4.679e-02    700.0  0.01714 *
56 geneCT_70:SiteHF:Jdaycont           2.317e-02
      1.383e-03  4.789e-02    700.0  0.05143 .
57 geneCT_83:SiteHF:Jdaycont           4.817e-02
      1.820e-02  7.441e-02    700.0  0.00286 **
58 geneCT_40:SiteDF:baittemp.ave:Jdaycont 7.573e-04
      3.373e-04  1.118e-03    700.0 < 0.001 **
59 geneCT_70:SiteDF:baittemp.ave:Jdaycont 3.523e-04
      -4.856e-05 7.660e-04    700.0  0.10857
60 geneCT_83:SiteDF:baittemp.ave:Jdaycont -9.246e-04
      -1.390e-03 -3.810e-04   700.0 < 0.001 **
61 geneCT_40:SiteHF:baittemp.ave:Jdaycont -1.025e-03
      -1.816e-03 -3.201e-04   700.0  0.00857 **
62 geneCT_70:SiteHF:baittemp.ave:Jdaycont -1.041e-03
      -1.925e-03 -2.063e-04   700.0  0.02571 *
63 geneCT_83:SiteHF:baittemp.ave:Jdaycont -1.651e-03
      -2.635e-03 -5.883e-04   700.0  0.00571 **
64 ---
65 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'.
                  0.1 ' ' 1

```

diagnostics



Model 5: Naive- most complex model- 3 way interaction between delta, site, Jdaycont

and 3 way interaction between bait, site, Jdaycont

Model testing for effect of Site by bait temp by julian day continuous; controlling for RIN and including chamber as a random effect

- Fixed effects: Site * bait* Jdaycont + Site * Delta* Jdaycont
- Global fixed effects: RIN value (rna quality)
- Random factor: Chamber

```
1 com1=mcmc.qpcr(  
2   fixed="Site*baittemp.ave*Jdaycont+Site*Delta*Jdaycont",  
3   random=c("Cham"),globalFixed=c("RIN_Value"),  
4   data=dd,pr=TRUE,vprior="iw",nitt=25000,geneSpecRes =  
  TRUE,pl=TRUE)  
5 summary(com1)  
6 diagnostic.mcmc(model=com1, col="grey50",cex=0.8)
```

Model 5 output: Naive- most complex model- 3 way interaction between delta, site, Jdaycont and 3 way interaction between bait, site, Jdaycont

```
1 Iterations = 3001:24991  
2 Thinning interval = 10  
3 Sample size = 2200  
4  
5 DIC: 15031.14  
6  
7 G-structure: ~sample  
8  
9      post.mean l-95% CI u-95% CI eff.samp
```

```

10 sample      1.837     1.466     2.245      2200
11
12 ~idh(gene):Cham
13
14 post.mean l-95% CI u-95% CI eff.samp
15 geneCT_18s.Cham    0.5911    0.2259    1.0935    2200
16 geneCT_40.Cham    0.3927    0.1621    0.6771    2200
17 geneCT_70.Cham    0.3983    0.1698    0.7059    2200
18 geneCT_83.Cham    0.4718    0.1777    0.8140    2200
19 geneCT_actin.Cham 0.4319    0.1643    0.7832    2200
20 geneCT_gapdh.Cham 0.5181    0.2231    0.9287    2200
21
22 R-structure: ~idh(gene):units
23
24 post.mean l-95% CI u-95% CI
eff.samp
25 geneCT_18s.units    4.4093    3.5220    5.2938
2027.8
26 geneCT_40.units    0.3753    0.2659    0.4949
1152.1
27 geneCT_70.units    0.5522    0.4221    0.6899
1626.8
28 geneCT_83.units    1.0948    0.8301    1.3632
528.7
29 geneCT_actin.units 1.1718    0.8997    1.4371
1410.1
30 geneCT_gapdh.units 1.0445    0.8184    1.3015
1325.4
31
32 Location effects: count ~ 0 + gene + RIN_Value +
+gene:Site * baittemp.ave * Jdaycont + gene:Site *
Delta * Jdaycont
33
34 post.mean
l-95% CI   u-95% CI eff.samp      pMCMC

```

35	geneCT_18s					-2.823e+00
	-1.120e+01	5.159e+00	2200	0.485455		
36	geneCT_40					-1.189e+01
	-1.608e+01	-7.850e+00	2010	< 5e-04	***	
37	geneCT_70					-1.565e+01
	-1.968e+01	-1.169e+01	2200	< 5e-04	***	
38	geneCT_83					-3.130e+01
	-3.643e+01	-2.673e+01	1330	< 5e-04	***	
39	geneCT_actin					-1.952e+01
	-2.438e+01	-1.490e+01	2200	< 5e-04	***	
40	geneCT_gapdh					-2.262e+01
	-2.759e+01	-1.794e+01	1919	< 5e-04	***	
41	RIN_Value					1.995e-01
	1.083e-01	2.974e-01	1773	< 5e-04	***	
42	baittemp.ave					-1.349e-01
	-6.804e-01	3.516e-01	1970	0.606364		
43	Jdaycont					-6.081e-05
	-4.225e-02	4.287e-02	2395	0.995455		
44	Delta					-2.896e-01
	-7.886e-01	1.875e-01	2200	0.236364		
45	geneCT_18s:SiteHF					1.077e+01
	-1.013e+01	3.064e+01	2200	0.318182		
46	geneCT_40:SiteHF					-3.127e+00
	-1.523e+01	9.476e+00	1859	0.630909		
47	geneCT_70:SiteHF					-3.848e+00
	-1.662e+01	9.315e+00	2200	0.567273		
48	geneCT_83:SiteHF					2.845e+00
	-1.191e+01	1.736e+01	1948	0.690909		
49	geneCT_actin:SiteHF					1.153e+01
	-3.734e+00	2.505e+01	2200	0.120909		
50	geneCT_gapdh:SiteHF					2.778e+01
	1.389e+01	4.227e+01	2015	0.000909	***	
51	baittemp.ave:Jdaycont					4.026e-04
	-1.357e-03	1.953e-03	2385	0.623636		

52	Jdaycont:Delta					4.409e-04
	-1.011e-03	1.869e-03	2200	0.561818		
53	geneCT_18s:SiteDF:baittemp.ave					6.757e-01
	1.042e-02	1.254e+00	2145	0.034545	*	
54	geneCT_40:SiteDF:baittemp.ave					6.270e-01
	7.783e-02	1.167e+00	2051	0.027273	*	
55	geneCT_70:SiteDF:baittemp.ave					8.921e-01
	3.226e-01	1.409e+00	2007	0.001818	**	
56	geneCT_83:SiteDF:baittemp.ave					1.386e+00
	8.521e-01	1.948e+00	1966	< 5e-04	***	
57	geneCT_actin:SiteDF:baittemp.ave					1.009e+00
	4.530e-01	1.552e+00	2200	0.001818	**	
58	geneCT_gapdh:SiteDF:baittemp.ave					1.098e+00
	5.528e-01	1.628e+00	2027	0.000909	***	
59	geneCT_18s:SiteHF:baittemp.ave					3.694e-01
	-3.419e-01	1.036e+00	2257	0.308182		
60	geneCT_40:SiteHF:baittemp.ave					7.578e-01
	3.526e-01	1.115e+00	1965	< 5e-04	***	
61	geneCT_70:SiteHF:baittemp.ave					1.087e+00
	6.721e-01	1.456e+00	2200	< 5e-04	***	
62	geneCT_83:SiteHF:baittemp.ave					1.249e+00
	7.937e-01	1.743e+00	1088	< 5e-04	***	
63	geneCT_actin:SiteHF:baittemp.ave					5.374e-01
	6.048e-02	1.010e+00	1737	0.028182	*	
64	geneCT_18s:SiteDF:Jdaycont					7.880e-02
	3.089e-02	1.248e-01	2298	0.005455	**	
65	geneCT_40:SiteDF:Jdaycont					4.986e-02
	9.422e-03	9.765e-02	2473	0.030000	*	
66	geneCT_70:SiteDF:Jdaycont					5.994e-02
	1.345e-02	1.040e-01	2385	0.012727	*	
67	geneCT_83:SiteDF:Jdaycont					9.934e-02
	5.596e-02	1.459e-01	2284	< 5e-04	***	
68	geneCT_actin:SiteDF:Jdaycont					7.985e-02
	3.446e-02	1.240e-01	2393	0.002727	**	

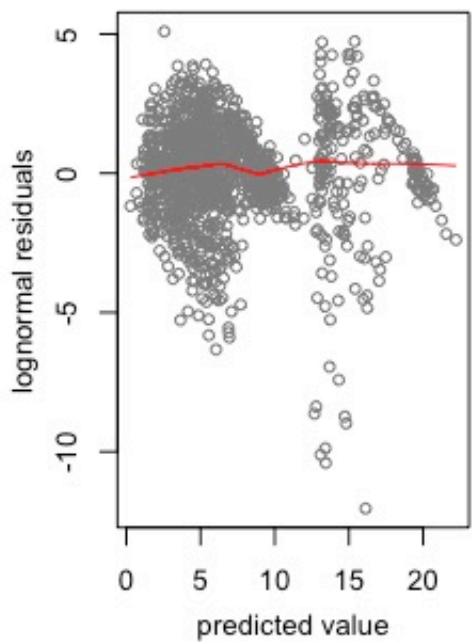
69	geneCT_gapdh:SiteDF:Jdaycont					8.841e-02
	3.979e-02	1.303e-01	2200	< 5e-04	***	
70	geneCT_18s:SiteHF:Jdaycont					1.270e-02
	-4.679e-02	7.036e-02	2518	0.690909		
71	geneCT_40:SiteHF:Jdaycont					3.339e-02
	8.214e-04	6.623e-02	1801	0.047273	*	
72	geneCT_70:SiteHF:Jdaycont					5.750e-02
	2.394e-02	9.013e-02	2200	0.000909	***	
73	geneCT_83:SiteHF:Jdaycont					5.783e-02
	1.919e-02	9.634e-02	1586	0.003636	**	
74	geneCT_actin:SiteHF:Jdaycont					2.431e-02
	-1.361e-02	6.343e-02	1747	0.221818		
75	geneCT_18s:SiteDF:Delta					8.948e-02
	-6.011e-01	7.888e-01	2200	0.804545		
76	geneCT_40:SiteDF:Delta					1.564e-01
	-4.777e-01	6.987e-01	2200	0.602727		
77	geneCT_70:SiteDF:Delta					1.317e-01
	-4.143e-01	7.291e-01	2200	0.644545		
78	geneCT_83:SiteDF:Delta					1.091e-01
	-4.431e-01	7.837e-01	2200	0.714545		
79	geneCT_actin:SiteDF:Delta					1.013e-01
	-5.002e-01	7.231e-01	2200	0.749091		
80	geneCT_gapdh:SiteDF:Delta					2.241e-01
	-3.804e-01	8.227e-01	2200	0.488182		
81	geneCT_18s:SiteHF:Delta					-5.227e-01
	-1.201e+00	8.997e-02	2200	0.110909		
82	geneCT_40:SiteHF:Delta					-2.061e-01
	-5.670e-01	1.574e-01	1595	0.267273		
83	geneCT_70:SiteHF:Delta					-4.682e-01
	-8.370e-01	-8.751e-02	2200	0.013636	*	
84	geneCT_83:SiteHF:Delta					-2.866e-01
	-7.655e-01	1.374e-01	1535	0.212727		
85	geneCT_actin:SiteHF:Delta					-1.495e-01
	-6.054e-01	3.167e-01	1826	0.530000		

86	geneCT_18s:SiteDF:baittemp.ave:Jdaycont	-3.128e-03 -4.892e-03 -1.247e-03	2289 0.003636 **
87	geneCT_40:SiteDF:baittemp.ave:Jdaycont	-2.150e-03 -3.735e-03 -3.373e-04	2456 0.016364 *
88	geneCT_70:SiteDF:baittemp.ave:Jdaycont	-2.535e-03 -4.252e-03 -7.815e-04	2376 0.006364 **
89	geneCT_83:SiteDF:baittemp.ave:Jdaycont	-3.902e-03 -5.534e-03 -2.096e-03	2286 < 5e-04 ***
90	geneCT_actin:SiteDF:baittemp.ave:Jdaycont	-3.288e-03 -4.920e-03 -1.450e-03	2407 0.000909 ***
91	geneCT_gapdh:SiteDF:baittemp.ave:Jdaycont	-3.545e-03 -5.112e-03 -1.627e-03	2200 < 5e-04 ***
92	geneCT_18s:SiteHF:baittemp.ave:Jdaycont	-6.323e-04 -2.889e-03 1.635e-03	2519 0.590000
93	geneCT_40:SiteHF:baittemp.ave:Jdaycont	-1.495e-03 -2.860e-03 -3.344e-04	1816 0.022727 *
94	geneCT_70:SiteHF:baittemp.ave:Jdaycont	-2.562e-03 -3.815e-03 -1.243e-03	2200 < 5e-04 ***
95	geneCT_83:SiteHF:baittemp.ave:Jdaycont	-2.173e-03 -3.659e-03 -6.606e-04	1935 0.003636 **
96	geneCT_actin:SiteHF:baittemp.ave:Jdaycont	-1.130e-03 -2.636e-03 3.316e-04	1762 0.150909
97	geneCT_18s:SiteDF:Jdaycont:Delta		6.662e-04
		-1.029e-03 2.752e-03	2200 0.480909
98	geneCT_40:SiteDF:Jdaycont:Delta		8.687e-05
		-1.483e-03 1.754e-03	2200 0.916364
99	geneCT_70:SiteDF:Jdaycont:Delta		-8.150e-05
		-1.645e-03 1.546e-03	2200 0.919091
100	geneCT_83:SiteDF:Jdaycont:Delta		5.699e-04
		-1.078e-03 2.176e-03	2200 0.487273
101	geneCT_actin:SiteDF:Jdaycont:Delta		1.675e-04
		-1.521e-03 1.799e-03	2200 0.839091
102	geneCT_gapdh:SiteDF:Jdaycont:Delta		4.072e-04
		-1.305e-03 2.010e-03	2200 0.650909

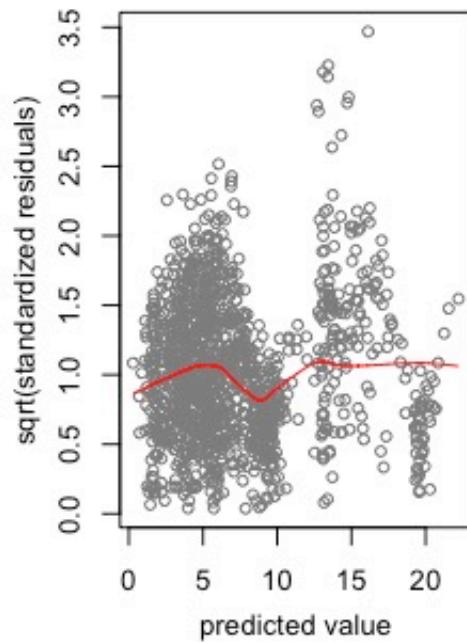
```
103 geneCT_18s:SiteHF:Jdaycont:Delta      8.713e-04  
     -9.942e-04  2.950e-03    2347 0.393636  
104 geneCT_40:SiteHF:Jdaycont:Delta      5.417e-04  
     -6.536e-04  1.532e-03    1787 0.334545  
105 geneCT_70:SiteHF:Jdaycont:Delta      1.767e-03  
     6.850e-04  2.950e-03    2200 0.002727 **  
106 geneCT_83:SiteHF:Jdaycont:Delta      7.647e-04  
     -6.120e-04  2.075e-03    1867 0.241818  
107 geneCT_actin:SiteHF:Jdaycont:Delta    8.130e-04  
     -5.026e-04  2.161e-03    1832 0.237273  
108 ---  
109 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'  
                      0.1 ' ' 1
```

diagnostics

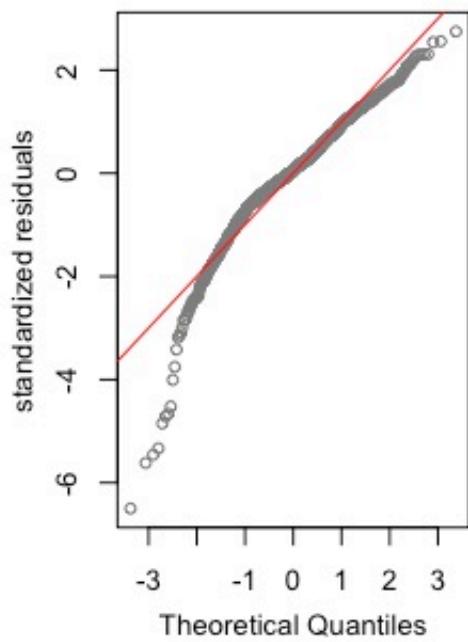
residuals vs predicted



Scale-Location



normal QQ plot



Model 6: Soft normalization- most complex model- 3 way interaction between delta,

site, Jdaycont and 3 way interaction between bait, site, Jdaycont

Model testing for effect of Site by bait temp by julian day continuous and Site by delta by julian day continuous; controlling for RIN and including chamber as a random effect

- Fixed effects: Sitebaittemp.aveJdaycont+SiteDeltaJdaycont
- Global fixed effects: RIN value (rna quality)
- Random factor: Chamber

```
1 comp2=mcmc.qpcr(fixed="Site*baittemp.ave*Jdaycont+Site*D  
elta*Jdaycont",random="Cham",data=dd,pr=TRUE,vprior="iw"  
,nitt=25000,geneSpecRes =  
TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_actin","CT_ga  
pdh","CT_18s"),globalFixed=c("RIN_Value"))  
2 summary(comp2)  
3 diagnostic.mcmc(model=comp2, col="grey50",cex=0.8)
```

Model 6 OUTPUT: Soft normalization- most complex model- 3 way interaction between delta, site, Jdaycont and 3 way interaction between bait, site, Jdaycont

```
1 Iterations = 3001:24991  
2 Thinning interval = 10  
3 Sample size = 2200  
4  
5 DIC: 9009.713  
6  
7 G-structure: ~sample  
8
```

```

9      post.mean l-95% CI u-95% CI eff.samp
10 sample     1.732     1.385     2.184     1501
11
12 ~idh(gene):Cham
13
14      post.mean l-95% CI u-95% CI eff.samp
15 geneNORM.Cham    0.4025    0.1483    0.7419    2503
16 geneCT_40.Cham   0.2984    0.1205    0.5389    2200
17 geneCT_70.Cham   0.3028    0.1228    0.5676    2200
18 geneCT_83.Cham   0.3957    0.1406    0.7272    2001
19
20 R-structure: ~idh(gene):units
21
22      post.mean l-95% CI u-95% CI eff.samp
23 geneNORM.units   0.8910    0.6979    1.1078    2007.8
24 geneCT_40.units  0.2969    0.2043    0.4117    1134.6
25 geneCT_70.units  0.5373    0.4173    0.6735    1201.4
26 geneCT_83.units  1.1912    0.9067    1.5100    621.4
27
28 Location effects: count ~ gene + RIN_Value + Site *
29                                * baittemp.ave * Jdaycont + Site * Delta * Jdaycont +
30                                gene:Site * baittemp.ave * Jdaycont + gene:Site * Delta
31                                * Jdaycont
32
33      post.mean l-
34      95% CI u-95% CI eff.samp pMCMC
35 (Intercept)           -8.798e+00
36 -1.361e+01 -4.331e+00    2015 < 5e-04 ***
37 geneCT_40              -1.588e+00
38 -5.868e+00  2.627e+00    1925 0.484545
39 geneCT_70              -5.515e+00
40 -1.008e+01 -9.670e-01    1960 0.018182 *
41 geneCT_83              -2.119e+01
42 -2.675e+01 -1.585e+01    1614 < 5e-04 ***

```

35	RIN_Value				1.910e-01	
	1.024e-01	3.002e-01	1908	< 5e-04	***	
36	SiteHF				1.349e+01	
	3.435e-01	2.695e+01	2089	0.045455	*	
37	baittemp.ave				5.616e-01	
	3.747e-01	7.367e-01	1991	< 5e-04	***	
38	Jdaycont				6.935e-02	
	5.715e-02	8.314e-02	2200	< 5e-04	***	
39	Delta				-1.194e-01	
	-4.827e-01	2.018e-01	2200	0.497273		
40	SiteHF:baittemp.ave				-5.084e-01	
	-1.016e+00	7.365e-04	2089	0.049091	*	
41	SiteHF:Jdaycont				-6.249e-02	
	-1.044e-01	-2.113e-02	2200	0.000909	***	
42	baittemp.ave:Jdaycont				-2.424e-03	
	-2.926e-03	-1.934e-03	2200	< 5e-04	***	
43	SiteHF:Delta				-2.491e-01	
	-7.974e-01	3.488e-01	2200	0.403636		
44	Jdaycont:Delta				7.583e-04	
	1.996e-05	1.463e-03	2200	0.044545	*	
45	geneCT_40:SiteHF				-1.740e+01	
	-2.860e+01	-7.230e+00	1983	< 5e-04	***	
46	geneCT_70:SiteHF				-1.846e+01	
	-3.004e+01	-7.750e+00	1856	0.000909	***	
47	geneCT_83:SiteHF				-1.163e+01	
	-2.530e+01	7.869e-01	1777	0.068182	.	
48	SiteHF:baittemp.ave:Jdaycont				2.444e-03	
	7.669e-04	4.034e-03	2200	0.000909	***	
49	SiteHF:Jdaycont:Delta				4.717e-05	
	-1.613e-03	1.689e-03	2200	0.953636		
50	geneCT_40:SiteDF:baittemp.ave				-1.212e-01	
	-2.943e-01	3.885e-02	1942	0.154545		
51	geneCT_70:SiteDF:baittemp.ave				1.481e-01	
	-3.409e-02	3.180e-01	1932	0.092727	.	

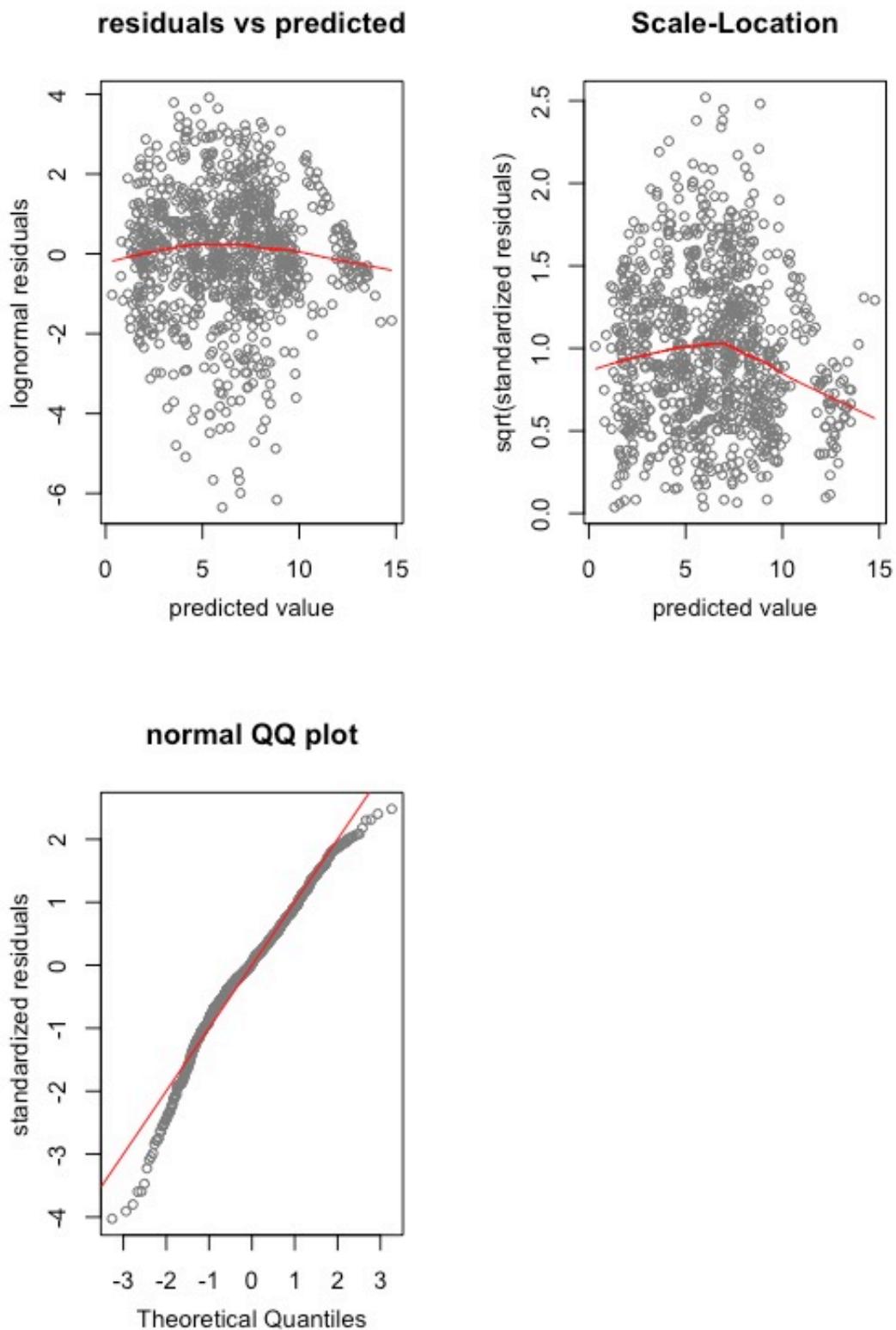
52	geneCT_83:SiteDF:baittemp.ave		6.444e-01	
	4.416e-01	8.651e-01	1657	< 5e-04 ***
53	geneCT_40:SiteHF:baittemp.ave		5.427e-01	
	1.903e-01	9.062e-01	1798	0.002727 **
54	geneCT_70:SiteHF:baittemp.ave		8.914e-01	
	4.999e-01	1.267e+00	2087	< 5e-04 ***
55	geneCT_83:SiteHF:baittemp.ave		1.051e+00	
	5.834e-01	1.500e+00	1633	< 5e-04 ***
56	geneCT_40:SiteDF:Jdaycont		-2.335e-02	
	-3.356e-02	-1.328e-02	1976	< 5e-04 ***
57	geneCT_70:SiteDF:Jdaycont		-1.270e-02	
	-2.331e-02	-1.892e-03	2200	0.019091 *
58	geneCT_83:SiteDF:Jdaycont		2.683e-02	
	1.438e-02	4.077e-02	1672	< 5e-04 ***
59	geneCT_40:SiteHF:Jdaycont		2.504e-02	
	-3.193e-03	5.420e-02	1867	0.092727 .
60	geneCT_70:SiteHF:Jdaycont		5.021e-02	
	1.877e-02	8.278e-02	1945	0.000909 ***
61	geneCT_83:SiteHF:Jdaycont		5.046e-02	
	1.094e-02	8.666e-02	2200	0.006364 **
62	geneCT_40:SiteDF:Delta		-2.968e-02	
	-2.735e-01	2.198e-01	2042	0.810000
63	geneCT_70:SiteDF:Delta		-5.073e-02	
	-3.200e-01	1.949e-01	2200	0.718182
64	geneCT_83:SiteDF:Delta		-8.435e-02	
	-3.878e-01	2.478e-01	1936	0.610909
65	geneCT_40:SiteHF:Delta		-9.843e-02	
	-4.475e-01	2.744e-01	1733	0.589091
66	geneCT_70:SiteHF:Delta		-3.744e-01	
	-7.534e-01	-3.034e-04	1868	0.042727 *
67	geneCT_83:SiteHF:Delta		-2.118e-01	
	-6.535e-01	2.681e-01	1508	0.342727
68	geneCT_40:SiteDF:baittemp.ave:Jdaycont	8.110e-04		
	4.423e-04	1.225e-03	1973	< 5e-04 ***

```

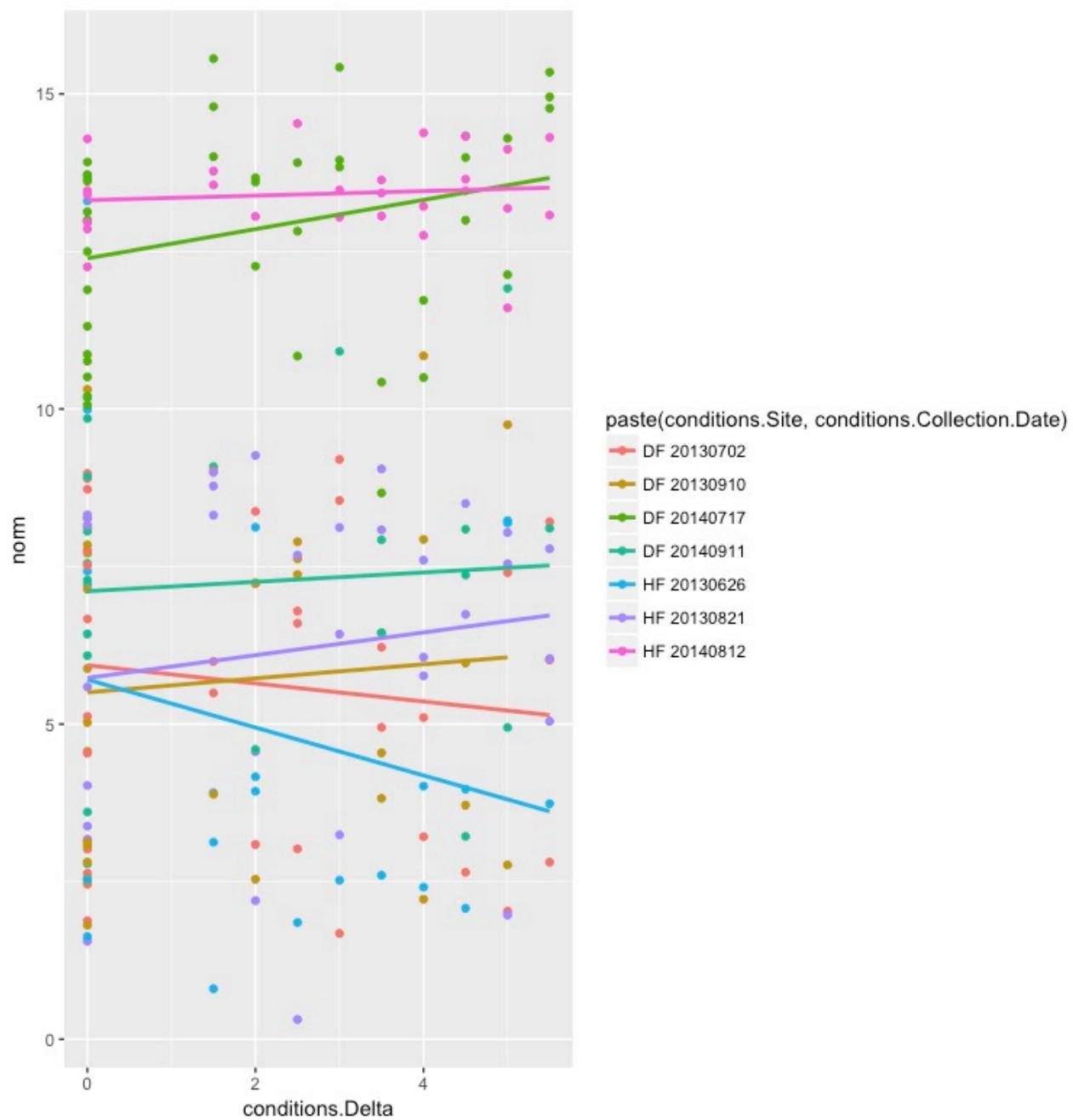
69 geneCT_70:SiteDF:baittemp.ave:Jdaycont 4.062e-04
    -1.848e-06 8.277e-04      2200 0.050909 .
70 geneCT_83:SiteDF:baittemp.ave:Jdaycont -9.647e-04
    -1.500e-03 -4.873e-04      1780 < 5e-04 ***
71 geneCT_40:SiteHF:baittemp.ave:Jdaycont -1.054e-03
    -2.134e-03 1.010e-04      1879 0.069091 .
72 geneCT_70:SiteHF:baittemp.ave:Jdaycont -2.162e-03
    -3.387e-03 -9.247e-04      1956 < 5e-04 ***
73 geneCT_83:SiteHF:baittemp.ave:Jdaycont -1.771e-03
    -3.105e-03 -1.986e-04      2200 0.016364 *
74 geneCT_40:SiteDF:Jdaycont:Delta          -1.827e-04
    -6.758e-04 3.058e-04      2200 0.485455
75 geneCT_70:SiteDF:Jdaycont:Delta         -3.358e-04
    -8.632e-04 2.170e-04      2522 0.224545
76 geneCT_83:SiteDF:Jdaycont:Delta         2.994e-04
    -3.849e-04 1.003e-03      2117 0.400000
77 geneCT_40:SiteHF:Jdaycont:Delta         1.086e-04
    -9.713e-04 1.120e-03      2200 0.841818
78 geneCT_70:SiteHF:Jdaycont:Delta         1.361e-03
    8.136e-05 2.380e-03      2009 0.014545 *
79 geneCT_83:SiteHF:Jdaycont:Delta         3.906e-04
    -9.783e-04 1.700e-03      2066 0.558182
80 ---
81 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

diagnostics



**plotting hsp70 for this model 6 against delta
and site and collection date**



Using a more simple approach: ANOVA

Testing the effect of gene by site by delta by season(julian day) and gene by site by bait temp by season(julian day)

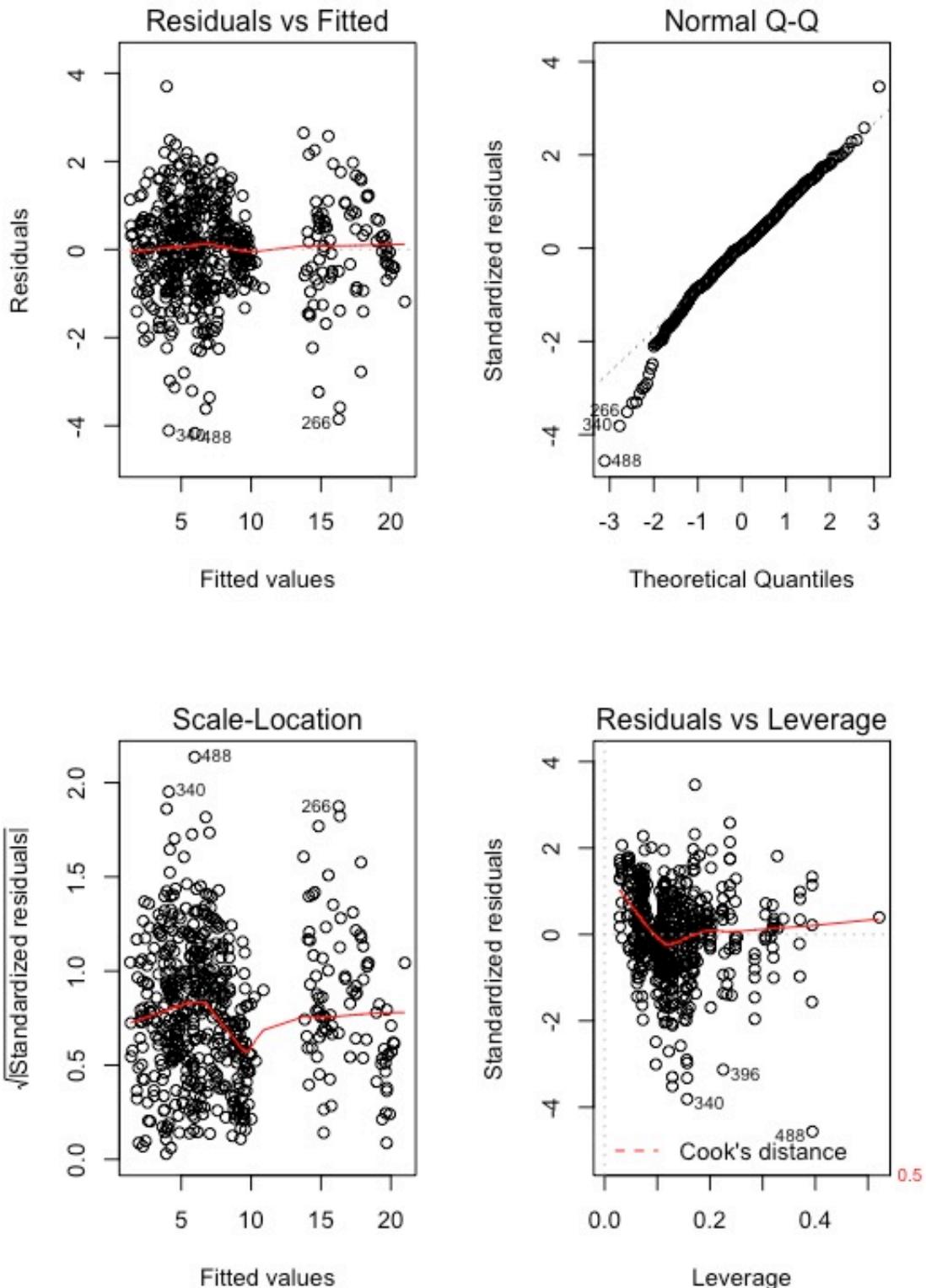
```
1 glob2<-
aov(log(count+1)~RIN_Value+gene*Site*Jdaycont*Delta+gene
*Site*Jdaycont*baittemp.ave,data=ddave)
```

output:

		Df	Sum Sq	Mean Sq	F
1	value Pr(>F)				
2	RIN_Value	1	1223	1223.4	
	887.552 < 2e-16 ***				
3	gene	5	9337	1867.3	
	1354.741 < 2e-16 ***				
4	Site	1	5	4.7	
	3.379 0.0666662 .				
5	Jdaycont	1	76	75.7	
	54.889 5.93e-13 ***				
6	Delta	1	1	1.1	
	0.828 0.363410				
7	baittemp.ave	1	4	4.2	
	3.054 0.081167 .				
8	gene:Site	5	36	7.3	
	5.287 9.74e-05 ***				
9	gene:Jdaycont	5	72	14.5	
	10.512 1.39e-09 ***				
10	Site:Jdaycont	1	30	29.7	
	21.582 4.40e-06 ***				
11	gene:Delta	5	2	0.4	
	0.301 0.912386				
12	Site:Delta	1	1	1.1	
	0.770 0.380517				
13	Jdaycont:Delta	1	4	3.6	
	2.623 0.105977				
14	gene:baittemp.ave	5	55	11.0	
	7.967 3.19e-07 ***				
15	Site:baittemp.ave	1	41	40.8	
	29.587 8.60e-08 ***				
16	Jdaycont:baittemp.ave	1	448	447.8	
	324.912 < 2e-16 ***				

17	gene:Site:Jdaycont	5	11	2.2
	1.596 0.159755			
18	gene:Site:Delta	5	3	0.7
	0.498 0.778013			
19	gene:Jdaycont:Delta	5	1	0.1
	0.092 0.993462			
20	Site:Jdaycont:Delta	1	20	19.8
	14.387 0.000168 ***			
21	gene:Site:baittemp.ave	5	9	1.9
	1.360 0.237899			
22	gene:Jdaycont:baittemp.ave	5	24	4.8
	3.480 0.004242 **			
23	Site:Jdaycont:baittemp.ave	1	14	14.2
	10.283 0.001434 **			
24	gene:Site:Jdaycont:Delta	5	6	1.2
	0.870 0.501010			
25	gene:Site:Jdaycont:baittemp.ave	5	9	1.8
	1.339 0.246378			
26	Residuals	472	651	1.4
27	---			
28	Signif. codes:	0 *** 0.001 ** 0.01 * 0.05 .		
	0.1 ' ' 1			

diagnostics:



Meeting with SHC and NJG

- split by site and do models because control genes are different

between sites

- do the traditional delta delta ct method
-

Page 76: 2017-05-04. Stressed in nature project: Analyses including technical replicates

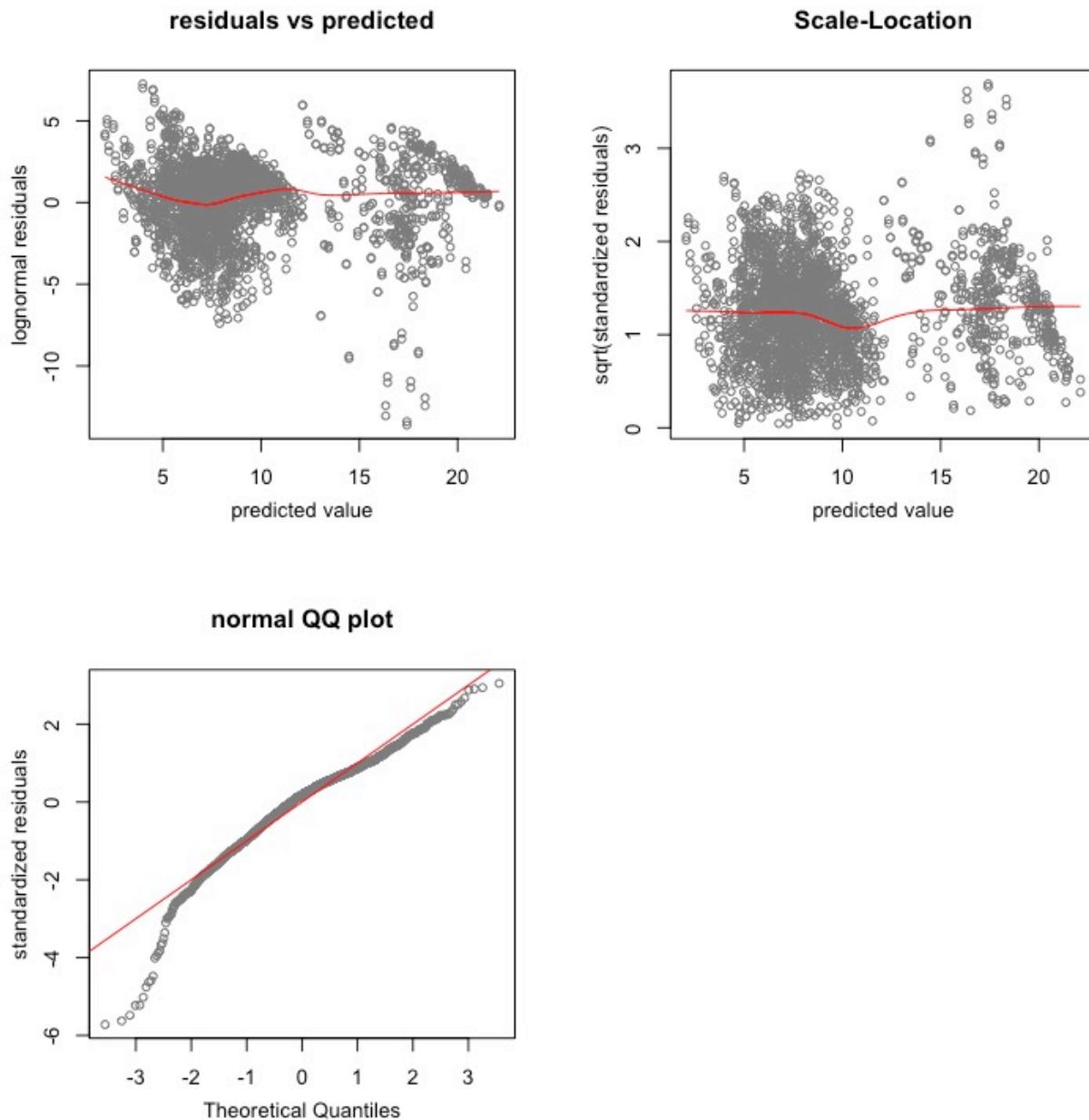
I have to recode the chamber data. It was coded as 15 level factors, but there is really 30 levels.

Model 1: Naive model- testing effect of Site * bait and Site * delta

- Fixed effects- Site * bait and Site * delta
- Global fixed - Rin value and Jdaycont
- Random effects - Chamber

```
1 mm=mcmc.qpcr(  
2   fixed="Site+baittemp.ave+Delta+Site:baittemp.ave+Site:De  
lta",  
3   globalRandom=c("Cham2"),globalFixed=c("RIN_Value","Jdayc  
ont"),controls=c("CT_actin","CT_gapdh","CT_18s"),include  
=0,data=dd,pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =  
TRUE,pl=TRUE)  
4 summary(mm)
```

diagnostics



Model 1 output : Naive model- testing effect of Site * bait and Site * delta

```

1
2 Iterations = 3001:49991
3 Thinning interval = 10
4 Sample size = 4700
5

```

```

6 DIC: 35139.52
7
8 G-structure: ~sample
9
10 post.mean l-95% CI u-95% CI eff.samp
11 sample      3.519     2.798     4.281     4508
12
13 ~Cham2
14
15 post.mean l-95% CI u-95% CI eff.samp
16 Cham2       0.04 6.823e-09   0.1558    1192
17
18 R-structure: ~idh(gene):units
19
20 post.mean l-95% CI u-95% CI eff.samp
21 geneCT_40.units      0.4136    0.3315    0.5069    4105
22 geneCT_70.units      0.4770    0.3988    0.5500    4055
23 geneCT_83.units      2.4430    2.1154    2.8343    3307
24 geneCT_18s.units     4.2449    3.6756    4.8468    4700
25 geneCT_gapdh.units  1.6591    1.4060    1.9019    4377
26 geneCT_actin.units  1.4672    1.2420    1.6840    4531
27
28 Location effects: count ~ 0 + gene + RIN_Value +
Jdaycont + +gene:Site + gene:baittemp.ave + gene:Delta
+ gene:Site:baittemp.ave + gene:Site:Delta
29
30 post.mean   l-95% CI
31          u-95% CI eff.samp   pMCMC
32 geneCT_40           12.842132 11.434150
33          14.428087 4447 < 2e-04 ***
34 geneCT_70           15.135790 13.698704
35          16.594134 4151 < 2e-04 ***
36 geneCT_83           10.650385 8.424201
37          12.787419 4700 < 2e-04 ***

```

34	geneCT_18s		33.007600	30.453198
	35.634718	4700 < 2e-04 ***		
35	geneCT_gapdh		17.997321	16.115197
	19.939950	4700 < 2e-04 ***		
36	geneCT_actin		14.635113	12.791848
	16.399952	4700 < 2e-04 ***		
37	RIN_Value		-0.179934	-0.272161
	-0.084120	4700 0.000426 ***		
38	Jdaycont		0.008706	0.007161
	0.010252	4700 < 2e-04 ***		
39	geneCT_40:SiteHF		-23.440970	-30.821937
	-16.978075	4700 < 2e-04 ***		
40	geneCT_70:SiteHF		-25.312961	-32.432097
	-18.447230	4700 < 2e-04 ***		
41	geneCT_83:SiteHF		-27.395793	-35.308069
	-19.350075	4700 < 2e-04 ***		
42	geneCT_18s:SiteHF		-30.831404	-39.288593
	-22.145284	4700 < 2e-04 ***		
43	geneCT_gapdh:SiteHF		-22.508387	-29.931357
	-14.709748	4700 < 2e-04 ***		
44	geneCT_actin:SiteHF		-17.976511	-25.242206
	-10.147747	4700 < 2e-04 ***		
45	geneCT_40:baittemp.ave		-0.383023	-0.439891
	-0.326289	4437 < 2e-04 ***		
46	geneCT_70:baittemp.ave		-0.350082	-0.404711
	-0.296560	4131 < 2e-04 ***		
47	geneCT_83:baittemp.ave		-0.239388	-0.322902
	-0.157701	4700 < 2e-04 ***		
48	geneCT_18s:baittemp.ave		-0.705876	-0.797390
	-0.601410	4700 < 2e-04 ***		
49	geneCT_gapdh:baittemp.ave		-0.486578	-0.563606
	-0.420150	4470 < 2e-04 ***		
50	geneCT_actin:baittemp.ave		-0.361752	-0.430539
	-0.293431	4700 < 2e-04 ***		

51	geneCT_40:Delta 0.340509 4700 0.077872 .	0.159633	-0.012466
52	geneCT_70:Delta 0.336399 4844 0.060426 .	0.167922	-0.014704
53	geneCT_83:Delta 0.422831 4700 0.025957 *	0.225763	0.036008
54	geneCT_18s:Delta 0.460552 4619 0.024255 *	0.246831	0.044045
55	geneCT_gapdh:Delta 0.353596 4700 0.057447 .	0.177491	-0.012900
56	geneCT_actin:Delta 0.331628 4393 0.109787	0.150764	-0.032281
57	geneCT_40:SiteHF:baittemp.ave 1.195277 4700 < 2e-04 ***	0.936244	0.675711
58	geneCT_70:SiteHF:baittemp.ave 1.290935 4700 < 2e-04 ***	1.016519	0.763996
59	geneCT_83:SiteHF:baittemp.ave 1.373619 4700 < 2e-04 ***	1.083009	0.771374
60	geneCT_18s:SiteHF:baittemp.ave 1.610912 4700 < 2e-04 ***	1.278770	0.963448
61	geneCT_gapdh:SiteHF:baittemp.ave 1.166238 4700 < 2e-04 ***	0.879381	0.588608
62	geneCT_actin:SiteHF:baittemp.ave 0.968543 4700 < 2e-04 ***	0.695124	0.396252
63	geneCT_40:SiteHF:Delta -0.114304 4932 0.006809 **	-0.428692	-0.736058
64	geneCT_70:SiteHF:Delta -0.167260 5201 0.004681 **	-0.456783	-0.794041
65	geneCT_83:SiteHF:Delta -0.203274 4700 0.001702 **	-0.532939	-0.893603
66	geneCT_18s:SiteHF:Delta -0.361775 4736 < 2e-04 ***	-0.720123	-1.100263
67	geneCT_gapdh:SiteHF:Delta 0.026377 4475 0.075319 .	-0.299594	-0.636480

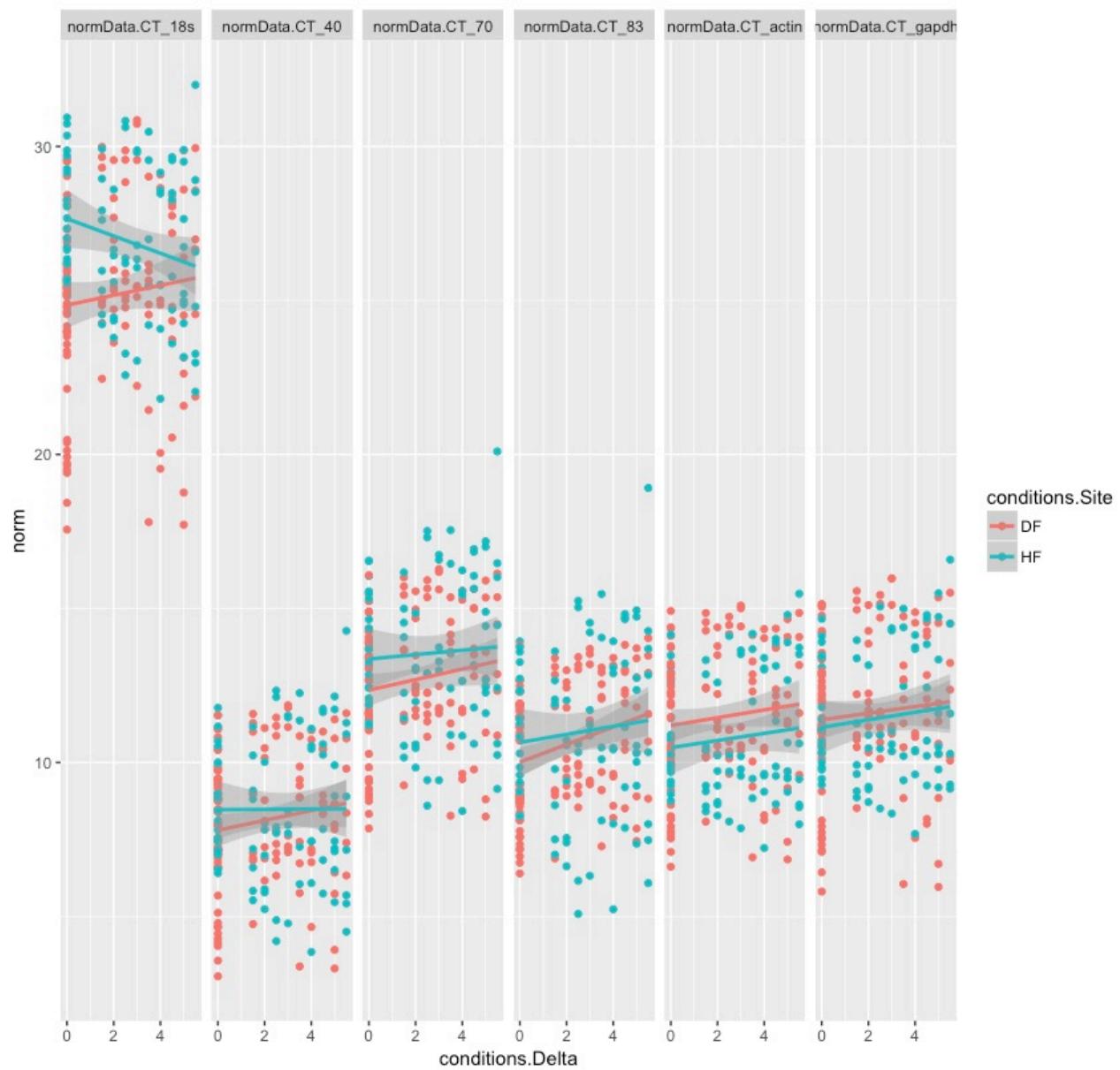
```

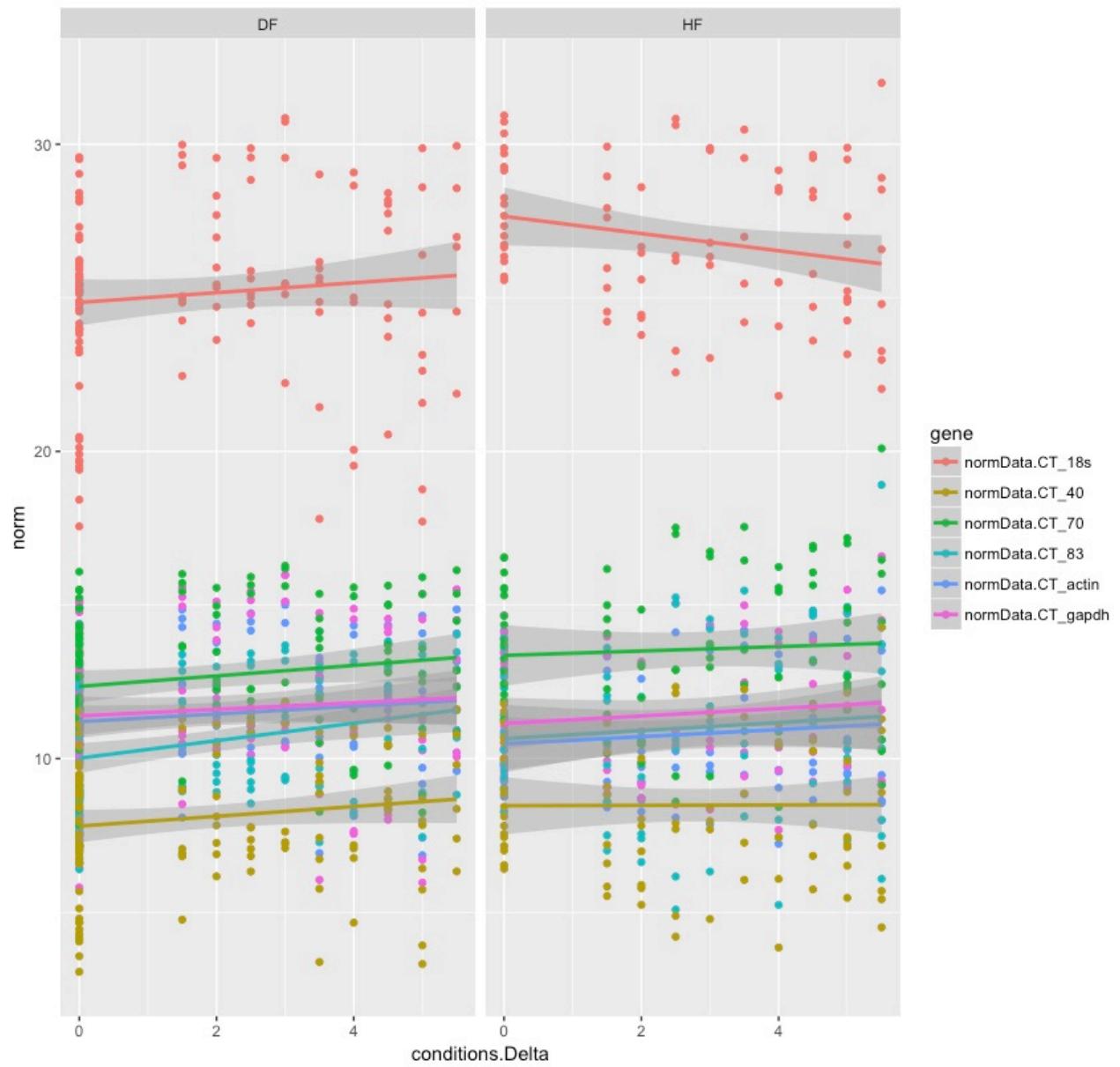
68 geneCT_actin:SiteHF:Delta      -0.271953  -0.622655
     0.037738      5139  0.097021 .
69 ---
70 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
   0.1 ' ' 1

```

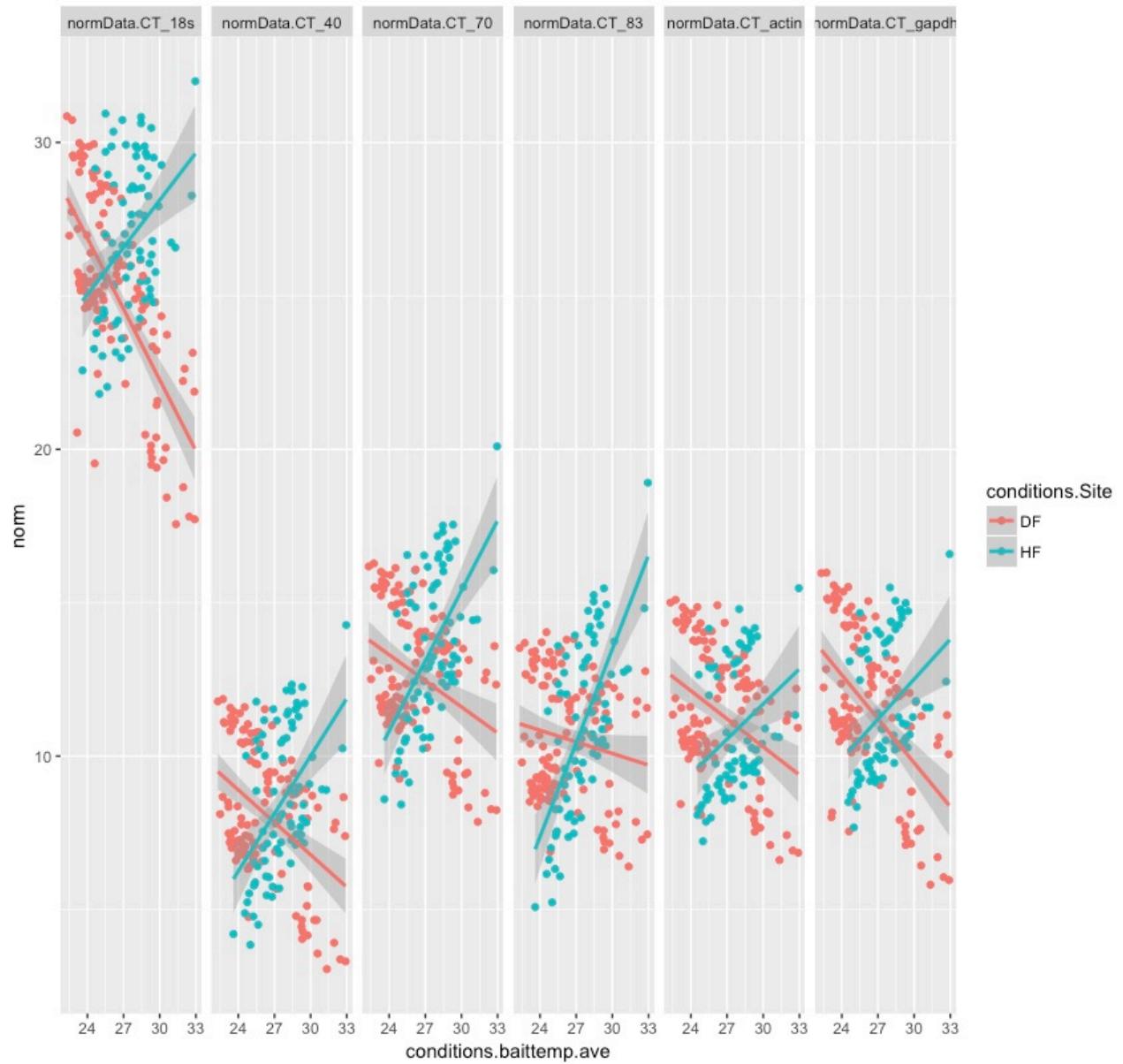
Model 1 plots

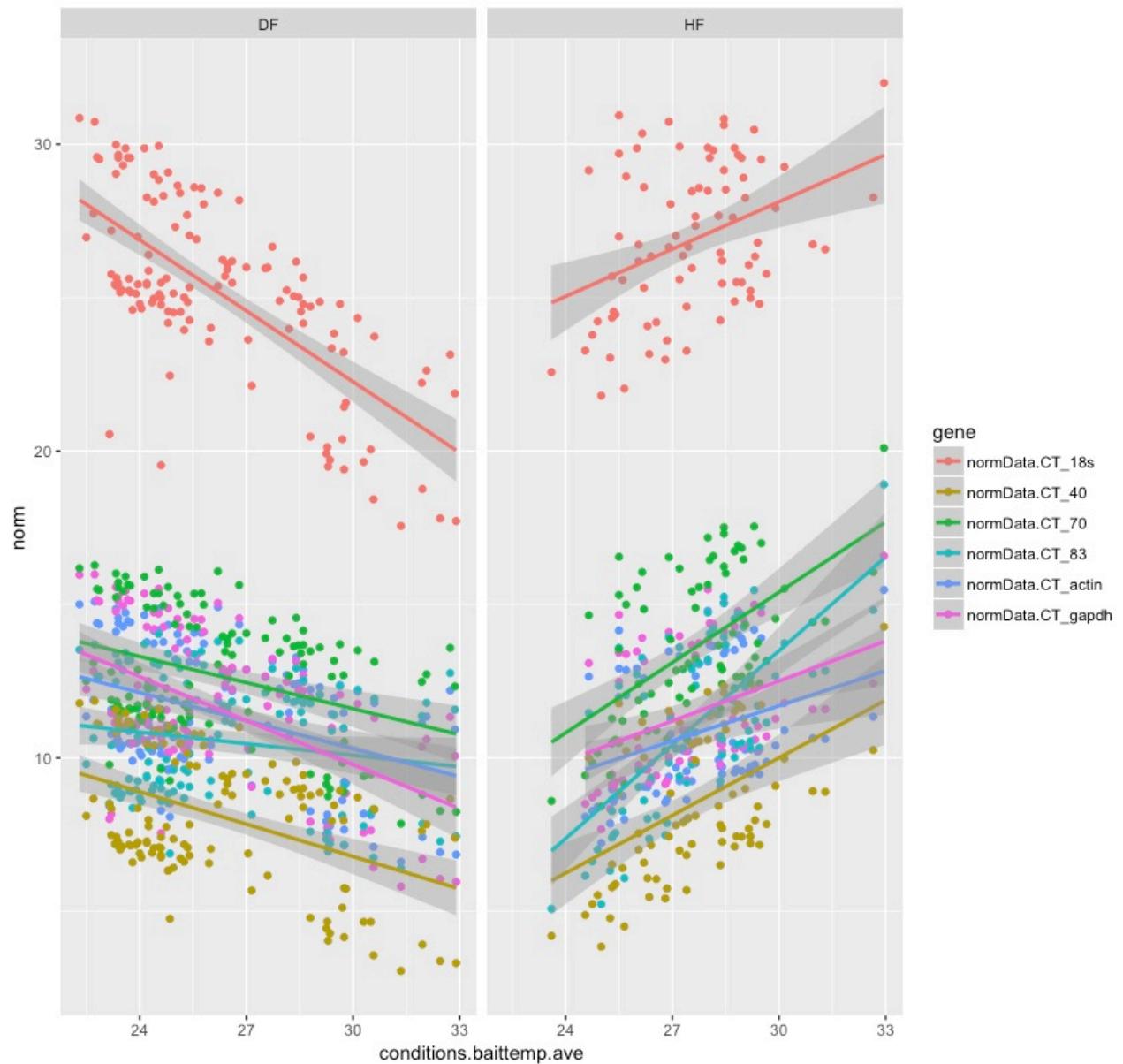
Normalized gene counts vs delta for each gene





Normalized gene counts vs bait temp for each gene





MOdel 2: Soft normalization-testing effect of Site * bait and Site * delta

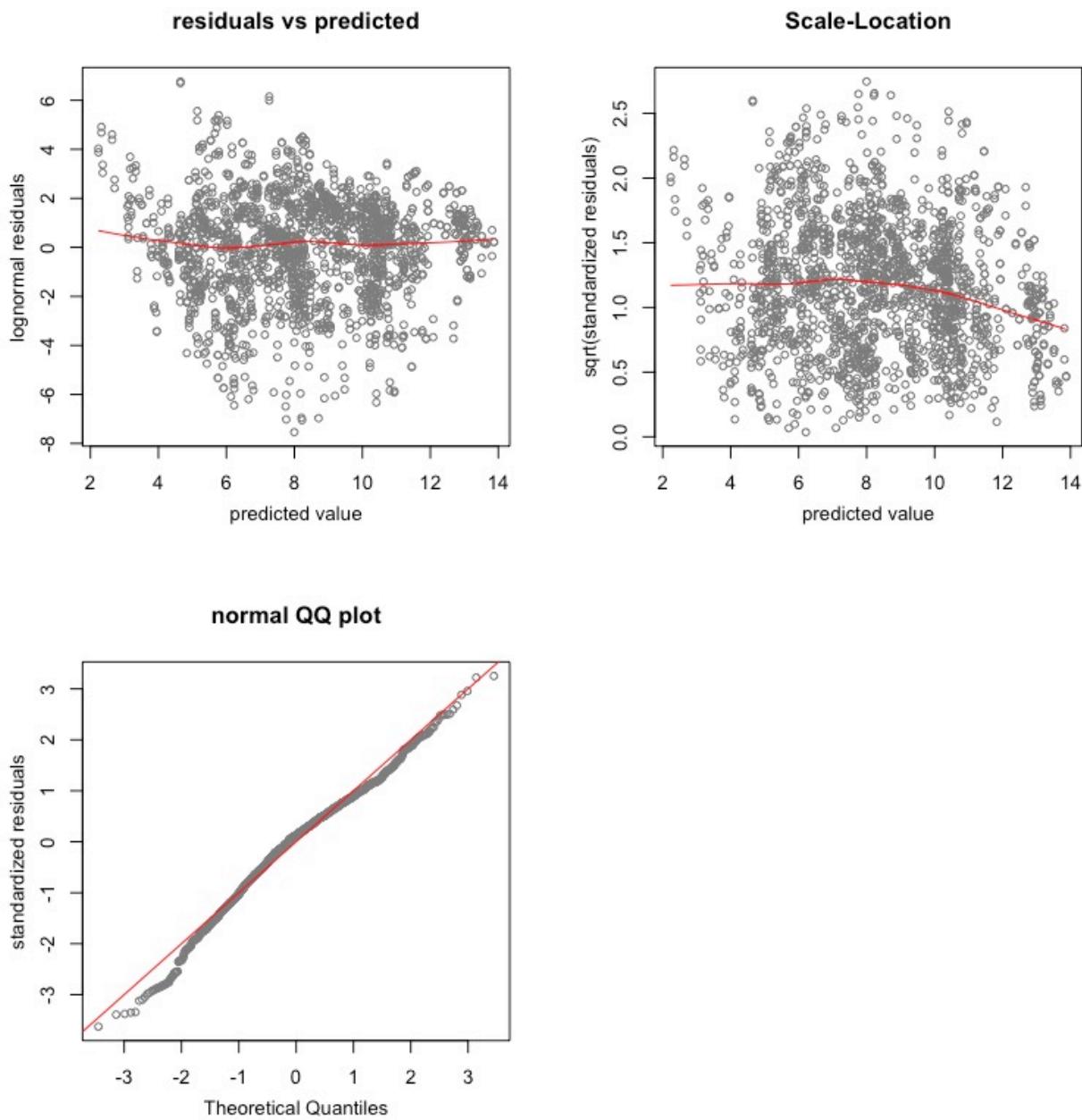
- Fixed effects- Site * bait and Site * delta
- Global fixed - Rin value and Jdaycont
- Random effects - Chamber

```

1 softnorm=mcmc.qpcr(fixed="Site+baittemp.ave+Delta+Site:baittemp.ave+Site:Delta",globalRandom="Cham",data=dd,pr=TRUE,vprior="iw",nitt=50000,geneSpecRes = TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_actin","CT_gapdh","CT_18s"),globalFixed=c("RIN_Value","Jdaycont"))
2 summary(softnorm)

```

diagnostics



MOdel 2 output: Soft normalization-testing effect of Site * bait and Site * delta

```
1
2 Iterations = 3001:49991
3 Thinning interval = 10
4 Sample size = 4700
5
6 DIC: 19495.47
7
8 G-structure: ~sample
9
10      post.mean l-95% CI u-95% CI eff.samp
11 sample     2.97    2.349     3.6     3995
12
13      ~Cham
14
15      post.mean l-95% CI u-95% CI eff.samp
16 Cham    0.04905 5.908e-09   0.1811    1031
17
18 R-structure: ~idh(gene):units
19
20      post.mean l-95% CI u-95% CI eff.samp
21 geneNORM.units    1.2051    1.000    1.4300    4700
22 geneCT_83.units   2.8181    2.410    3.2174    3450
23 geneCT_70.units   0.4852    0.408    0.5646    4700
24 geneCT_40.units   0.2422    0.184    0.3026    3298
25
26 Location effects: count ~ gene + RIN_Value + Jdaycont
+ Site + baittemp.ave + Delta + Site:baittemp.ave +
Site:Delta + gene:Site + gene:baittemp.ave + gene:Delta
+ gene:Site:baittemp.ave + gene:Site:Delta
27
```

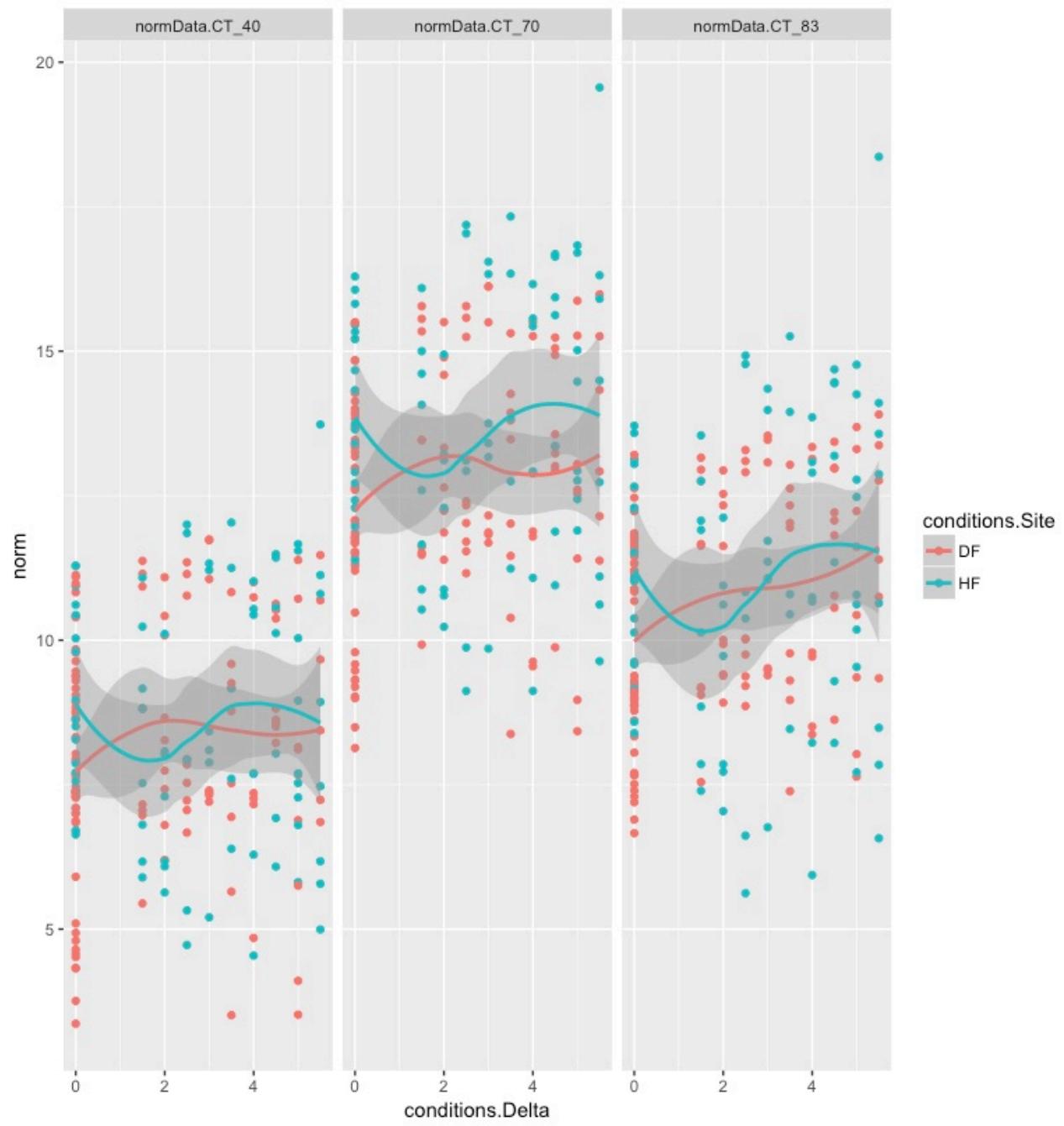
			post.mean	1-95% CI
28		u-95% CI eff.samp	pMCMC	
29	(Intercept)		19.404647	17.634494
	21.200173	4700 < 2e-04	***	
30	geneCT_83		-8.516793	-11.106321
	-6.249250	4700 < 2e-04	***	
31	geneCT_70		-4.025889	-5.613463
	-2.454108	4700 < 2e-04	***	
32	geneCT_40		-6.284329	-7.796068
	-4.786409	4700 < 2e-04	***	
33	RIN_Value		-0.127996	-0.216956
	-0.036955	4700 0.005532	**	
34	Jdaycont		0.007721	0.006342
	0.009334	5202 < 2e-04	***	
35	SiteHF		-14.568427	-22.367099
	-7.392627	4700 < 2e-04	***	
36	baittemp.ave		-0.420104	-0.489386
	-0.352497	4700 < 2e-04	***	
37	Delta		0.099998	-0.091511
	0.267359	4700 0.256596		
38	SiteHF:baittemp.ave		0.584561	0.307684
	0.871625	4700 < 2e-04	***	
39	SiteHF:Delta		-0.145545	-0.469774
	0.165354	4700 0.377872		
40	geneCT_83:SiteHF		-10.844994	-16.903600
	-4.867328	4700 0.000426	***	
41	geneCT_70:SiteHF		-8.739658	-12.731594
	-4.338587	3724 < 2e-04	***	
42	geneCT_40:SiteHF		-6.926685	-10.760303
	-2.717615	4700 0.001702	**	
43	geneCT_83:baittemp.ave		0.177069	0.085766
	0.269983	4700 < 2e-04	***	
44	geneCT_70:baittemp.ave		0.066504	0.005830
	0.125771	4700 0.028936	*	

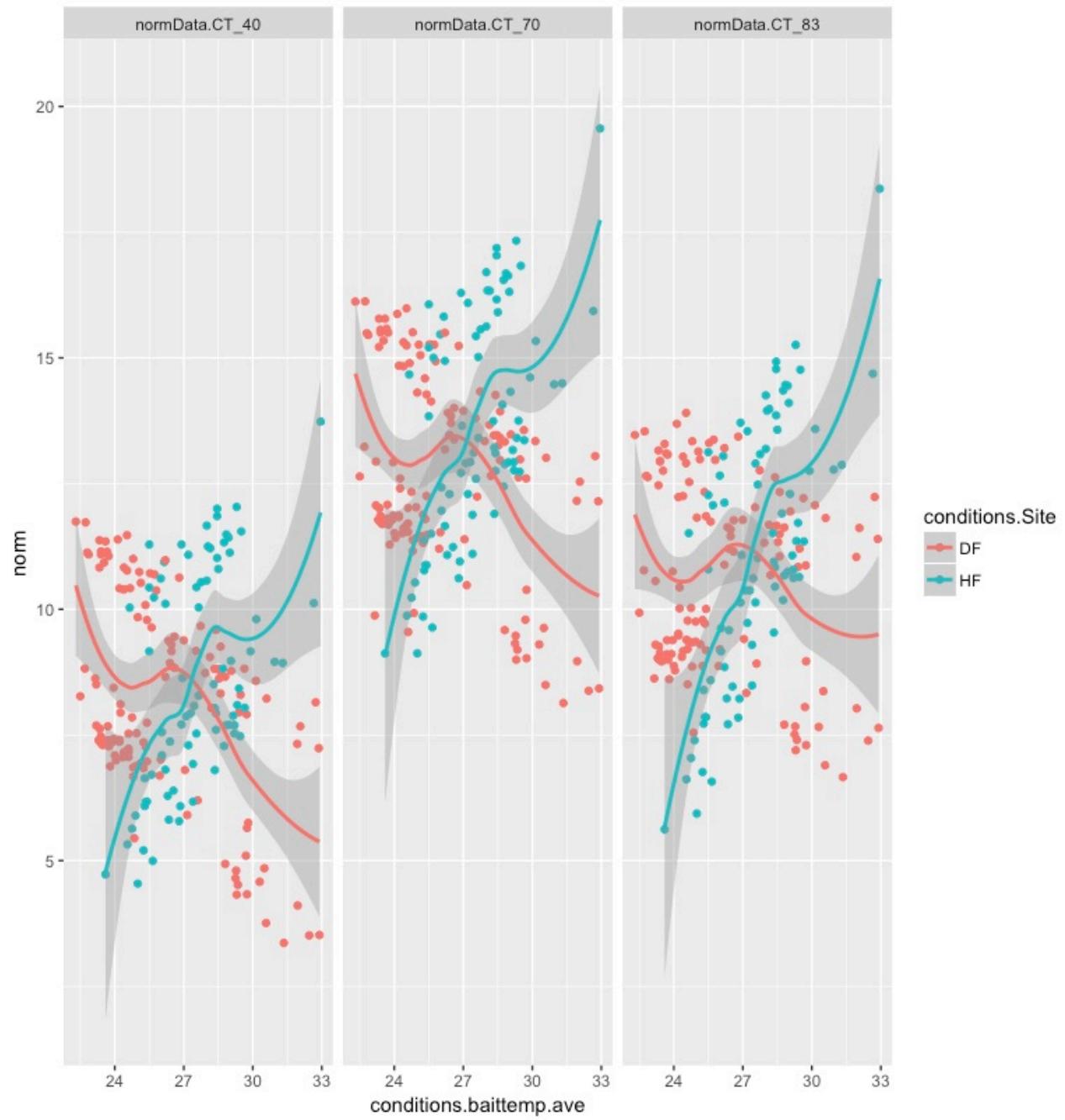
```

45 geneCT_40:baittemp.ave      0.032970 -0.022780
    0.091430     4700 0.260851
46 geneCT_83:Delta            0.116540 -0.005252
    0.241345     4459 0.062128 .
47 geneCT_70:Delta            0.055542 -0.029111
    0.140927     4700 0.200851
48 geneCT_40:Delta            0.047400 -0.031901
    0.128024     4700 0.248511
49 geneCT_83:SiteHF:baittemp.ave 0.420970 0.185736
    0.641951     4700 0.000426 ***
50 geneCT_70:SiteHF:baittemp.ave 0.353379 0.185695
    0.501578     3709 < 2e-04 ***
51 geneCT_40:SiteHF:baittemp.ave 0.274517 0.114185
    0.418024     4700 0.000851 ***
52 geneCT_83:SiteHF:Delta       -0.332825 -0.565332
    -0.094994     4462 0.005106 **
53 geneCT_70:SiteHF:Delta       -0.253365 -0.414750
    -0.080858     4700 0.002979 **
54 geneCT_40:SiteHF:Delta       -0.225062 -0.388829
    -0.072069     4099 0.007660 **
55 ---
56 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'.
    0.1 ' ' 1

```

Model 2 figures soft normalization:





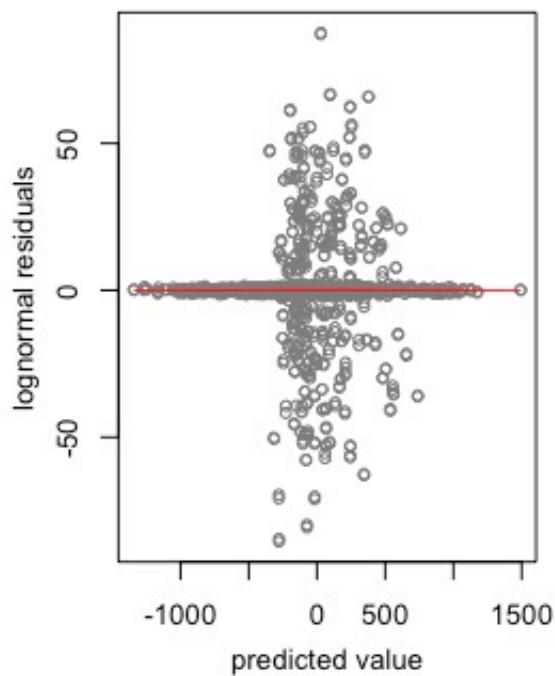
MOdel 3: Classic model-- delta delta ct method

- Fixed effects- Site * bait and Site * delta
- Global fixed - Rin value and Jdaycont
- Random effects - Chamber

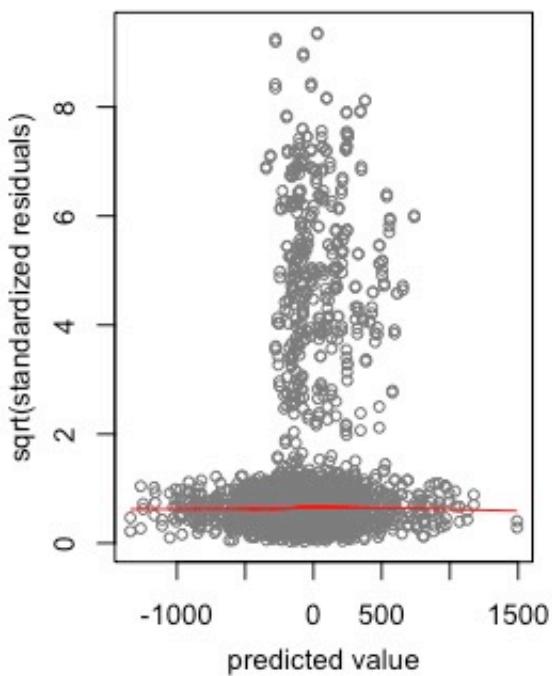
```
1 classicmod<-  
2 mcmc.qpcr.classic(fixed="Site+baittemp.ave+Delta+Site:ba  
ittemp.ave+Site:Delta",  
3 globalRandom="Cham",globalFixed=c("RIN_Value","Jdaycont"  
) ,data=ctmeth,controls=c("CT_18s","CT_actin","CT_gapdh")  
,pr=T,pl=T,nitt=50000,singular.ok=TRUE)  
3 summary(classicmod)  
4 diagnostic.mcmc(model=classicmod,col="grey50",cex=0.8)
```

diagnostics

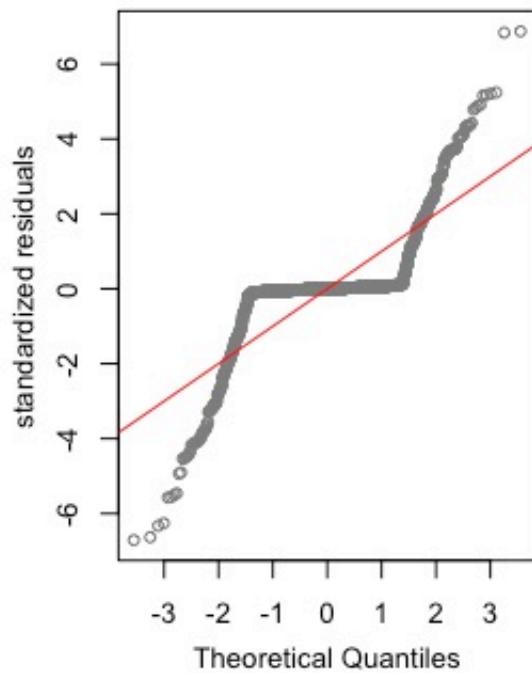
residuals vs predicted



Scale-Location



normal QQ plot



```
1 Iterations = 3001:49991  
2 Thinning interval = 10
```

```

3 Sample size = 4700
4
5 DIC: 0
6
7 G-structure: ~Cham
8
9     post.mean l-95% CI u-95% CI eff.samp
10 Cham    0.02966 2.118e-40    0.0993      65.2
11
12             ~idh(gene):sample
13
14                 post.mean l-95% CI   u-95% CI
15
16                 eff.samp
17 geneCT_40.sample    1.739e-01 2.254e-09 5.603e-01
18 133.59
19 geneCT_70.sample    5.918e-02 1.993e-22 2.380e-01
20 79.10
21 geneCT_83.sample    1.299e+00 7.980e-13 3.519e+00
22 12.37
23 geneCT_gapdh.sample 1.302e+04 8.917e-06 7.305e+04
24 10.73
25 geneCT_actin.sample 6.696e-01 1.086e-07 5.680e+00
26 25.63
27 geneCT_18s.sample    1.497e-01 7.383e-23 1.012e+00
28 39.48
29
30 R-structure: ~units
31
32     post.mean l-95% CI u-95% CI eff.samp
33 units      1705    0.0905    12211      15.75
34
35 Location effects: count ~ 0 + gene + RIN_Value +
36 Jdaycont + +gene:Site + gene:baittemp.ave + gene:Delta
37 + gene:Site:baittemp.ave + gene:Site:Delta
38

```

			post.mean	l-95%	CI
29		u-95% CI eff.samp pMCMC			
30	geneCT_40		1.053e+03	-4.823e+02	
	7.010e+03	1.389 0.9013			
31	geneCT_70		1.047e+03	-3.535e+01	
	4.417e+03	3.277 0.1221			
32	geneCT_83		1.008e+03	-1.311e+03	
	6.766e+03	2.829 0.6357			
33	geneCT_gapdh		-1.128e+03	-5.719e+03	
	2.849e+02	2.809 0.5119			
34	geneCT_actin		5.292e+02	-6.079e+02	
	6.708e+03	6.344 0.7885			
35	geneCT_18s		8.093e+02	-1.768e+02	
	5.053e+03	5.752 0.5260			
36	RIN_Value		-3.818e+01	-1.425e+02	
	5.092e+00	2.314 0.6038			
37	Jdaycont		2.761e-01	-7.798e-02	
	1.353e+00	4.687 0.7302			
38	geneCT_40:SiteHF		8.392e+02	-6.353e+03	
	5.895e+03	8.269 0.2515			
39	geneCT_70:SiteHF		-7.961e+02	-5.570e+03	
	7.284e+02	9.757 0.8664			
40	geneCT_83:SiteHF		-7.636e+03	-3.515e+04	
	1.954e+03	1.364 0.6060			
41	geneCT_gapdh:SiteHF		1.018e+03	-2.240e+03	
	7.038e+03	6.232 0.8770			
42	geneCT_actin:SiteHF		4.782e+03	-1.220e+02	
	1.899e+04	2.330 0.0532	.		
43	geneCT_18s:SiteHF		-2.538e+03	-1.348e+04	
	3.907e+02	3.557 0.6443			
44	geneCT_40:baittemp.ave		-2.699e+01	-2.396e+02	
	1.678e+01	1.746 0.8979			
45	geneCT_70:baittemp.ave		-3.338e+01	-1.550e+02	
	1.780e+00	4.442 0.1553			

46	geneCT_83:baittemp.ave		-4.613e+01	-2.689e+02
	5.122e+01	2.536	0.6370	
47	geneCT_gapdh:baittemp.ave		4.419e+01	-1.346e+01
	2.116e+02	2.447	0.4494	
48	geneCT_actin:baittemp.ave		-3.250e+01	-2.685e+02
	1.354e+01	5.426	0.6026	
49	geneCT_18s:baittemp.ave		-3.435e+01	-2.003e+02
	4.960e+00	5.850	0.3877	
50	geneCT_40:Delta		-1.091e+01	-1.783e+02
	6.033e+01	8.267	0.8221	
51	geneCT_70:Delta		-7.888e+01	-3.479e+02
	4.164e+01	1.740	0.2813	
52	geneCT_83:Delta		7.219e+01	-3.586e+01
	4.061e+02	2.130	0.6072	
53	geneCT_gapdh:Delta		9.920e+00	-1.528e+02
	1.424e+02	10.051	0.7213	
54	geneCT_actin:Delta		7.125e+01	-4.334e+00
	2.508e+02	3.086	0.5511	
55	geneCT_18s:Delta		3.309e+00	-9.722e+01
	1.049e+02	8.610	0.9102	
56	geneCT_40:SiteHF:baittemp.ave		-5.552e+01	-3.086e+02
	1.975e+02	6.021	0.2191	
57	geneCT_70:SiteHF:baittemp.ave		2.943e+01	-2.739e+01
	2.042e+02	10.090	0.8438	
58	geneCT_83:SiteHF:baittemp.ave		3.213e+02	-7.232e+01
	1.482e+03	1.311	0.6077	
59	geneCT_gapdh:SiteHF:baittemp.ave		-4.338e+01	-2.892e+02
	8.391e+01	5.204	0.7523	
60	geneCT_actin:SiteHF:baittemp.ave		-1.486e+02	-5.989e+02
	4.479e+00	2.944	0.0489 *	
61	geneCT_18s:SiteHF:baittemp.ave		1.049e+02	-1.349e+01
	5.464e+02	3.254	0.6511	
62	geneCT_40:SiteHF:Delta		-4.646e+01	-3.431e+02
	5.487e+01	7.411	0.5566	

63	geneCT_70:SiteHF:Delta	9.252e+01	-5.145e+01
	4.685e+02 2.925 0.3766		
64	geneCT_83:SiteHF:Delta	-3.837e+02	-1.855e+03
	6.381e+01 1.315 0.7043		
65	geneCT_gapdh:SiteHF:Delta	1.357e+02	-2.403e+01
	7.192e+02 3.095 0.7166		
66	geneCT_actin:SiteHF:Delta	2.340e+01	-1.008e+02
	2.067e+02 29.595 0.5136		
67	geneCT_18s:SiteHF:Delta	-4.692e+00	-1.318e+02
	3.094e+02 8.322 0.6728		

HF specific model

1

DF specific model

1

Page 77: 2017-05-05. stressed in nature project: Analyses with recoded data and splitting by site

Recoded data such that empty ones are -1 to indicate no amplification.

Approach: For each site fit a 2 model to test for effect of bait temp and Delta, while controlling for RNA quality and Julia Day.

Models:

1. Naive- include all genes and test their effects
 - Identify the ones that change and then do a soft normalization for the ones without a significant effect on log(counts)
2. Soft normalization - include genes that are "stable"

Harvard forest naive model

```

1 HFnonnorm=mcmc.qpcr(fixed="baittemp.ave+Delta",random="Cham",
2 data=ddHF,pr=TRUE,
3 vprior="iw",nitt=50000,geneSpecRes
4 =FALSE,pl=TRUE,include=0,controls=c("CT_18s","CT_actin",
5 "CT_gapdh"),
6 globalFixed=c("RIN_Value","Jdaycont"),m.fix=1.5)
7 summary(HFsoftnorm)

```

Harvard forest naive model output and figures

```

1
2 Iterations = 3001:49991
3 Thinning interval = 10
4 Sample size = 4700
5
6 DIC: 13813.2
7
8 G-structure: ~sample
9
10 post.mean l-95% CI u-95% CI eff.samp
11 sample     3.152      2.12     4.266     3129
12
13 ~idh(gene):Cham
14

```

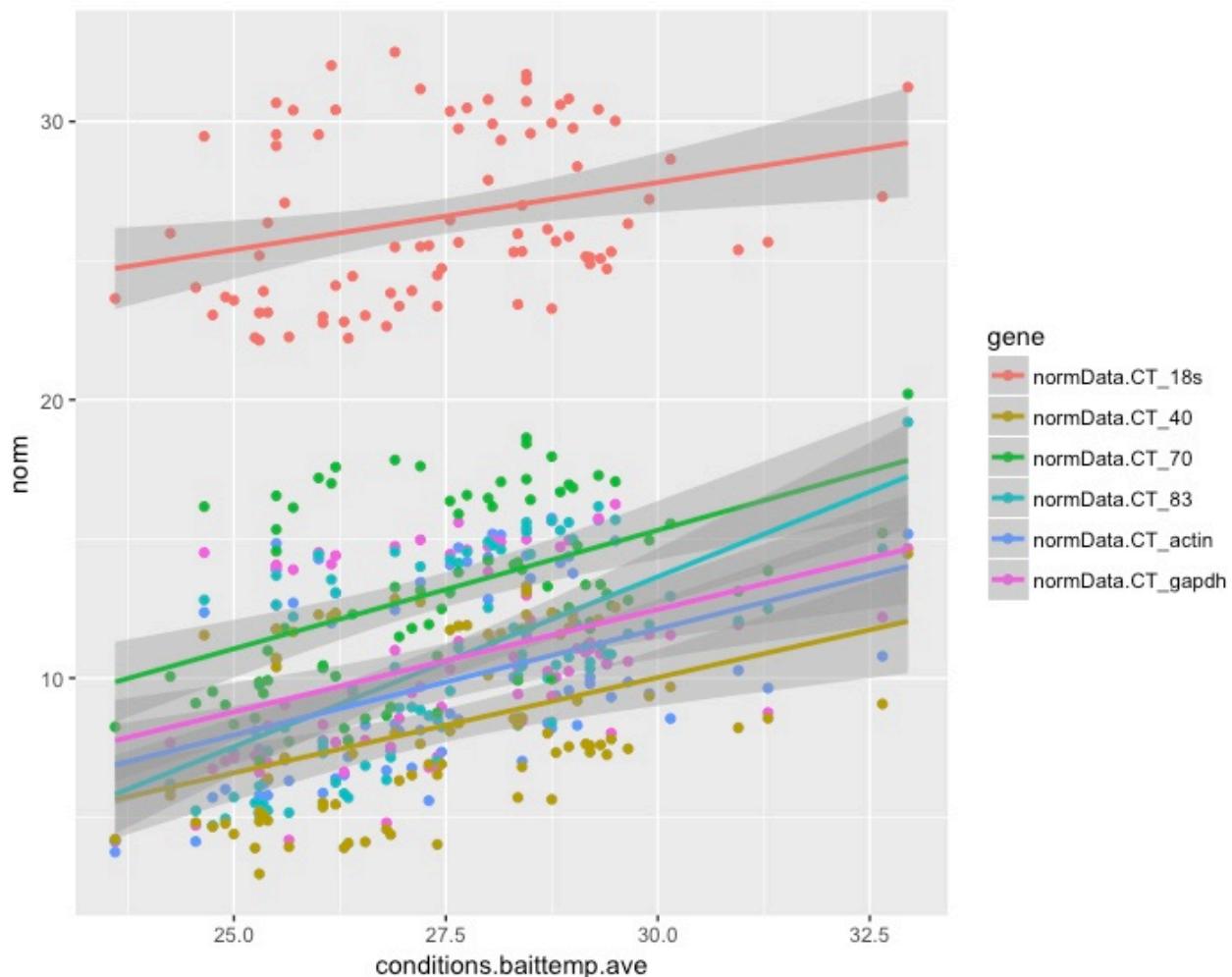
15		post.mean	l-95% CI	u-95% CI	eff.samp
16	geneCT_40.Cham	0.6428	0.2149	1.219	4700
17	geneCT_70.Cham	0.6768	0.2146	1.301	5346
18	geneCT_83.Cham	0.6275	0.2179	1.190	4980
19	geneCT_gapdh.Cham	0.9583	0.2747	1.876	4913
20	geneCT_actin.Cham	1.0032	0.2746	1.992	4239
21	geneCT_18s.Cham	0.8761	0.2524	1.717	4700
22					
23	R-structure: ~units				
24					
25		post.mean	l-95% CI	u-95% CI	eff.samp
26	units	3.635	3.274	3.997	2666
27					
28	Location effects: count ~ 0 + gene + RIN_Value + Jdaycont + +gene:baittemp.ave + gene:Delta				
29					
30		post.mean	l-95% CI	u-95%	
	CI eff.samp	pMCMC			
31	geneCT_40		-7.908805	-15.787958	
	0.375144	4380	0.056596	.	
32	geneCT_70		-9.255603	-17.346088	
	-0.988940	3824	0.025957	*	
33	geneCT_83		-15.547922	-23.769932	
	-7.254690	3660	< 2e-04	***	
34	geneCT_gapdh		-6.248468	-14.853084	
	1.718760	4118	0.140851		
35	geneCT_actin		-3.414440	-11.069957	
	5.373458	4181	0.417447		
36	geneCT_18s		8.630935	1.012052	
	17.079760	4700	0.035745	*	
37	RIN_Value		0.181456	-0.071108	
	0.419909	4700	0.144681		
38	Jdaycont		0.008945	0.006094	
	0.011705	4700	< 2e-04	***	

```

39 geneCT_40:baittemp.ave      0.377825   0.067289
     0.698200    4383 0.014468 *
40 geneCT_70:baittemp.ave      0.563959   0.250113
     0.891069    3642 0.000851 ***
41 geneCT_83:baittemp.ave      0.710939   0.396867
     1.032394    3646 < 2e-04 ***
42 geneCT_gapdh:baittemp.ave  0.373791   0.065401
     0.698623    4181 0.019574 *
43 geneCT_actin:baittemp.ave  0.220159   -0.081049
     0.546881    3740 0.174043
44 geneCT_18s:baittemp.ave    0.230615   -0.080949
     0.542483    4322 0.159574
45 geneCT_40:Delta           -0.264703   -0.639966
     0.094455    4700 0.159574
46 geneCT_70:Delta           -0.402686   -0.778682
     -0.026056    4116 0.036170 *
47 geneCT_83:Delta           -0.228574   -0.584389
     0.162200    4700 0.233191
48 geneCT_gapdh:Delta        -0.238826   -0.652636
     0.161600    4700 0.249787
49 geneCT_actin:Delta        0.088106   -0.327377
     0.508461    4700 0.673191
50 geneCT_18s:Delta          -0.204585   -0.582160
     0.196536    4488 0.321702
51 ---
52 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Relationship between normalized log2 abundance vs bait temp.



Harvard forest soft normalization

based on naive model, I should normalize based on actin and 18s

```

1 HFsoftnorm2=mcmc.qpcr(fixed="baittemp.ave+Delta",random=
  "Cham",
2 data=ddHF,pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =
  TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_18s","CT_acti
  n","CT_gapdh"),
3 globalFixed=c("RIN_Value","Jdaycont"),m.fix=1.5)

```

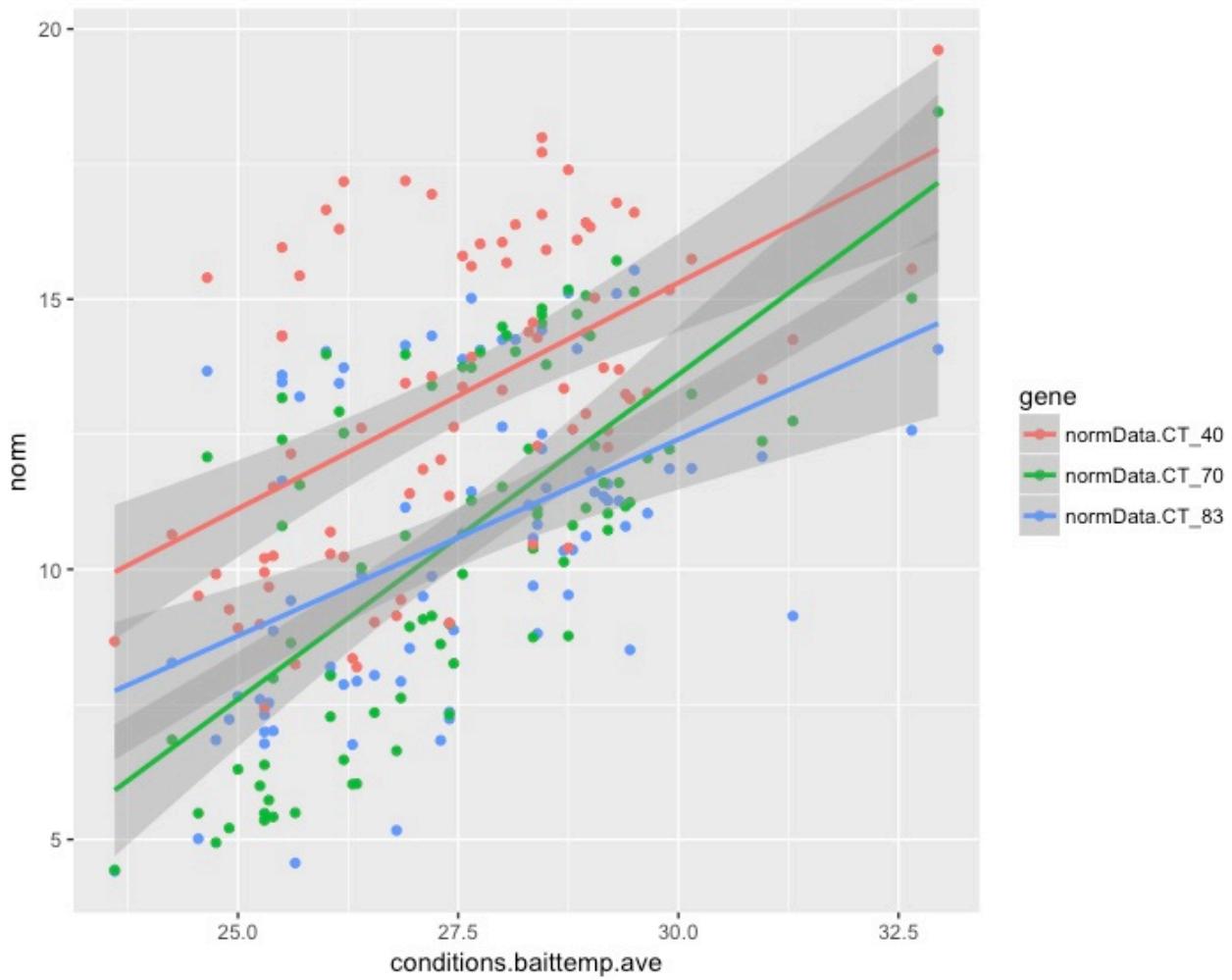
Harvard forest soft norm based on actin and 18s output and figures

```

1 Iterations = 3001:49991
2 Thinning interval = 10
3 Sample size = 4700
4
5 DIC: 9572.318
6
7 G-structure: ~sample
8
9          post.mean l-95% CI u-95% CI eff.samp
10 sample      2.321     1.606     3.237     1579
11
12          ~idh(gene):Cham
13
14          post.mean l-95% CI u-95% CI eff.samp
15 geneNORM.Cham      0.6719    0.2049    1.336    5204
16 geneCT_gapdh.Cham  0.9482    0.2624    1.921    4484
17 geneCT_83.Cham     0.5887    0.1860    1.131    3916
18 geneCT_70.Cham     0.5626    0.1961    1.096    4700
19 geneCT_40.Cham     0.5433    0.1796    1.033    3812
20
21 R-structure: ~idh(gene):units
22
23          post.mean l-95% CI u-95% CI eff.samp
24 geneNORM.units      2.9320    2.1486    3.7880   4700.0
25 geneCT_gapdh.units  5.1292    3.8858    6.4769   915.2
26 geneCT_83.units     3.5455    2.6909    4.4898   2183.6
27 geneCT_70.units     0.9661    0.7120    1.2359   1840.6
28 geneCT_40.units     0.3372    0.1898    0.4848   2068.1
29
30 Location effects: count ~ gene + RIN_Value + Jdaycont
31 + baittemp.ave + Delta + gene:baittemp.ave + gene:Delta
32          post.mean    l-95% CI    u-95%
33          CI eff.samp    pMCMC

```

33	(Intercept)		6.743917	-0.948957
	14.211602	4700	0.07830	.
34	geneCT_gapdh		-12.334777	-19.925190
	-4.609552	3219	0.00128	**
35	geneCT_83		-21.397507	-28.127556
	-14.221628	4304	< 2e-04	***
36	geneCT_70		-15.016724	-20.784617
	-9.630071	4700	< 2e-04	***
37	geneCT_40		-14.328587	-19.830423
	-9.416300	3147	< 2e-04	***
38	RIN_Value		0.236854	0.023824
	0.431321	4700	0.02128	*
39	Jdaycont		0.006872	0.004391
	0.009112	4700	< 2e-04	***
40	baittemp.ave		0.076859	-0.221565
	0.365524	4700	0.60596	
41	Delta		-0.026863	-0.417063
	0.321977	4700	0.90000	
42	geneCT_gapdh:baittemp.ave		0.285499	-0.008747
	0.567074	3222	0.04936	*
43	geneCT_83:baittemp.ave		0.614563	0.360369
	0.879077	4236	< 2e-04	***
44	geneCT_70:baittemp.ave		0.461884	0.265655
	0.675489	3532	< 2e-04	***
45	geneCT_40:baittemp.ave		0.301720	0.119143
	0.503916	3176	0.00128	**
46	geneCT_gapdh:Delta		-0.203782	-0.678987
	0.270997	4700	0.38809	
47	geneCT_83:Delta		-0.185188	-0.606977
	0.228282	4177	0.38128	
48	geneCT_70:Delta		-0.331567	-0.752459
	0.025757	4069	0.09021	.
49	geneCT_40:Delta		-0.218815	-0.588207
	0.167011	4486	0.26340	



Harvard forest soft normalization based on 3 ref sets: 18s, actin, gapdh

```

1 HFsoftnorm2=mcmc.qpcr(fixed="baittemp.ave+Delta",random=
  "Cham",data=ddHF,
2 pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =
  TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_18s","CT_acti
  n","CT_gapdh"),
3 globalFixed=c("RIN_Value","Jdaycont"),m.fix=1.5)#+Delta
4 summary(HFsoftnorm2)

```

Output hf soft norm 18s actin gapdh

```

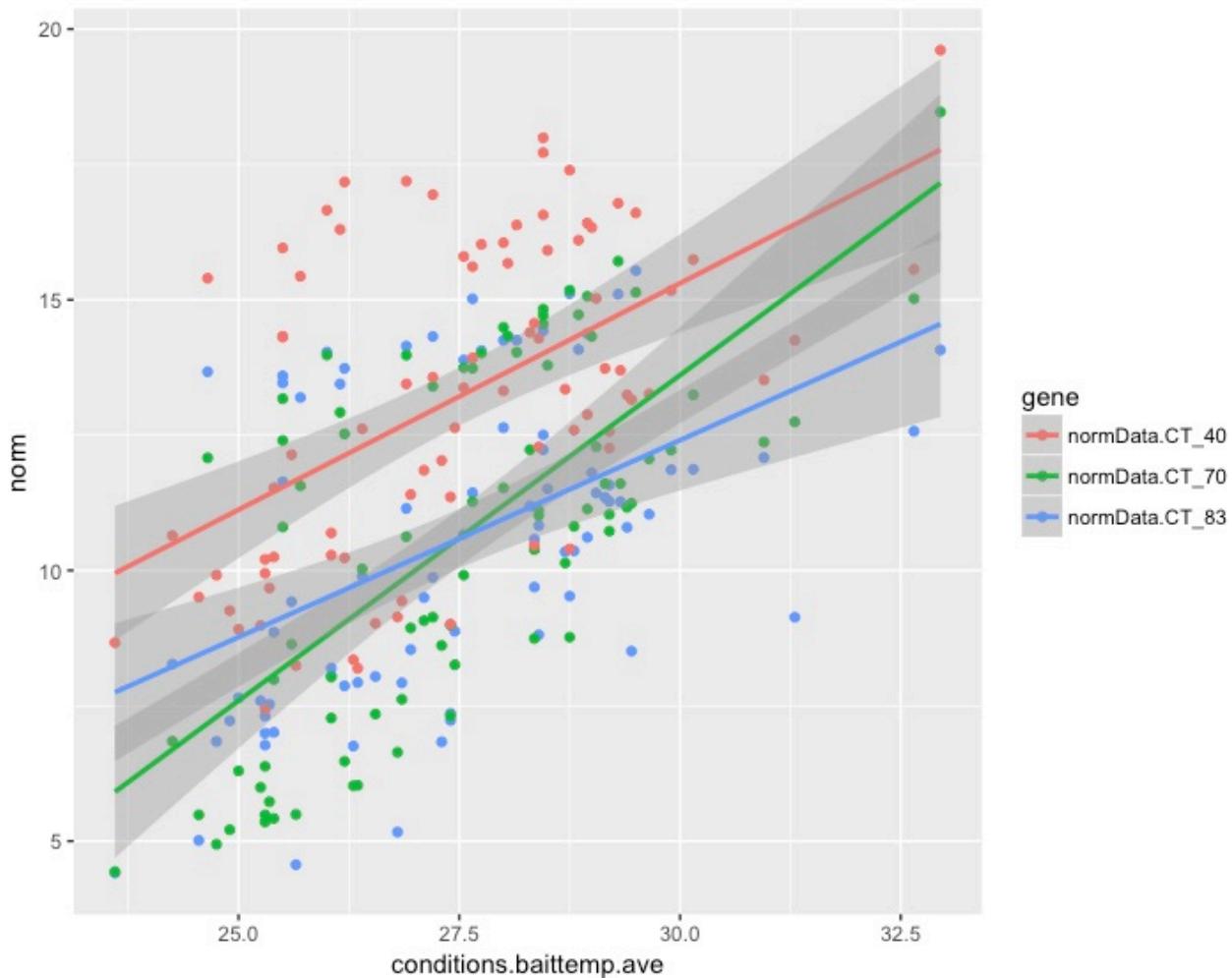
1 Iterations = 3001:49991
2 Thinning interval = 10
3 Sample size = 4700
4
5 DIC: 7485.264
6
7 G-structure: ~sample
8
9          post.mean l-95% CI u-95% CI eff.samp
10 sample      2.174     1.458     2.981     820.6
11
12          ~idh(gene):Cham
13
14          post.mean l-95% CI u-95% CI eff.samp
15 geneNORM.Cham    0.7162    0.1672    1.4609    4700
16 geneCT_83.Cham   0.5460    0.1449    1.0985    4700
17 geneCT_70.Cham   0.4833    0.1494    0.9621    4278
18 geneCT_40.Cham   0.4580    0.1279    0.8886    4700
19
20 R-structure: ~idh(gene):units
21
22          post.mean l-95% CI u-95% CI eff.samp
23 geneNORM.units   2.1269    1.4886    2.7597    3656
24 geneCT_83.units  3.6440    2.7244    4.5392    2250
25 geneCT_70.units  1.0194    0.7435    1.2882    2187
26 geneCT_40.units  0.2667    0.1538    0.3879    2073
27
28 Location effects: count ~ gene + RIN_Value + Jdaycont
29 + baittemp.ave + Delta + gene:baittemp.ave + gene:Delta
30
31          post.mean    l-95% CI    u-95% CI
eff.samp      pMCMC
(Intercept)        4.218044   -2.788091   11.013168
4386 0.235745

```

```

32 geneCT_83           -18.818872 -25.510002 -12.596046
      3065 < 2e-04 ***
33 geneCT_70           -12.321787 -17.449208 -7.164157
      4153 < 2e-04 ***
34 geneCT_40           -11.642306 -16.547428 -7.329180
      3682 < 2e-04 ***
35 RIN_Value           0.228410  0.032208  0.424987
      4464 0.019574 *
36 Jdaycont            0.006782  0.004683  0.009070
      4700 < 2e-04 ***
37 baittemp.ave         0.112626 -0.162648  0.375277
      4382 0.410213
38 Delta                -0.052040 -0.414500  0.287017
      4700 0.762553
39 geneCT_83:baittemp.ave 0.579150  0.340771  0.814229
      3073 < 2e-04 ***
40 geneCT_70:baittemp.ave 0.422447  0.236734  0.609659
      3977 < 2e-04 ***
41 geneCT_40:baittemp.ave 0.263062  0.093489  0.429304
      3806 0.000851 ***
42 geneCT_83:Delta       -0.155925 -0.577289  0.249848
      4700 0.455319
43 geneCT_70:Delta       -0.297292 -0.677188  0.070924
      4700 0.115745
44 geneCT_40:Delta       -0.188711 -0.552140  0.177476
      4700 0.294468
45 ---
46 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
      0.1 ' ' 1

```



DF naive model

```

1 DFnaive=mcmc.qpcr(fixed="baittemp.ave+Delta",random=c("C
ham"),
2 data=ddDF,pr=TRUE,vprior="iw",nitt=50000,pl=TRUE,
3 include=0,controls=c("CT_18s","CT_actin","CT_gapdh"),
4 globalFixed=c("RIN_Value","Jdaycont"),m.fix=2,geneSpecRe
s=FALSE)

```

DF naive model output

```

1 Iterations = 3001:49991
2 Thinning interval = 10
3 Sample size = 4700

```

```

4
5 DIC: 23250.17
6
7 G-structure: ~sample
8
9         post.mean l-95% CI u-95% CI eff.samp
10 sample      6.548     4.926     8.282      4700
11
12             ~idh(gene):Cham
13
14         post.mean l-95% CI u-95% CI eff.samp
15 geneCT_40.Cham      0.5158    0.2058    0.9504      4700
16 geneCT_70.Cham      0.4734    0.1829    0.8395      4700
17 geneCT_83.Cham      0.4935    0.1951    0.8691      4481
18 geneCT_gapdh.Cham   0.5111    0.1877    0.9270      4463
19 geneCT_actin.Cham   0.4577    0.1761    0.8226      4700
20 geneCT_18s.Cham     0.6851    0.2522    1.2444      4932
21
22 R-structure: ~units
23
24         post.mean l-95% CI u-95% CI eff.samp
25 units       2.511     2.327      2.7        3020
26
27 Location effects: count ~ 0 + gene + RIN_Value +
Jdaycont + +gene:baittemp.ave + gene:Delta
28
29         post.mean l-95% CI u-95% CI
eff.samp pMCMC
30 geneCT_40           12.172657  9.546441 14.869301
3528 <2e-04 ***
31 geneCT_70           14.130384 11.554427 16.748759
4700 <2e-04 ***
32 geneCT_83           9.021455  6.457620 11.670320
4137 <2e-04 ***

```

33	geneCT_gapdh 4700 <2e-04 ***	15.475238	12.901377	18.073151
34	geneCT_actin 4700 <2e-04 ***	12.853326	10.301988	15.450867
35	geneCT_18s 4700 <2e-04 ***	30.161942	27.519678	32.654623
36	RIN_Value 4435 0.0234 *	-0.169156	-0.324264	-0.020386
37	Jdaycont 4700 <2e-04 ***	0.009466	0.007021	0.012421
38	geneCT_40:baittemp.ave 3242 <2e-04 ***	-0.382081	-0.483717	-0.285517
39	geneCT_70:baittemp.ave 4700 <2e-04 ***	-0.327654	-0.425382	-0.231107
40	geneCT_83:baittemp.ave 3904 <2e-04 ***	-0.195752	-0.296805	-0.100031
41	geneCT_gapdh:baittemp.ave 4700 <2e-04 ***	-0.415716	-0.514251	-0.321747
42	geneCT_actin:baittemp.ave 4478 <2e-04 ***	-0.305161	-0.403253	-0.211346
43	geneCT_18s:baittemp.ave 4700 <2e-04 ***	-0.620586	-0.722245	-0.529481
44	geneCT_40:Delta 4700 0.2749	0.163161	-0.131068	0.462919
45	geneCT_70:Delta 4700 0.3421	0.145821	-0.143871	0.433128
46	geneCT_83:Delta 4700 0.1915	0.196108	-0.091740	0.497139
47	geneCT_gapdh:Delta 4632 0.1238	0.232599	-0.079664	0.522301
48	geneCT_actin:Delta 5679 0.3226	0.146889	-0.151989	0.423387
49	geneCT_18s:Delta 4700 0.1123	0.257939	-0.047918	0.580536
50	---			

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

All genes are going down...I'll try fitting a soft norm with 3 ref genes

Duke forest soft normalization with 3 ref genes; 18s, actin, gapdh

So here i'd test, whether hsp's are changing at a faster/slower rate than the ref set along bait temp.

```
1 DFsoftnorm2=mcmc.qpcr(fixed="baittemp.ave+Delta",random=  
"Cham",data=ddDF,  
2 pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =  
TRUE,pl=TRUE,normalize=TRUE,controls=c("CT_18s","CT_acti  
n","CT_gapdh"),  
3 globalFixed=c("RIN_Value","Jdaycont"),m.fix=1.5)
```

Duke forest soft norm with 3 refs output and figs

```
1 Iterations = 3001:49991  
2 Thinning interval = 10  
3 Sample size = 4700  
4  
5 DIC: 12844.79  
6  
7 G-structure: ~sample  
8  
9 post.mean l-95% CI u-95% CI eff.samp  
10 sample 5.534 4.151 7.043 2809  
11  
12 ~idh(gene):Cham
```

```

13
14          post.mean l-95% CI u-95% CI eff.samp
15 geneNORM.Cham      0.4929   0.1500   0.9190    4700
16 geneCT_83.Cham    0.4214   0.1431   0.8095    4700
17 geneCT_70.Cham    0.3526   0.1263   0.6596    4700
18 geneCT_40.Cham    0.4154   0.1451   0.7777    4700
19
20 R-structure: ~idh(gene):units
21
22          post.mean l-95% CI u-95% CI eff.samp
23 geneNORM.units     1.3954   1.1122   1.7147    2642
24 geneCT_83.units    3.2881   2.6955   3.9084    1686
25 geneCT_70.units    0.7968   0.6352   0.9718    1786
26 geneCT_40.units    0.9899   0.7637   1.2331    1357
27
28 Location effects: count ~ gene + RIN_Value + Jdaycont
+ baittemp.ave + Delta + gene:baittemp.ave + gene:Delta
29
30          post.mean l-95% CI u-95% CI
31 eff.samp pMCMC
32 (Intercept)        22.140225 19.657121 24.435802
33 3341 < 2e-04 ***
32 geneCT_83           -9.647201 -12.324195 -6.867827
33 4700 < 2e-04 ***
33 geneCT_70           -4.653489 -6.550787 -2.575684
34 4700 < 2e-04 ***
34 geneCT_40           -6.618051 -8.703288 -4.501595
35 3900 < 2e-04 ***
35 RIN_Value            -0.124016 -0.266650  0.014639
36 3039 0.08383 .
36 Jdaycont             0.008426  0.005884  0.010823
37 4626 < 2e-04 ***
37 baittemp.ave         -0.548965 -0.636750 -0.462001
36 3619 < 2e-04 ***

```

```

38 Delta 0.183909 -0.117180 0.453974
      4700 0.19830
39 geneCT_83:baittemp.ave 0.226478 0.129376 0.329830
      4700 < 2e-04 ***
40 geneCT_70:baittemp.ave 0.099461 0.025047 0.167024
      4700 0.00511 **
41 geneCT_40:baittemp.ave 0.047210 -0.027670 0.120743
      3815 0.21787
42 geneCT_83:Delta 0.024721 -0.231461 0.319838
      4700 0.85106
43 geneCT_70:Delta -0.018846 -0.282361 0.227798
      4486 0.88255
44 geneCT_40:Delta -0.001807 -0.260435 0.259638
      4989 0.99617
45 ---
46 Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 .
   0.1 ' ' 1

```

Lets look at combined SOFT Normalized model: test for site * bait temp interaction

model construction:

```

1 mm=mcmc.qpcr(
2   fixed="Site+baittemp.ave+Delta+Site:baittemp.ave+Site:De
lta",
3   random=c("Cham2"),globalFixed=c("RIN_Value","Jdaycont"),
4   controls=c("CT_actin","CT_gapdh","CT_18s"),include=0,dat
a=dd,pr=TRUE,vprior="iw",nitt=50000,geneSpecRes =
FALSE,pl=TRUE)
5 summary(mm)

```

```

1
2 Iterations = 3001:49991

```

```

3 Thinning interval = 10
4 Sample size = 4700
5
6 DIC: 37218.42
7
8 G-structure: ~sample
9
10 post.mean l-95% CI u-95% CI eff.samp
11 sample 6.274 5.005 7.636 672.2
12
13 ~idh(gene):Cham2
14
15 post.mean l-95% CI u-95% CI eff.samp
16 geneCT_40.Cham2 0.4176 0.1896 0.7132 4700
17 geneCT_70.Cham2 0.3976 0.1855 0.6655 4700
18 geneCT_83.Cham2 0.3950 0.1747 0.6653 4700
19 geneCT_18s.Cham2 0.5834 0.2599 1.0168 4700
20 geneCT_gapdh.Cham2 0.5542 0.2432 0.9563 4700
21 geneCT_actin.Cham2 0.4558 0.1924 0.7927 4700
22
23 R-structure: ~units
24
25 post.mean l-95% CI u-95% CI eff.samp
26 units 2.898 2.713 3.064 3033
27
28 Location effects: count ~ 0 + gene + RIN_Value +
Jdaycont + +gene:Site + gene:baittemp.ave + gene:Delta
+ gene:Site:baittemp.ave + gene:Site:Delta
29
30 post.mean l-95% CI
31 u-95% CI eff.samp pMCMC
32 geneCT_40 11.822062 9.005032
14.432795 3697 < 2e-04 ***
geneCT_70 13.794258 11.133172
16.508019 3917 < 2e-04 ***

```

33	geneCT_83			8.620266	5.721065
	11.238058	4442	< 2e-04	***	
34	geneCT_18s			29.771465	27.161697
	32.606825	4700	< 2e-04	***	
35	geneCT_gapdh			15.096474	12.192159
	17.705313	3620	< 2e-04	***	
36	geneCT_actin			12.489472	9.796855
	15.269530	4465	< 2e-04	***	
37	RIN_Value			-0.113658	-0.265027
	0.028872	4482	0.123404		
38	Jdaycont			0.010066	0.007892
	0.012371	4700	< 2e-04	***	
39	geneCT_40:SiteHF			-21.975881	-32.487062
	-11.289222	4700	< 2e-04	***	
40	geneCT_70:SiteHF			-25.065039	-35.036285
	-14.541056	4700	< 2e-04	***	
41	geneCT_83:SiteHF			-25.060015	-35.575939
	-14.371653	4700	< 2e-04	***	
42	geneCT_18s:SiteHF			-23.153648	-33.472287
	-12.514714	4955	< 2e-04	***	
43	geneCT_gapdh:SiteHF			-23.363184	-33.619975
	-12.762416	4700	< 2e-04	***	
44	geneCT_actin:SiteHF			-17.705744	-28.052297
	-7.287786	4700	0.000426	***	
45	geneCT_40:baittemp.ave			-0.388470	-0.485090
	-0.282392	3288	< 2e-04	***	
46	geneCT_70:baittemp.ave			-0.335051	-0.433817
	-0.235877	4026	< 2e-04	***	
47	geneCT_83:baittemp.ave			-0.200785	-0.296581
	-0.091502	4205	< 2e-04	***	
48	geneCT_18s:baittemp.ave			-0.625554	-0.726443
	-0.527732	4700	< 2e-04	***	
49	geneCT_gapdh:baittemp.ave			-0.421660	-0.523804
	-0.322599	3495	< 2e-04	***	

50	geneCT_actin:baittemp.ave -0.209647 4582 < 2e-04 ***	-0.311388	-0.411320
51	geneCT_40:Delta 0.431712 4700 0.306809	0.149857	-0.141126
52	geneCT_70:Delta 0.429957 4700 0.328511	0.140155	-0.153728
53	geneCT_83:Delta 0.467758 4700 0.196170	0.187249	-0.098200
54	geneCT_18s:Delta 0.554538 4700 0.119574	0.248577	-0.058683
55	geneCT_gapdh:Delta 0.525936 4700 0.148511	0.227091	-0.084731
56	geneCT_actin:Delta 0.414983 4700 0.358298	0.139773	-0.171241
57	geneCT_40:SiteHF:baittemp.ave 1.308730 4700 < 2e-04 ***	0.892454	0.505511
58	geneCT_70:SiteHF:baittemp.ave 1.401237 4550 < 2e-04 ***	1.016732	0.626500
59	geneCT_83:SiteHF:baittemp.ave 1.386713 4700 < 2e-04 ***	0.998540	0.591417
60	geneCT_18s:SiteHF:baittemp.ave 1.357887 4991 < 2e-04 ***	0.970596	0.565018
61	geneCT_gapdh:SiteHF:baittemp.ave 1.315875 4700 < 2e-04 ***	0.913561	0.529752
62	geneCT_actin:SiteHF:baittemp.ave 1.062279 4700 0.000851 ***	0.654713	0.280649
63	geneCT_40:SiteHF:Delta 0.053141 4480 0.058723 .	-0.490708	-0.949807
64	geneCT_70:SiteHF:Delta -0.106584 4700 0.017447 *	-0.609615	-1.101972
65	geneCT_83:SiteHF:Delta -0.056325 4924 0.034043 *	-0.542245	-1.046443
66	geneCT_18s:SiteHF:Delta -0.012797 4700 0.052340 .	-0.518337	-1.054346

```

67 geneCT_gapdh:SiteHF:Delta      -0.536793 -1.043260
     0.003550    4700 0.049787 *
68 geneCT_actin:SiteHF:Delta      -0.194960 -0.696403
     0.310583    4700 0.459574
69 ---
70 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
     0.1 ' ' 1

```

Lets look at combined SOFT Normalized model: test for site * bait temp interaction

...while controlling for 3 reference genes.... actin, 18s, gapdh

Model construction

```

1 softnorm=mcmc.qpcr(fixed="Site+baittemp.ave+Delta+Site:baittemp.ave+Site:Delta",
2 globalRandom="Cham2",data=dd,pr=TRUE,vprior="iw",nitt=50000,
3 geneSpecRes
  =FALSE,pl=TRUE,normalize=TRUE,controls=c("CT_actin","CT_gapdh","CT_18s"),
4 globalFixed=c("RIN_Value","Jdaycont"),m.fix=2)

```

output:

```

1 Iterations = 3001:49991
2 Thinning interval = 10
3 Sample size = 4700
4
5 DIC: 20307.05
6
7 G-structure: ~sample

```

```

8
9      post.mean l-95% CI u-95% CI eff.samp
10 sample     4.488     3.603     5.441     4700
11
12 ~Cham2
13
14      post.mean l-95% CI u-95% CI eff.samp
15 Cham2    0.05307 2.383e-09   0.2132    1238
16
17 R-structure: ~units
18
19      post.mean l-95% CI u-95% CI eff.samp
20 units    1.597     1.482     1.725     2090
21
22 Location effects: count ~ gene + RIN_Value + Jdaycont
+ Site + baittemp.ave + Delta + Site:baittemp.ave +
Site:Delta + gene:Site + gene:baittemp.ave + gene:Delta
+ gene:Site:baittemp.ave + gene:Site:Delta
23
24
25      post.mean     l-95% CI
26      u-95% CI eff.samp     pMCMC
27 (Intercept)           20.765616 18.518837
28      23.010194    3192 < 2e-04 ***
29 geneCT_83            -9.101701 -11.154752
30      -7.040275    4334 < 2e-04 ***
31 geneCT_70            -4.481281 -6.436953
32      -2.265938    4700 0.000426 ***
33 geneCT_40            -6.981607 -9.163210
34      -4.942460    3594 < 2e-04 ***
35 RIN_Value            -0.009514 -0.147358
36      0.121073    4269 0.881702
37 Jdaycont             0.008601  0.006671
38      0.010532    3976 < 2e-04 ***
39 SiteHF               -20.033476 -29.109914
40      -11.837664   4176 < 2e-04 ***

```

32	baittemp.ave		-0.518049	-0.608928
	-0.436984	2785 < 2e-04	***	
33	Delta		0.160806	-0.037144
	0.380270	4700 0.124681		
34	SiteHF:baittemp.ave		0.792163	0.465078
	1.113011	4215 < 2e-04	***	
35	SiteHF:Delta		-0.272132	-0.629893
	0.103980	4700 0.155319		
36	geneCT_83:SiteHF		-8.761300	-13.821726
	-3.439695	4046 0.000426	***	
37	geneCT_70:SiteHF		-6.710382	-11.733099
	-1.388099	4244 0.007234	**	
38	geneCT_40:SiteHF		-3.267049	-8.625278
	1.738584	3598 0.216596		
39	geneCT_83:baittemp.ave		0.204517	0.127550
	0.283634	4319 < 2e-04	***	
40	geneCT_70:baittemp.ave		0.090717	0.009584
	0.167971	4700 0.028511	*	
41	geneCT_40:baittemp.ave		0.059822	-0.017174
	0.143952	3497 0.146383		
42	geneCT_83:Delta		0.037275	-0.076766
	0.147296	4403 0.512766		
43	geneCT_70:Delta		-0.010336	-0.125017
	0.093081	4700 0.853191		
44	geneCT_40:Delta		-0.005442	-0.115911
	0.108092	4444 0.934894		
45	geneCT_83:SiteHF:baittemp.ave		0.335731	0.128712
	0.518240	4106 0.000426	***	
46	geneCT_70:SiteHF:baittemp.ave		0.286206	0.090584
	0.478491	4299 0.003404	**	
47	geneCT_40:SiteHF:baittemp.ave		0.147327	-0.038703
	0.354963	3555 0.130638		
48	geneCT_83:SiteHF:Delta		-0.183220	-0.391793
	0.014840	3846 0.081277	.	

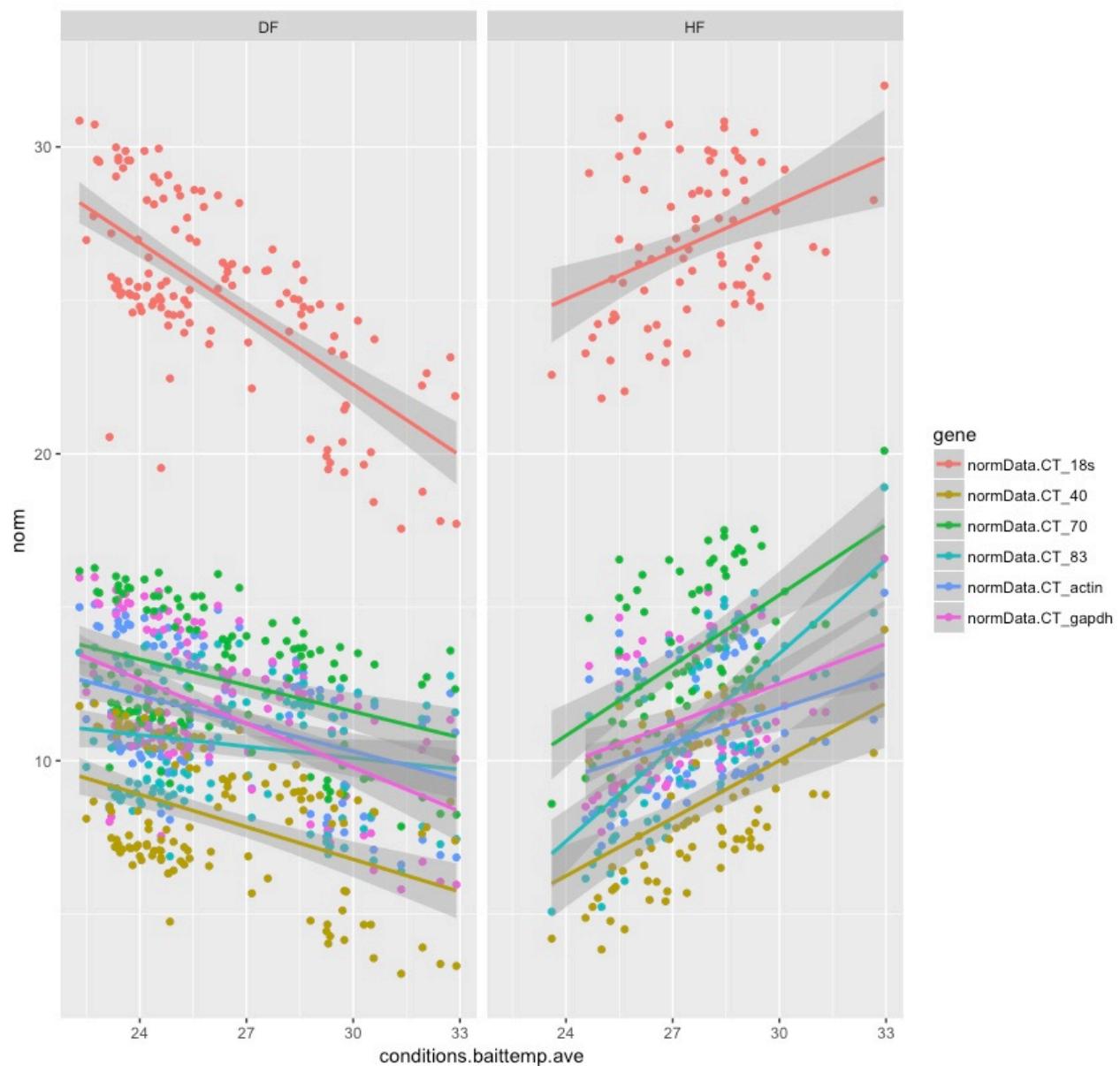
```

49 geneCT_70:SiteHF:Delta      -0.281034  -0.475357
     -0.071999      3531 0.006809 ** 
50 geneCT_40:SiteHF:Delta      -0.165264  -0.366331
     0.039326      3576 0.109787
51 ---
52 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
     0.1 ' ' 1

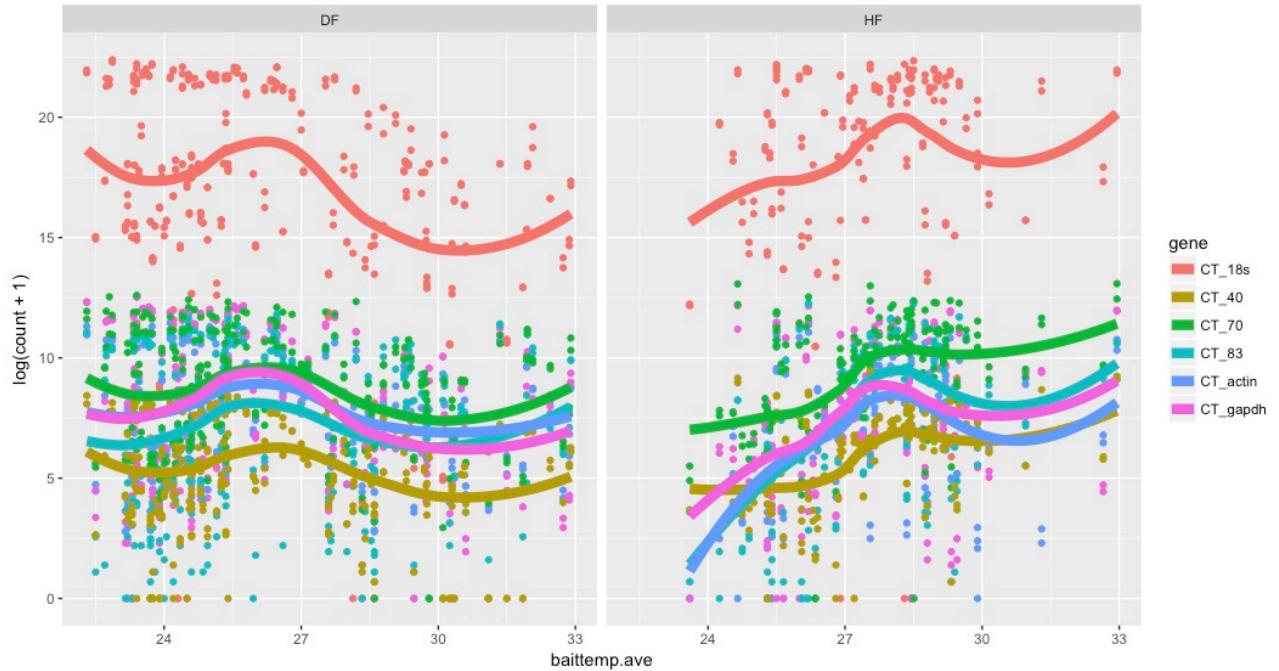
```

Showing effect of bait temp on gxp

Linear fit



Loess fit



Refitting classic mod

Parsing data

```
1 ctmeth<-
  cq2log(data=warm, genecols=c(14:19), condcols=c(13,5:6,1:4
  ,20:27), noamp=37, effic=amp)
2 ctmeth$RIN_Value<-
  as.numeric(as.character(ctmeth$RIN_Value))
3 ctmeth$Delta<-as.numeric(as.character(ctmeth$Delta))
4 ctmeth$baittemp.ave<-
  as.numeric(as.character(ctmeth$baittemp.ave))
5 ctmeth$Jdaycont<-
  as.numeric(as.character(ctmeth$Jdaycont))
6 str(ctmeth)
```

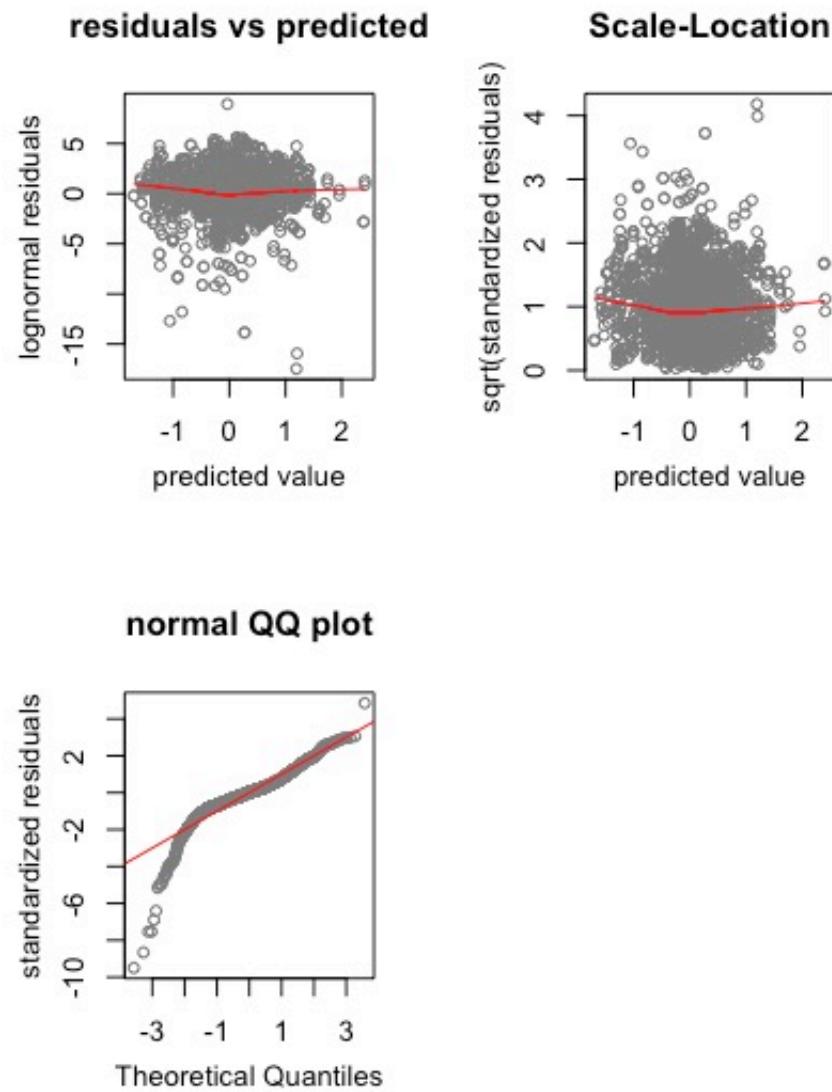
Model fitting

```

1 classicmod<-
mcmc.qpcr.classic(fixed="Site+baittemp.ave+Delta+Site:ba
ittemp.ave+Site:Delta",globalRandom="Cham",globalFixed=c
("RIN_Value","Jdaycont"),data=ctmeth,controls=c("CT_acti
n","CT_gapdh"),pr=T,pl=T,nitt=50000)

```

diagnostics



output:

```

1 Iterations = 3001:49991
2 Thinning interval = 10

```

```

3 Sample size = 4700
4
5 DIC: 10316.62
6
7 G-structure: ~idh(gene):Cham
8
9 post.mean l-95% CI u-95% CI
eff.samp
10 geneCT_40.Cham 2.168e-03 2.843e-63 1.743e-08
134.26
11 geneCT_70.Cham 5.678e-02 5.842e-49 3.045e-01
78.55
12 geneCT_83.Cham 6.201e-18 6.381e-108 4.589e-33
0.00
13 geneCT_gapdh.Cham 1.519e-03 1.681e-39 4.673e-03
364.01
14 geneCT_actin.Cham 7.960e-05 2.641e-49 2.683e-07
615.45
15 geneCT_18s.Cham 3.017e-05 2.951e-56 1.972e-09
401.62
16
17 ~idh(gene):sample
18
19 post.mean l-95% CI u-95% CI
eff.samp
20 geneCT_40.sample 1.8769 1.3605 2.436
4700.0
21 geneCT_70.sample 1.5604 1.0790 2.072
706.6
22 geneCT_83.sample 2.1509 1.5688 2.778
4700.0
23 geneCT_gapdh.sample 0.7484 0.4714 1.062
4004.1
24 geneCT_actin.sample 1.7336 1.2430 2.231
5279.4

```

```

25 geneCT_18s.sample      2.6897   2.0524   3.357
26 4700.0
27 R-structure: ~units
28
29       post.mean l-95% CI u-95% CI eff.samp
30 units      1.676     1.55     1.797     4988
31
32 Location effects: count ~ 0 + gene + RIN_Value +
Jdaycont + +gene:Site + gene:baittemp.ave + gene:Delta
+ gene:Site:baittemp.ave + gene:Site:Delta
33
34                               post.mean    l-95% CI
35                         u-95% CI eff.samp    pMCMC
36 geneCT_40                      1.399e+00 -7.993e-01
37             3.874e+00 4700.0 0.24511
38 geneCT_70                      7.141e-01 -1.508e+00
39             2.982e+00 4700.0 0.52851
40 geneCT_83                      -1.465e+00 -4.054e+00
41             8.896e-01 4700.0 0.23617
42 geneCT_gapdh                  3.971e-02 -1.760e+00
43             1.890e+00 4955.7 0.97447
44 geneCT_actin                  -1.757e+00 -3.932e+00
45             3.514e-01 4700.0 0.10468
46 geneCT_18s                     5.140e+00  2.722e+00
47             7.495e+00 5052.1 < 2e-04 ***
48 RIN_Value                      -3.748e-02 -8.486e-02
49             1.419e-02 4700.0 0.14894
50 Jdaycont                      -4.573e-05 -7.047e-04
51             6.020e-04 4700.0 0.87787
52 geneCT_40:SiteHF              -3.176e+00 -9.216e+00
53             2.793e+00 4700.0 0.29957
54 geneCT_70:SiteHF              -7.646e+00 -1.329e+01
55             -1.757e+00 4700.0 0.01064 *

```

45	geneCT_83:SiteHF		-1.059e+01	-1.697e+01
	-4.130e+00	4542.6	0.00128	**
46	geneCT_gapdh:SiteHF		-2.480e+00	-7.021e+00
	2.227e+00	4700.0	0.29702	
47	geneCT_actin:SiteHF		1.344e+00	-4.488e+00
	7.702e+00	4924.3	0.65957	
48	geneCT_18s:SiteHF		-3.452e+00	-1.021e+01
	3.342e+00	4700.0	0.32809	
49	geneCT_40:baittemp.ave		-5.011e-02	-1.392e-01
	3.826e-02	4700.0	0.27532	
50	geneCT_70:baittemp.ave		-2.405e-02	-1.078e-01
	6.303e-02	4700.0	0.57532	
51	geneCT_83:baittemp.ave		5.755e-02	-3.511e-02
	1.517e-01	4700.0	0.22085	
52	geneCT_gapdh:baittemp.ave		1.086e-02	-6.153e-02
	7.812e-02	4940.6	0.74426	
53	geneCT_actin:baittemp.ave		9.179e-02	1.157e-02
	1.746e-01	4700.0	0.02766	*
54	geneCT_18s:baittemp.ave		-2.113e-01	-3.014e-01
	-1.217e-01	4891.3	< 2e-04	***
55	geneCT_40:Delta		-4.476e-02	-1.834e-01
	9.682e-02	4700.0	0.52851	
56	geneCT_70:Delta		-7.246e-02	-2.236e-01
	8.242e-02	587.3	0.33064	
57	geneCT_83:Delta		1.839e-02	-1.273e-01
	1.684e-01	4700.0	0.79489	
58	geneCT_gapdh:Delta		5.717e-03	-1.051e-01
	1.103e-01	4700.0	0.92681	
59	geneCT_actin:Delta		-6.490e-02	-2.030e-01
	7.087e-02	4700.0	0.34085	
60	geneCT_18s:Delta		1.006e-01	-5.750e-02
	2.662e-01	4700.0	0.22383	
61	geneCT_40:SiteHF:baittemp.ave		1.466e-01	-7.800e-02
	3.808e-01	4700.0	0.20468	

```

62 geneCT_70:SiteHF:baittemp.ave      3.269e-01  1.113e-01
      5.461e-01  4700.0 0.00426 **

63 geneCT_83:SiteHF:baittemp.ave      4.085e-01  1.596e-01
      6.436e-01  4700.0 0.00170 **

64 geneCT_gapdh:SiteHF:baittemp.ave  8.618e-02 -9.165e-02
      2.573e-01  4700.0 0.33617

65 geneCT_actin:SiteHF:baittemp.ave -1.129e-01 -3.417e-01
      1.177e-01  4924.1 0.31745

66 geneCT_18s:SiteHF:baittemp.ave   2.030e-01 -5.703e-02
      4.546e-01  4700.0 0.12128

67 geneCT_40:SiteHF:Delta          -7.663e-02 -3.230e-01
      1.895e-01  5225.8 0.54766

68 geneCT_70:SiteHF:Delta          -2.160e-01 -4.589e-01
      2.999e-02  4700.0 0.08766 .

69 geneCT_83:SiteHF:Delta          -1.278e-01 -3.876e-01
      1.310e-01  4700.0 0.33447

70 geneCT_gapdh:SiteHF:Delta      -9.679e-02 -2.957e-01
      9.375e-02  4700.0 0.31234

71 geneCT_actin:SiteHF:Delta      2.710e-01  2.563e-02
      5.100e-01  4700.0 0.02979 *

72 geneCT_18s:SiteHF:Delta          -2.297e-01 -5.112e-01
      6.177e-02  4700.0 0.11957

73 ---
74 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Manually calculating delta delta CT method

data parsing

```

1 x<-
2   read.csv("../Data/20170427_final_dataset_2013_2014_year
3     s_HF_DF_del.csv",skip=10)
4 str(x)
5 x$baittemp.ave<-apply(x[,8:11],1,mean,na.rm=TRUE)
6 x$Cham<-as.factor(as.character(x$Cham))
7 x2<-inner_join(x,chd,by=c("Cham","Site"))
8
9 #julian day
10 v2<-paste(substr(x2$Collection.Date,1,4),"-
11   ",substr(x2$Collection.Date,5,6),"-
12   ",substr(x2$Collection.Date,7,8))
13 x2$date<-as.character(gsub(" ", "", v2, fixed = TRUE))
14 x2$JulianDay<-as.numeric(format(as.Date(x2$date), "%j"))
15 x2$Jdaycont<-
16 ifelse(x2$Year_collect==2013,x2$JulianDay,x2$JulianDay+
17   365)
18 x2<-x2[!is.na(x2$RIN_Value),]
19 x3<-na.omit(x2[,-8:-11])
20
21 hkg<-cbind(x3$CT_18s,x3$CT_actin,x3$CT_gapdh)
22 x3$geomean<-apply(hkg,1,function(x){exp(mean(log(x)))})
23
24 globalmean<-apply(x3[,10:15],2,mean)
25 deltaCTnum<-data.frame(t(apply(x3[,10:15],1,function(x)
26   {globalmean-x})))
27 deltaCTden<-mean(x3$geomean)-x3$geomean
28
29 ##calculating delta delta ct
30 gxp<-data.frame(log2(apply(deltaCTnum,2,function(x)
31   {2^(x)/2^(deltaCTden)})))
32 names(gxp)<-
33 c("fc18s","fchsp40","fchsp70","fchsp83","fcactin","fcga
34 pdh")

```

```

26 findat<-data.frame(x3,gxp)
27 str(findat)
28
29 ##convert to long format
30 findat.long<-gather(findat,gene,FC,fc18s:fcgapdh)
31

```

Anovas

```

1 fullmod2<-
2 aov(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+gene*S
ite*Delta,data=findat.long)
3 tep: AIC=1611.21
4 FC ~ RIN_Value + Jdaycont + gene + Site + baittemp.ave
+ Delta +
      gene:Site + gene:baittemp.ave + Site:baittemp.ave +
Site:Delta
5
6
7 <none>
8   Df Sum of Sq    RSS    AIC
9 - RIN_Value           1  7.679 4392.9 1611.4
10 - Site:baittemp.ave  1 14.268 4399.4 1613.2
11 - Site:Delta         1 14.417 4399.6 1613.3
12 + gene:Site:baittemp.ave 5 24.236 4360.9 1614.4
13 + gene:Delta          5  2.085 4383.1 1620.6
14 - gene:Site           5 185.321 4570.5 1652.4
15 - Jdaycont            1 197.915 4583.1 1663.8
16 - gene:baittemp.ave   5 306.347 4691.5 1684.7
17
18   Df Sum Sq Mean Sq F value    Pr(>F)
19 RIN_Value             1     46    46.0  12.719 0.000376
20 ***

21 Jdaycont              1   115   114.8  31.747 2.18e-08
22 ***

```

```

19 gene 5 0 0.0 0.000 1.000000
20 Site 1 63 63.3 17.505 3.07e-05
***  

21 baittemp.ave 1 460 460.4 127.347 < 2e-16
***  

22 Delta 1 1 0.9 0.262 0.608759
23 gene:Site 5 161 32.1 8.889 2.66e-08
***  

24 gene:baittemp.ave 5 306 61.3 16.948 3.18e-16
***  

25 Site:baittemp.ave 1 5 5.4 1.505 0.220175
26 Site:Delta 1 14 14.4 3.988 0.046046
*
27 Residuals 1213 4385 3.6
28 ---
29 Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 .
0.1 ' ' 1

```

Statistics: linear models

```

1 fullmod3<-
lm(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+
2 gene*Site*Delta,data=findat.long)
3
4 Call:
5 lm(formula = FC ~ RIN_Value + Jdaycont + gene * Site *
baittemp.ave +
6     gene * Site * Delta, data = findat.long)
7
8 Residuals:

```

9		Min	1Q	Median	3Q	Max
10		-9.8693	-1.0781	-0.0503	1.0091	9.5413
11	 Coefficients:					
12	 Estimate Std. Error t					
13		value	Pr(> t)			
14	(Intercept)				2.453e+00	1.625e+00
15	1.510	0.131335				
16	RIN_Value				4.992e-02	3.426e-02
17	1.457	0.145368				
18	Jdaycont				-3.134e-03	4.237e-04
19	-7.397	2.61e-13 ***				
20	genefcactin				-5.834e+00	2.294e+00
21	-2.543	0.011114 *				
22	genefcgapdh				-7.245e+00	2.294e+00
23	-3.158	0.001627 **				
24	genefchsp40				-9.305e+00	2.294e+00
25	-4.056	5.31e-05 ***				
26	genefchsp70				-1.310e+01	2.294e+00
27	-5.709	1.43e-08 ***				
28	genefchsp83				-1.592e+01	2.294e+00
29	-6.939	6.46e-12 ***				
30	SiteHF				1.644e+00	4.042e+00
31	0.407	0.684228				
32	baittemp.ave				-7.262e-02	6.232e-02
33	-1.165	0.244167				
34	Delta				5.386e-02	8.458e-02
35	0.637	0.524364				
36	genefcactin:SiteHF				-1.749e+00	5.672e+00
37	-0.308	0.757798				
38	genefcgapdh:SiteHF				8.873e+00	5.672e+00
39	1.565	0.117957				
40	genefchsp40:SiteHF				8.031e-01	5.672e+00
41	0.142	0.887418				

28	genefchsp70:SiteHF 0.359 0.719756	2.035e+00	5.672e+00
29	genefchsp83:SiteHF -1.003 0.316148	-5.688e+00	5.672e+00
30	genefcactin:baittemp.ave 2.958 0.003156 **	2.601e-01	8.791e-02
31	genefcgapdh:baittemp.ave 3.578 0.000361 ***	3.145e-01	8.791e-02
32	genefchsp40:baittemp.ave 4.303 1.82e-05 ***	3.783e-01	8.791e-02
33	genefchsp70:baittemp.ave 5.873 5.55e-09 ***	5.163e-01	8.791e-02
34	genefchsp83:baittemp.ave 7.156 1.45e-12 ***	6.291e-01	8.791e-02
35	SiteHF:baittemp.ave -0.122 0.903285	-1.865e-02	1.534e-01
36	genefcactin:Delta -1.126 0.260301	-1.346e-01	1.195e-01
37	genefcgapdh:Delta -0.880 0.378968	-1.052e-01	1.195e-01
38	genefchsp40:Delta -1.022 0.306847	-1.222e-01	1.195e-01
39	genefchsp70:Delta -1.060 0.289178	-1.268e-01	1.195e-01
40	genefchsp83:Delta -0.837 0.402580	-1.001e-01	1.195e-01
41	SiteHF:Delta -1.061 0.288975	-1.694e-01	1.597e-01
42	genefcactin:SiteHF:baittemp.ave -0.313 0.754003	-6.741e-02	2.151e-01
43	genefcgapdh:SiteHF:baittemp.ave -2.129 0.033428 *	-4.580e-01	2.151e-01
44	genefchsp40:SiteHF:baittemp.ave -0.525 0.599933	-1.128e-01	2.151e-01

```

45 genechsp70:SiteHF:baittemp.ave -1.472e-01 2.151e-01
   -0.684 0.493905
46 genechsp83:SiteHF:baittemp.ave 1.194e-01 2.151e-01
   0.555 0.578921
47 genefactin:SiteHF:Delta      5.114e-01 2.255e-01
   2.267 0.023544 *
48 genefgapdh:SiteHF:Delta     4.643e-01 2.255e-01
   2.058 0.039760 *
49 genechsp40:SiteHF:Delta     3.145e-01 2.255e-01
   1.394 0.163497
50 genechsp70:SiteHF:Delta     3.197e-01 2.255e-01
   1.418 0.156581
51 genechsp83:SiteHF:Delta     1.927e-01 2.255e-01
   0.854 0.393119
52 ---
53 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
   0.1 ' ' 1
54
55 Residual standard error: 1.902 on 1198 degrees of
   freedom
56 Multiple R-squared: 0.2203,    Adjusted R-squared:
   0.1962
57 F-statistic: 9.147 on 37 and 1198 DF,  p-value: < 2.2e-
   16

```

With model selection

```

1 Call:
2 lm(formula = FC ~ RIN_Value + Jdaycont + gene + Site +
   baittemp.ave +
   Delta + gene:Site + gene:baittemp.ave +
   Site:baittemp.ave +
   Site:Delta, data = findat.long)
5
6 Residuals:

```

	Min	1Q	Median	3Q	Max	
7	-10.2877	-1.0909	-0.0054	1.0415	10.0486	
8	9					
10	Coefficients:					
11				Estimate	Std. Error	t value
12	Pr(> t)					
13	(Intercept)			2.403e+00	1.476e+00	1.628
14	0.103837					
15	RIN_Value			4.992e-02	3.425e-02	1.457
16	0.145257					
17	Jdaycont			-3.134e-03	4.236e-04	-7.399
18	2.55e-13 ***					
19	genefcactin			-6.521e+00	2.039e+00	-3.199
20	0.001416 **					
21	genefcgapdh			-5.735e+00	2.039e+00	-2.813
22	0.004988 **					
23	genefchsp40			-9.209e+00	2.039e+00	-4.517
24	6.89e-06 ***					
25	genefchsp70			-1.281e+01	2.039e+00	-6.284
26	4.60e-10 ***					
27	genefchsp83			-1.682e+01	2.039e+00	-8.249
28	4.11e-16 ***					
29	SiteHF			3.943e+00	1.732e+00	2.277
30	0.022962 *					
31	baittemp.ave			-6.304e-02	5.609e-02	-1.124
32	0.261332					
33	Delta			-4.429e-02	3.462e-02	-1.280
34	0.200961					
35	genefcactin:SiteHF			-2.324e+00	4.035e-01	-5.761
36	1.06e-08 ***					
37	genefcgapdh:SiteHF			-2.496e+00	4.035e-01	-6.185
38	8.46e-10 ***					
39	genefchsp40:SiteHF			-1.520e+00	4.035e-01	-3.767
40	0.000173 ***					

```

26  genefchsp70:SiteHF      -1.215e+00  4.035e-01 -3.012
    0.002647 **

27  genefchsp83:SiteHF      -1.964e+00  4.035e-01 -4.868
    1.28e-06 ***

28  genefcactin:baittemp.ave 2.758e-01  7.720e-02  3.572
    0.000368 ***

29  genefcgapdh:baittemp.ave 2.487e-01  7.720e-02  3.222
    0.001306 **

30  genefchsp40:baittemp.ave 3.651e-01  7.720e-02  4.730
    2.51e-06 ***

31  genefchsp70:baittemp.ave 4.955e-01  7.720e-02  6.419
    1.96e-10 ***

32  genefchsp83:baittemp.ave 6.556e-01  7.720e-02  8.493
    < 2e-16 ***

33  SiteHF:baittemp.ave     -1.296e-01  6.526e-02 -1.987
    0.047184 *

34  SiteHF:Delta            1.310e-01  6.560e-02  1.997
    0.046046 *

35  ---
36  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
    0.1 ' ' 1
37
38  Residual standard error: 1.901 on 1213 degrees of
   freedom
39  Multiple R-squared:  0.2109,    Adjusted R-squared:
   0.1966
40  F-statistic: 14.74 on 22 and 1213 DF,  p-value: < 2.2e-
   16

```

Statistics with random effects

```

1 fullmod5<-
  lme(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+gene*S
  ite*Delta,random=~1|Cham2,data=findat.long)

```

```

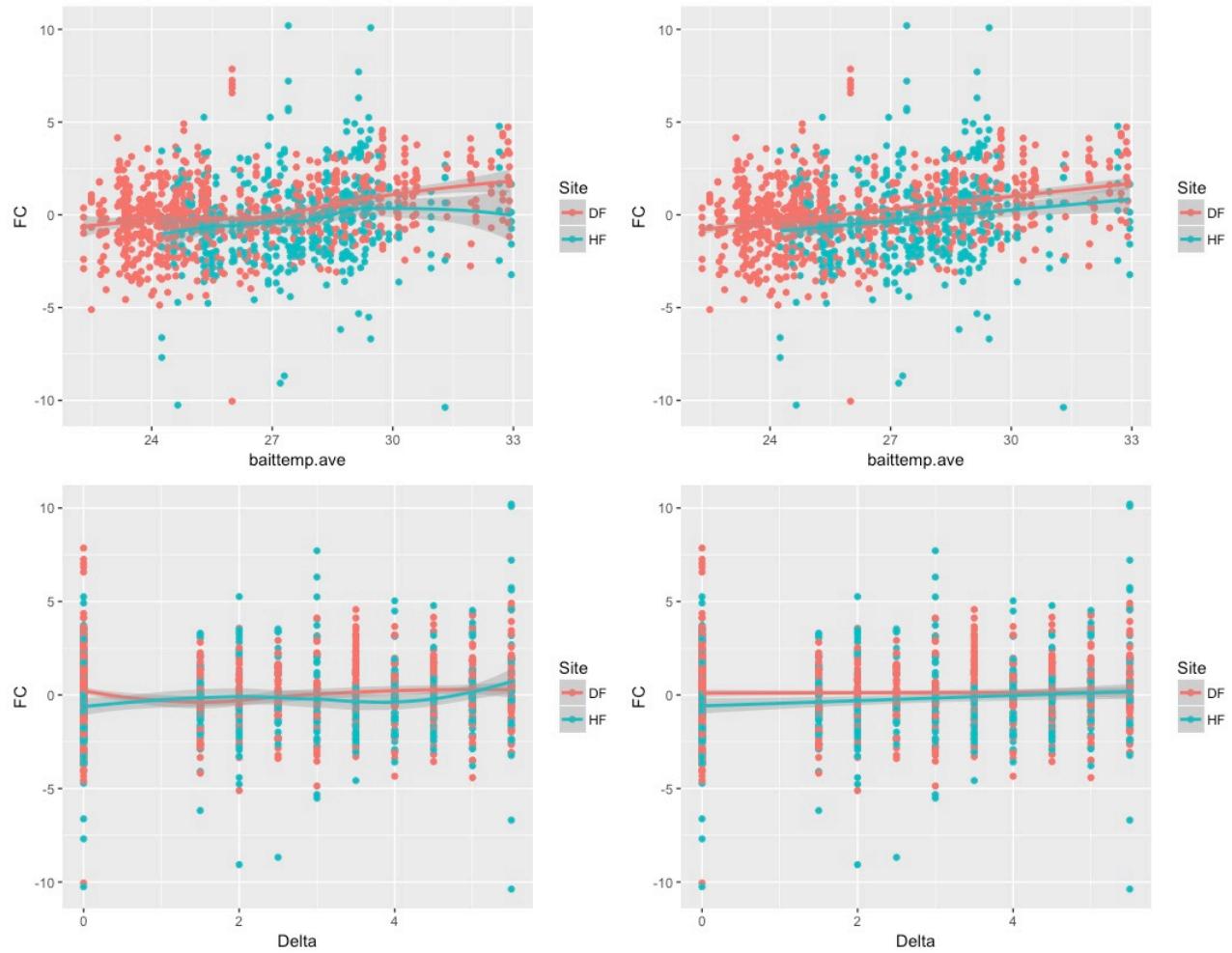
2 summary(fullmod5)
3 anova(fullmod5)

4                               numDF  denDF   F-value p-value
5 (Intercept)                  1    1175  0.03058  0.8612
6 RIN_Value                   1    1175 13.28986  0.0003
7 Jdaycont                     1    1175 35.65154 <.0001
8 gene                          5    1175  0.00000  1.0000
9 Site                         1      23  12.21883  0.0019
10 baittemp.ave                 1    1175 122.07718 <.0001
11 Delta                        1      23  0.13595  0.7157
12 gene:Site                    5    1175  8.95546 <.0001
13 gene:baittemp.ave            5    1175 17.07517 <.0001
14 Site:baittemp.ave            1    1175  2.17829  0.1402
15 gene:Delta                   5    1175  0.11621  0.9888
16 Site:Delta                   1      23  2.95385  0.0991
17 gene:Site:baittemp.ave       5    1175  1.39975  0.2216
18 gene:Site:Delta              5    1175  1.37753  0.2300

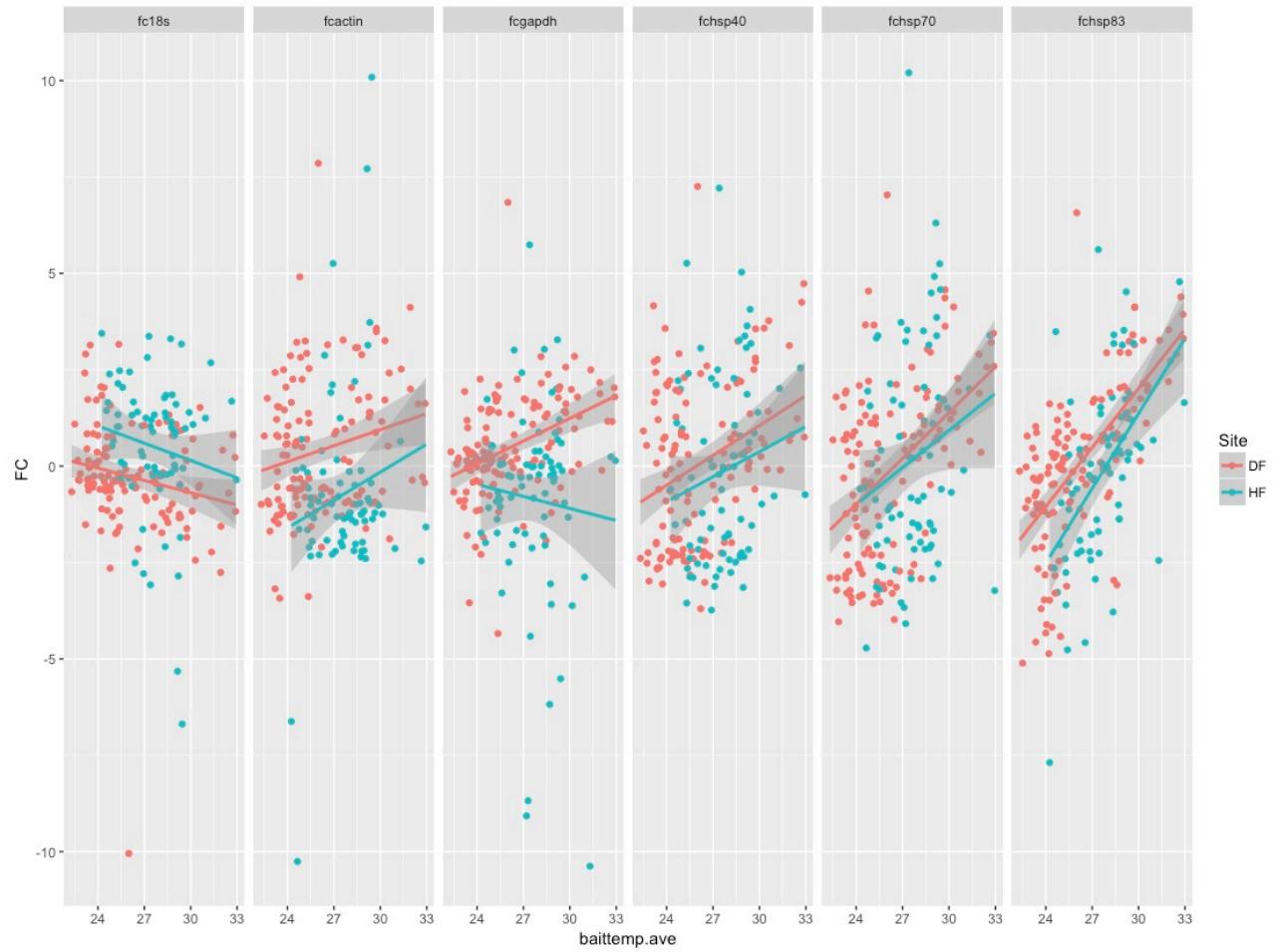
```

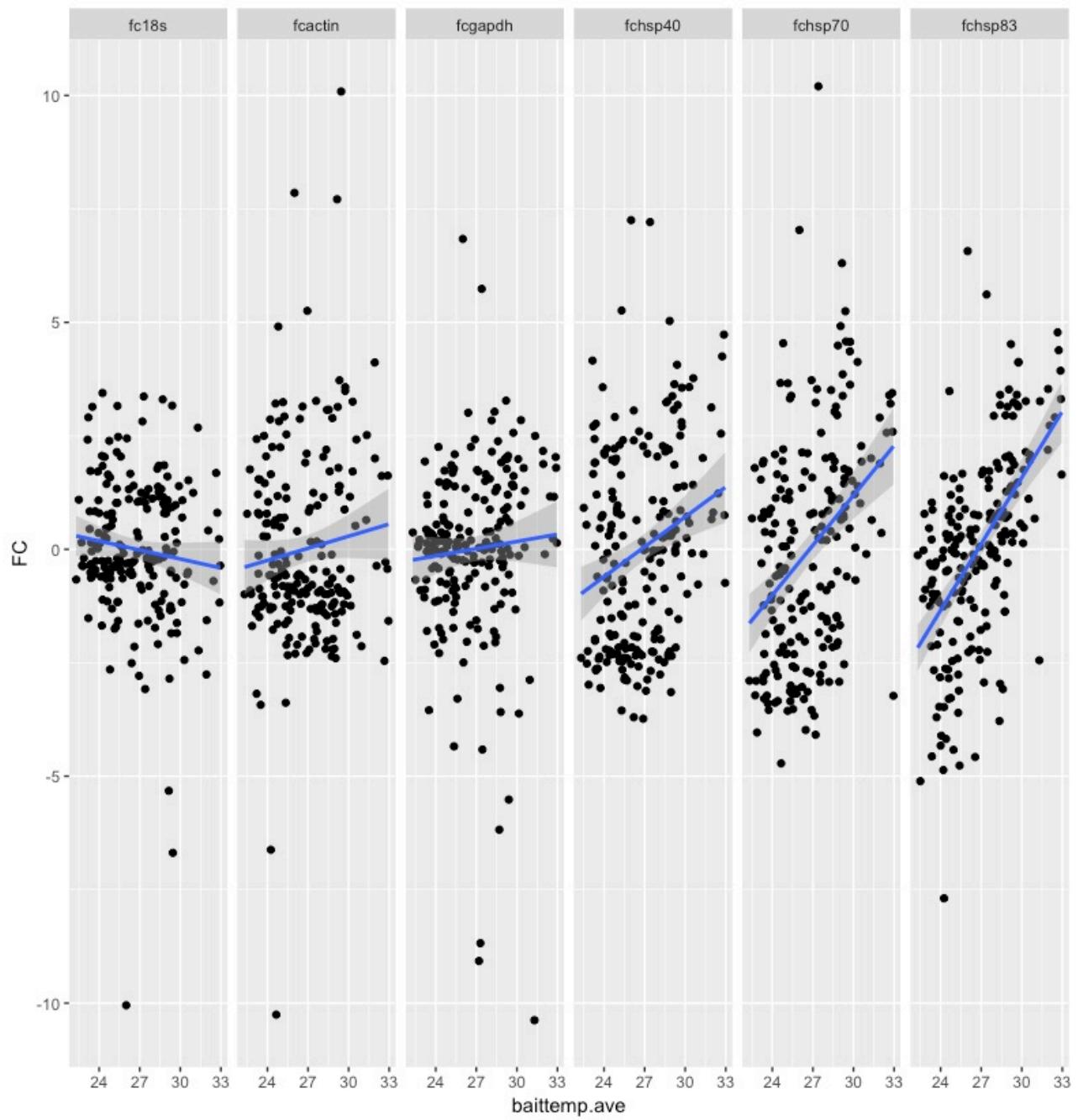
Figures

Differences between hf and df with bait and delta



Parse out by gene

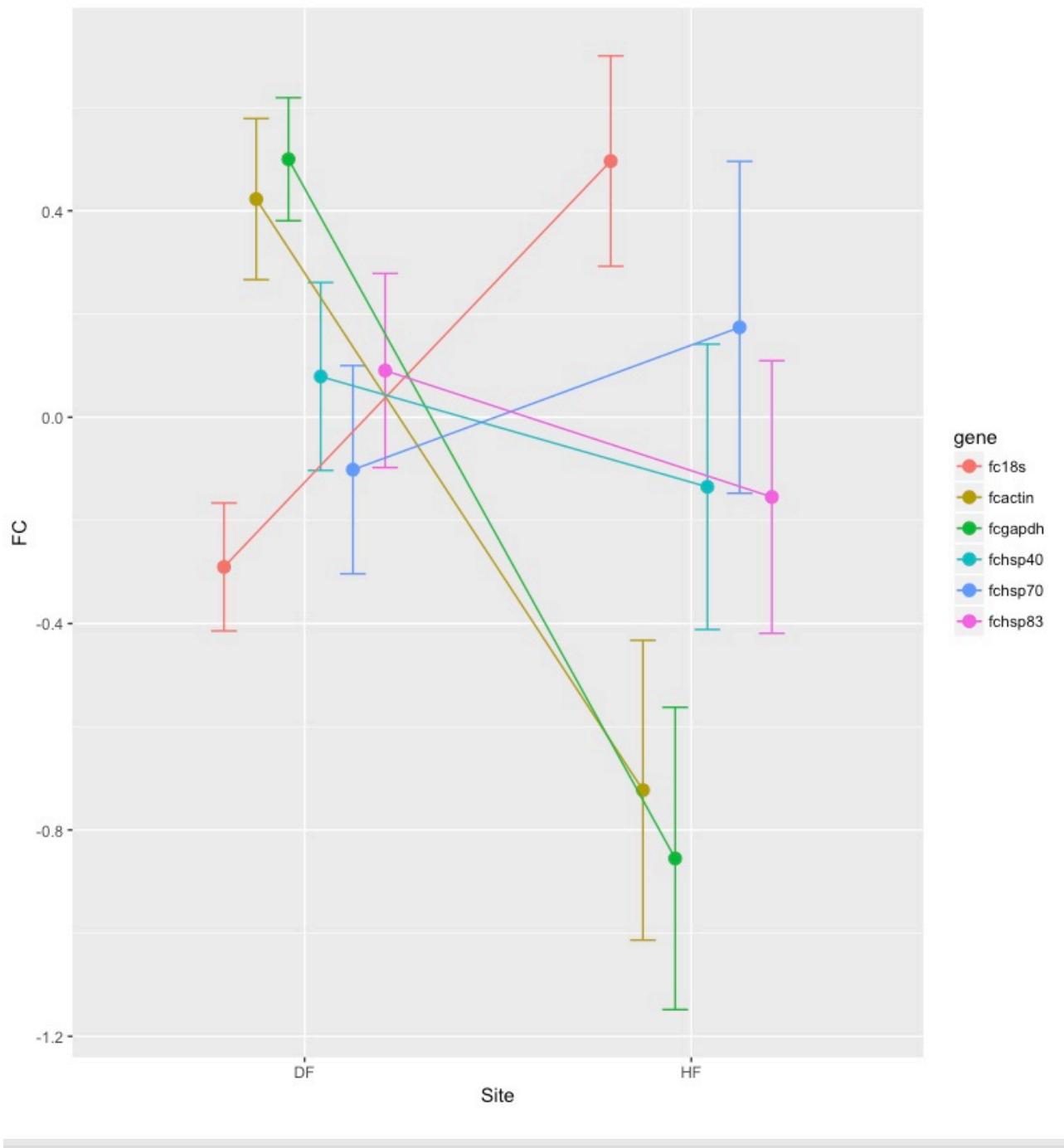




loess



Difference in FC of different genes between sites



Page 78: 2017-05-08. quantifying gene stability with geNorm in the SLqPCR package

MCMC.qpcr package can convert raw ct values into a table for geNORM. geNORM evaluates gene expression stability. Below is the code....

- for the 3 reference genes

```
1 stable<-
cq2genorm(warm,genes=c("CT_actin","CT_gapdh","CT_18s"),n
oamp=37,effic=amp)
2 head(stable)
```

- for all 6 genes

```
1 all<-
cq2genorm(warm,genes=c("CT_actin","CT_gapdh","CT_18s","C
T_70","CT_83","CT_40"),noamp=37,effic=amp)
```

The selectHKgenes function:

```
1 This function can be used to determine a set of
reference/housekeeping (HK) genes for gene expression
experiments. The default method "Vandesompele" was
proposed by Vandesompele et al. (2002).
2
3 If method = "Vandesompele" a list with the following
components is returned
4 ranking ranking of genes from best to worst where the
two most stable genes cannot be ranked
5 variation pairwise variation during stepwise selection
6 meanM average expression stability M
```

What is stability M? How is it measured?

From their paper (Vandesompele); they determine the pairwise variation for every control gene as the SD of log transformed expression ratios. They define this as gene-stability measure M -- average variation of a particular gene with all other control genes.

- LOW M = most stable.

1. Evaluating stability for 3 ref sets:

```
1 selectHKgenes(stable,geneSymbol=c("actin","gapdh","18s"))
2 #####
3 #####
4 Step 1 :
5 gene expression stability values M:
6      gapdh      actin      18s
7      4.109859  4.417064  5.242474
8 average expression stability M:  4.589799
9 gene with lowest stability (largest M value):    18s
10 Pairwise variation, ( 2 / 3 ) :   1.662695
11 #####
12 #####
13 Step 2 :
14 gene expression stability values M:
15      actin      gapdh
16      3.28445  3.28445
17 average expression stability M:  3.28445
18 $ranking
19      1          1          3
20      "actin"   "gapdh"   "18s"
21
22 $variation
23      2/3
24      1.662695
25
```

```
24 | $meanM  
25 | 3 2  
26 | 4.589799 3.284450
```

2. Eval all 6 genes

```
1 selectHKgenes(all,geneSymbol=c("CT_actin","CT_gapdh","C  
T_18s","CT_70","CT_83","CT_40"))  
2 #####  
3 Step 1 :  
4 gene expression stability values M:  
5 CT_70 CT_40 CT_83 CT_gapdh CT_actin CT_18s  
6 3.213355 3.274018 3.618803 3.692725 3.713486 4.898314  
7 average expression stability M: 3.735117  
8 gene with lowest stability (largest M value): CT_18s  
9 Pairwise variation, ( 5 / 6 ): 0.7470254  
10 #####  
11 Step 2 :  
12 gene expression stability values M:  
13 CT_70 CT_40 CT_actin CT_83 CT_gapdh  
14 2.880148 2.963412 3.254438 3.287504 3.382089  
15 average expression stability M: 3.153518  
16 gene with lowest stability (largest M value):  
CT_gapdh  
17 Pairwise variation, ( 4 / 5 ): 0.5661937  
18 #####  
19 Step 3 :  
20 gene expression stability values M:  
21 CT_70 CT_40 CT_83 CT_actin  
22 2.758816 2.838390 3.162911 3.244434  
23 average expression stability M: 3.001138
```

```

24 gene with lowest stability (largest M value): CT_actin
25 Pairwise variation, ( 3 / 4 ): 0.702729
26 #####
27 #####
28 Step 4 :
29 gene expression stability values M:
30   CT_70    CT_40    CT_83
31 2.498380 2.659493 3.115653
32 average expression stability M: 2.757842
33 gene with lowest stability (largest M value): CT_83
34 Pairwise variation, ( 2 / 3 ): 0.9826598
35 #####
36 Step 5 :
37 gene expression stability values M:
38   CT_70    CT_40
39 2.04222 2.04222
40 $ranking
41      1          1          3          4          5
42      6
43      "CT_70"    "CT_40"    "CT_83"  "CT_actin" "CT_gapdh"
44      "CT_18s"
45
46 $variation
47      5/6        4/5        3/4        2/3
48 0.7470254 0.5661937 0.7027290 0.9826598
49
50 $meanM
51      6          5          4          3          2
52 3.735117 3.153518 3.001138 2.757842 2.042220

```

It looks like the ref set is more variable than the gene set we're interested in. Anyway, when looking at the control set, gapdh and actin are the most stable. I should try to calculate relative expression using the geometric mean of these 2.

OK, lets look at the model without including 18s as a HKG.

data parsing: i just need to alter this line of code

```
1 #hkg<-cbind(x3$CT_18s,x3$CT_actin,x3$CT_gapdh)
2 hkg<-cbind(x3$CT_actin,x3$CT_gapdh)
3
```

Statistics: linear mixed effects models

```
1 fullmod5<-
  lme(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+gene*S
  ite*Delta,random=~1|Cham2/Vial.me,data=findat.long)
2 anova(fullmod5)
3
4             numDF  denDF   F-value p-value
5 (Intercept)        1  1006  0.097464  0.7550
6 RIN_Value          1  1006  0.002386  0.9611
7 Jdaycont           1  1006  7.053760  0.0080
8 gene               5  1006  0.000000  1.0000
9 Site                1      23 18.638167  0.0003
10 baittemp.ave       1  1006  4.349922  0.0373
11 Delta              1      23  0.094197  0.7617
12 gene:Site          5  1006  9.628759 <.0001
13 gene:baittemp.ave  5  1006 18.358946 <.0001
14 Site:baittemp.ave  1     169  1.410071  0.2367
15 gene:Delta          5  1006  0.124950  0.9868
16 Site:Delta          1      23  3.457577  0.0758
17 gene:Site:baittemp.ave  5  1006  1.504986  0.1855
```

```
17 | gene:Site:Delta      5  1006  1.481099  0.1933
```

Ok, let's do model selection: forward and backwards:

```
1 anova(summary(stepAIC(fullmod5,direction="both")))
2
3             numDF  denDF   F-value p-value
4 (Intercept)       1  1022  0.061735  0.8038
5 Jdaycont         1  1022  4.757535  0.0294
6 gene            5  1022  0.000000  1.0000
7 Site             1     23 24.925993 <.0001
8 baittemp.ave     1  1022  4.687937  0.0306
9 Delta            1     23  0.080154  0.7796
10 gene:Site        5  1022  9.591095 <.0001
11 gene:baittemp.ave 5  1022 18.287133 <.0001
12 Site:baittemp.ave 1    169  1.137005  0.2878
13 Site:Delta        1     23  3.889363  0.0607
```

Ok, let's look at the parameter estimates:

```
1 summary(stepAIC(fullmod5,direction="both"))
2 Fixed effects: FC ~ Jdaycont + gene + Site +
3                                baittemp.ave + Delta + gene:Site +
4                                gene:baittemp.ave + Site:baittemp.ave + Site:Delta
5
6             Value Std.Error DF t-
7             value p-value
8 (Intercept) 7.079159 1.5735006 1022
9 4.498987 0.0000
10 Jdaycont -0.000706 0.0004481 1022
11 -1.575720 0.1154
12 genefcactin -6.521205 1.9626275 1022
13 -3.322691 0.0009
14 genefcgapdh -5.734743 1.9626275 1022
15 -2.921972 0.0036
16 genefchsp40 -9.208631 1.9626275 1022
17 -4.691991 0.0000
```

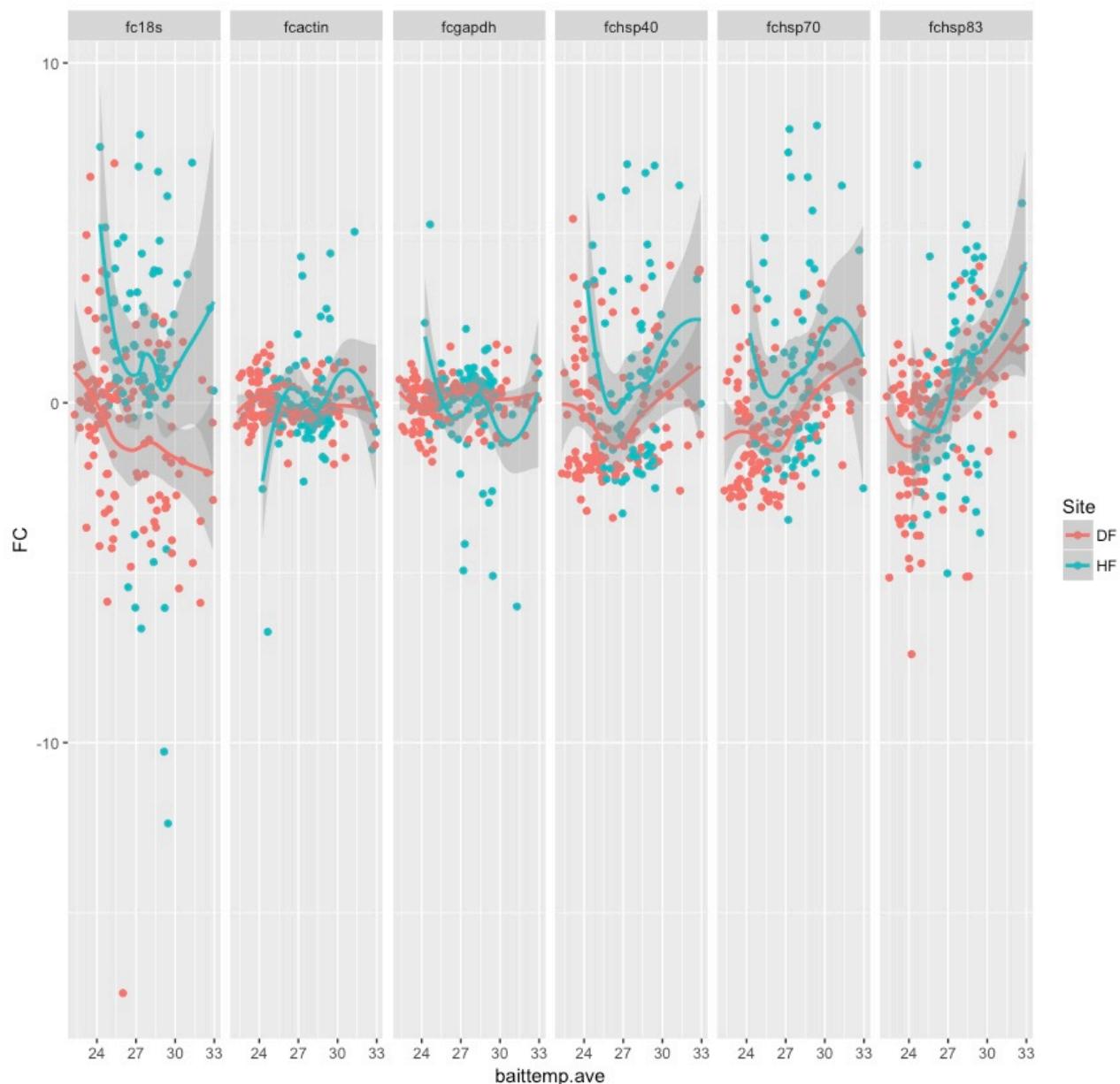
9	genefchsp70	-12.810381	1.9626275	1022
	-6.527159	0.0000		
10	genefchsp83	-16.817619	1.9626275	1022
	-8.568931	0.0000		
11	SiteHF	-1.669959	2.4053577	23
	-0.694266	0.4945		
12	baittemp.ave	-0.289514	0.0600672	1022
	-4.819841	0.0000		
13	Delta	0.033287	0.0527495	23
	0.631046	0.5342		
14	genefcactin:SiteHF	-2.324439	0.3884348	1022
	-5.984117	0.0000		
15	genefcgapdh:SiteHF	-2.495652	0.3884348	1022
	-6.424894	0.0000		
16	genefchsp40:SiteHF	-1.519840	0.3884348	1022
	-3.912729	0.0001		
17	genefchsp70:SiteHF	-1.215410	0.3884348	1022
	-3.128994	0.0018		
18	genefchsp83:SiteHF	-1.964134	0.3884348	1022
	-5.056534	0.0000		
19	genefcactin:baittemp.ave	0.275771	0.0743151	1022
	3.710831	0.0002		
20	genefcgapdh:baittemp.ave	0.248739	0.0743151	1022
	3.347080	0.0008		
21	genefchsp40:baittemp.ave	0.365116	0.0743151	1022
	4.913071	0.0000		
22	genefchsp70:baittemp.ave	0.495528	0.0743151	1022
	6.667932	0.0000		
23	genefchsp83:baittemp.ave	0.655617	0.0743151	1022
	8.822116	0.0000		
24	SiteHF:baittemp.ave	0.163719	0.0907195	169
	1.804667	0.0729		
25	SiteHF:Delta	-0.191388	0.0970456	23
	-1.972147	0.0607		
26				

It looks like there is an overall site effect such that HF has lower expression in all 6 genes than DF. Each gene has a sig slope/relationship with bait temp, but they differ in magnitudie. This relationship does not differ sig between sites.

Some plots

Plotting FC vs bait temp

Loess fits for fold change vs bait temp for all genes colored by sites

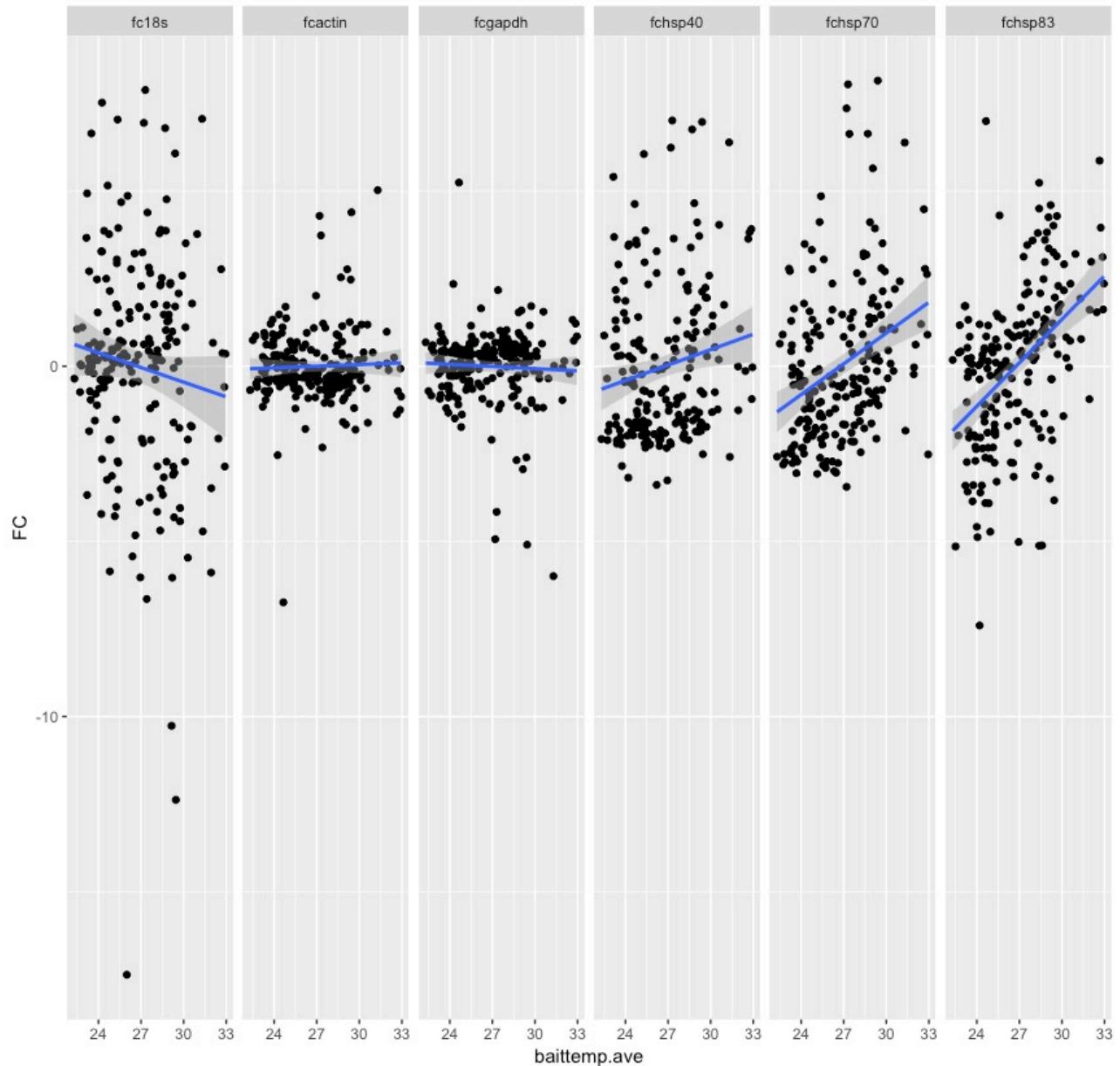


Linear fits for fold change vs bait temp for all genes colored by sites



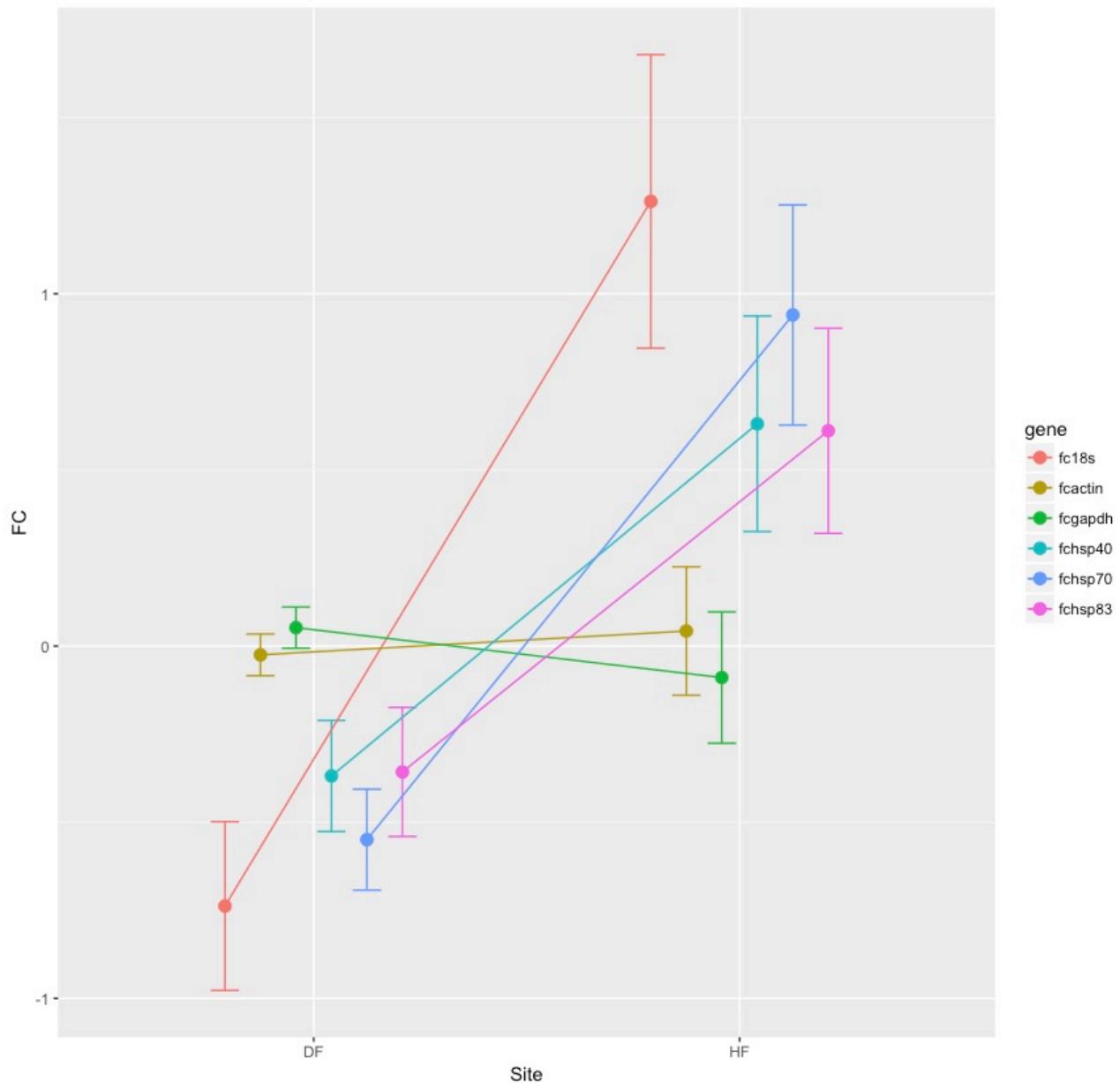
The site level differences are not sig....

Linear fits for fold change vs bait temp for all genes



Plotting FC vs gene for each site

```
1 findat.long$gene<-as.factor(findat.long$gene)
2 gg<-summarySE(findat.long,measurevar
= "FC",groupvars=c("Site","gene"))
3 ggplot(gg,aes(x=Site,y=FC,colour=site,group=site ))+geom_
errorbar(aes(ymin=FC-
se,ymax=FC+se),width=.4,position=position_dodge(.5))+geo
m_point(size=3,position=position_dodge(.5))+geom_line(po
sition=position_dodge(.5))
```



Redoing stats with more precise efficiencies this time

data parsing:

```
1 gxp1<-
  data.frame(log2(apply(deltaCTnum[,1:2], 2, function(x)
  {2^(x)/1.93^(deltaCTden)}))) # 18 eff = 2
```

```

2 gxp2<-
  data.frame(log2(apply(deltaCTnum[,2:4],2,function(x)
  {1.97^(x)/1.93^(deltaCTden)})))) # hsp effic = 1.97
3 gxp3<-
  data.frame(log2(apply(deltaCTnum[,5:6],2,function(x)
  {1.93^(x)/1.93^(deltaCTden)})))) # gapdh and actin eff =
  1.93
4 gxp<-data.frame(CT_18s=gxp1[,1],gxp2,gxp3)
5 names(gxp)<-
  c("fc18s","fchsp40","fchsp70","fchsp83","fcactin","fcga
  pdh")
6
7 head(gxp)
8      fc18s      fchsp40      fchsp70      fchsp83
9      fcactin      fcgapdh
9 1 -3.35598720  0.08869005 -0.196357 -5.10857584
  -0.2974979  0.33337142
10 2 -4.22295350 -1.71440629 -1.573938 -3.08692726
  -0.3648462  0.39693007
11 3  1.15712657  3.13068058  2.927802  0.79659331
  -0.5778360  0.60227319
12 4 -4.22385927  0.17671276  1.997721  1.55194331
  1.0106763 -0.98797586
13 5  0.04264102  3.54676875  2.529053  3.67969085
  -0.6733599  0.69267514
14 6  4.67288683  0.14506857  1.144392  0.03931094
  0.1147508 -0.06922802

```

same model construction

```

1 fullmod5<-
2   lme(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+gene*Site*Delta,
3   random=~1|Cham2/Vial.me,data=findat.long,method="ML")
4 anova(summary(stepAIC(fullmod5,direction="both")))

```

model output

```

1 Step: AIC=5133.09
2 FC ~ gene + Site + baittemp.ave + Delta + gene:Site +
  gene:baittemp.ave +
  Site:baittemp.ave + Site:Delta
4
5           Df      AIC
6 <none>          5133.1
7 + Jdaycont       1 5134.4
8 + RIN_Value      1 5134.8
9 - Site:baittemp.ave  1 5135.1
10 - Site:Delta     1 5135.4
11 + gene:Site:baittemp.ave 5 5135.8
12 + gene:Delta     5 5142.5
13 - gene:Site      5 5178.8
14 - gene:baittemp.ave 5 5213.8
15           numDF denDF   F-value p-value
16 (Intercept)      1 1023  0.077408  0.7809
17 gene            5 1023  0.000000  1.0000
18 Site             1    23 27.835461 <.0001
19 baittemp.ave     1 1023  4.074081  0.0438
20 Delta            1    23  0.115459  0.7371
21 gene:Site        5 1023  9.620273 <.0001
22 gene:baittemp.ave 5 1023 18.634274 <.0001
23 Site:baittemp.ave 1   169  1.592226  0.2087
24 Site:Delta        1    23  4.344712  0.0484

```

Ok lets look at parameter estimates

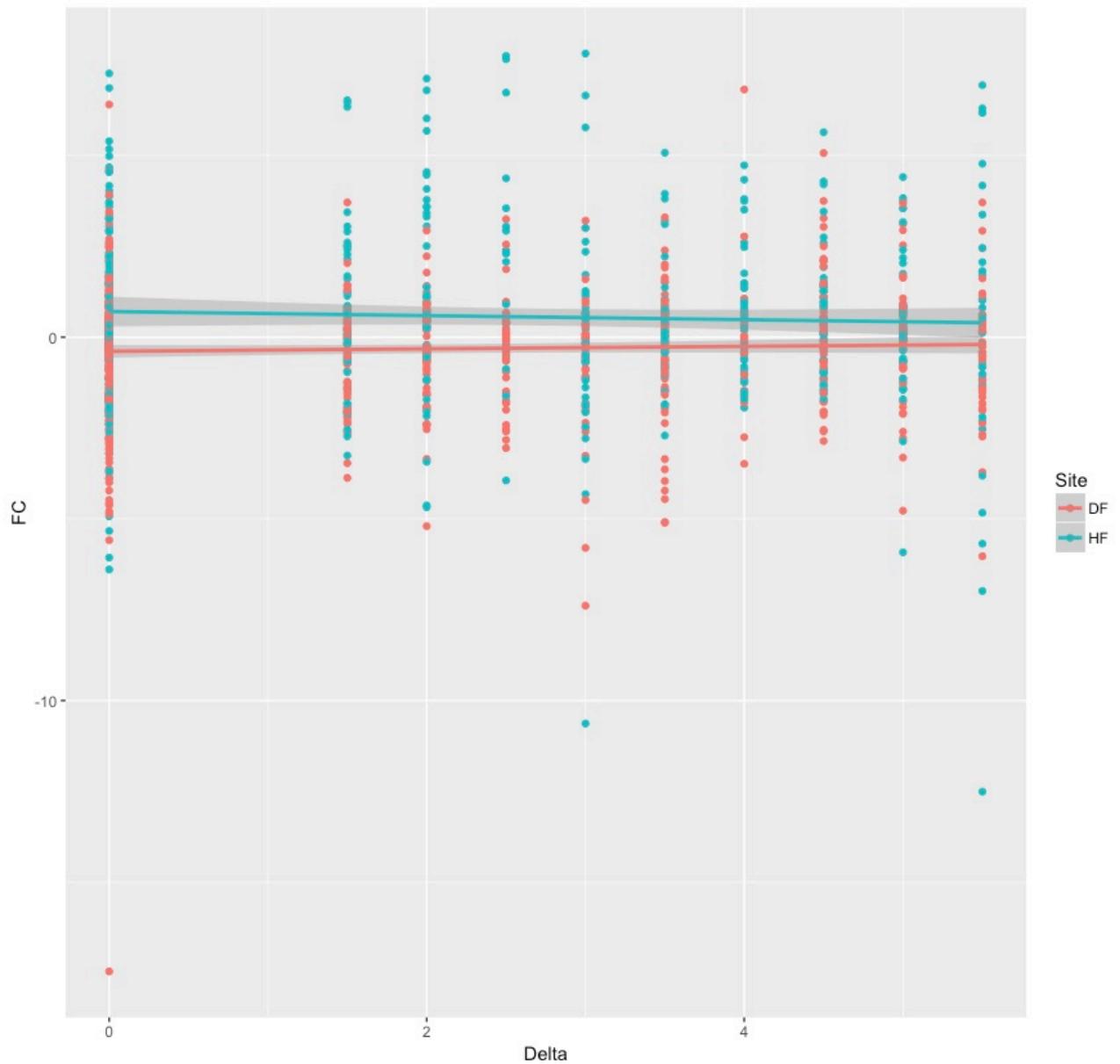
1	Fixed effects: FC ~ gene + Site + baittemp.ave + Delta + gene:Site + gene:baittemp.ave + Site:baittemp.ave + Site:Delta	Value	Std.Error	DF	t-
2	value	p-value			
3	(Intercept)	6.958985	1.5282717	1023	
	4.553500	0.0000			
4	genefcactin	-6.449641	1.9157589	1023	
	-3.366624	0.0008			
5	genefcgapdh	-5.703602	1.9157589	1023	
	-2.977202	0.0030			
6	genefchsp40	-9.119674	1.9157589	1023	
	-4.760346	0.0000			
7	genefchsp70	-12.642891	1.9157589	1023	
	-6.599417	0.0000			
8	genefchsp83	-16.562754	1.9157589	1023	
	-8.645531	0.0000			
9	SiteHF	-1.966949	2.3055972	23	
	-0.853119	0.4024			
10	baittemp.ave	-0.294889	0.0580370	1023	
	-5.081056	0.0000			
11	Delta	0.031595	0.0512302	23	
	0.616729	0.5435			
12	genefcactin:SiteHF	-2.292505	0.3791587	1023	
	-6.046293	0.0000			
13	genefcgapdh:SiteHF	-2.454918	0.3791587	1023	
	-6.474644	0.0000			
14	genefchsp40:SiteHF	-1.523837	0.3791587	1023	
	-4.018994	0.0001			
15	genefchsp70:SiteHF	-1.226045	0.3791587	1023	
	-3.233592	0.0013			
16	genefchsp83:SiteHF	-1.958443	0.3791587	1023	
	-5.165232	0.0000			

```

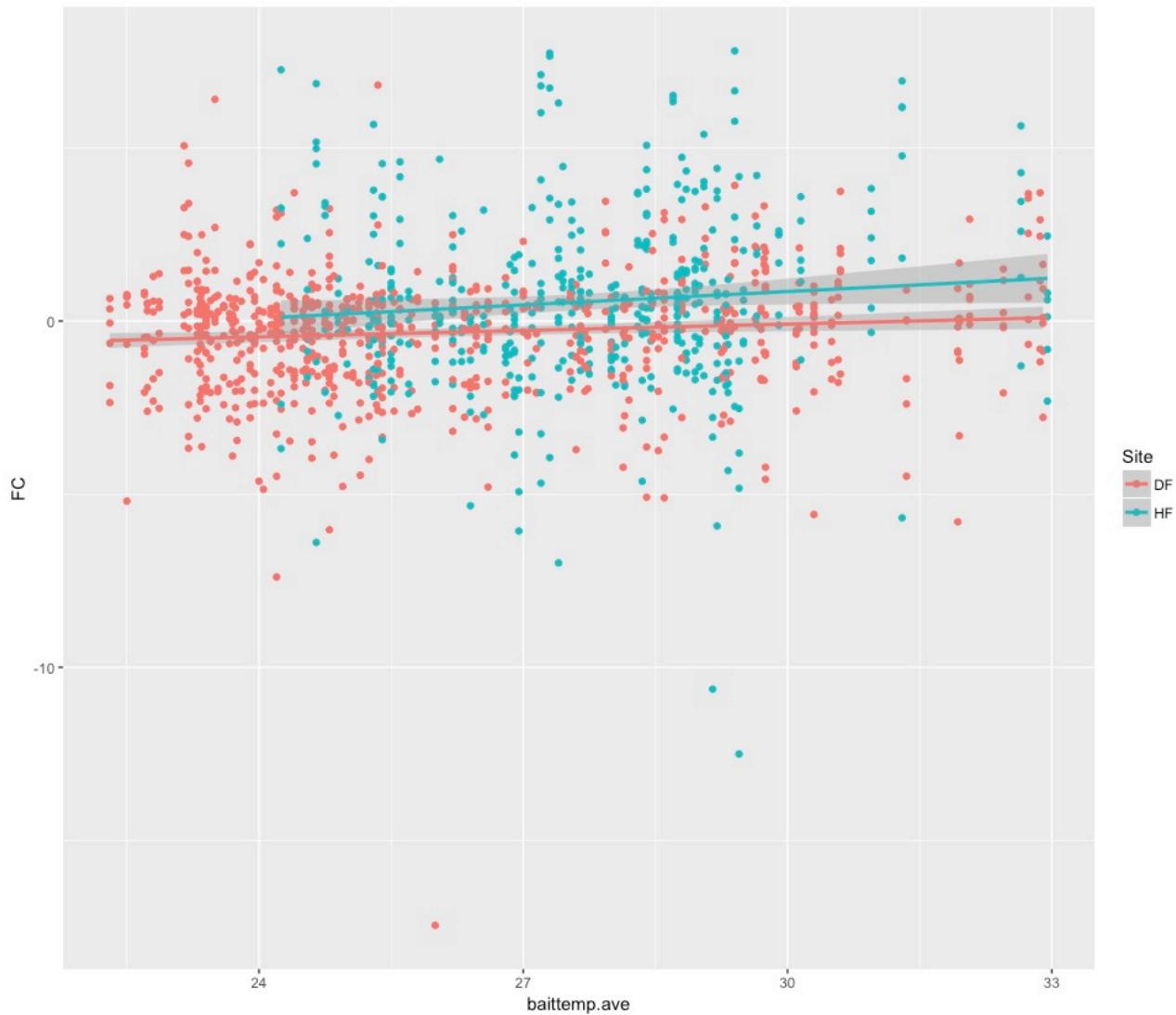
17  genefcactin:baittemp.ave    0.272656 0.0725404 1023
      3.758675 0.0002
18  genefcgapdh:baittemp.ave    0.247013 0.0725404 1023
      3.405179 0.0007
19  genefchsp40:baittemp.ave    0.361846 0.0725404 1023
      4.988197 0.0000
20  genefchsp70:baittemp.ave    0.489415 0.0725404 1023
      6.746790 0.0000
21  genefchsp83:baittemp.ave    0.646013 0.0725404 1023
      8.905556 0.0000
22  SiteHF:baittemp.ave        0.175617 0.0868661 169
      2.021694 0.0448
23  SiteHF:Delta                -0.195618 0.0938489 23
      -2.084397 0.0484
24  summary(stepAIC(fullmod5,direction="both"))

```

delta * site interaction significant this time.

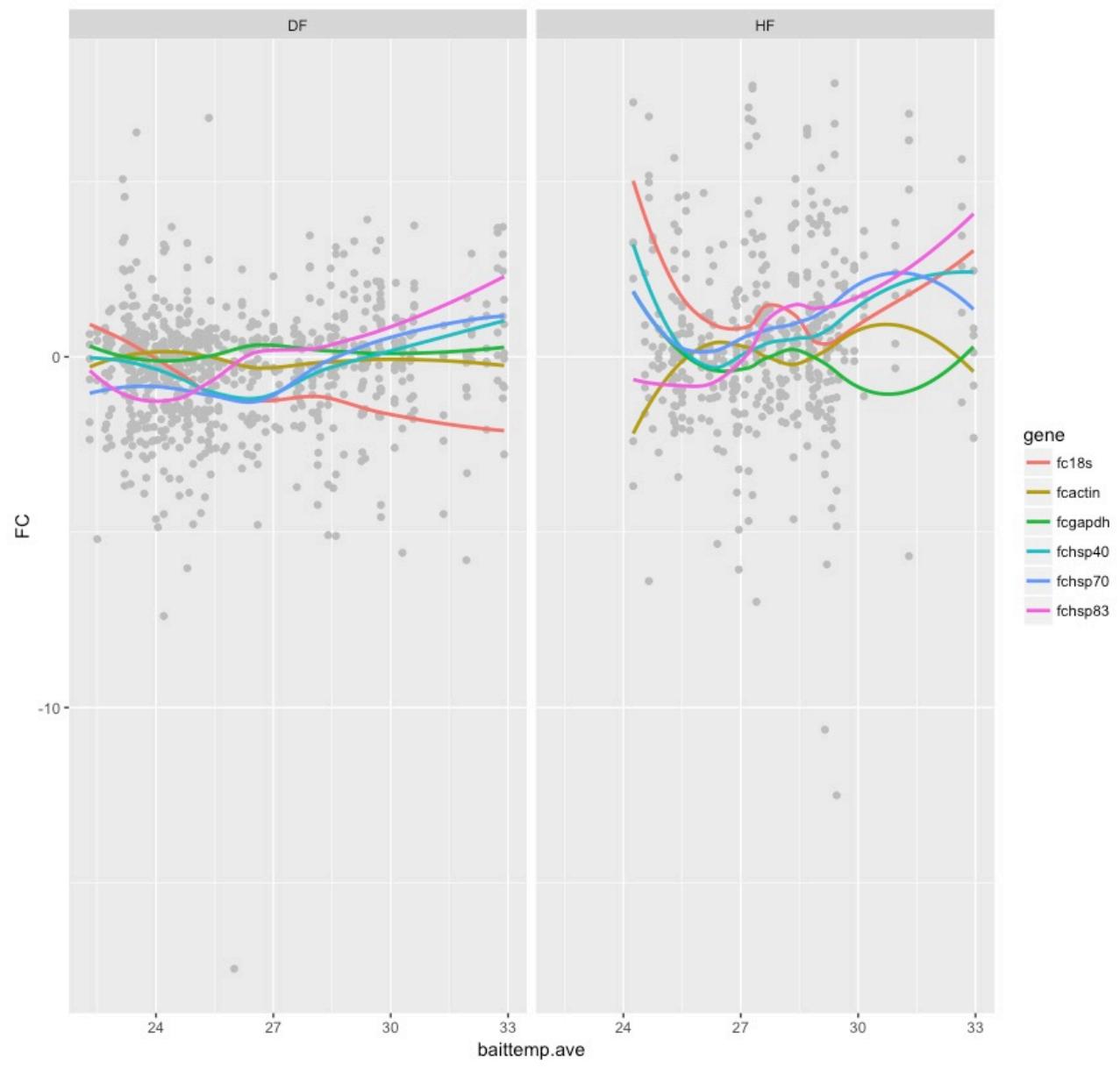


bait temp * site interaction sig

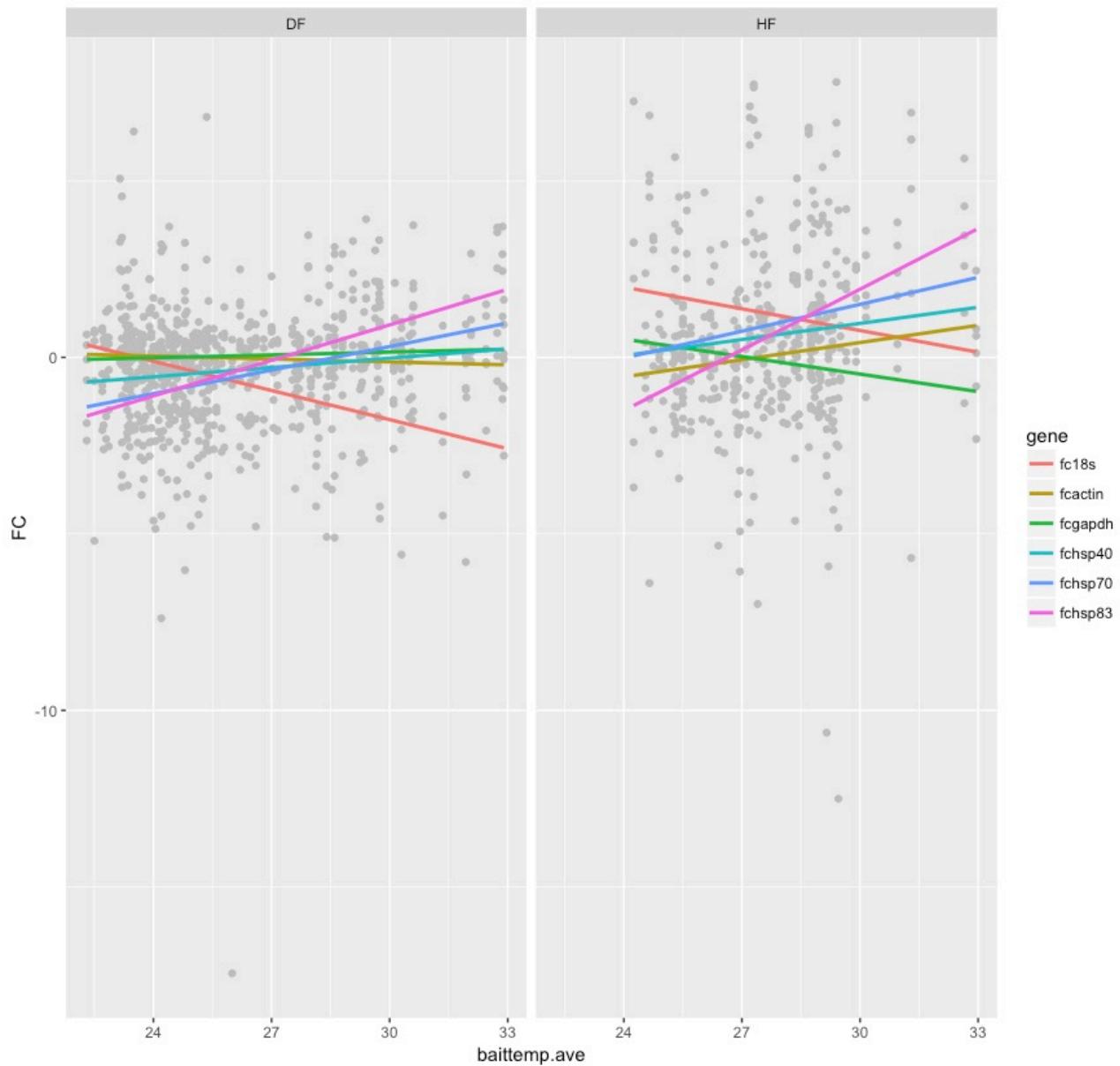


Plots: FC vs bait, 2 panel sites

loess



linear



Meeting notes with SHC, 1pm

- do analysis for each gene, take out hkg
 - hkg is good first start, but take them out, get slapped!
- Consider taking residuals from regression model including julian day and rin values
- Need info on summer temperatures

- mean, min, max monthly temperatures in growing season.
- get experienced temperature from chambers (average across 5 years)

Model without HKGs

```

1 findat.long<-gather(findat,gene,FC,fchsp40:fchsp83)
#taking out control genes
2 fullmod5<-
lme(FC~RIN_Value+Jdaycont+gene*Site*baittemp.ave+gene*S
ite*Delta,random=~1|Cham2/Vial.me,data=findat.long)
3
4 anova(fullmod5)
5
6 (Intercept)           numDF  denDF   F-value p-value
7 RIN_Value             1       406   0.069148 0.7927
8 Jdaycont              1       406   0.412211 0.5212
9 gene                   2       406   0.000000 1.0000
10 Site                  1       23    14.303692 0.0010
11 baittemp.ave           1       406   27.960798 <.0001
12 Delta                 1       23    0.227509 0.6379
13 gene:Site              2       406   1.682853 0.1871
14 gene:baittemp.ave      2       406   10.552477 <.0001
15 Site:baittemp.ave     1       169   1.437611 0.2322
16 gene:Delta              2       406   0.017785 0.9824
17 Site:Delta              1       23    2.201969 0.1514
18 gene:Site:baittemp.ave  2       406   1.054346 0.3494
19 gene:Site:Delta         2       406   0.301835 0.7396

```

model selection based on AIC

```

1 Formula: ~1 | Vial.me %in% Cham2
2             (Intercept) Residual
3 StdDev:      1.246812 1.510167
4
5 Fixed effects: FC ~ Jdaycont + gene + Site +
6                                baittemp.ave + Delta + gene:Site +
7                                gene:baittemp.ave + Site:baittemp.ave + Site:Delta
8                                         Value Std.Error DF   t-
9                                         value p-value
10 (Intercept)           -1.642948 1.549218 413
11 -1.060502 0.2895
12 Jdaycont            -0.001323 0.000605 413
13 -2.187980 0.0292
14 gene:fchsp70         -3.523216 1.636554 413
15 -2.152826 0.0319
16 gene:fchsp83         -7.443079 1.636554 413
17 -4.548019 0.0000
18 SiteHF               -5.118212 3.216726 23
19 -1.591124 0.1252
20 baittemp.ave          0.067622 0.059515 413
21 1.136224 0.2565
22 Delta                 0.020335 0.070936 23
23 0.286661 0.7769
24 gene:fchsp70:SiteHF  0.297792 0.323900 413
25 0.919396 0.3584
26 gene:fchsp83:SiteHF -0.434606 0.323900 413
27 -1.341793 0.1804
28 gene:fchsp70:baittemp.ave 0.127569 0.061968 413
29 2.058619 0.0402
30 gene:fchsp83:baittemp.ave 0.284167 0.061968 413
31 4.585681 0.0000
32 SiteHF:baittemp.ave  0.233840 0.121765 169
33 1.920421 0.0565

```

```

19 SiteHF:Delta           -0.213972  0.130397  23
20 -1.640922  0.1144
21
22
23 anova(summary(stepAIC(fullmod5,direction="both")))
24
25 Step: AIC=2519.29
26 FC ~ Jdaycont + gene + Site + baittemp.ave + Delta +
27 gene:Site +
28     gene:baittemp.ave + Site:baittemp.ave + Site:Delta
29
30             Df      AIC
31 <none>          2519.3
32 - Site:Delta      1 2520.0
33 - gene:Site       2 2520.5
34 + RIN_Value      1 2520.8
35 - Site:baittemp.ave 1 2521.0
36 + gene:Site:baittemp.ave 2 2521.1
37 - Jdaycont       1 2522.1
38 + gene:Delta      2 2523.2
39 - gene:baittemp.ave 2 2536.3
40
41             numDF denDF   F-value p-value
42 (Intercept)      1    413  0.044025  0.8339
43 Jdaycont         1    413  4.633790  0.0319
44 gene              2    413  0.000000  1.0000
45 Site              1    23   21.076865  0.0001
46 baittemp.ave      1    413  29.836209  <.0001
47 Delta             1    23   0.169841  0.6841
48 gene:Site         2    413  1.682588  0.1872
49 gene:baittemp.ave 2    413  10.550810 <.0001
50 Site:baittemp.ave 1    169   1.819052  0.1792
51 Site:Delta         1    23   2.692626  0.1144

```

Statistics: linear mixed effects models for each gene

hsp70

```
1 hsp70<-subset(findat.long,findat.long$gene=="fchsp70")
2 fullmod6<-
3   lme(FC~RIN_Value+Jdaycont+Site*baittemp.ave+Site*Delta,
4     random=~1|Cham2/Vial.me,data=hsp70,method="ML")
5 summary(stepAIC(fullmod6,direction="both"))
6
7 Formula: ~1 | Vial.me %in% Cham2
8           (Intercept) Residual
9 StdDev: 0.0002263996 1.677996
10
11 Fixed effects: FC ~ Jdaycont + Site + baittemp.ave
12             Value Std.Error DF t-value p-value
13 (Intercept) -6.364371 1.3026365 170 -4.885761 0e+00
14 Jdaycont    -0.004589 0.0006572    7 -6.981991 2e-04
15 SiteHF      0.788088 0.2736238   25  2.880187 8e-03
16 baittemp.ave 0.289517 0.0499854    7  5.792034 7e-04
17 Correlation:
18             (Intr) Jdycnt SiteHF
19 Jdaycont    -0.025
20 SiteHF      0.178  0.173
21 baittemp.ave -0.974 -0.168 -0.281
22
23 Standardized Within-Group Residuals:
24             Min          Q1          Med          Q3          Max
25 -2.1274266 -0.6098562 -0.1545382  0.3930239  3.7842314
26
27 Number of Observations: 206
28 Number of Groups:
29           Cham2 Vial.me %in% Cham2
```

hsp83

```

1 hsp83<-subset(findat.long,findat.long$gene=="fchsp83")
2 fullmod7<-
3 lme(FC~RIN_Value+Jdaycont+Site*baittemp.ave+Site*Delta,
4 random=~1|Cham2/Vial.me,data=hsp83,method="ML")
5 #summary(fullmod7)
6 summary(stepAIC(fullmod7,direction="both"))

7 Fixed effects: FC ~ RIN_Value + Jdaycont + Site *
   baittemp.ave + Site * Delta
8
9             Value Std.Error DF t-value
10            p-value
11 (Intercept) -9.915097 1.528096 169 -6.488532
12            0.0000
13 RIN_Value      0.149878 0.082116   6  1.825203
14            0.1178
15 Jdaycont       0.002731 0.001028   6  2.655774
16            0.0377
17 SiteHF        -7.726510 4.135596  23 -1.868294
18            0.0745
19 baittemp.ave    0.295245 0.058582   6  5.039819
20            0.0024
21 Delta          0.008764 0.091198  23  0.096100
22            0.9243
23 SiteHF:baittemp.ave 0.329418 0.157427 169  2.092515
24            0.0379
25 SiteHF:Delta    -0.292487 0.166271  23 -1.759101
26            0.0919
27 Correlation:
28
29             (Intr) RIN_Vl Jdycnt SiteHF bttmp.
30 Delta     StHF:.
```

```

19 RIN_Value           -0.147
20 Jdaycont            0.089 -0.691
21 SiteHF              -0.404  0.291 -0.225
22 baittemp.ave         -0.964  0.057 -0.164  0.377
23 Delta                0.017 -0.070  0.040 -0.022 -0.129
24 SiteHF:baittemp.ave  0.401 -0.305  0.246 -0.993 -0.386
25                         0.068
25 SiteHF:Delta          -0.019  0.111 -0.099  0.343  0.081
25                         -0.553 -0.427
26
27 Standardized Within-Group Residuals:
28             Min          Q1          Med          Q3
28             Max
29 -1.90917202 -0.24335576  0.07299595  0.29028657
29                         1.92310971
30
31 Number of Observations: 206
32 Number of Groups:
33             Cham2 Vial.me %in% Cham2
34                         27                      197

```

hsp40

```
1 hsp40<-subset(findat.long,findat.long$gene=="fchsp40")
2 fullmod8<-
lme(FC~RIN_Value+Jdaycont+Site*baittemp.ave+Site*Delta,
random=~1|Cham2/Vial.me,data=hsp40,method="ML")
3
4 summary(stepAIC(fullmod8,direction="both"))
5
```

```

6 Fixed effects: FC ~ Jdaycont + Site + baittemp.ave
7
8 (Intercept) -2.8557557 1.4687690 170 -1.944319 0.0535
9 Jdaycont -0.0032511 0.0007497 7 -4.336718 0.0034
10 SiteHF 0.5809013 0.2956002 25 1.965159 0.0606
11 baittemp.ave 0.1430383 0.0564215 7 2.535174 0.0389
12 Correlation:
13 (Intr) Jdycnt SiteHF
14 Jdaycont -0.025
15 SiteHF 0.195 0.183
16 baittemp.ave -0.974 -0.170 -0.295
17
18 Standardized Within-Group Residuals:
19 Min Q1 Med Q3 Max
20 -2.1019055 -0.6332219 -0.2321619 0.4740407 3.0532535
21
22 Number of Observations: 206
23 Number of Groups:
24 Cham2 Vial.me %in% Cham2
25 27 197

```

Page 79: 2017-05-09. Proteome stability project: addressing notes from 2017-04-26 climate cascade meeting

Notes from SHC and NGotelli:

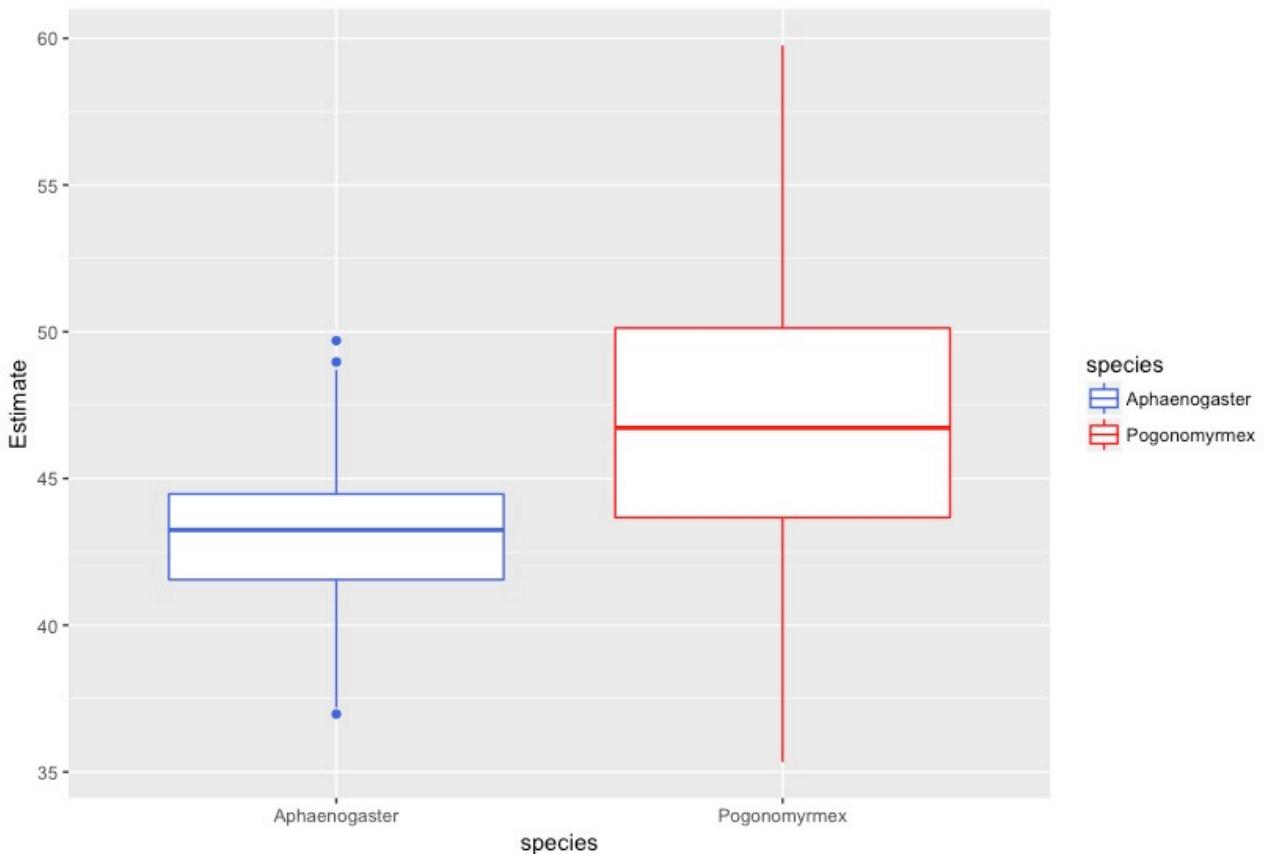
1. Do a randomization test:
 - Shuffle the labels-- go to the colony level
 - Compare different medians

For dataset with unique and overlapping peptides between aphaeno and pogo

```
1 ddply(combined,.  
  (species),summarize,median=median(Estimate))  
  species    median  
1 Aphaenogaster 43.24474  
2 Pogonomyrmex 46.72500
```

Accompanying plot

```
1 Tmbox<-ggplot(combined,  
  aes(x=species,y=Estimate,colour=species))+geom_boxplot()  
+ggcol  
2 Tmbox
```



2. Do KS test to determine difference in cumulative density function (Kolmogorov-Smirnov test)

- What is KS test? Null hypothesis is 2 vectors , x and y are drawn from the same continuous distribution. Alternative is that they are not.

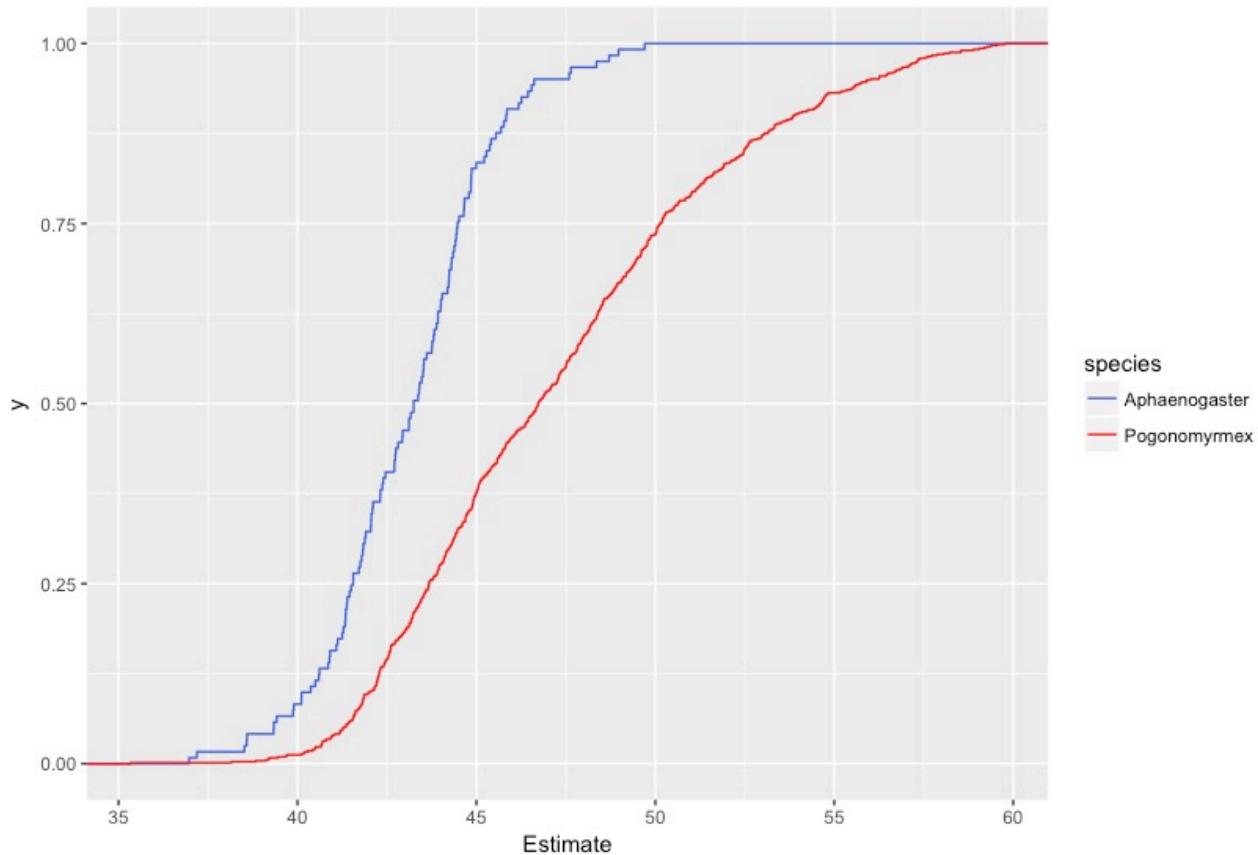
For dataset with unique and overlapping peptides between aphaeno and pogo

```

1 ks.test(Tmaph.ave$Estimate,Tmpog.ave$Estimate)
2
3   Two-sample Kolmogorov-Smirnov test
4
5 data: Tmaph.ave$Estimate andTmpog.ave$Estimate
6 D = 0.47254, p-value < 2.2e-16
7 alternative hypothesis: two-sided

```

associated cumulative prob plot



Shuffling species IDs for 1 case

```
1 spec<-c("Aphaenogaster", "Pogonomyrmex")
2 combined$shuffle<-
3   sample(spec,length(combined$Sequence),replace=TRUE)
4
5 shufaph<-
6   subset(combined,combined$shuffle=="Aphaenogaster")
7 shufpog<-
8   subset(combined,combined$shuffle=="Pogonomyrmex")
9
10 ks.test(shufpog$Estimate,shufaph$Estimate)
11 Two-sample Kolmogorov-Smirnov test
12
13 data: shufpog$Estimate and shufaph$Estimate
14 D = 0.072021, p-value = 0.2211
15 alternative hypothesis: two-sided
```

Shuffling species IDs and replicating it for 1000 cases

Wrote a function: It is a little specific to our data

It basically randomizes names to our dataset and computes the KS test
D statistic

```

1 rando<-function(data=combined){
2   data$shuffle<-
3   sample(spec,length(data$Sequence),replace=TRUE)
4   shufaph<-subset(data,data$shuffle=="Aphaenogaster")
5   shufpog<-subset(data,data$shuffle=="Pogonomyrmex")
6   y=ks.test(shufaph$Estimate,shufpog$Estimate)
7   #print(round(y$p.value,3))
8   print(as.vector(round(y[[1]],3)))
9 }
10 rando()

```

Used the replicate function to repeat rando()

```

1 distks<-replicate(1000,rando())
2 distks<-data.frame(distks)

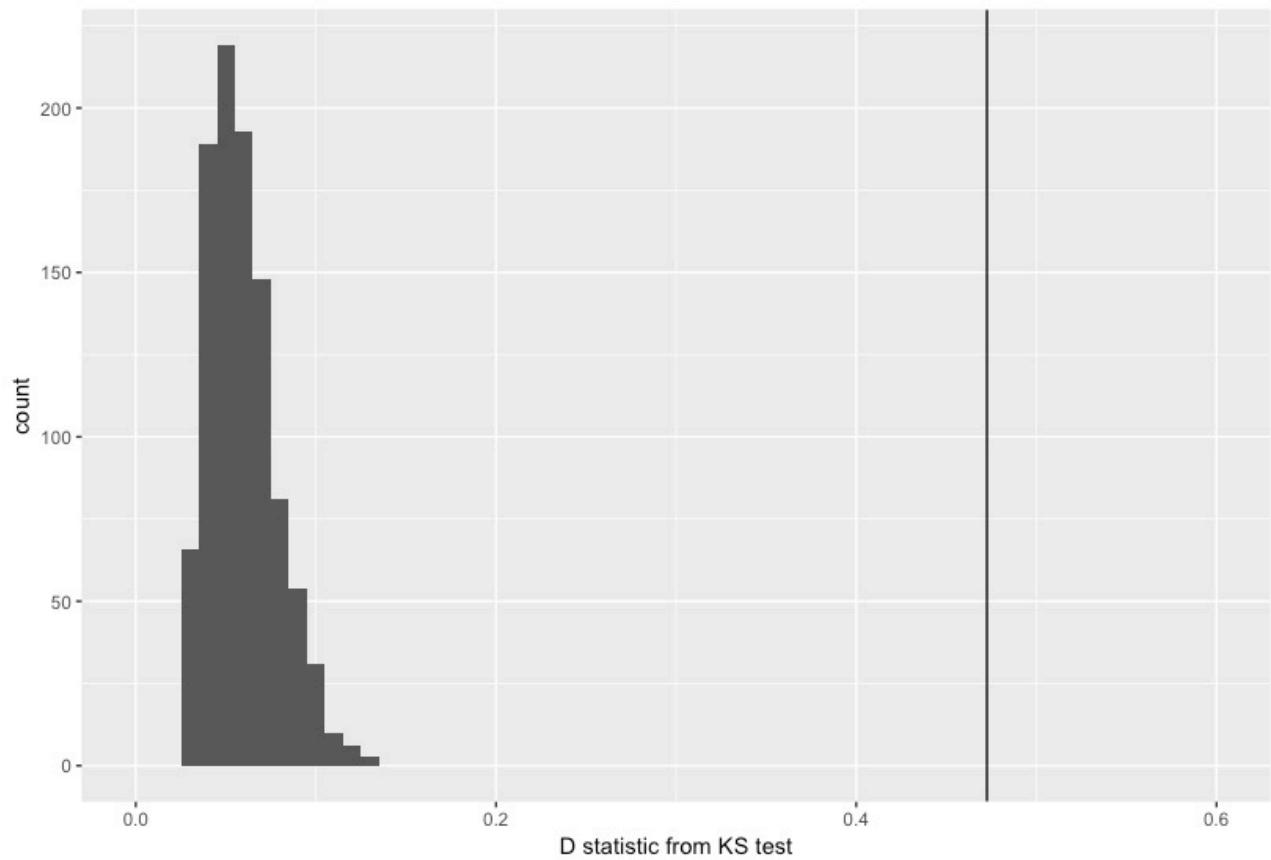
```

Accompanying figure: gray is the randomized distribution of D statistics, the line is the D statistic from our analysis

```

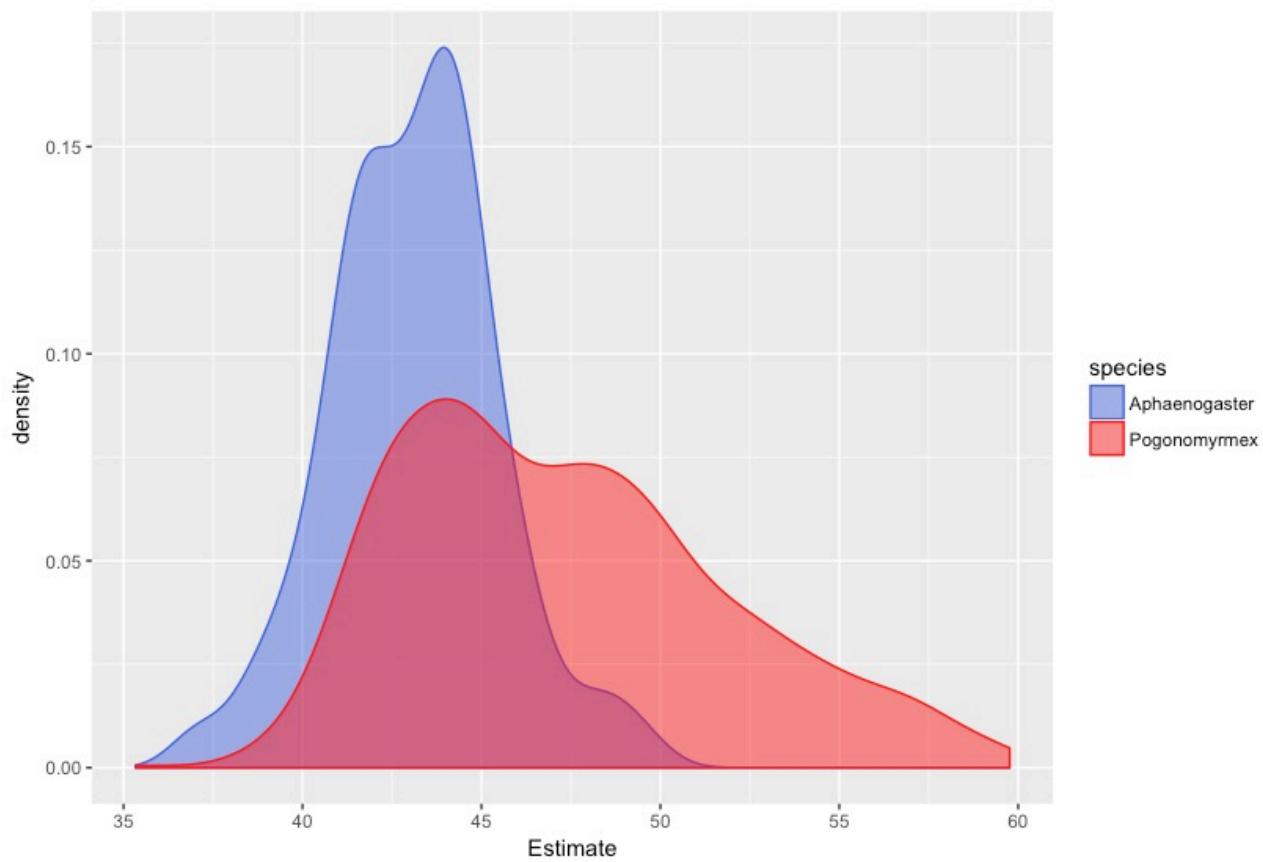
1 ggplot(distks,aes(x=distks))+geom_histogram(binwidth=.01
)+
2 xlim(0,.6)+geom_vline(xintercept=0.47254)+xlab("D
statistic from KS test")
3

```



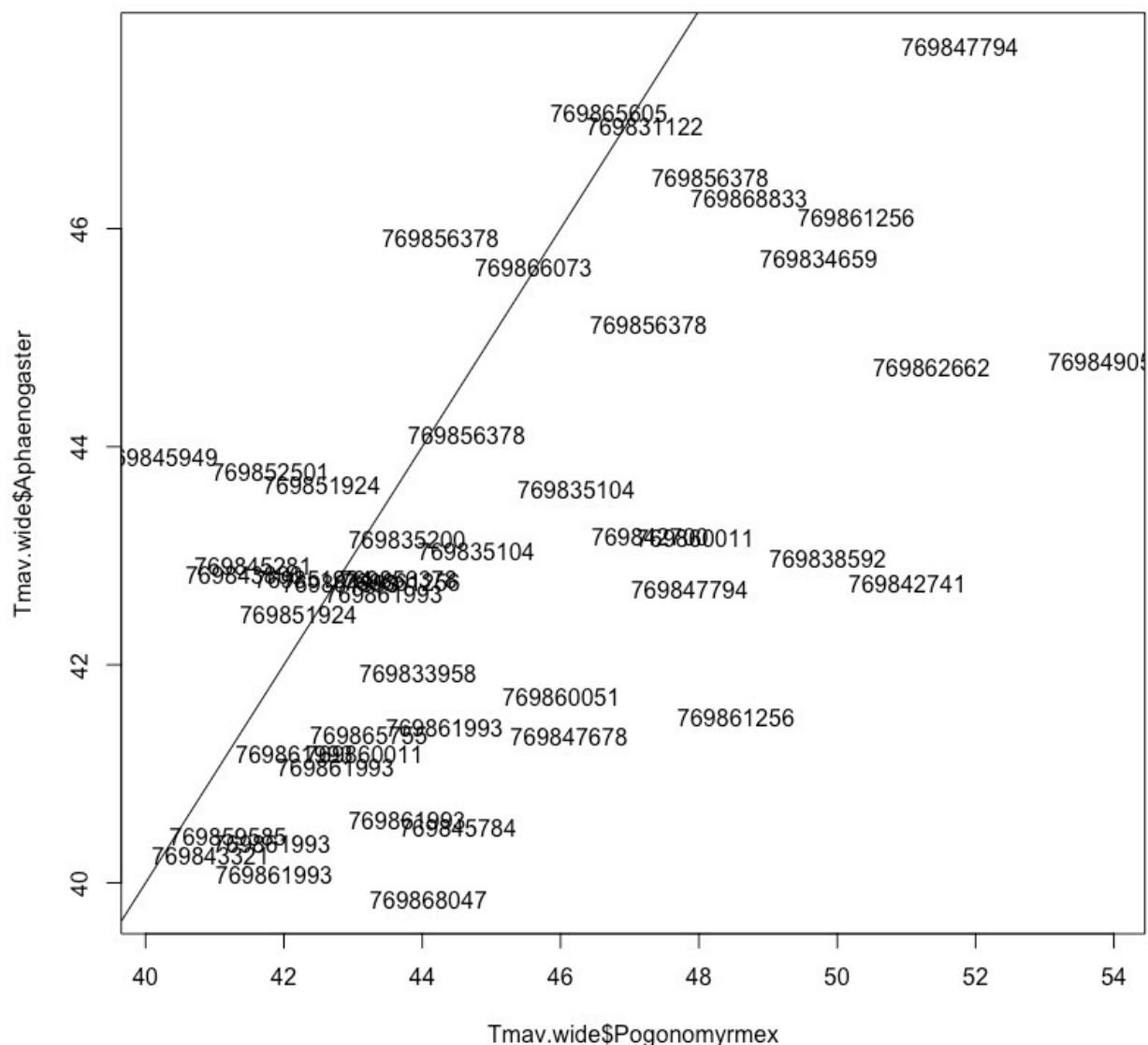
3. Plot density curves

**For dataset with unique and overlapping peptides between
aphaeno and pogo**

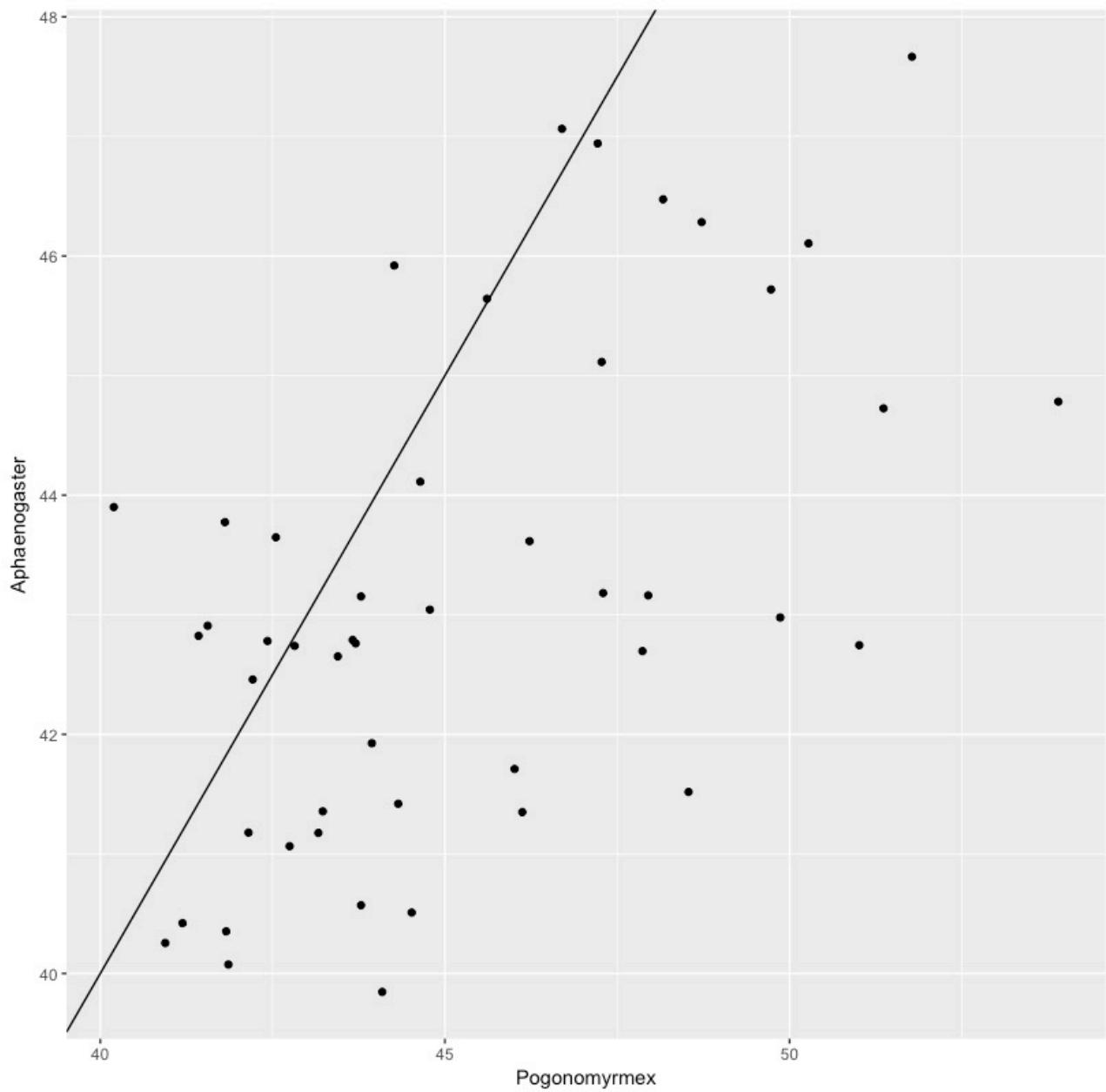


4. Plot Aph Tm vs Pogo Tm; COMMON PEPTIDES!

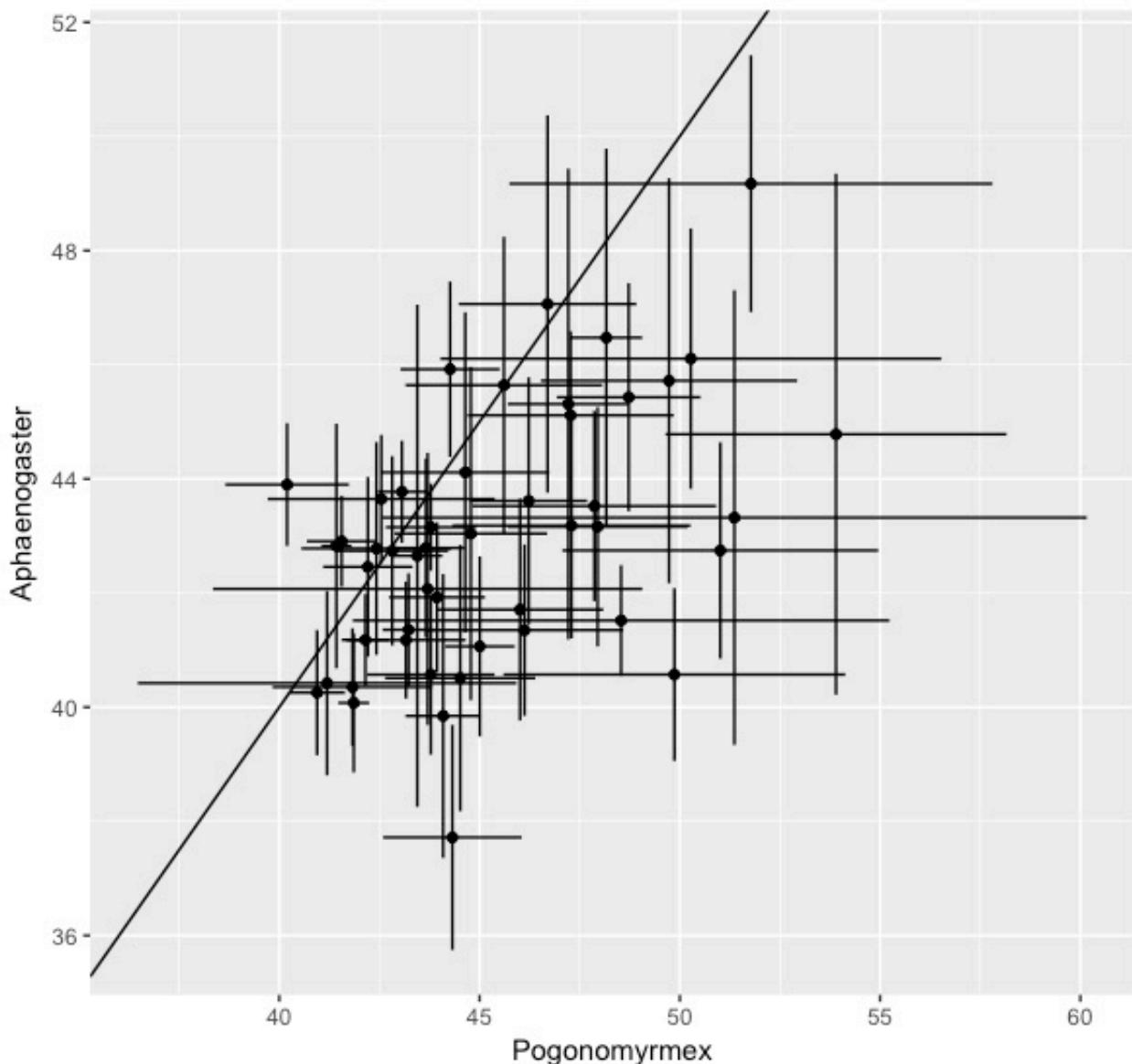
With accession numbers!



With just points

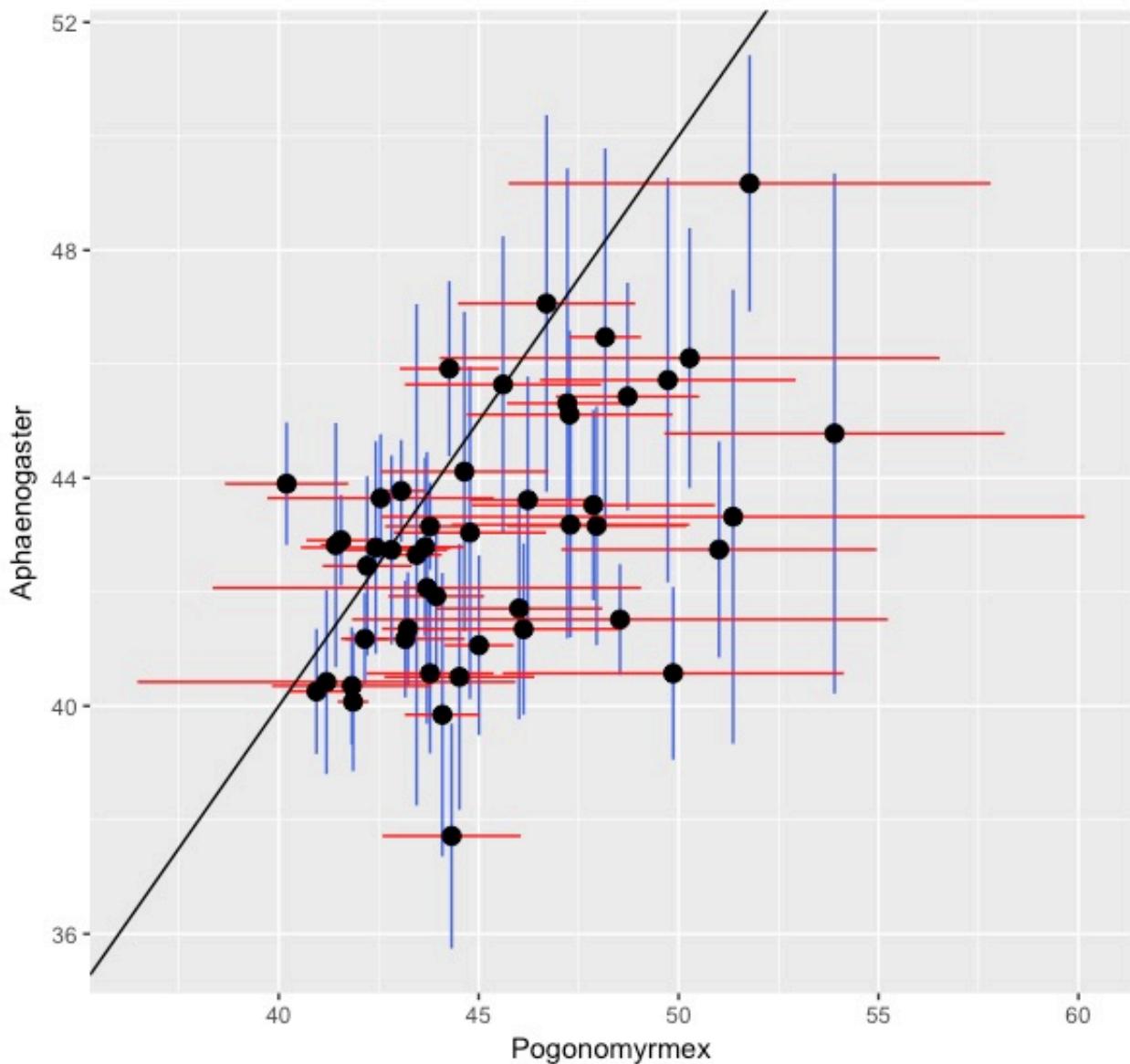


Plot with SEs



Plot with SEs with color for error bars

```
1 ggplot(TT,aes(y=Aphaenogaster,x=Pogonomyrmex)) +geom_errorbar(aes(ymax=Aphaenogaster+Ase,ymin=Aphaenogaster-Ase),colour="royalblue") +geom_errorbarh(aes(xmax=Pogonomyrmex+Pse,xmin=Pogonomyrmex-Pse),colour="red") +geom_point(size=3) +geom_abline(slope=1)
```



Calculating offsets and showing the largest offsets:

```

1 TT$offset<-TT$Pogonomyrmex-TT$Aphaenogaster
2 knitr::kable(subset(TT,TT$offset>8))
3

```

	Protein.Group.Accessions	Sequence	Aphaenogaster	Pogonomyrmex	Ase	Pse	offset	randoff
7	769838592	APFDPPGPPGTPK	40.57204	49.86469	1.514251	4.267933	9.292642	NA
9	769842741	DAGTISGLVVMR	42.74541	51.01181	1.896613	3.944651	8.266397	NA
19	769849050	GASLQDLMISK	44.78217	53.89990	4.565298	4.257023	9.117733	0.0332116
47	769862662	FFDMVEYFFHR	43.32144	51.36331	3.983255	8.796392	8.041870	-4.2309084

Create story board for manuscript ; focus on narrative and then construct figures accordingly

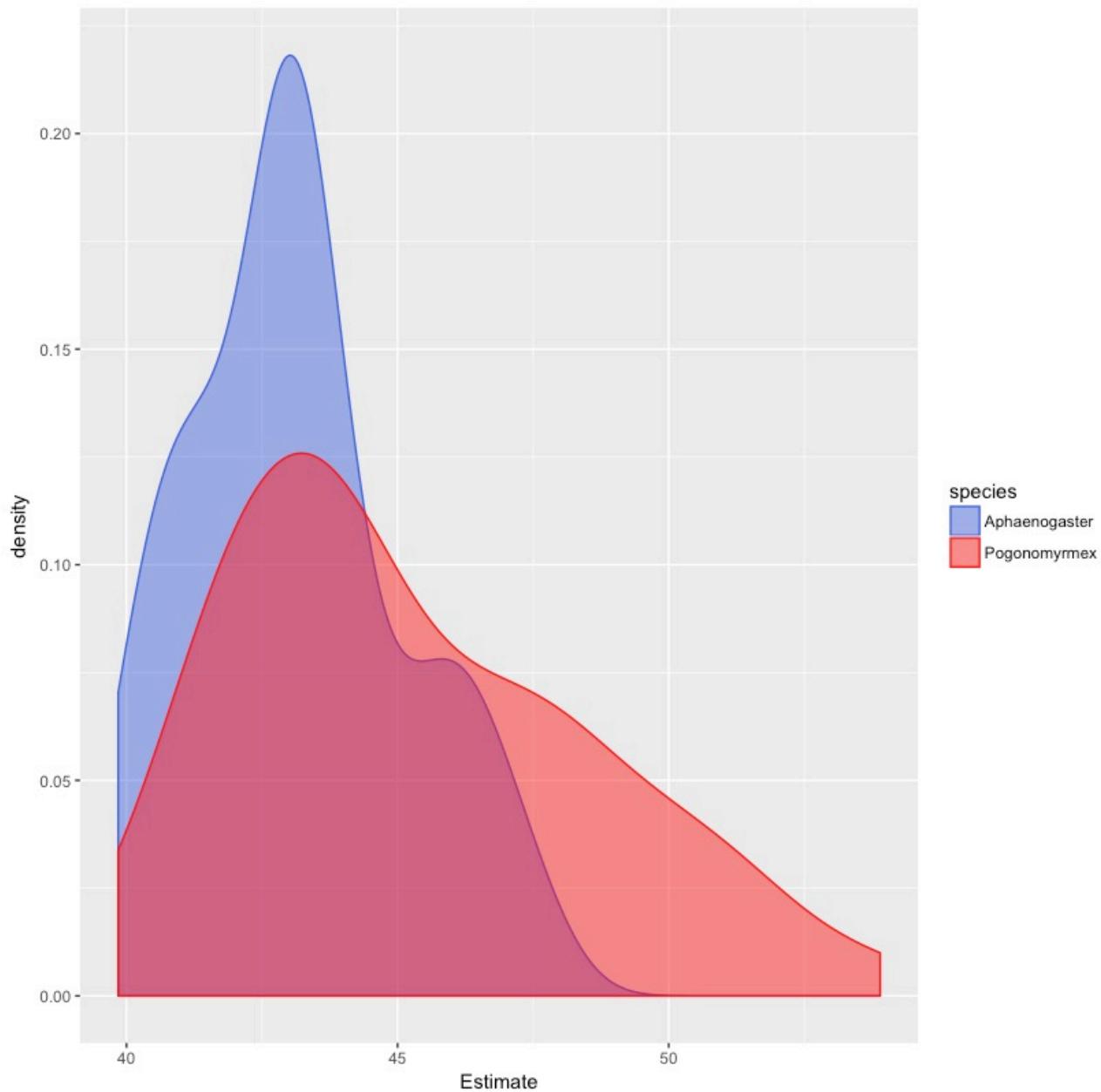
Try to analyze by block which might account for variation intraspecific variation

NGotelli comment: Aph Tms are within the range of pogo Tms

Redoing KS test with common peptides

```
1 suba<-subset(Tmav,Tmav$species=="Aphaenogaster")
2 subp<-subset(Tmav,Tmav$species=="Pogonomyrmex")
3 ks.test(suba$Estimate,subp$Estimate)
4 Two-sample Kolmogorov-Smirnov test
5
6 data: suba$Estimate and subp$Estimate
7 D = 0.31891, p-value = 0.009158
8 alternative hypothesis: two-sided
```

accompanying density plot



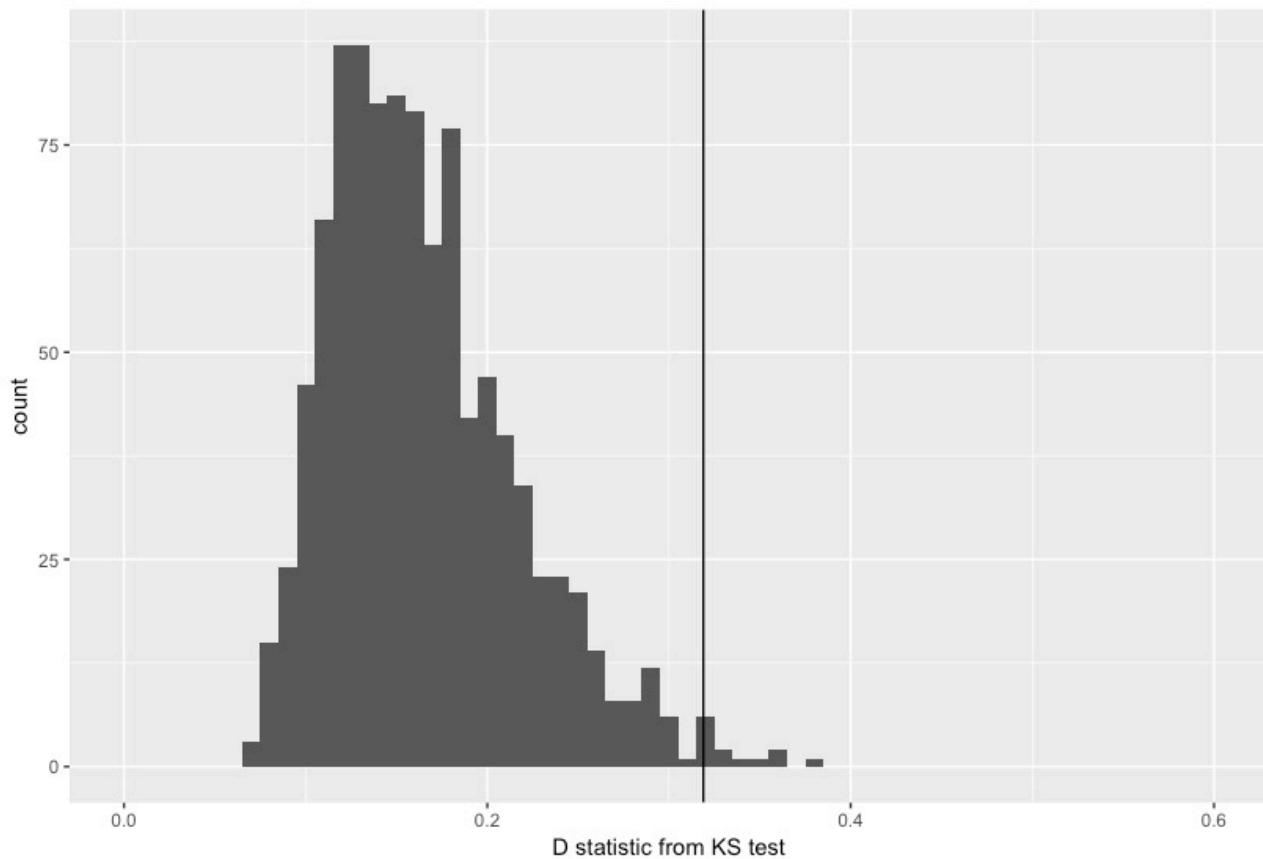
try randomization and t test?

```
1 uniq<-replicate(1000,rando(data=Tmav))
2 hist(uniq)
3 uniq<-data.frame(uniq)
4 # do t test 1 sided?
5 t.test(uniq$uniq, mu=0.31891, alternative="less")
6 One Sample t-test
7
8 data: uniq$uniq
9 t = -93.229, df = 999, p-value < 2.2e-16
```

```
10 alternative hypothesis: true mean is less than 0.31891
11 95 percent confidence interval:
12     -Inf 0.1681551
13 sample estimates:
14 mean of x
15 0.165445
```

Accmpanying Figure

```
1 ggplot(uniq,aes(x=uniq))+geom_histogram(binwidth=.01)+xli
  im(0,.6)+geom_vline(xintercept= 0.31891)+xlab("D
statistic from KS test")
```



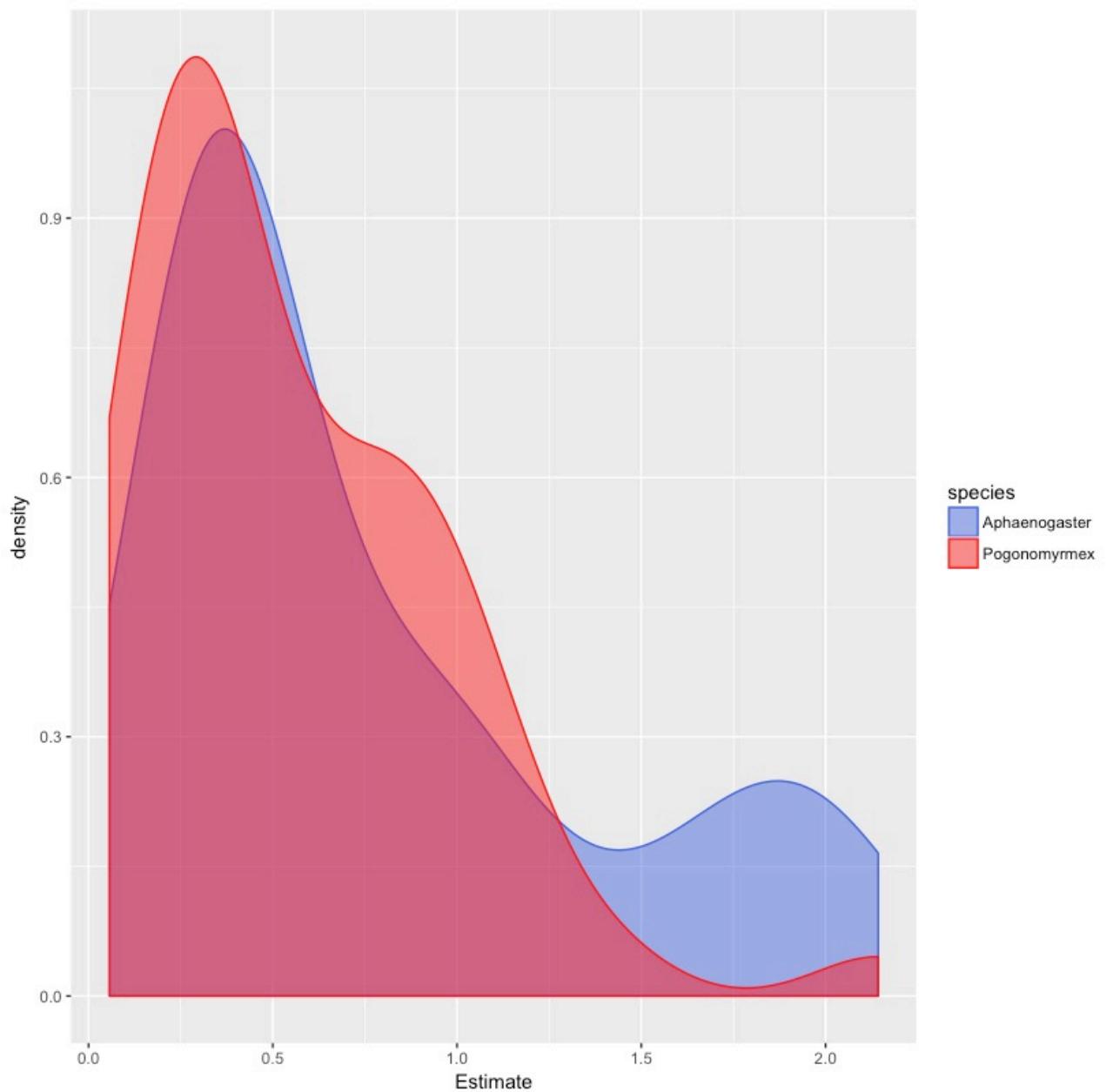
**common peptides, but doign KS test for
slope and min**

slope

```
1 slav.wide<-spread(slav, species,Estimate )#cool way to
  integrate data into ks test by converting from long to
  wide format
2 ks.test(slav.wide$Aphaenogaster,slav.wide$Pogonomyrmex)
3
4   Two-sample Kolmogorov-Smirnov test
5
6 data: slav.wide$Aphaenogaster and
  slav.wide$Pogonomyrmex
7 D = 0.19231, p-value = 0.2935
8 alternative hypothesis: two-sided
9
```

slope density plot

```
1 slplot3<-ggplot(slav, aes(x=Estimate,colour=species,
  fill=species)) +geom_density(alpha=.5,
  position="identity")+ggcol+ggcol2
2 slplot3
```



min

```

1 ks.test(minav.wide$Aphaenogaster,minav.wide$Pogonomyrmex
)
2
3   Two-sample Kolmogorov-Smirnov test
4
5 data: minav.wide$Aphaenogaster and
minav.wide$Pogonomyrmex
6 D = 0.46314, p-value = 2.204e-05
7 alternative hypothesis: two-sided

```

Figure

```

1 minplot3<-ggplot(minav, aes(x=Estimate,
colour=species,fill=species)) +
  geom_density(alpha=.5,
position="identity")+ggcol+ggcol2
3 minplot3

```

Ok, KS tests for common and unique peptides for SLOPE AND MIN

SLOPE

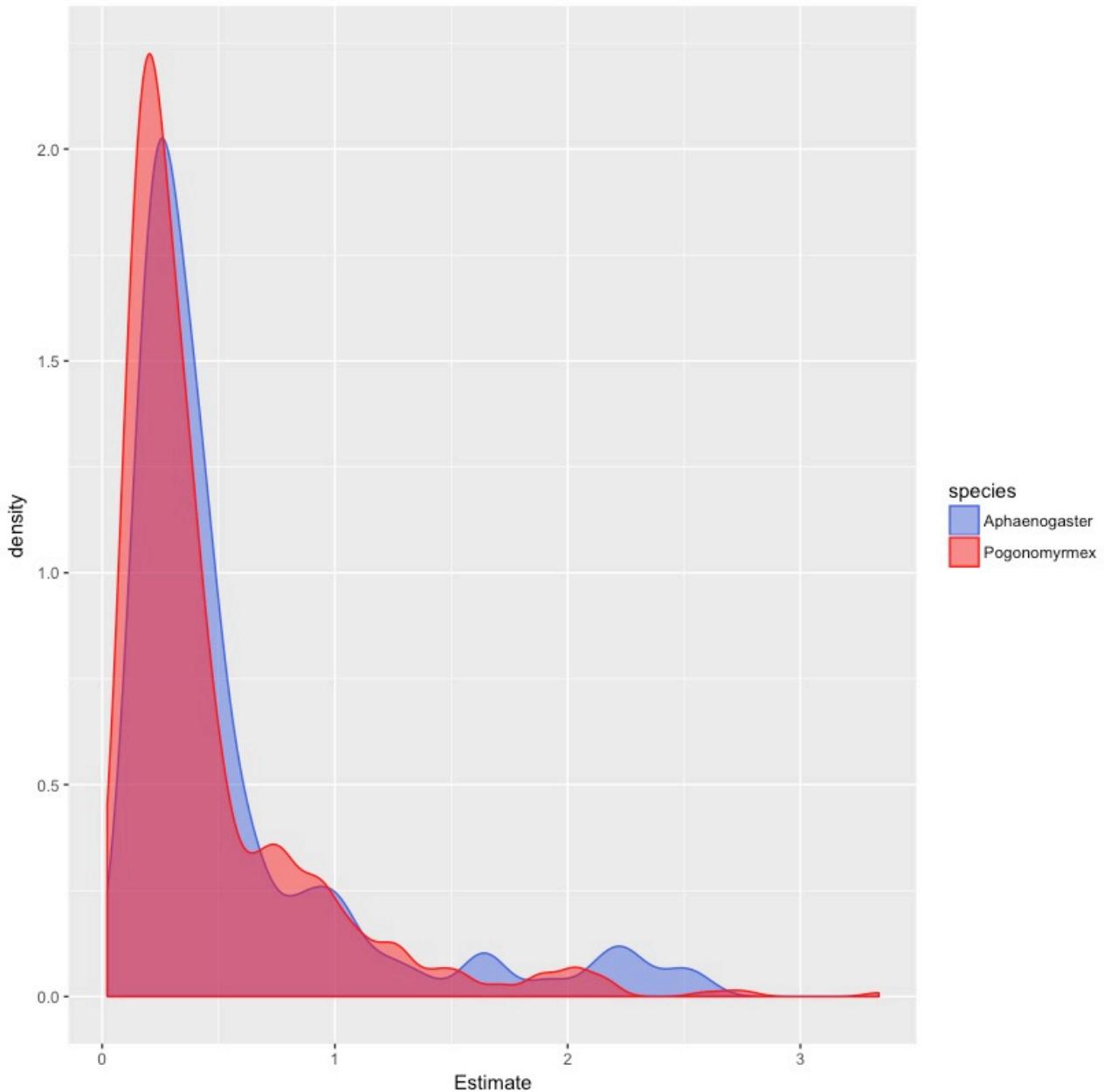
```

1 ks.test(minav.wide$Aphaenogaster,minav.wide$Pogonomyrmex
)
2
3   Two-sample Kolmogorov-Smirnov test
4
5 data: minav.wide$Aphaenogaster and
minav.wide$Pogonomyrmex
6 D = 0.46314, p-value = 2.204e-05
7 alternative hypothesis: two-sided

```

slope figure

```
1 slplot3<-ggplot(sl.ave, aes(x=Estimate, colour=species,  
fill=species)) +geom_density(alpha=.5,  
position="identity")+ggcol+ggcol2  
2 slplot3
```

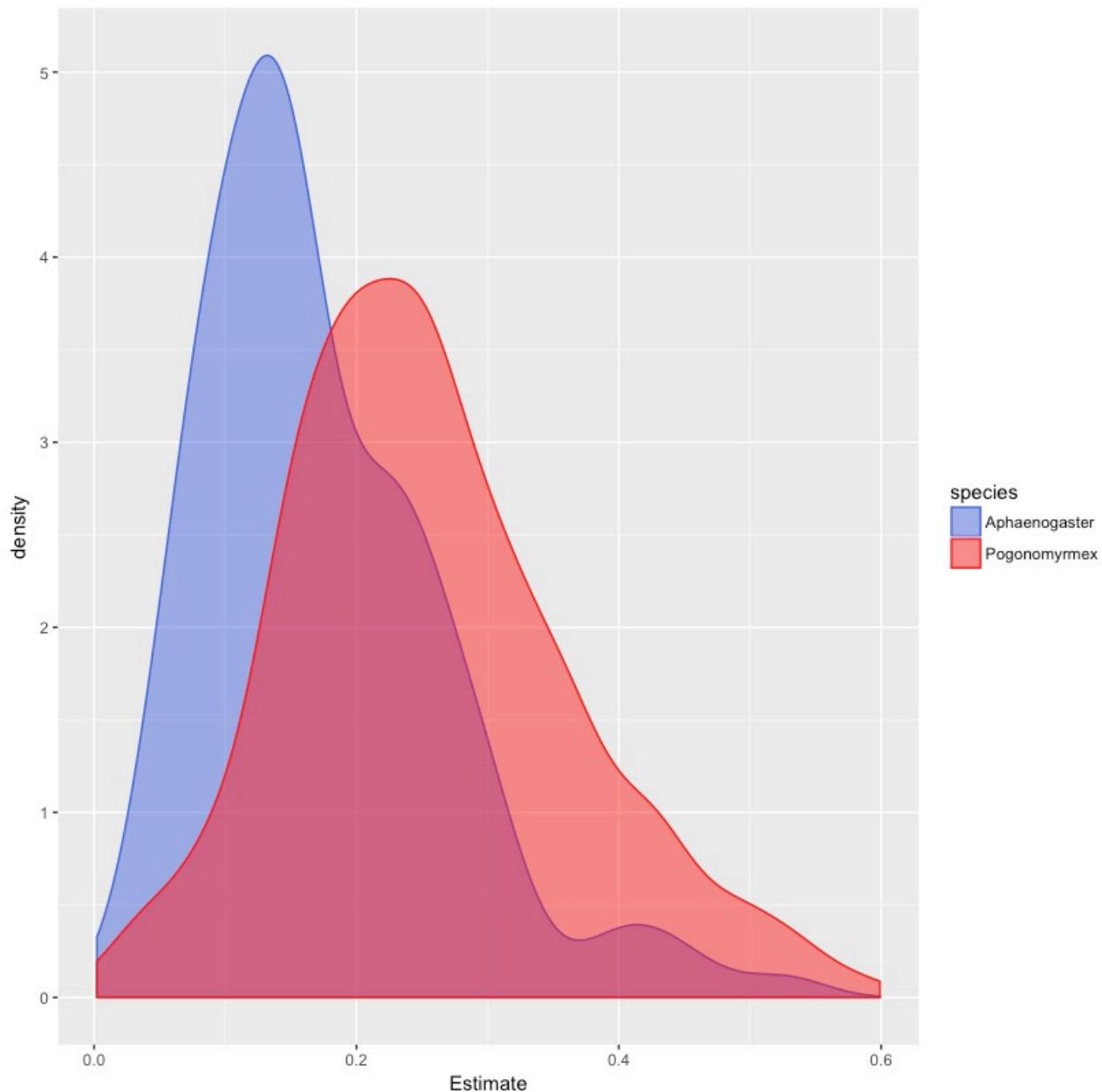


MIN

```
1 ks.test(minav.wide$Aphaenogaster,minav.wide$Pogonomyrmex
)
2
3   Two-sample Kolmogorov-Smirnov test
4
5 data: minav.wide$Aphaenogaster and
minav.wide$Pogonomyrmex
6 D = 0.22756, p-value = 0.003121
7 alternative hypothesis: two-sided
```

min fig

```
1 minplot3<-ggplot(min.ave, aes(x=Estimate, colour=species,
fill=species)) +geom_density(alpha=.5,
position="identity")+ggcol+ggcol2
2 minplot3
```



Example of unfolding curves between species for TWITCHIN-like

```
1 aphfit$species<-
  rep("Aphaenogaster",length(aphfit$colony))
2 pogfit$species<-
  rep("Pogonomyrmex",length(pogfit$colony))
3 f<-rbind(aphfit,pogfit)
4 names(f)
```

```

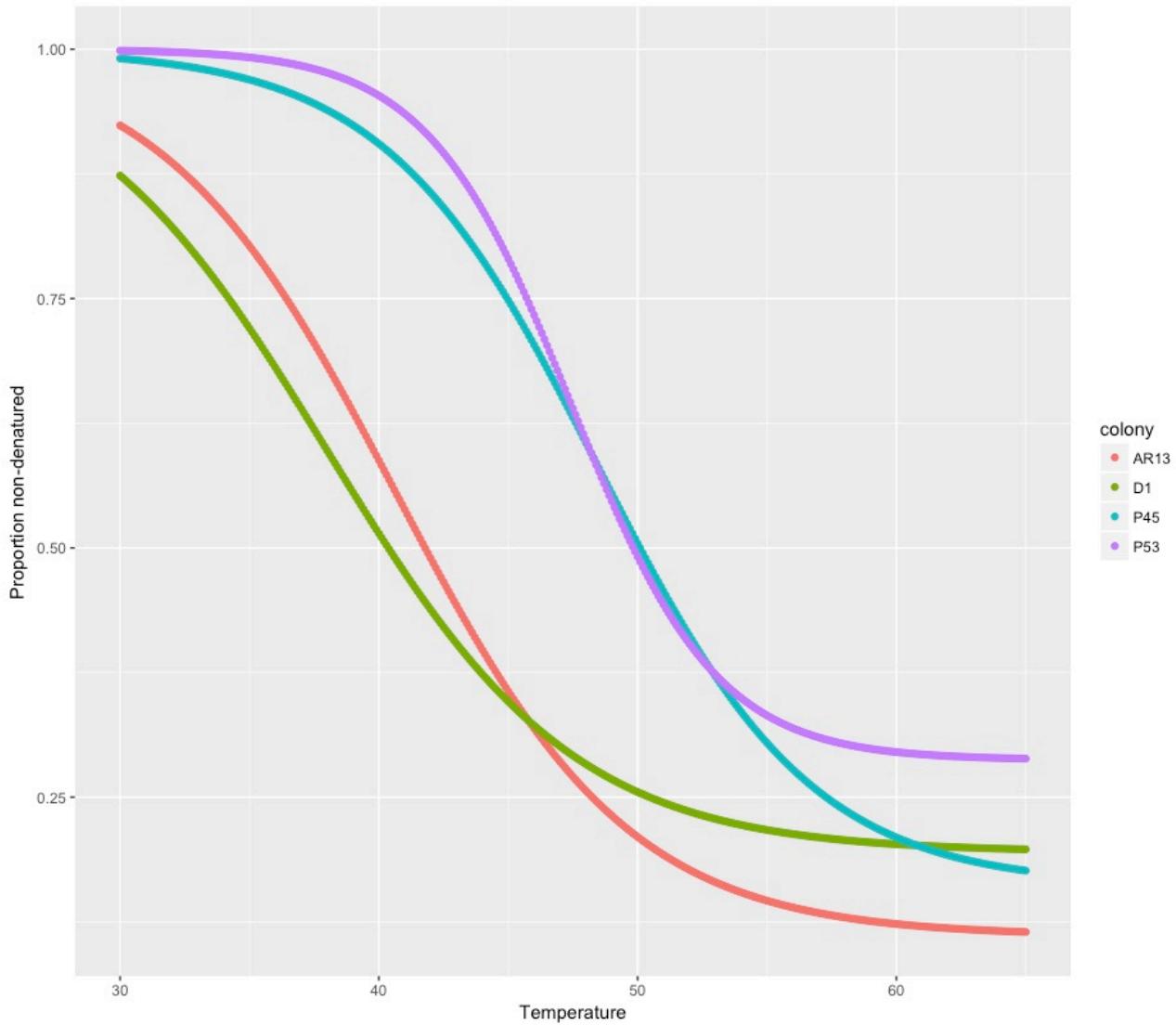
5 pep1<-subset(f,f$Sequence=="SDPSEVTPLITTK")
6
7 feeder<-
8   dcast(pep1,species+colony~param,value.var="Estimate")
9
10 list_predictions<-
11   sapply(split(feeder,list(feeder$colony,feeder$species)),
12     function(x)
13       {fud(T=seq(30,65,.1),Tm=x$Tm,slope=x$slope,min=x$min)}) )
14
15 predi<-as.data.frame(do.call("rbind",
16   list_predictions),stringAsFactors=FALSE)
17 predi$Sample<-row.names(predi)
18
19 nom<-
20   as.data.frame(matrix(unlist(strsplit(predi$Sample,
21     "[ . ]")),ncol=2,byrow=TRUE))
22 names(nom)<-c("colony","species")
23 predictions<-cbind(predi,nom)
24
25 conv<-gather(predictions,colony,uf,V1:V351)
26 conv<-conv[order(conv$Sample),]#sort
27
28 conv$T<-rep(seq(30,65,.1),nrow(predi))
29
30
31 ### overlay actual points
32
33 atwitch<-subset(aph,aph$Sequence=="SDPSEVTPLITTK" &
34   aph$colony!="ARY")
35
36 ptwitch<-subset(pog,pog$Sequence=="SDPSEVTPLITTK")
37
38
39 ##without points

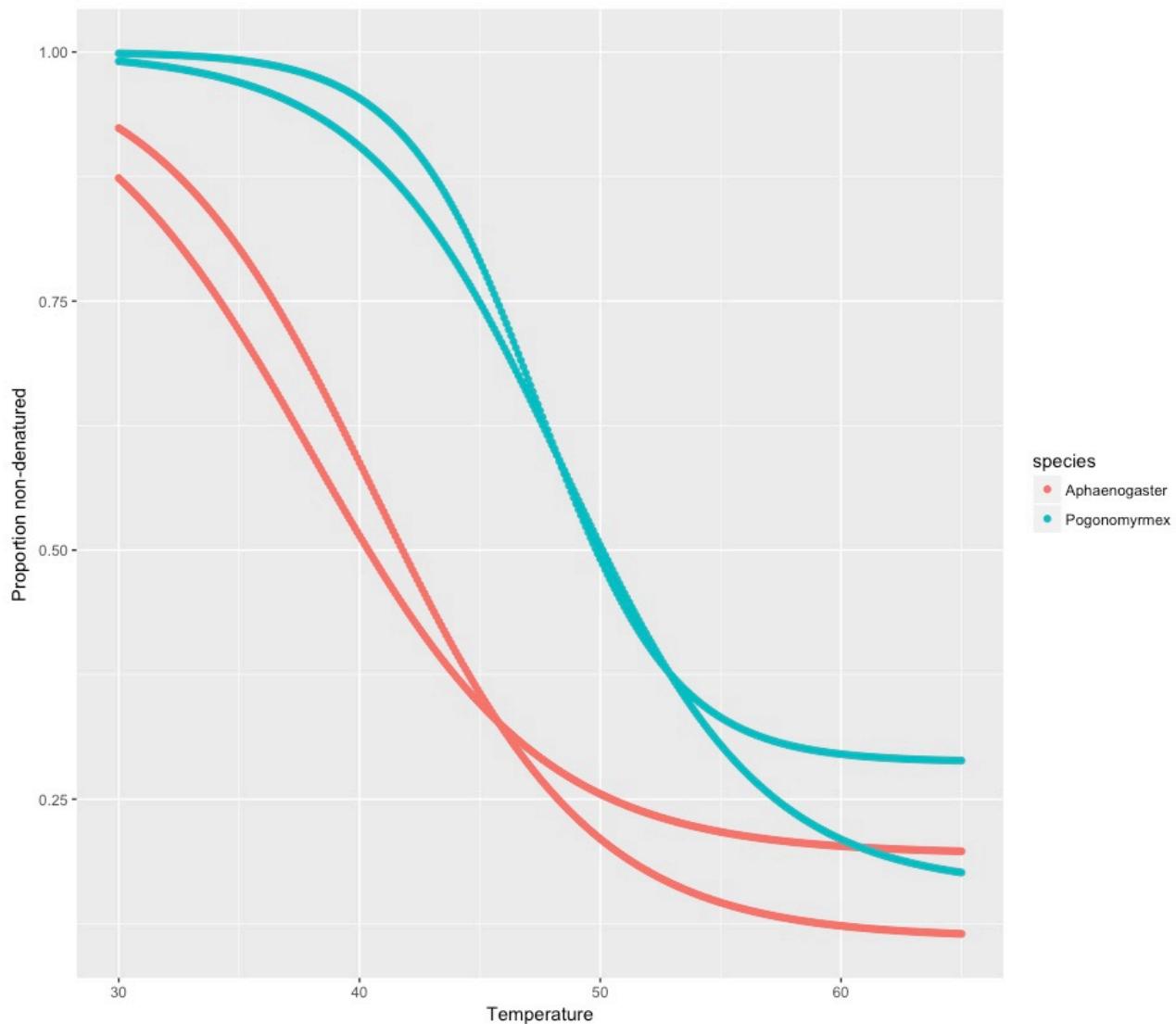
```

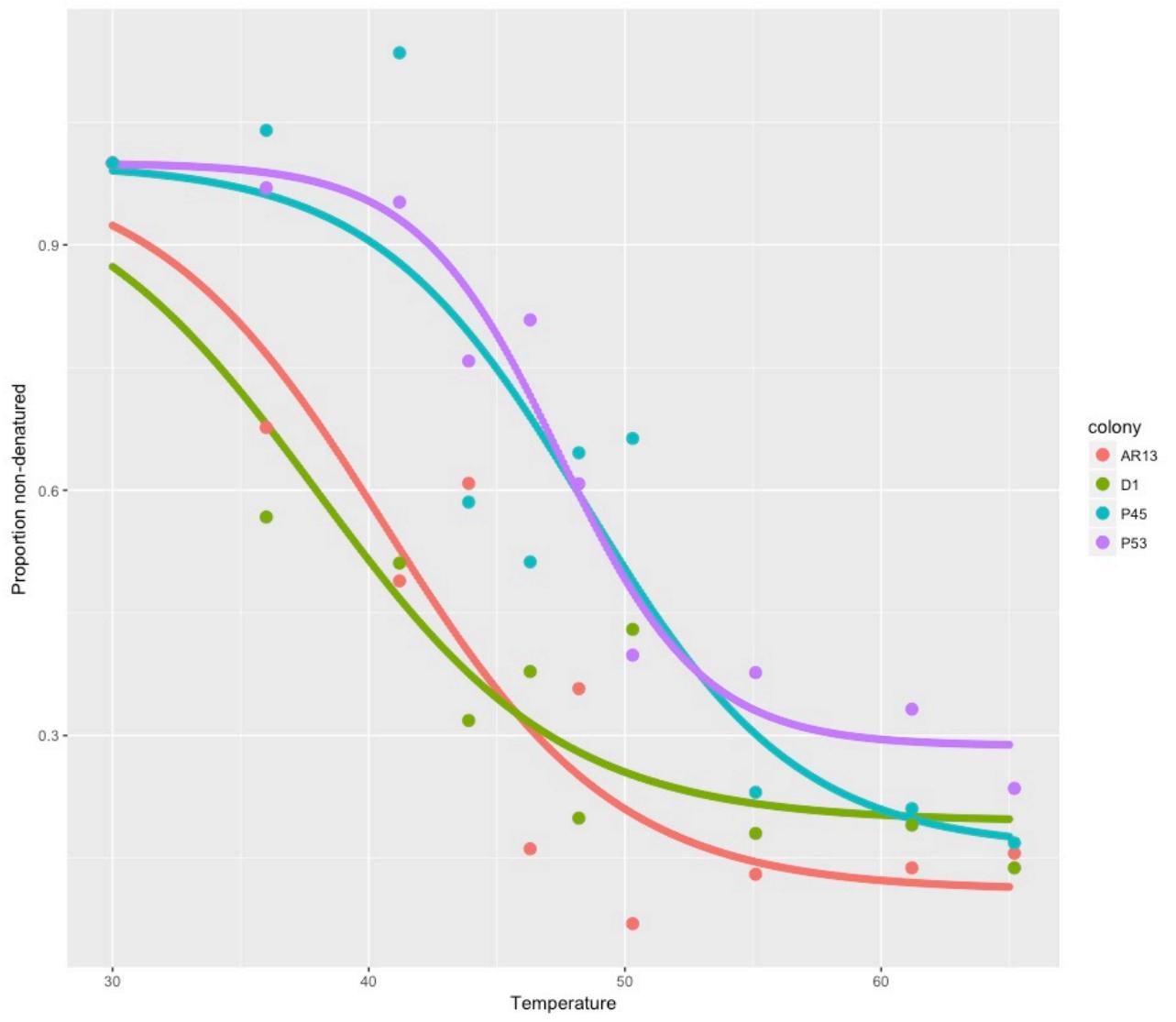
```

32 ggplot(conv,aes(y=uf,x=T,colour=colony))+geom_point()+
  lab("Temperature")+ylab("Proportion non-
denatured")#+geom_point(data=atwitch,aes(x=temperature,
y=unfolding,colour=colony),size=3)+geom_point(data=ptwi
tch,aes(x=temperature,y=unfolding,colour=colony),size=3
)

```







Page 80: 2017-05-11 Stress in nature project: parsing out chamber data

From meeting with SHC from (Page 77: 2017-05-05.):

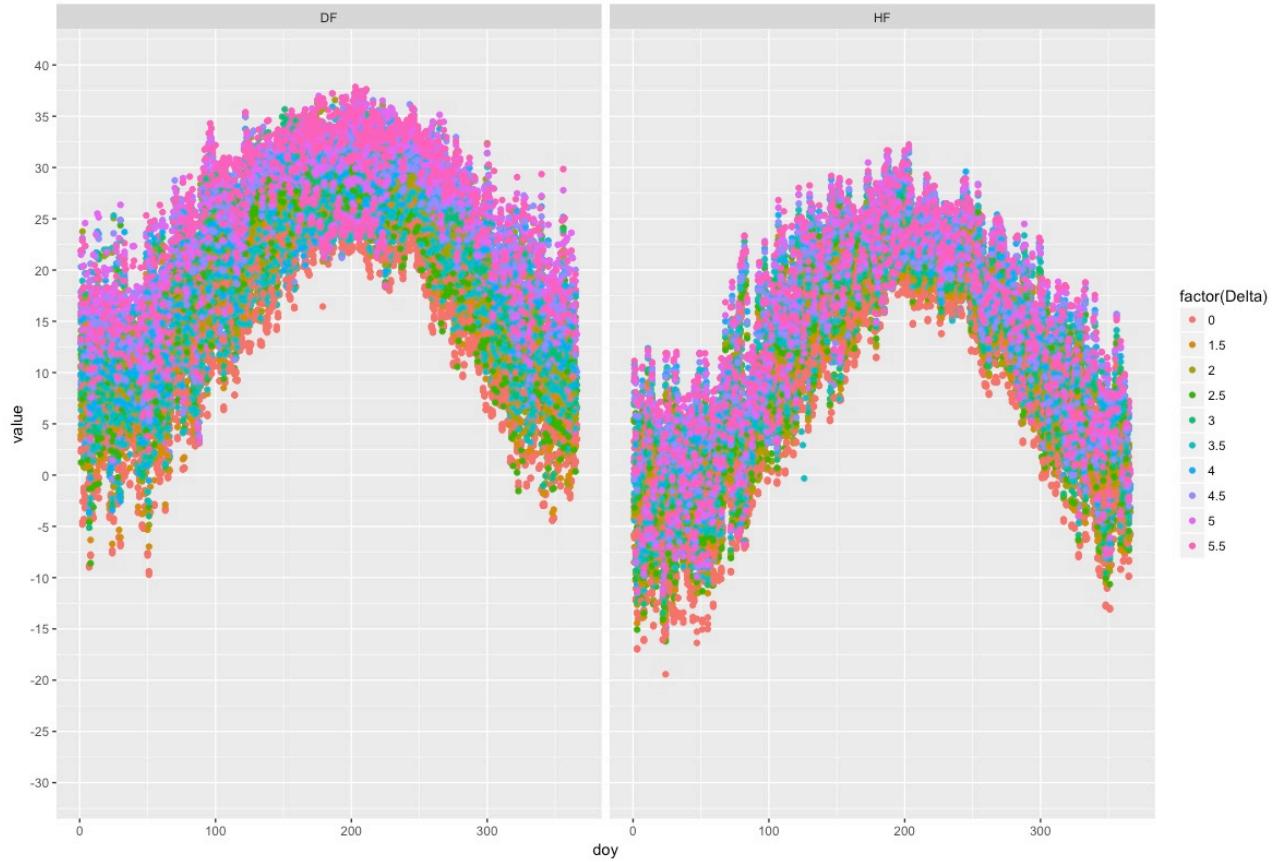
- Need info on summer temperatures
 - mean, min, max monthly temperatures in growing season.
 - get experienced temperature from chambers (average across 5 years)

The parsed dataset, I'll show script on website. I basically turned the columns into a factor so I can take averages across them and then just subset the variables I want

```
1
2 cdave<-ddply(cdlong,.
  (Site,year,Delta,Cham,envfac,doy),summarize,value=mean(
  value,na.rm=TRUE))
3 dim(cdave)
4 str(cdave)
5 'data.frame': 380312 obs. of 7 variables:
6   $ Site : chr "DF" "DF" "DF" "DF" ...
7   $ year : num 2010 2010 2010 2010 2010 2010 2010 2010
2010 2010 ...
8   $ Delta : num 0 0 0 0 0 0 0 0 0 0 ...
9   $ Cham : chr "11" "11" "11" "11" ...
10  $ envfac: chr "avecat" "avecat" "avecat" "avecat" ...
11  $ doy   : int 2 3 4 5 6 7 8 9 10 54 ...
12  $ value : num -4.82 -4.43 -3.23 -2.85 -1.13 ...
```

For example...for the average chamber air temp...

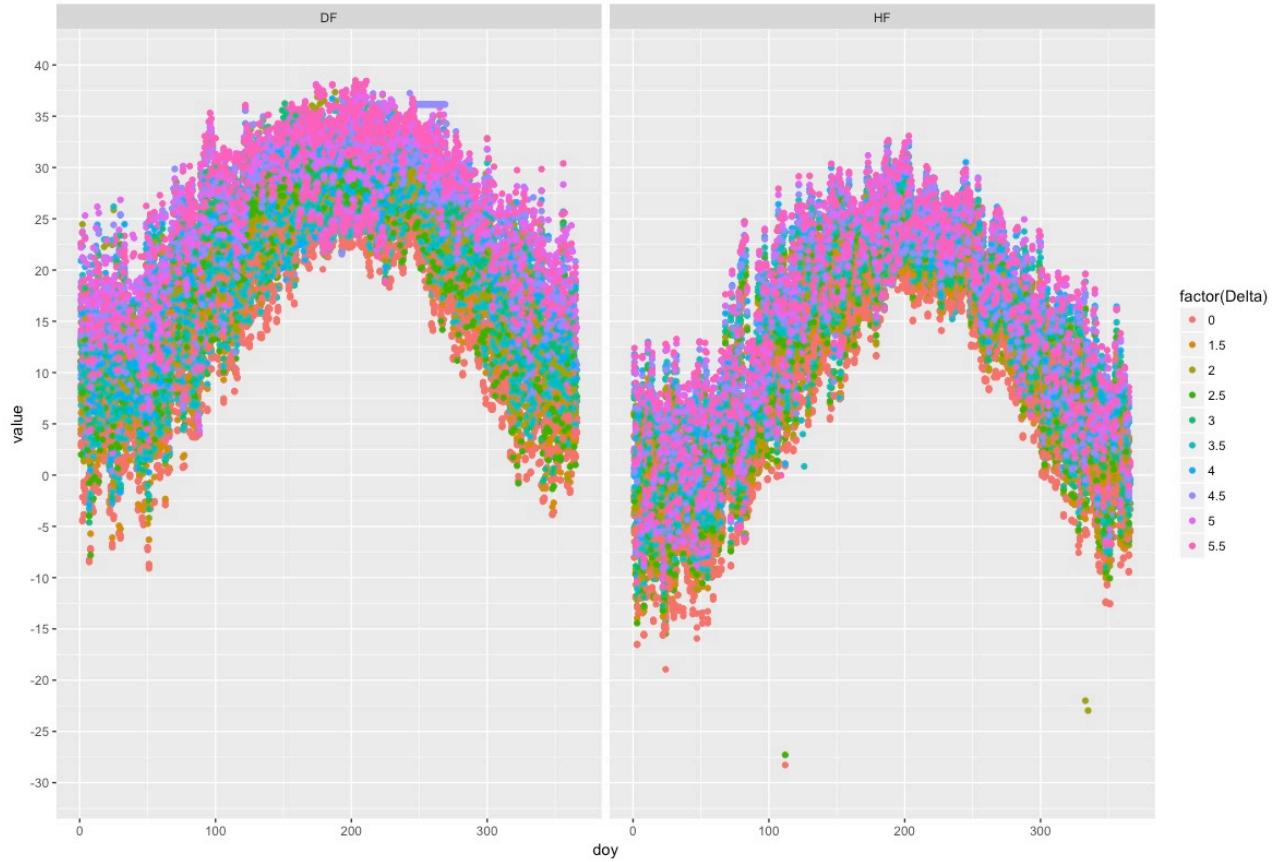
```
1 ave<-subset(cdave,cdave$envfac=="avecat")
```



I'll learn tidyverse one of these days...

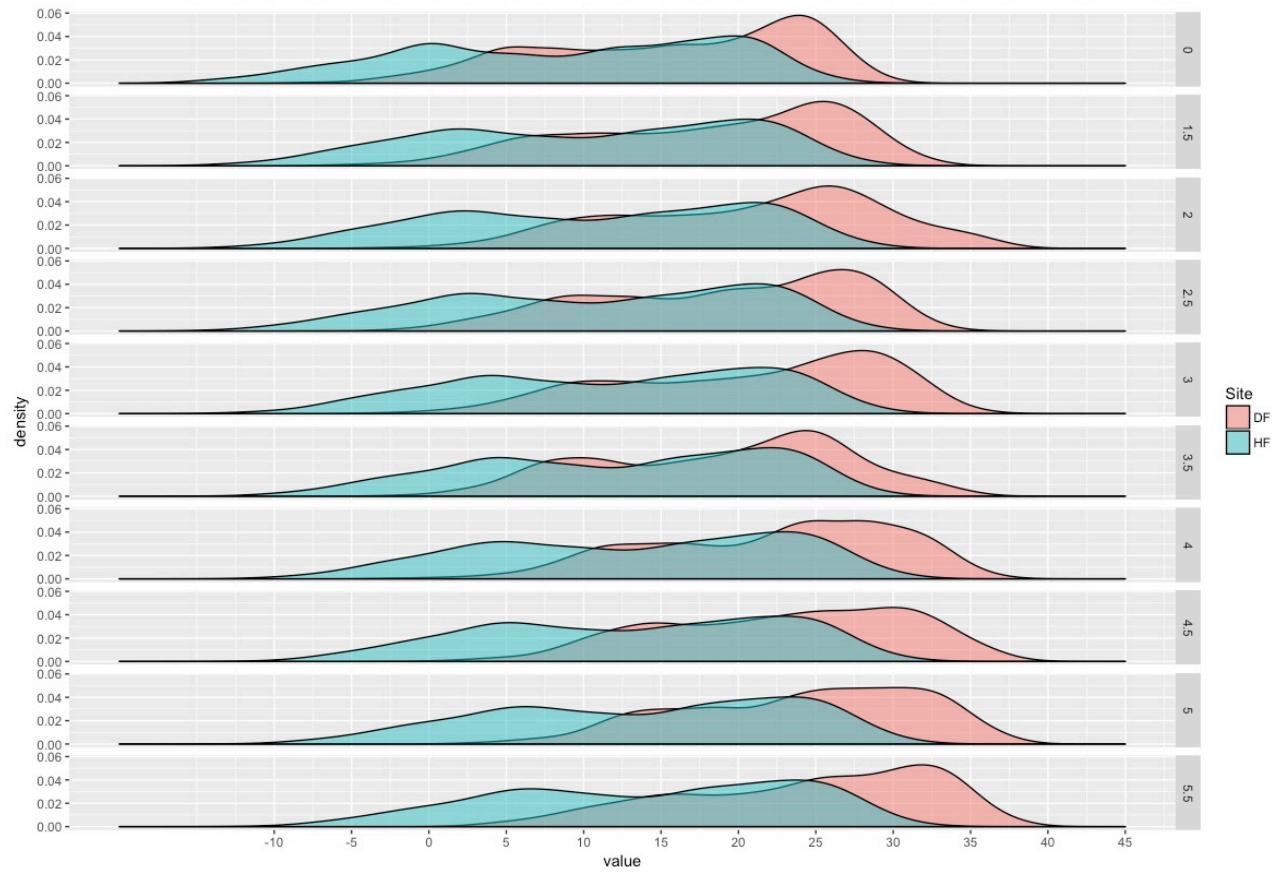
and for max(max has weird extreme low values):

```
1 | max<-subset(cdave,cdave$envfac=="maxcat" &
| cdave$value>-50)
```



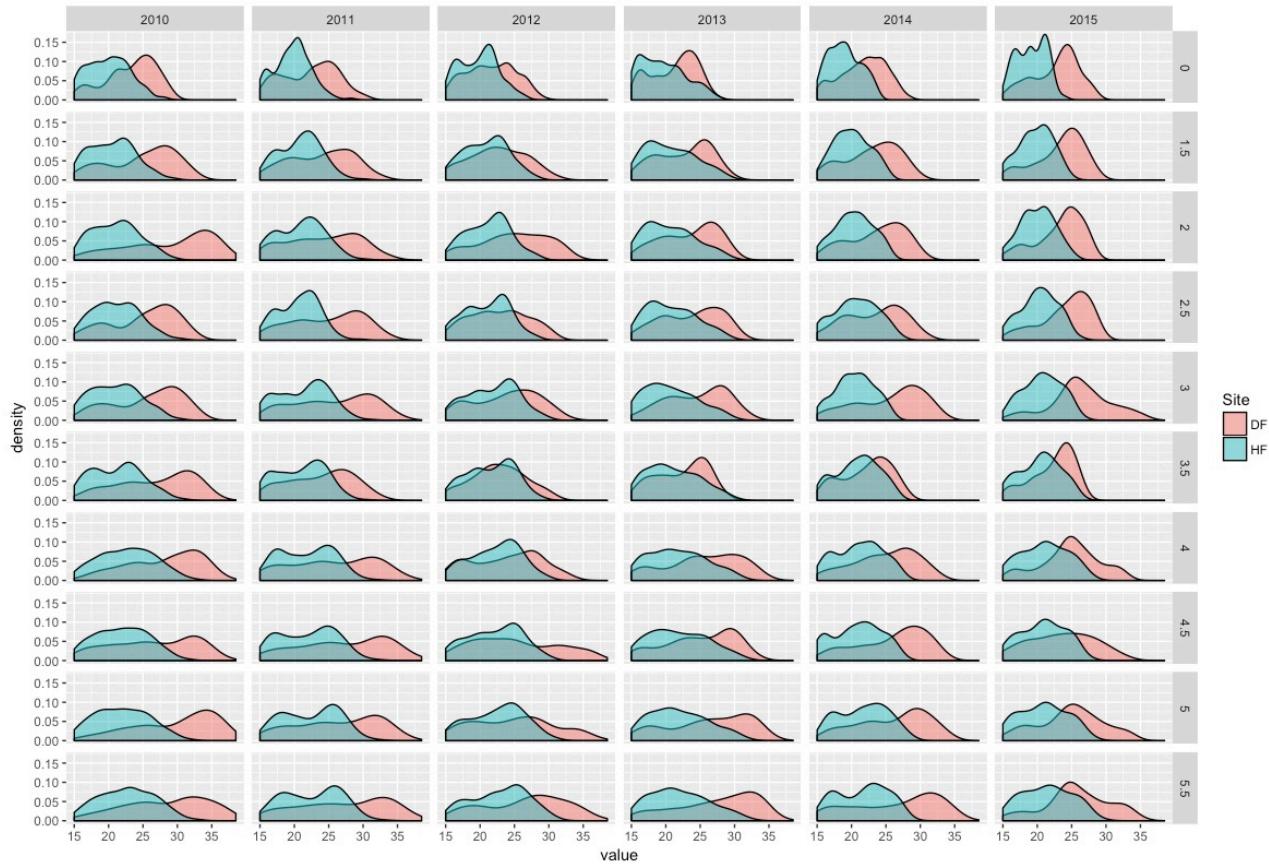
density plots for max, whole range of temps

```
1 ggplot(max,aes(x=value,fill=Site))+geom_density(alpha=.5
)+facet_grid(Delta~.)+scale_x_continuous(breaks=seq(-10,
45,5),labels=seq(-10,45,5),limits=c(-20,45))
```



plots from 15 C cutoff (GSL) TEMPERATURE

Densities of temperatures vs julian day (doy) for each year and site



Let's look at the soil organic layer

```

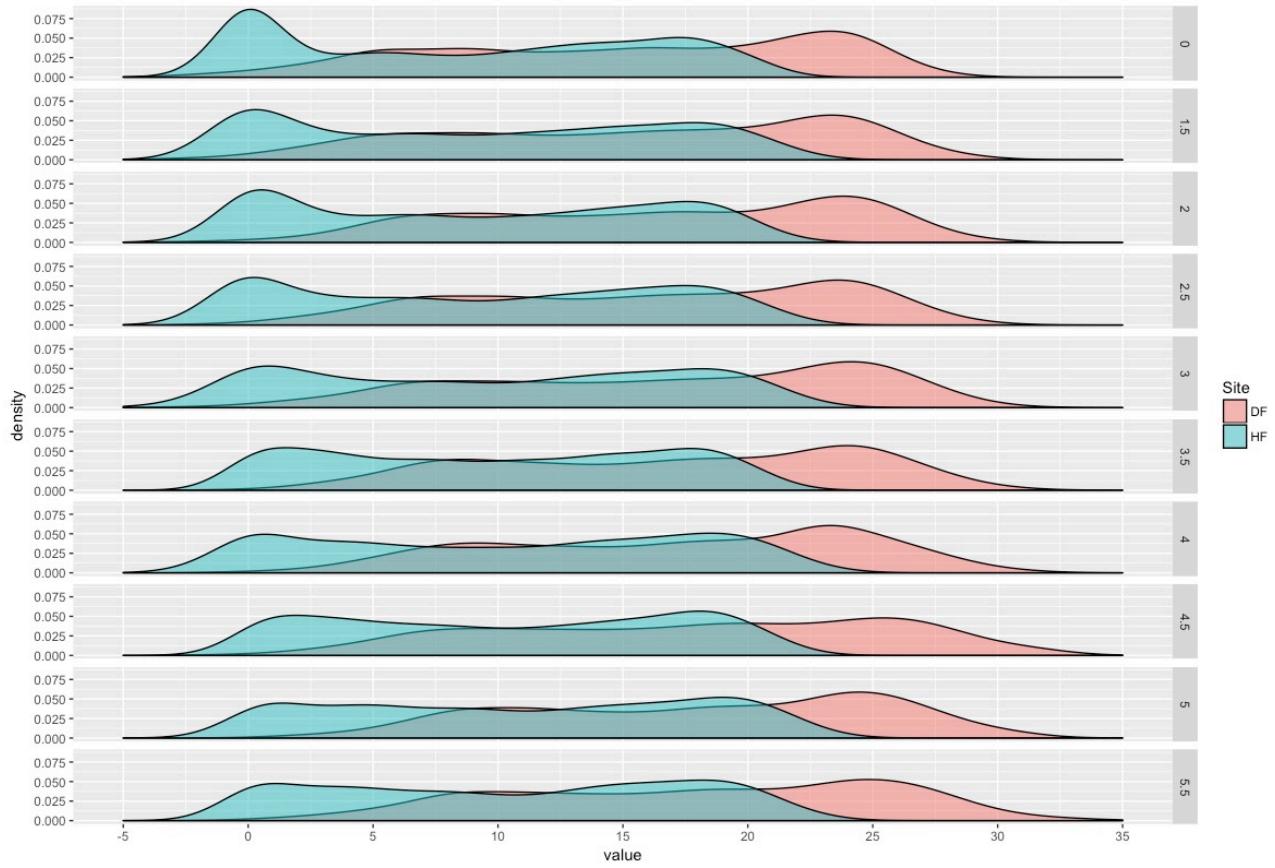
1 soilorg.max<-subset(cdave,cdave$envfac=="SOmax")
2 ### plotting
3 O<-
  scale_x_continuous(breaks=seq(-10,45,5),labels=seq(-10,4
  5,5),limits=c(-20,45))

```

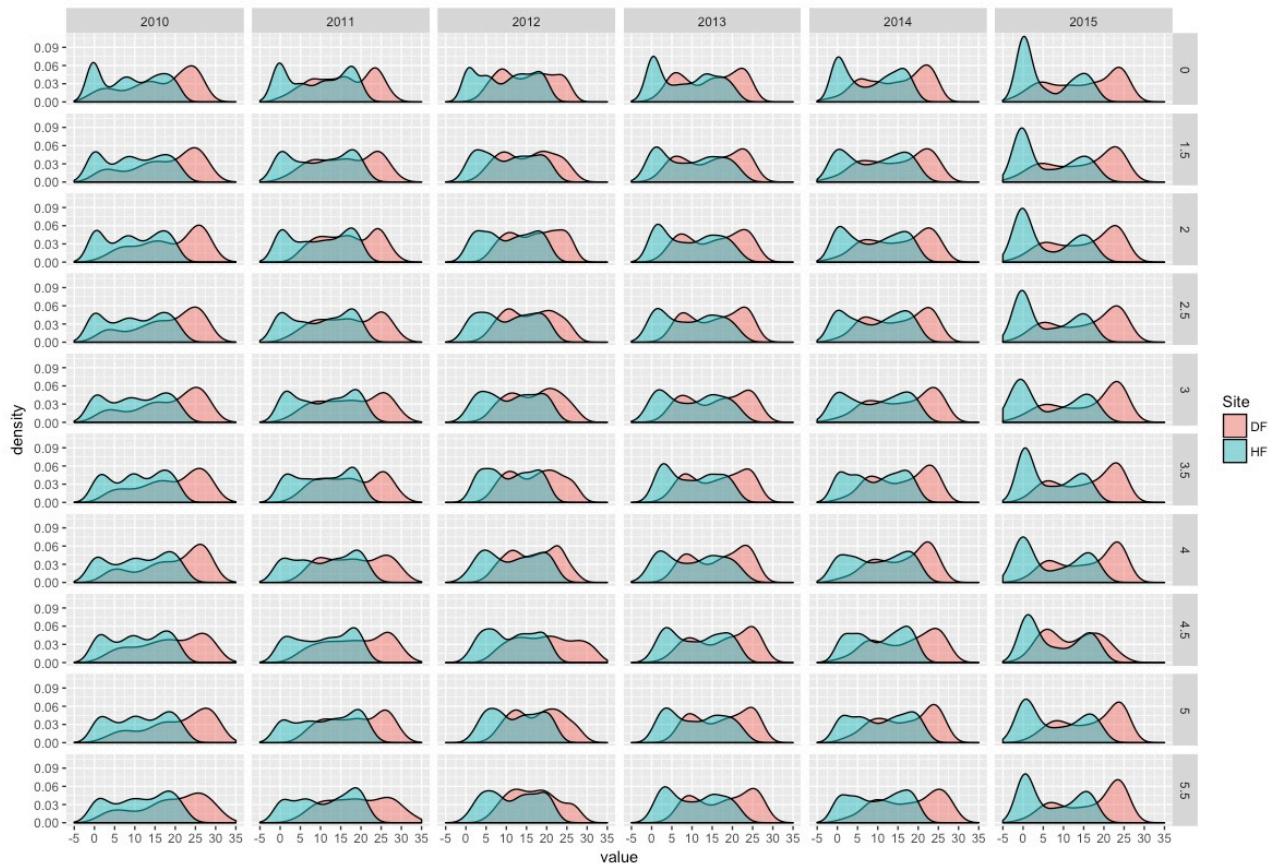
histogram



density



separate out by year



averaging soil organic layer across all years

parsing it out: Taking averages and SD for all env factors then
subsetting

```
1 cdave.deltas<-ddply(cdave,.(envfac,Site,Delta),  
2 summarize,dave=mean(value,na.rm=TRUE),  
3 sd=sd(value,na.rm=TRUE)/sqrt(length(Delta)))  
4  
5 S0maxDelta<-  
6 subset(cdave.deltas,cdave.deltas$envfac=="S0max" &  
7 cdave.deltas$dave > -25)  
8 head(S0maxDelta)
```

Fig

```
1 ggplot(S0maxDelta,aes(x=Delta,y=dave,colour=Site))+geom  
2 _point(size=3)+stat_smooth(method="lm")  
3 ! [ ]  
4 (https://cloud.githubusercontent.com/assets/4654474/259  
5 51498/ee490d7a-362b-11e7-9a69-07705dd05a1c.jpeg)  
6  
7 stats: has experimental warming done anything? Yes  
8  
9 R  
10 m2<-lm(dave~Delta*Site,data=S0maxDelta)  
11 summary(m2)  
12  
13 Call:  
14 lm(formula = dave ~ Delta * Site, data = S0maxDelta)  
15 Residuals:  
16 Min 1Q Median 3Q Max
```

```

16 -0.31033 -0.20202  0.03446  0.12334  0.35954
17
18 Coefficients:
19             Estimate Std. Error t value Pr(>|t|)
20 (Intercept) 15.04818   0.15514  97.000 < 2e-16
21 Delta        0.56336   0.04384  12.852 7.57e-10
22 SiteHF      -6.49663   0.21940 -29.612 2.11e-15
23 Delta:SiteHF -0.15103   0.06199 -2.436   0.0269
24
25 (Intercept) ***
26 Delta        ***
27 SiteHF      ***
28 Delta:SiteHF *
29 ---
30 Signif. codes:
31 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
32
33 Residual standard error: 0.2236 on 16 degrees of
freedom
34 Multiple R-squared:  0.9969 ,    Adjusted R-squared:
0.9963
35 F-statistic: 1705 on 3 and 16 DF,  p-value: < 2.2e-16
36

```

Let's explore all temperature realted factors

```

1 m3<-aov(dave~Delta*Site*envfac,data=cDave.deltas)
2 summary(m3)
3
4             Df Sum Sq Mean Sq  F value    Pr(>F)
5 Delta           1 168.4   168.4 542.359 < 2e-16
6 ***

7 Site            1 1917.8  1917.8 6176.185 < 2e-16
8 ***

```

```

6 envfac             7  356.2      50.9   163.869 < 2e-16
7
8 Delta:Site         1    1.9       1.9    6.153   0.0144
9
10 Delta:envfac      7   80.2      11.5   36.913 < 2e-16
11 Site:envfac      7   90.1      12.9   41.469 < 2e-16
12 Residuals        128  39.7      0.3
13
14
15
16 m4<-lm(dave~Delta*Site*envfac,data=cDave.deltas)
17
18 Call:
19 lm(formula = dave ~ Delta * Site * envfac, data =
20 cDave.deltas)
21
22 Residuals:
23
24
25 Coefficients:
26
27
28
```

		Estimate	Std. Error	t value
Pr(> t)				
(Intercept)	15.46048	0.38658	39.993	< 2e-16 ***
Delta	1.46191	0.10923	13.384	< 2e-16 ***

29	SiteHF	-7.06618	0.54670	-12.925
	< 2e-16 ***			
30	envfacmaxcat	0.45721	0.54670	0.836
	0.404538			
31	envfacmincat	-0.62281	0.54670	-1.139
	0.256743			
32	envfacSIave	-0.29921	0.54670	-0.547
	0.585132			
33	envfacSImax	-0.39879	0.54670	-0.729
	0.467064			
34	envfacSOave	-0.95765	0.54670	-1.752
	0.082226 .			
35	envfacSOmax	-0.41230	0.54670	-0.754
	0.452141			
36	envfacSOmin	-1.13297	0.54670	-2.072
	0.040238 *			
37	Delta:SiteHF	-0.54810	0.15448	-3.548
	0.000543 ***			
38	Delta:envfacmaxcat	0.03926	0.15448	0.254
	0.799809			
39	Delta:envfacmincat	-0.02588	0.15448	-0.168
	0.867195			
40	Delta:envfacSIave	-1.44304	0.15448	-9.341
	3.89e-16 ***			
41	Delta:envfacSImax	-1.38286	0.15448	-8.952
	3.44e-15 ***			
42	Delta:envfacSOave	-1.17475	0.15448	-7.605
	5.39e-12 ***			
43	Delta:envfacSOmax	-0.89854	0.15448	-5.817
	4.53e-08 ***			
44	Delta:envfacSOmin	-1.17982	0.15448	-7.637
	4.53e-12 ***			
45	SiteHF:envfacmaxcat	-0.03313	0.77316	-0.043
	0.965886			

```

46 SiteHF:envfacmincat      0.18968    0.77316    0.245
     0.806592
47 SiteHF:envfacSIave      0.50090    0.77316    0.648
     0.518237
48 SiteHF:envfacSImax      1.08374    0.77316    1.402
     0.163424
49 SiteHF:envfacSOave      1.07415    0.77316    1.389
     0.167151
50 SiteHF:envfacSOmax      0.56955    0.77316    0.737
     0.462679
51 SiteHF:envfacSOmin      1.16451    0.77316    1.506
     0.134486
52 Delta:SiteHF:envfacmaxcat 0.04352    0.21846    0.199
     0.842412
53 Delta:SiteHF:envfacmincat -0.07099   0.21846   -0.325
     0.745751
54 Delta:SiteHF:envfacSIave 0.88124    0.21846    4.034
     9.38e-05 ***
55 Delta:SiteHF:envfacSImax 0.70702    0.21846    3.236
     0.001541 **
56 Delta:SiteHF:envfacSOave 0.67109    0.21846    3.072
     0.002599 **
57 Delta:SiteHF:envfacSOmax 0.39708    0.21846    1.818
     0.071465 .
58 Delta:SiteHF:envfacSOmin 0.67203    0.21846    3.076
     0.002564 **
59 ---
60 Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 
     0.1 ' ' 1
61
62 Residual standard error: 0.5572 on 128 degrees of
     freedom
63 Multiple R-squared:  0.9851,    Adjusted R-squared:
     0.9815

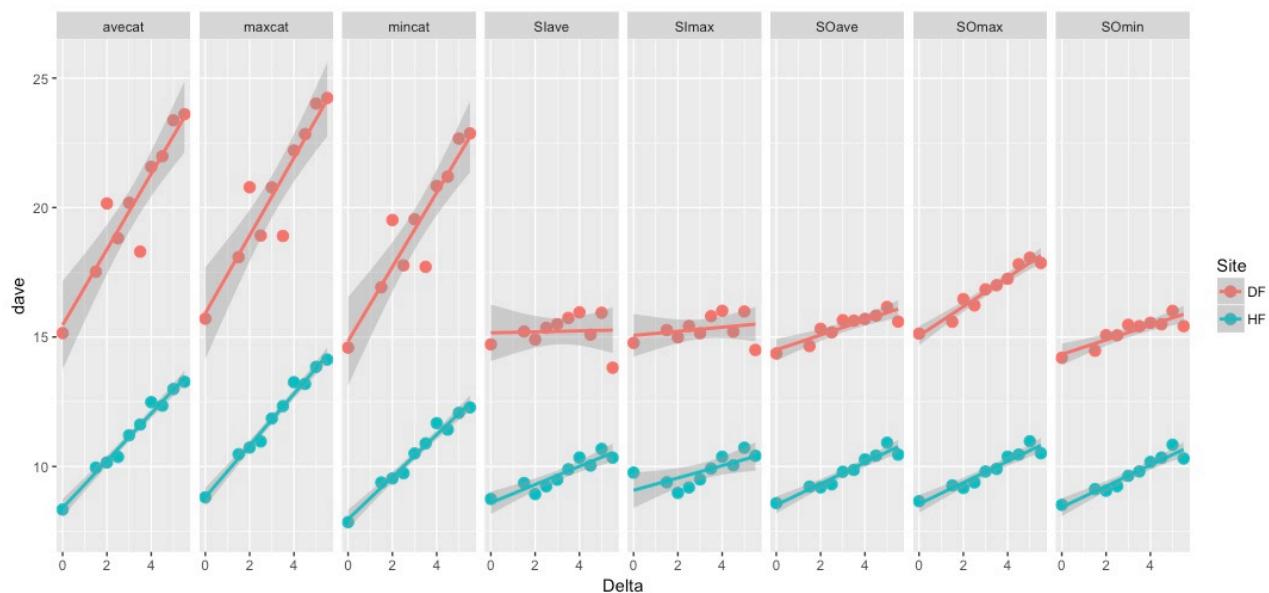
```

```
64 F-statistic: 273 on 31 and 128 DF, p-value: < 2.2e-16
```

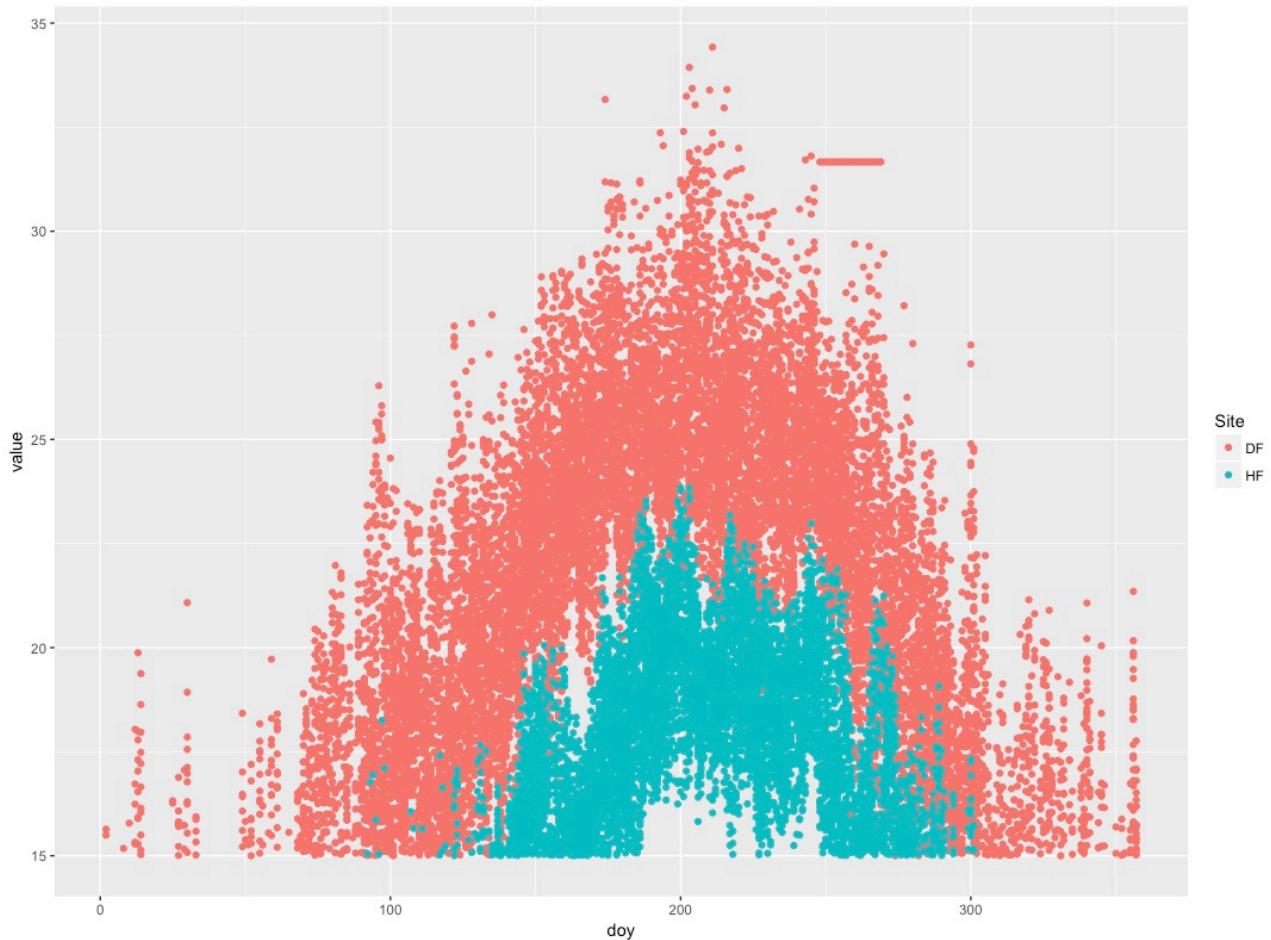
```
65
```

Let's plot this out

```
1 ggplot(cdave.deltas,aes(x=Delta,y=dave,colour=Site))+geom_point(size=3)+stat_smooth(method="lm")+facet_grid(.~envfac)
```



Focus: on GSL of soil organic temperature (MAX) that is above 15 C



**ok given the temps experienced in SO layer,
how much stress are they experiencing?**

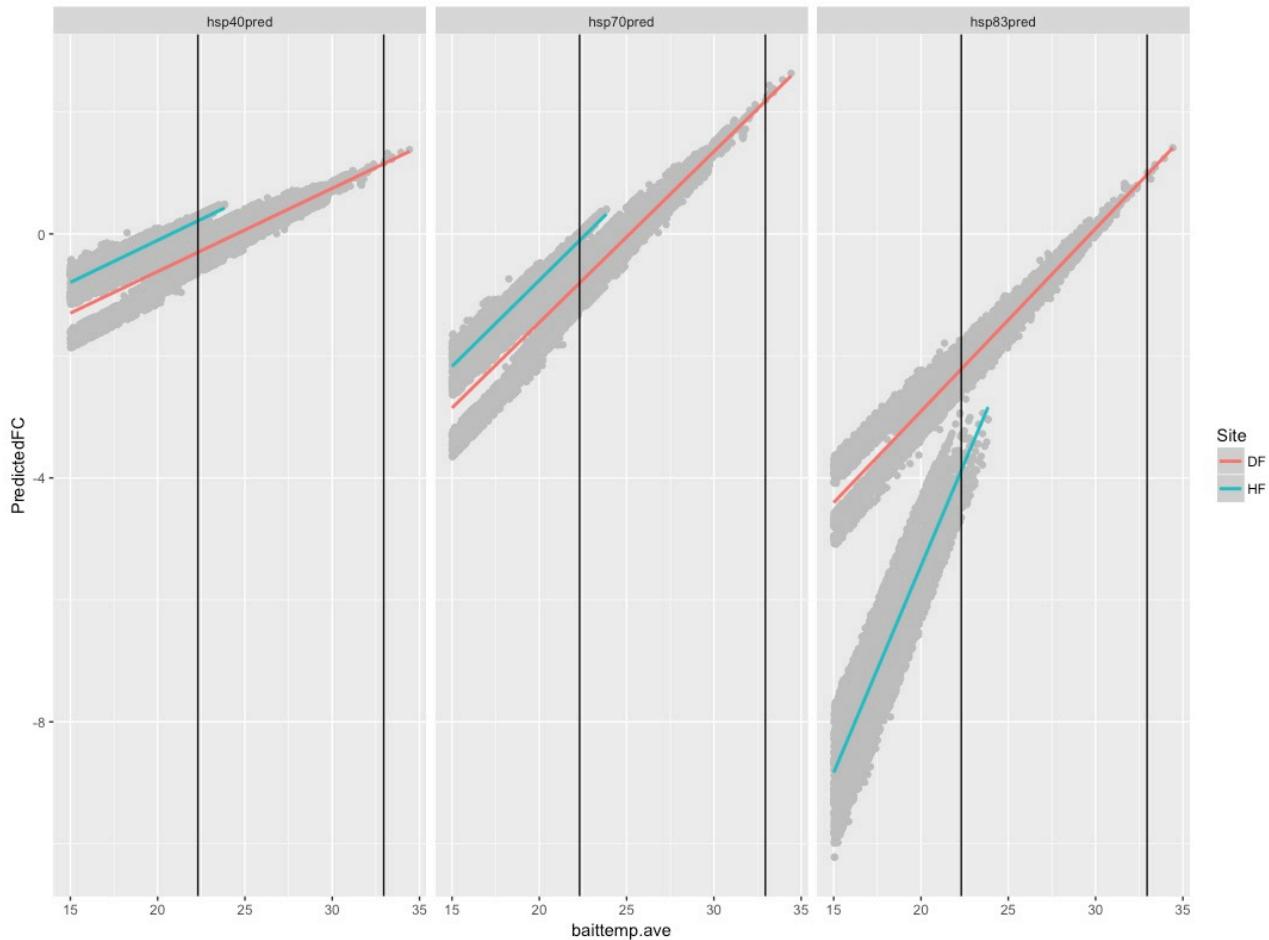
```

1 GSL.long<-
2   gather(GSL,gene,PredictedFC,hsp83pred:hsp40pred)
3 str(GSL.long)
4
5 'data.frame': 63027 obs. of 11 variables:
6   $ X           : int 1700 1710 1711 1712 1713 1714
7   $ Site        : Factor w/ 2 levels "DF","HF": 1 1 1 1
8   $ year        : int 2010 2010 2010 2010 2010 2010
9   $ Delta        : num 0 0 0 0 0 0 0 0 0 ...
10  $ Cham         : int 11 11 11 11 11 11 11 11 11 ...
11 
```

```
9  $ envfac      : Factor w/ 1 level "S0max": 1 1 1 1 1 1  
10 $ 1 1 1 1 ...  
11 $ doy         : int 71 89 90 91 92 93 94 95 96 97 ...  
12 $ baittemp.ave: num 15.8 18.2 16.6 19.2 21 ...  
13 $ Jdaycont    : int 71 89 90 91 92 93 94 95 96 97 ...  
14 $ gene        : chr "hsp83pred" "hsp83pred"  
"hsp83pred" "hsp83pred" ...  
15 $ PredictedFC : num -4.67 -3.91 -4.37 -3.61 -3.06 ...
```

scatter plot with predicted lines and the region where we constructed the model

```
1 ggplot(GSL.long,aes(x=baittemp.ave,y=PredictedFC,colour=  
Site))+geom_point(colour="gray")+stat_smooth(method="lm"  
) +facet_grid(.~gene)+geom_vline(xintercept=range(findat.  
long$baittemp.ave))  
2
```



for hsp40, estimate temp for FC at 0

```

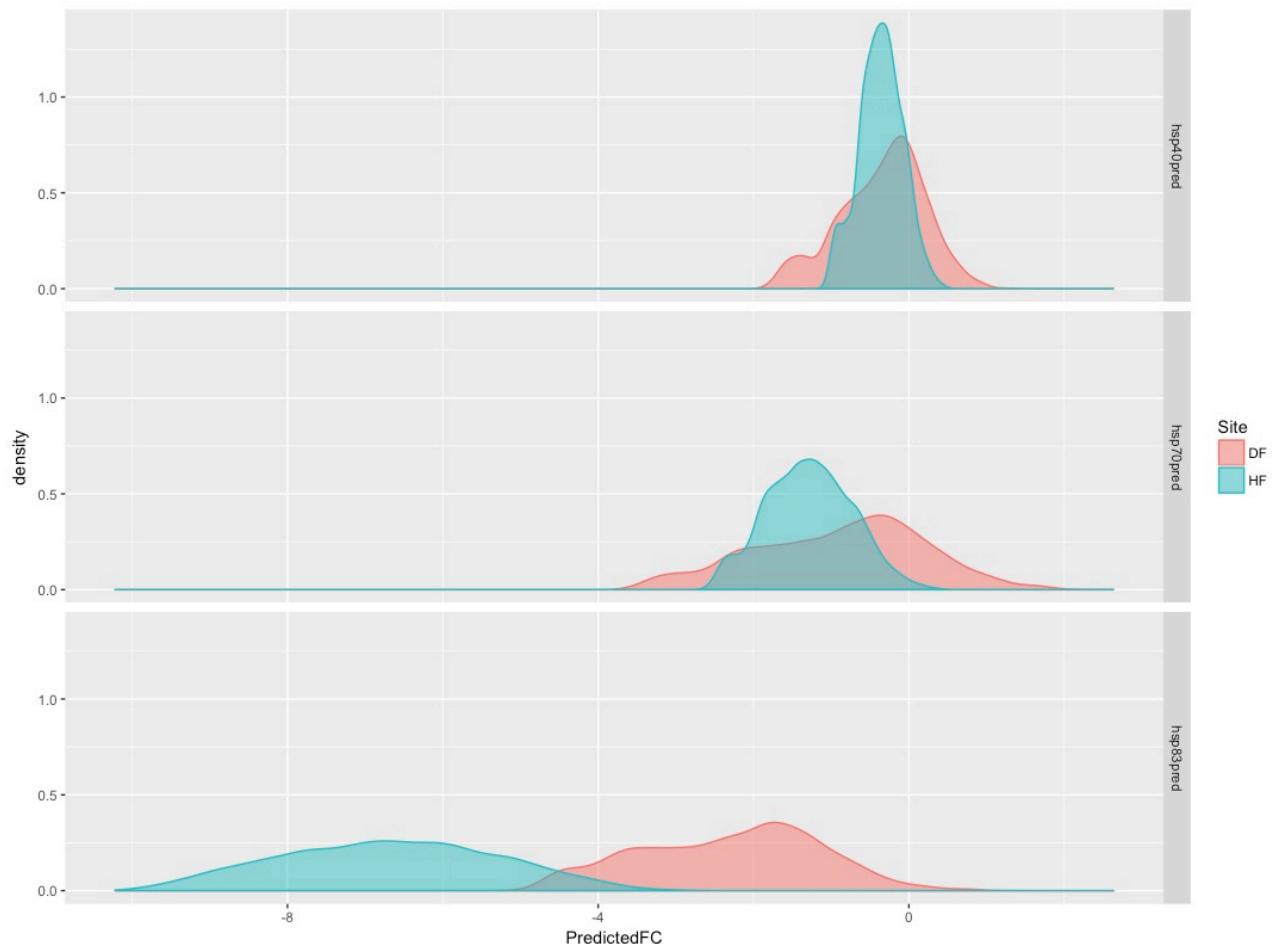
1 ##predict bait temp from 0 FC
2 #HF
3 HFzero_hsp40<-subset(GSL,GSL$Site=="HF" &
GSL$hsp40pred<0.01 & GSL$hsp40pred>-.01)
4 mean(HFzero_hsp40$baittemp.ave)
5 ### DF
6 DFzero_hsp40<-subset(GSL,GSL$Site=="DF" &
GSL$hsp40pred<0.01 & GSL$hsp40pred>-.01)
7 mean(DFzero_hsp40$baittemp.ave)
8

```

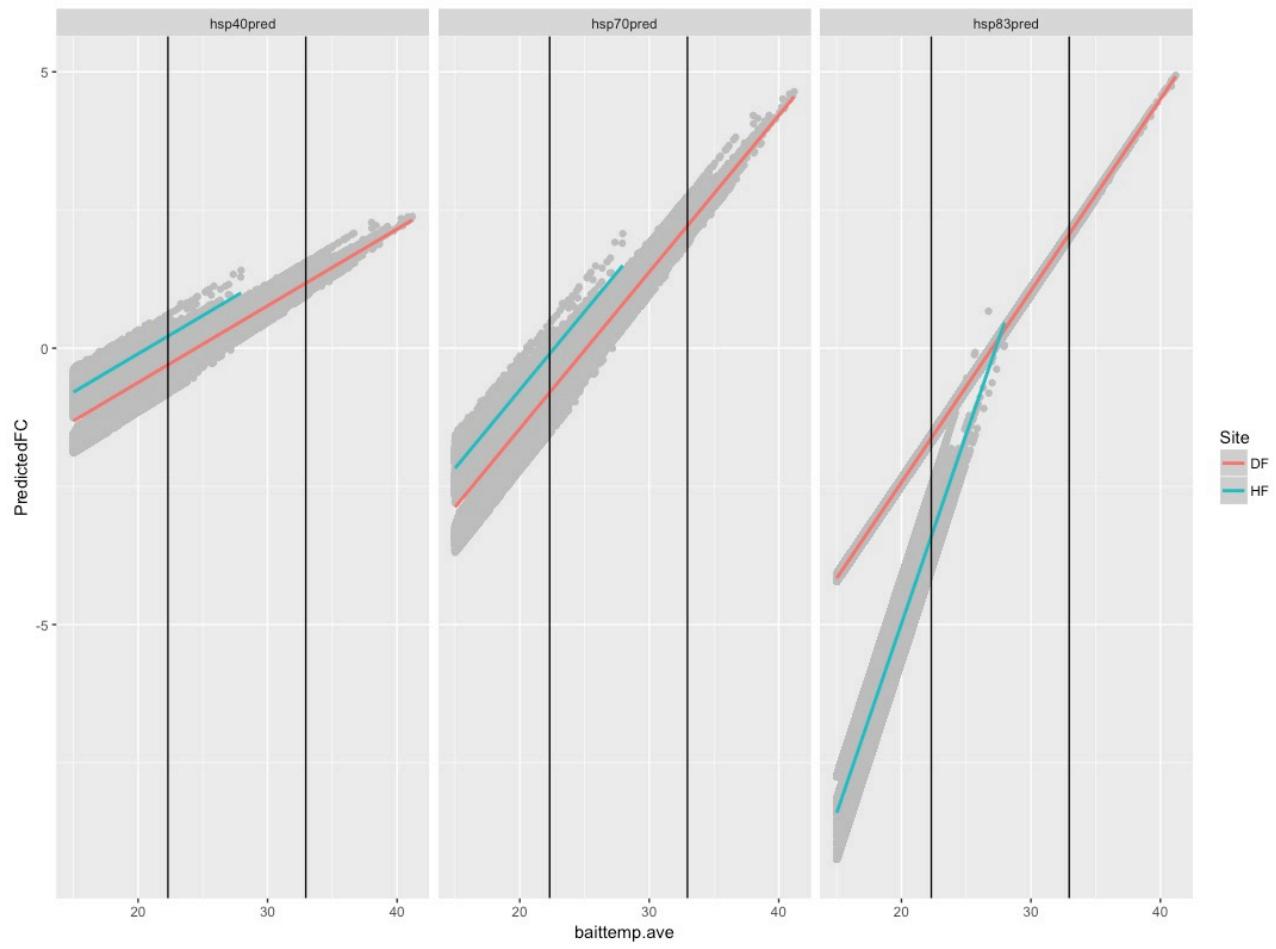
- HF = 20.49022
- DF = 24.31707

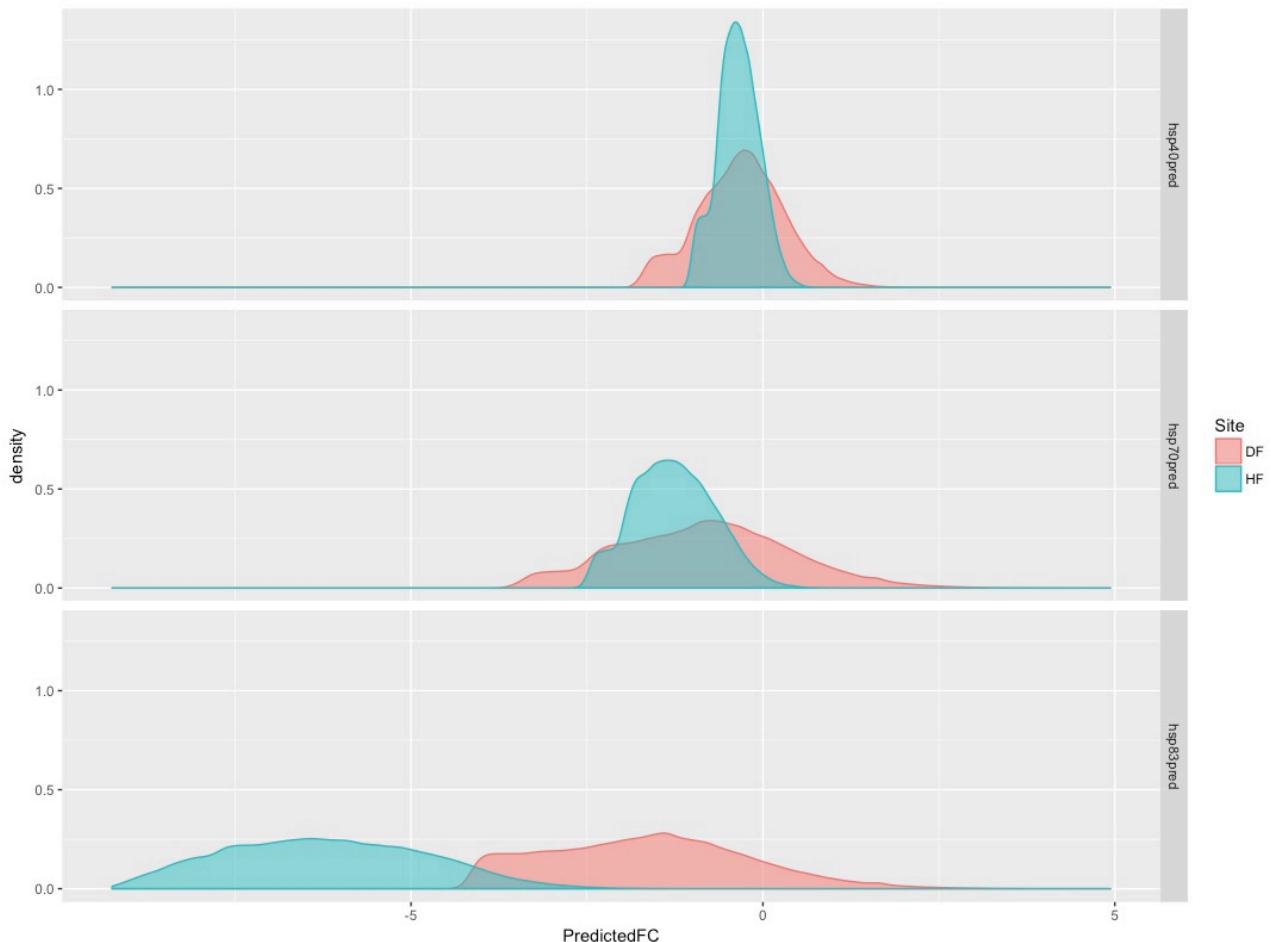
Density plot of "Stress"

```
1 ##histograms/desnity
2 ggplot(GSL.long,aes(x=PredictedFC,colour=Site,fill=Site)
) +geom_density(alpha=0.5)+facet_grid(gene~.)
```



plots with all bait temps experienced for 5 years(days are not averaged)





Meeting with SHC and NJG

- NJG: show figures with delta having no effect as the first figure
- Ignore IR data, use a correction factor. Plot censor vs bait temp
- SHC: consider taking residuals without julian day and then regression with julian day

Proteomics project:

- use quantile function to get tail probability
- shuffle colonies



Page 81: 2017-05-12. Stress in nature project: More Parsing chamber data

I have to relate chamber data with bait temps(which were taken with an IR gun).

Took gxp data set with year, julian day, bait temp, site, and deltas

```
1 names(chamcoll)<-
2   c("year", "doy", "baittemp.ave", "Site", "Delta")
3 str(chamcoll)
4 'data.frame': 206 obs. of 6 variables:
5   $ year           : int  2014 2014 2014 2013 2014 2013
6   2013 2013 2013 2013 ...
7   $ doy            : int  254 254 254 253 254 177 253 253
8   183 177 ...
9   $ baittemp.ave   : num  28.6 28.1 28.6 29.8 32.7 ...
10  $ Site           : Factor w/ 2 levels "DF", "HF": 1 1 1 1
11  1 2 1 1 1 2 ...
12  $ Delta          : num  3.5 0 0 3.5 5 0 0 2 0 4 ...
```

Merg with soil organic temp

data structure

```

1 str(cdlongSO)
2 'data.frame': 1130623 obs. of 10 variables:
3   $ datetime: Factor w/ 105782 levels "1/1/10 0:00",...
4   $ 1       : int 1 1 1 1 2 2 2 2 13 13 ...
5   $ doy     : int 1 1 1 1 1 1 1 1 1 1 ...
6   $ Cham    : chr "1" "2" "3" "4" ...
7   $ year    : num 2010 2010 2010 2010 2010 2010 2010
8   $ 2010    : num 2010 2010 2010 ...
9   $ SImin  : num 7.32 5.69 7.22 7.49 7.38 ...
10  $ Site   : chr "DF" "DF" "DF" "DF" ...
11  $ Delta   : num 3.5 0 4.5 2 3.5 0 4.5 2 3.5 0 ...
12  $ Cham2  : int 16 17 18 19 16 17 18 19 16 17 ...
13  $ envfac : chr "S0max" "S0max" "S0max" "S0max" ...
14  $ value   : num 10.2 5.96 10.16 9.39 9.99 ...

```

actual merger code

```

1 mergT<-
2   inner_join(cdlongSO,chamcoll,by=c("year","doy","Site",
3   "Delta"))
4 dim(mergT)
5 str(mergT)
6 'data.frame': 8352 obs. of 12 variables:
7   $ datetime      : Factor w/ 105782 levels "1/1/10
8   $ 0:00",...: 38795 38795 38795 38795 38795 38795 38795
9   $ 38795       : 38795 38795 38795 38795 38795 38795 38795 ...
10  $ doy          : int 183 183 183 183 183 183 183 183 183
11  $ 183          : 183 183 ...
12  $ Cham         : chr "1" "1" "2" "2" ...
13  $ year         : num 2013 2013 2013 2013 2013 ...
14  $ SImin        : num 22.6 22.6 21.6 21.6 21.6 ...
15  $ Site          : chr "DF" "DF" "DF" "DF" ...
16  $ Delta         : num 3.5 3.5 0 0 0 0 0 0 0 ...
17  $ Cham2        : int 16 16 17 17 17 17 17 17 17 ...
18  $ ...           : ...

```

```

13 $ envfac      : chr  "SOmax" "SOmax" "SOmax" "SOmax"
...
14 $ value       : num  24 24 22.5 22.5 22.5 ...
15 $ baittemp.ave : num  25.1 24.6 23.8 23.2 24.1 ...
16 $ bait_to_sensor: num  25.9 25.6 25.2 24.9 25.4 ...

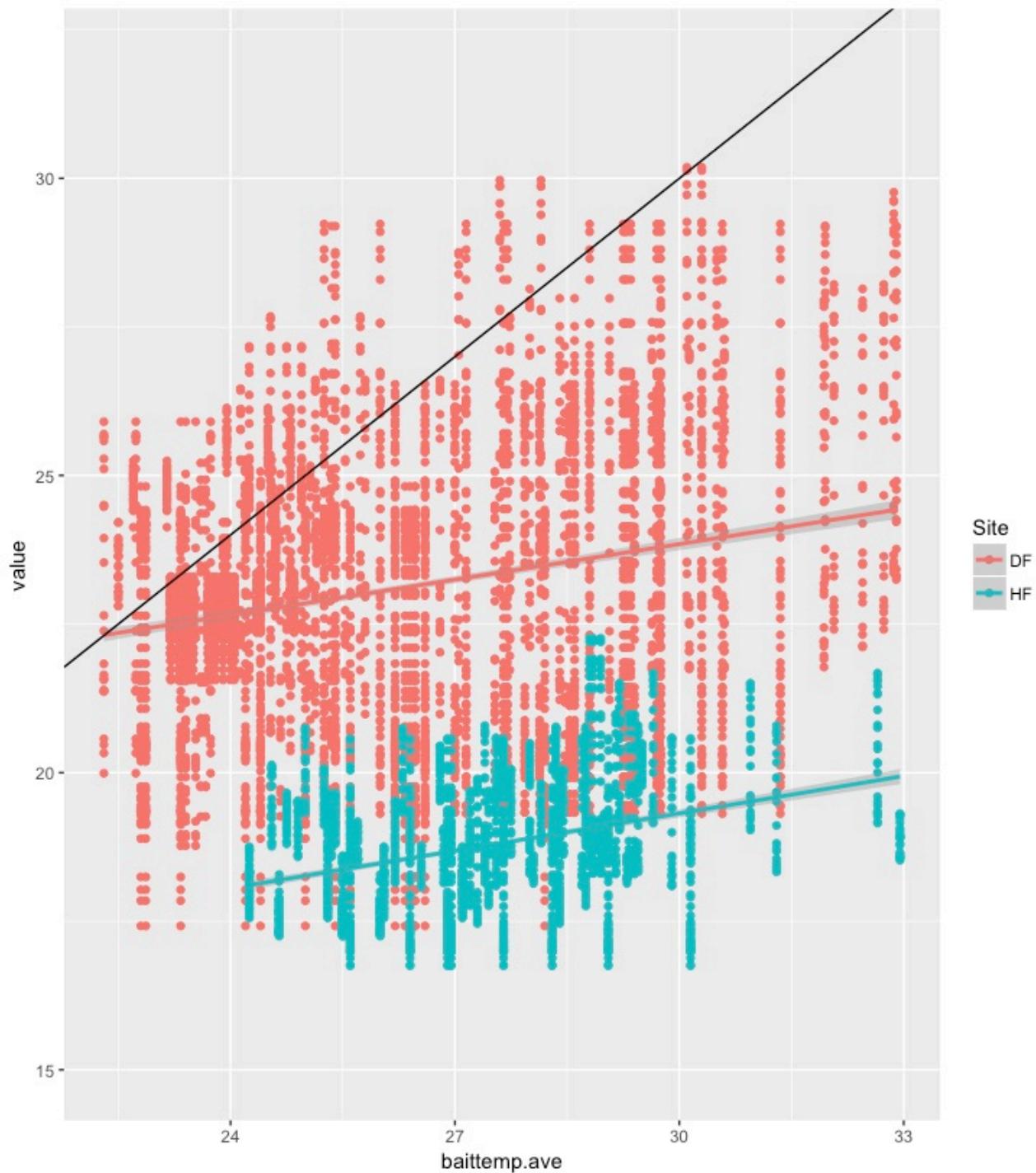
```

regression analysis

```

1 temppredmod<-lm(value~baittemp.ave+Site,mergT)
2 summary(temppredmod)
3 Call:
4 lm(formula = value ~ baittemp.ave + Site, data = mergT)
5
6 Residuals:
7   Min     1Q Median     3Q    Max
8 -6.0705 -1.0548  0.0664  1.1669  6.6001
9
10 Coefficients:
11                               Estimate Std. Error t value Pr(>|t|)
12 (Intercept)  17.820942    0.236386   75.39 <2e-16 ***
13 baittemp.ave  0.201048    0.008956   22.45 <2e-16 ***
14 SiteHF      -4.561699    0.046961  -97.14 <2e-16 ***
15 ---
16 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
17
18 Residual standard error: 1.952 on 8349 degrees of freedom
19 Multiple R-squared:  0.5307,    Adjusted R-squared:
20                      0.5306
21 F-statistic: 4720 on 2 and 8349 DF,  p-value: < 2.2e-
22          16

```



regressions when I take the average julian day for that temp.

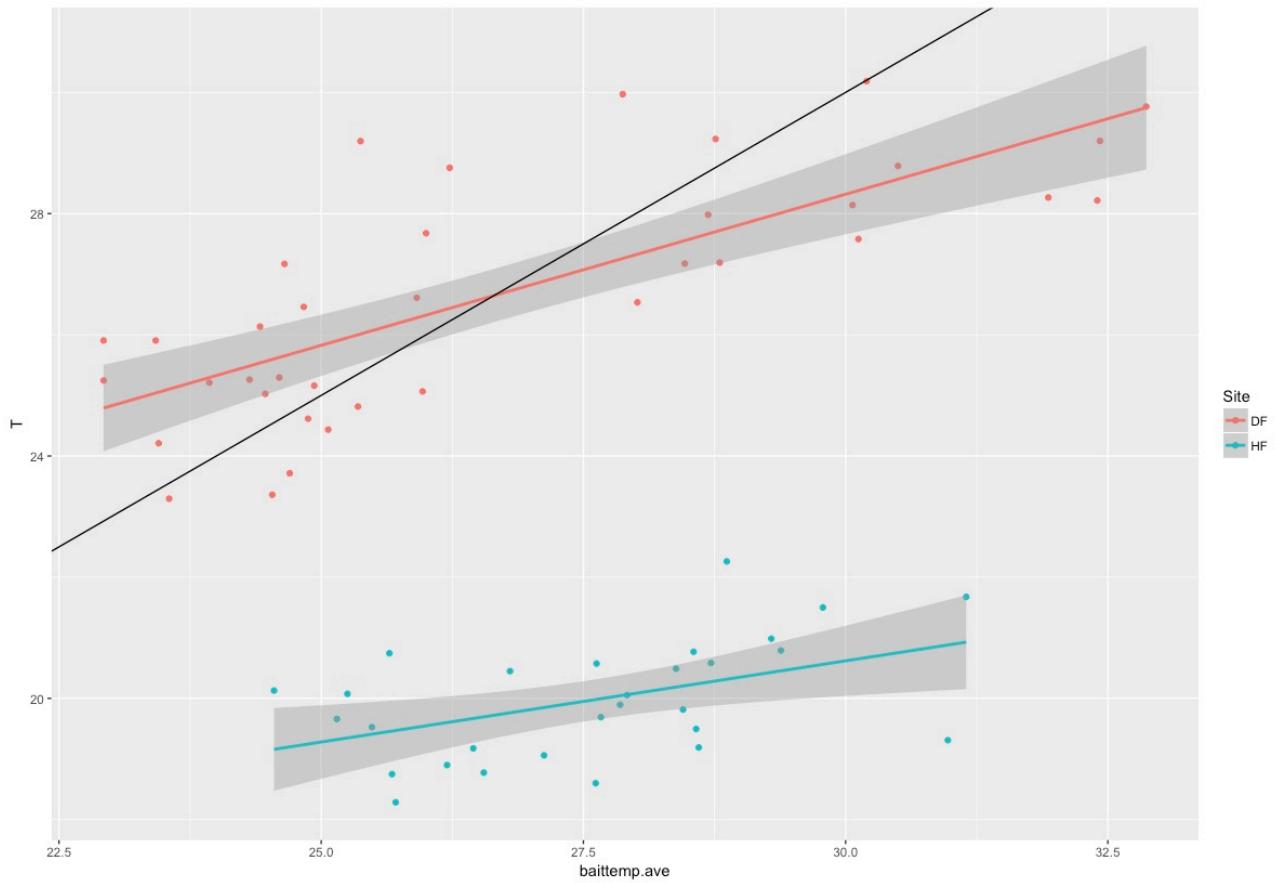
```

1 dayavemergT<-ddply(mergT,.
  (Site,doy,Delta,year),summarize,T=max(value),baittemp.av
  e=mean(baittemp.ave))
2

```

Fitting model to test effect of site by bait temp interaction on chamber sensor soil organic max temperature

```
1 temppredmod<-lm(T~baittemp.ave+Site,dayavemergT)
2 summary(temppredmod)
3
4 Call:
5 lm(formula = T ~ baittemp.ave + Site, data =
6 dayavemergT)
7
8 Residuals:
9   Min       1Q   Median       3Q      Max
10 -2.33662 -0.72673 -0.05928  0.57283  3.11921
11
12 Coefficients:
13             Estimate Std. Error t value Pr(>|t| )
14 (Intercept) 14.64434   1.55389   9.424 1.20e-13 ***
15 baittemp.ave  0.45050   0.05778   7.797 8.09e-11 ***
16 SiteHF       -7.09908   0.29178 -24.330 < 2e-16 ***
17 ---
18 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 
19 0.1 ' ' 1
20
21 Residual standard error: 1.158 on 63 degrees of freedom
22 Multiple R-squared:  0.9056,    Adjusted R-squared:
23                      0.9026
24 F-statistic: 302.2 on 2 and 63 DF,  p-value: < 2.2e-16
```



OK projecting predictions: Predict sensor dating given the site and bait temp measured in the field

Redoing stats with bait to sensor temperatures

global model

```

1 fullmod5<-
  lme(FC~RIN_Value+Jdaycont+gene*Site*bait_to_sensor+gene
  *Site*Delta,random=~1|Cham2/Vial.me,data=findat.long,me
  thod="ML")
2 aicmod1<-
  anova(summary(stepAIC(fullmod5,direction="both")))

```

```

3 aicmod1
4
5      numDF denDF   F-value p-value
6 (Intercept)          1    413  0.044025  0.8339
7 Jdaycont            1    413  4.633789  0.0319
8 gene                 2    413  0.000000  1.0000
9 Site                 1     23 21.076861  0.0001
10 bait_to_sensor      1    413 29.836210 <.0001
11 Delta                1     23  0.169841  0.6841
12 gene:Site            2    413  1.682588  0.1872
13 gene:bait_to_sensor  2    413 10.550810 <.0001
14 Site:bait_to_sensor 1    169  1.819053  0.1792
15 Site:Delta            1     23  2.692626  0.1144

```

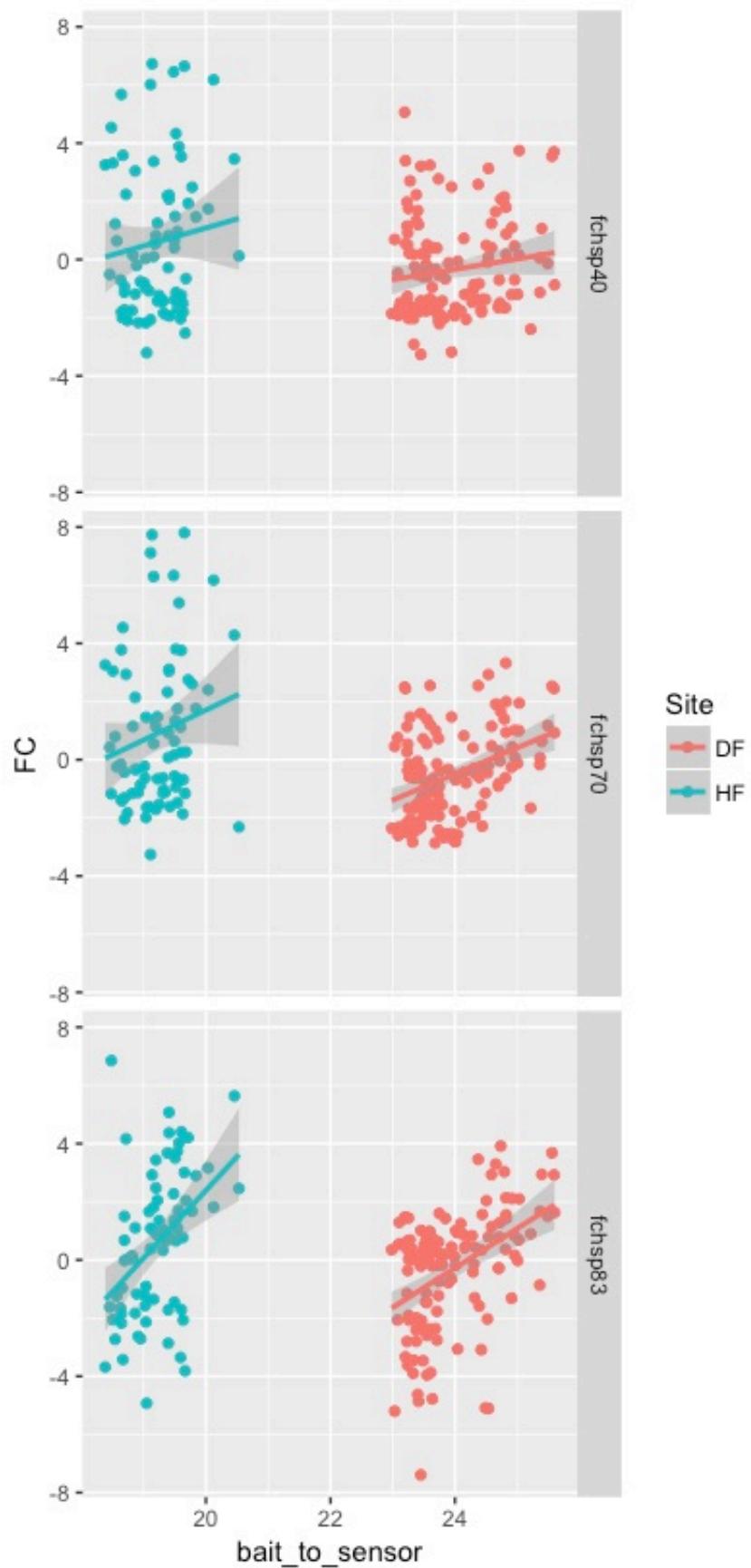
Similar results: gene by temp(sensor) interaction. Site effect

figures by each gene

```

1 ggplot(findat.long,aes(x=bait_to_sensor,y=FC,colour=Site
 ))+geom_point()+stat_smooth(method="lm")+facet_grid(gene
 ~ .)

```



Regression models for each gene:

Hsp83

```
1 fullmod7<-
  lme(FC~RIN_Value+Jdaycont+Site*bait_to_sensor+Site*Delta,
  random=~1|Cham2/Vial.me,data=hsp83,method="ML")
2 fullmod7sel<-
  summary(stepAIC(fullmod7,direction="both"))
3 fullmod7sel
4 Fixed effects: FC ~ RIN_Value + Jdaycont + Site *
  bait_to_sensor + Site * Delta
5
6             Value Std.Error DF t-
7             value p-value
8 (Intercept) -30.817003 5.635039 169
9 -5.468818 0.0000
10 RIN_Value 0.149878 0.082116 6
11 1.825203 0.1178
12 Jdaycont 0.002731 0.001028 6
13 2.655774 0.0377
14 SiteHF -18.194313 12.534387 23
15 -1.451552 0.1601
16 bait_to_sensor 1.196000 0.237310 6
17 5.039819 0.0024
18 Delta 0.008764 0.091198 23
19 0.096100 0.9243
20 SiteHF:bait_to_sensor 1.334430 0.637716 169
21 2.092515 0.0379
22 SiteHF:Delta -0.292487 0.166271 23
23 -1.759101 0.0919
```

hsp70

```

1 fullmod6<-
2   lme(FC~RIN_Value+Jdaycont+Site*bait_to_sensor+Site*Delt
3     a,random=~1|Cham2/Vial.me,data=hsp70,method="ML")
4 fullmod6sel<-
5   summary(stepAIC(fullmod6,direction="both"))
6 fullmod6sel
7 Fixed effects: FC ~ Jdaycont + Site + bait_to_sensor
8                               Value Std.Error DF t-value p-
9 value
10 (Intercept)      -26.860864  4.816424 170 -5.576931
11 0e+00
12 Jdaycont        -0.004589  0.000657    7 -6.981999  2e-
13 04
14 SiteHF          6.745361  0.987155   25  6.833136
15 0e+00
16 bait_to_sensor  1.172801  0.202485    7  5.792038  7e-
17 04
18 Correlation:

```

hsp40

```

1 hsp40<-subset(findat.long,findat.long$gene=="fchsp40")
2 #fullmod8<-
3 lme(FC~RIN_Value+Jdaycont+Site*baittemp.ave+Site*Delta,
random=~1|Cham2/Vial.me,data=hsp40,method="ML")
4 fullmod8<-
5 lme(FC~RIN_Value+Jdaycont+Site*bait_to_sensor+Site*Delta,
random=~1|Cham2/Vial.me,data=hsp40,method="ML")
6 fullmod8sel<-
7 summary(stepAIC(fullmod8,direction="both"))
8 fullmod8sel
9
10 Fixed effects: FC ~ Jdaycont + Site + bait_to_sensor
11                               Value Std.Error DF t-value p-
12                               value
13 (Intercept)      -12.982176  5.435435 170 -2.388434
14               0.0180
15 Jdaycont        -0.003251  0.000750     7 -4.336718
16               0.0034
17 SiteHF          3.524131  1.110366    25  3.173847
18               0.0040
19 bait_to_sensor   0.579430  0.228556     7  2.535174
20               0.0389

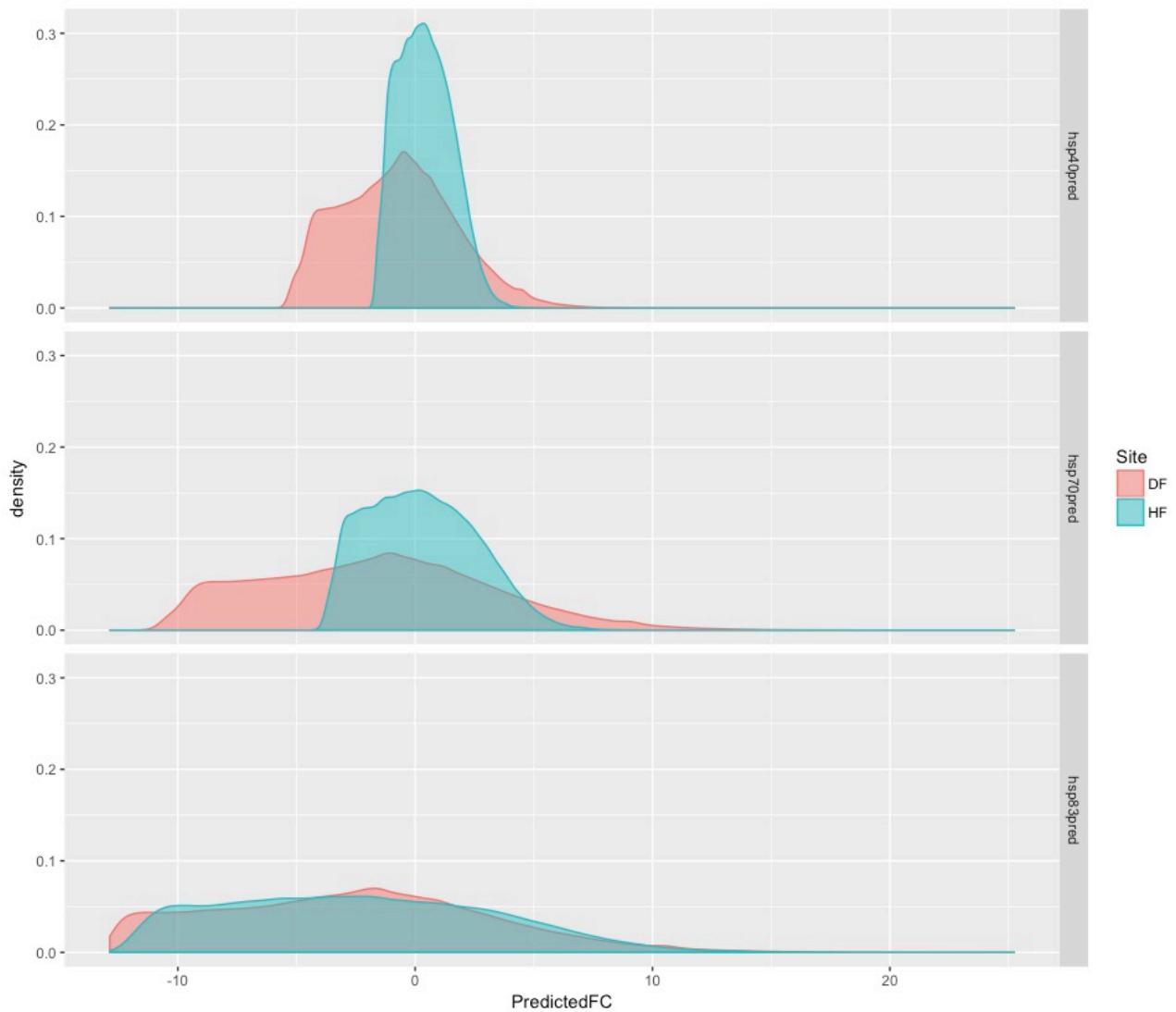
```

OK, new sets of predictions: Predict every time of day for every year

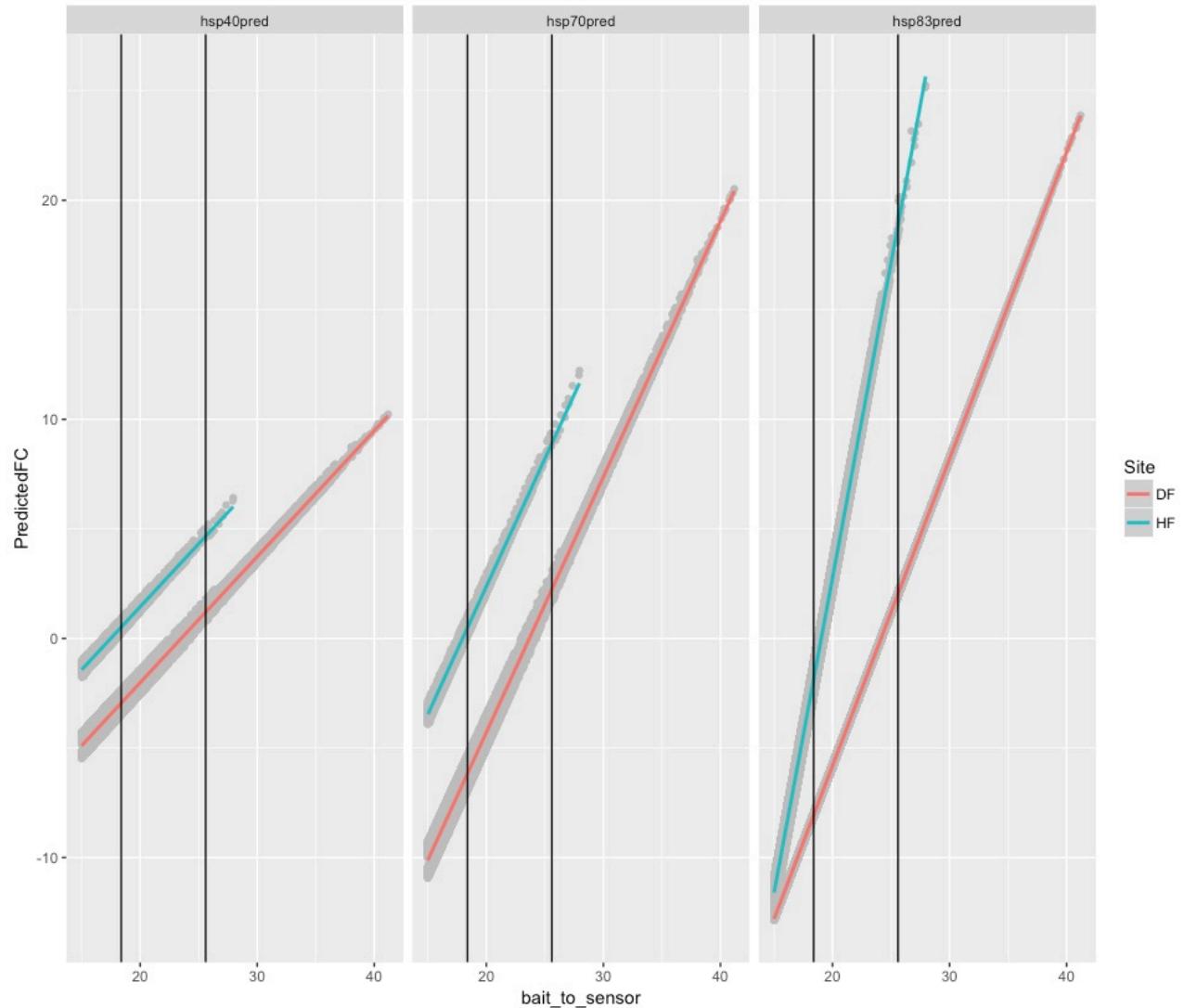
GSL data, defined as temps throughout the year above 15C for soil organic sensor temp (max):

```
1 str(GSL)
2 'data.frame': 500848 obs. of 10 variables:
3   $ X      : int 5687362 5687364 5687366 5687368
4     5687370 5687372 5687376 5687456 5687460 5687464 ...
5   $ datetime: Factor w/ 50384 levels "1/1/11 11:00",...
6     135 135 136 136 137 137 138 142 143 144 ...
7   $ doy     : int 18 18 18 18 18 18 18 19 19 19 ...
8   $ Cham    : int 1 3 1 3 1 3 3 3 3 3 ...
9   $ year    : int 2010 2010 2010 2010 2010 2010 2010
10    2010 2010 2010 ...
11   $ Site    : Factor w/ 2 levels "DF", "HF": 1 1 1 1 1 1
12     1 1 1 ...
13   $ Delta   : num 3.5 4.5 3.5 4.5 3.5 4.5 4.5 4.5 4.5 4.5
14     4.5 ...
15   $ Cham2   : int 16 18 16 18 16 18 18 18 18 18 ...
16   $ envfac  : Factor w/ 1 level "SOmax": 1 1 1 1 1 1 1 1
17     1 1 ...
18   $ value   : num 17.1 18.4 16 17.8 16.1 ...
```

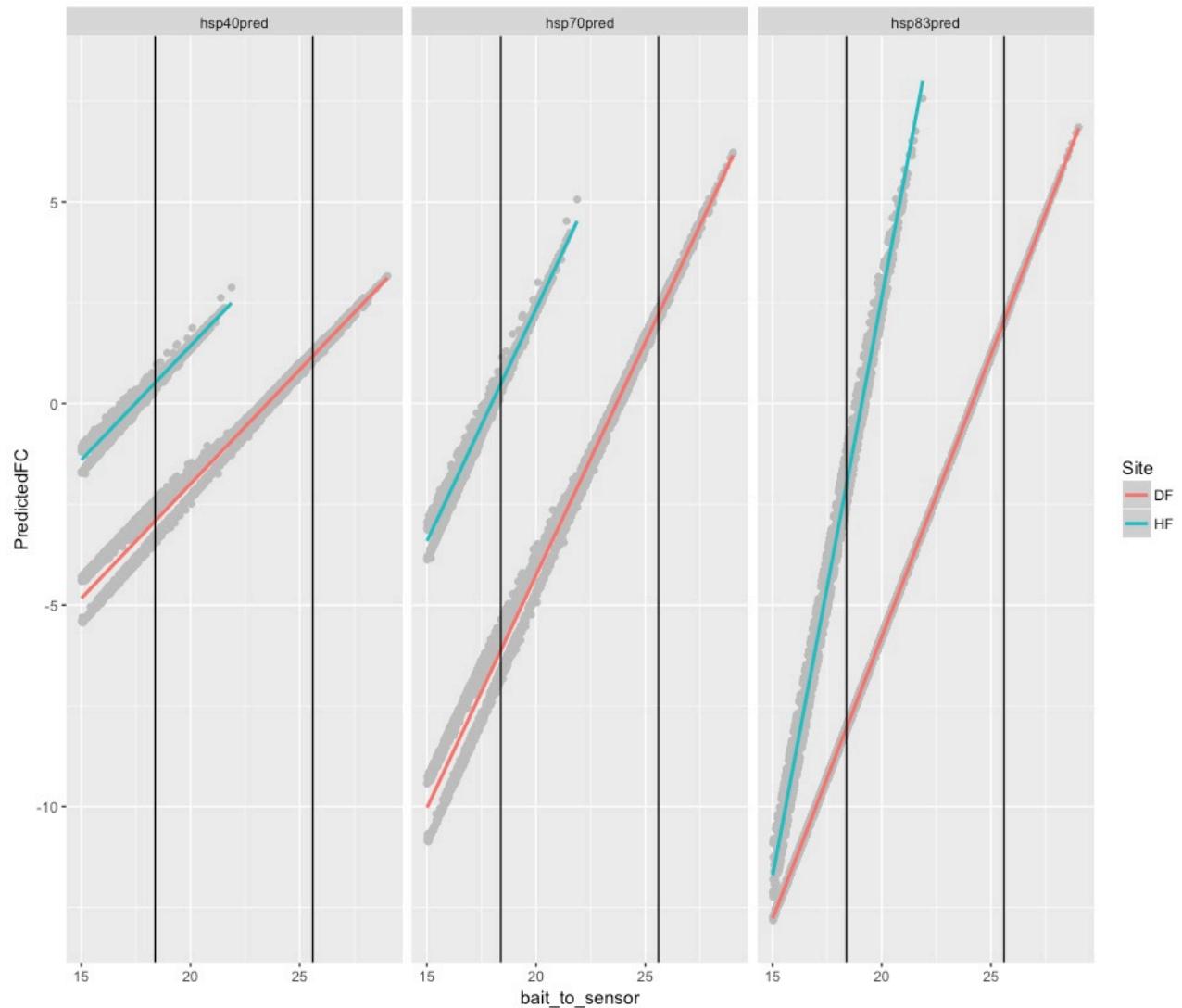
Density plots: Predict every time of day for every year

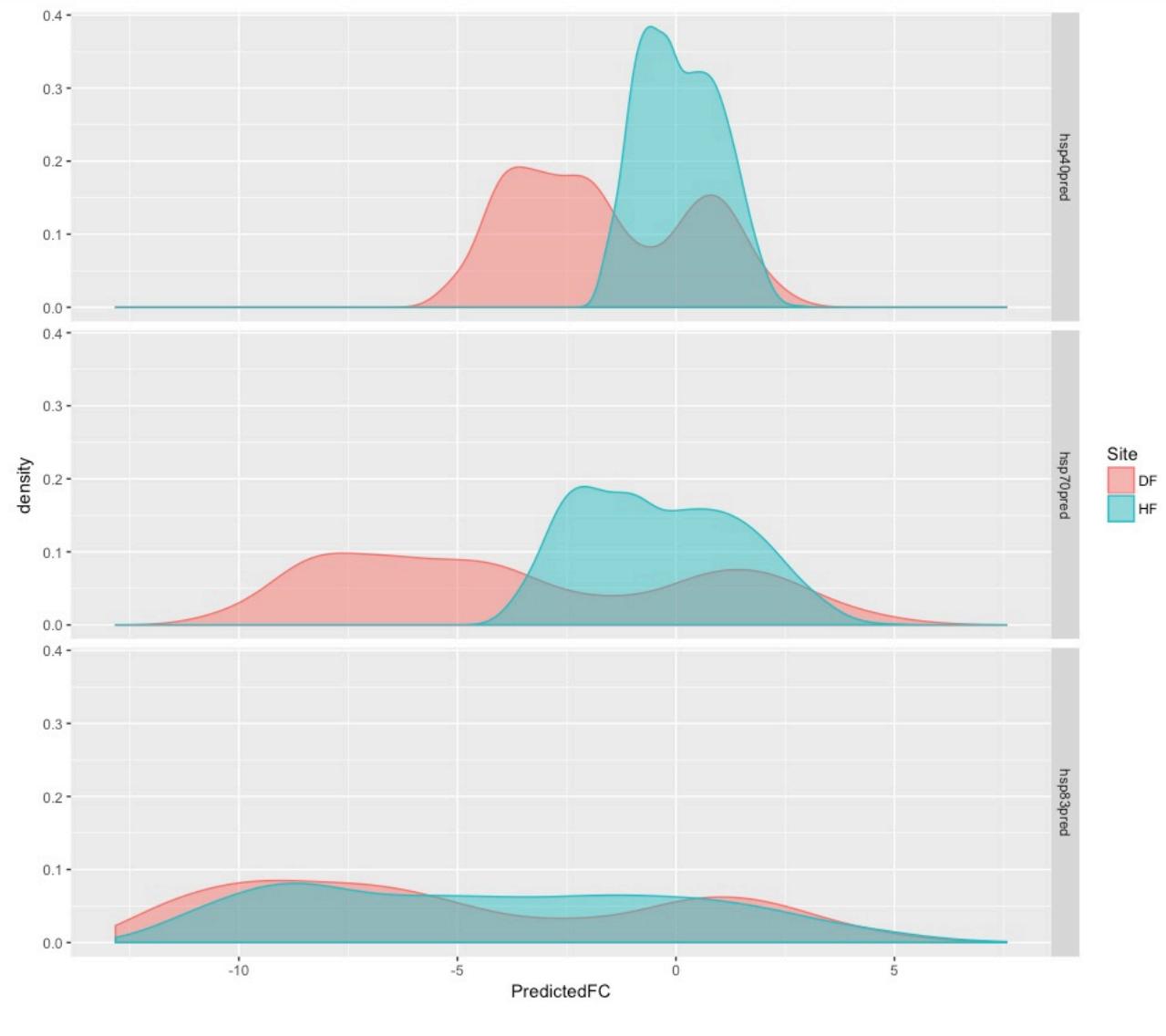


Checking the temperature range of where we constructed our model to temperature predictions of fold change (FC).



Predict every day temp, averaged for each year and time of day





Page 82: 2017-05-16. Removing the history of large files in github through terminal

Link <http://stackoverflow.com/questions/16877530/completely-remove-a-file-from-whole-git-repository>

```

1 $ git filter-branch --tree-filter 'rm -f passwords.txt'
HEAD
2 Rewrite 6b9b3cf04e7c5686a9cb838c3f36a8cb6a0fc2bd (21/21)
3 Ref 'refs/heads/master' was rewritten

```

Page 83: 2017-05-16 Stress in nature project: meeting with SHC

Convert bait temp to sensor temperature by the timing of sampling

Also, check the IR gun used for HF to see if it is overestimating temperatures.

2017-05-18: Meeting with SHC

Checked IR gun with thermometer:

1. IR gun
 - Ice = 0.2-0.4
 - Aphaenogaster Benchtop = 20.2
 - Pogonomyrmex benchtop = 28.3
2. thermometer
 - Ice = 0.2
 - Aphaenogaster Benchtop = 18.8
 - Pogonomyrmex benchtop = 27.3

At ice, they are similar, but not at higher temperatures. Real test would be to compare IR guns...

More data to explore

- Make figures similar to Kingsolver et al. 2013 Functional Ecology:
Measure thermal safety margin(thermal buffer, SDiamond prefers this!!) across the whole season!
 - I need CTmax. Where to get it? From Lacy's field survey
 - Make figure of density of growing season length!
-

Page 84: 2017-05-22. Meeting with SHC and NJG, stressed in nature project

Go over figures:

Make outline with key take homes in a narrative form. Create abstract.

1. create a figure with sites and chambers and plots. Mimick Diamond (green and orange) Science advances pictures. Have bait card inset photo in figure.
2. figure 2, y axis --take out calcs, and then log2(fold change). X -axis: "Experimental Temperature Increase","Experimental warming amount".
3. figure 3. Density plots for predicted log2(fold change). Insets with summed gxp values.
4. Anova tables. Does functional ecology let anova tables in?
- 5.

hsp gxp rxn

1. for figure 1, place C past D to show active vs passive mechanisms
2. get rid of fig 2 phylogeny and just have panel B
3. fig 2 c move to fig 3. Fig 3, make numbers more readable and make shorten width.

4. fig 4, move stress resistance to the end to keep it consistent with fig
1. Make fig 4 into 3 panel fig- top with 12 panels of ctmax vs gxp.
Bottom rows split between pc biplot and then regression with
ctmax and and pc1.
5. getr id of fig s2.
6. move fig s4 (see point # 4).

2017-05-23: figure insets:

cool stack overflow thread

<https://stackoverflow.com/questions/5219671/it-is-possible-to-create-inset-graphs>

```

1 require(ggplot2)
2
3 plot1 = qplot(1.00*mpg, 1.00*wt, data=mtcars) # Make
4   sure x and y values are floating values in plot 1
5 plot2 = qplot(hp, cyl, data=mtcars)
6 plot(plot1)
7
8 # Specify position of plot2 (in percentages of plot1)
9 # This is in the top left and 25% width and 25% height
10 xleft    = 0.05
11 xright   = 0.30
12 ybottom  = 0.70
13 ytop     = 0.95
14
15 # Calculate position in plot1 coordinates
16 # Extract x and y values from plot1
17 l1 = ggplot_build(plot1)
18 x1 = l1$panel$ranges[[1]]$x.range[1]
19 x2 = l1$panel$ranges[[1]]$x.range[2]
20 y1 = l1$panel$ranges[[1]]$y.range[1]
21 y2 = l1$panel$ranges[[1]]$y.range[2]
22 xdif = x2-x1

```

```
22 ydif = y2-y1
23 xmin = x1 + (xleft*xdif)
24 xmax = x1 + (xright*xdif)
25 ymin = y1 + (ybottom*ydif)
26 ymax = y1 + (ytop*ydif)
27
28 # Get plot2 and make grob
29 g2 = ggplotGrob(plot2)
30 plot3 = plot1 + annotation_custom(grob = g2, xmin=xmin,
31                                     xmax=xmax, ymin=ymin, ymax=ymax)
32 plot(plot3)
33
34
35
36 # Try and make a weird combination of plots
37 g1 <- ggplotGrob(plot1)
38 g2 <- ggplotGrob(plot2)
39 g3 <- ggplotGrob(plot3)
40
41 require("gridExtra")
42 require("grid")
43
44
45 t1 = arrangeGrob(g1, ncol=1, left = textGrob("A", y = 1,
46                                         vjust=1, gp=gpar(fontsize=20)))
47 t2 = arrangeGrob(g2, ncol=1, left = textGrob("B", y = 1,
48                                         vjust=1, gp=gpar(fontsize=20)))
49 t3 = arrangeGrob(g3, ncol=1, left = textGrob("C", y = 1,
50                                         vjust=1, gp=gpar(fontsize=20)))
51
52 final = arrangeGrob(t1,t2,t3, layout_matrix =
53                      cbind(c(1,2), c(3,3)))
54 grid.arrange(final)
```

```
52 | ggsave(filename = "test2.png", plot = final)
```

Page 85: 2017-04-24. Harvard + Duke Forest chamber info

```
[1] "| | Cham| Longitude| Latitude| Site | Delta| Cham2| col |" [2] "|:--  
|----:|-----:|-----:|----:|----:|-----|" [3] "|2 | 2| -79.07716|  
36.03567|DF | 0.0| 17|#FF8C00 |" [4] "|5 | 5| -79.07718|  
36.03559|DF | 0.0| 20|#FA8F0D |" [5] "|11 | 11| -79.07739|  
36.03538|DF | 0.0| 26|#F5931B |" [6] "|13 | 13| -79.07692|  
36.03547|DF | 0.0| 28|#F19628 |" [7] "|14 | 14| -79.07692|  
36.03532|DF | 0.0| 29|#EC9A36 |" [8] "|15 | 15| -79.07726|  
36.03527|DF | 0.0| 30|#E79D43 |" [9] "|6 | 6| -79.07710|  
36.03554|DF | 1.5| 21|#E3A151 |" [10] "|4 | 4| -79.07719|  
36.03574|DF | 2.0| 19|#DEA55F |" [11] "|10 | 10| -79.07732|  
36.03542|DF | 2.5| 25|#D9A86C |" [12] "|7 | 7| -79.07731|  
36.03558|DF | 3.0| 22|#D5AC7A |" [13] "|1 | 1| -79.07727|  
36.03570|DF | 3.5| 16|#D0AF87 |" [14] "|12 | 12| -79.07717|  
36.03541|DF | 4.0| 27|#CBB395 |" [15] "|3 | 3| -79.07700|  
36.03571|DF | 4.5| 18|#C7B6A2 |" [16] "|8 | 8| -79.07728|  
36.03552|DF | 5.0| 23|#C2BAB0 |" [17] "|9 | 9| -79.07749|  
36.03554|DF | 5.5| 24|#BEBEBE |"
```

from KMiller:

1 through 12 were chambers, 13 through 15 didn't have chambers, the chambers (in 1-12) which had temperature treatments that were 0 (i.e. control chambers) were as follows: HF = chambers 4, 6, and 11; DF = chambers 2, 5, and 11.

Page 86: 2017-05-30. Multi Multi-panel figures in R

(<https://stackoverflow.com/questions/30299529/ggplot2-define-plot-layout-with-grid-arrange-as-argument-of-do-call>)

[<https://stackoverflow.com/questions/30299529/ggplot2-define-plot-layout-with-grid-arrange-as-argument-of-do-call>]

```
1 require(ggplot2)
2 require(gridExtra)
3
4 df <- data.frame(value1 = rnorm(200),
5                     value2 = rnorm(200),
6                     value3 = rnorm(200),
7                     value4 = rnorm(200))
8
9 p1 <- ggplot(df) + geom_density(aes(x=value1))
10 p2 <- ggplot(df) + geom_density(aes(x=value2))
11 p3 <- ggplot(df) + geom_density(aes(x=value3))
12 p4 <- ggplot(df) + geom_density(aes(x=value4))
13
14 grid.arrange(p1, arrangeGrob(p2,p3,p4, ncol=3),
heights=c(2.5/4, 1.5/4), ncol=1)
```

Page 87: 2017-06-01. Protein stability thoughts

Trying to understand the relationship between stability (3D structure of protein) and kinetics (how a reaction is occurring). Both should be related (structure + function).

Reading Brent and George's paper:

Lockwood BL, Somero GN. 2012. Functional Determinants of Temperature Adaptation in Enzymes of Cold- versus Warm-Adapted Mussels (Genus *Mytilus*). *Mol Biol Evol* 29:3061–3070.

Specifically, I'm looking at figure 2, where they find adaptive variation in Km for **IDH** between cool (*M. trossulus*) vs warm adapted (*M. galloprovincialis*) mussel. There was no differences in Km between the two species for PK or PEPCK.

1. I think the relationship between Km (substrate in units of uM or M) and temperature pretty interesting.
 - Km is lower at lower temperatures.
 - In other words, Substrate-enzyme affinity is higher at lower temperatures.
 - This relationship differs between enzymes. Why?
 - Do they differ between reversible and non-reversible reactions?
2. Warm adapted species has lower Km than cool adapted species for IDH. So, under higher temperatures, the warm adapted species' IDH enzyme can still bind to substrate. This may be related to the stability between these enzymes.

What is Kcat? It is the turnover number: expressed as the number of molecules turned into product per minute.

Kcat can be calculated as Vmax/enzyme concentration.

Page 88: 2017-06-06. Hsp gxp rxn norm paper: writing notes

To help write my ms better, here is a list of the updated figures and supplemental figs:

Body Figures

1. Predictions figure: Fold induction vs temperature for each parameter
2. Non-phylogenetic controlled analysis: CTmax vs environment
3. Phylogenetic analysis: ancestral trait reconstruction and variance partitioning
4. CTmax vs hsp params
5. hsp params vs habitat type
6. Methods figure: example of fitting boltzmann to hsp gxp

Supplemental

- Fig. S1. - map of collection sites with pie charts showing sample size
- Fig. S2. - CTmax mapped onto phylogeny with colonies. So phylogeny + barplot with CTmax values for all colonies
- Fig. S3. - boxplot of habitat(yaxis) vs each gxp parameter
- Fig S4. - boxplot of hsp pca1 (yaxis) vs forest type

Ok, now tables:

- Table S1. Variance partitioning table at species and colony level
- Table S2. Regression models: relationship between CTmax and hsp parameters
- Table S3. Regression models: relationship between CTmax and PC1
- Table S4. Regression models: relationship between hsp parameters

and environment

I need to test the significance of variance components for phylo and local env in vegan package

Reference code for me to work off of

```
1 #testing the significance for each fraction
2 #fraction a+ b + c:
3 rda.all2<-
4   rda(Aph.dat$KO_temp_worker~bio1+bio5+habitat_v2+Axis.1
5     + Axis.2+ Axis.3
6     +Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9,data=Aph.dat
7   )
8 anova(rda.all2)
9
10 #fraction a: Phylogeny
11 rda.phy_eco2<-rda(Aph.dat$KO_temp_worker~ Axis.1 +
12   Axis.2+ Axis.3
13   +Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9+Condition(bi
14   o1+bio5+habitat_v2),data=Aph.dat)
15 anova(rda.phy_eco2)
16
17 #fraction c: Ecology
18 rda.eco_phy2<-
19   rda(Aph.dat$KO_temp_worker~bio1+bio5+habitat_v2+Conditi
20   on(Axis.1 + Axis.2+ Axis.3
21   +Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9),data=Aph.da
22   t)
23 anova(rda.eco_phy2)
24
25 ## a + b: All phylogeny
```

```

16 rda.phy2<-rda(Aph.dat$KO_temp_worker~Axis.1 + Axis.2+
  Axis.3
  +Axis.4+Axis.5+Axis.6+Axis.7+Axis.8+Axis.9,data=Aph.dat
  )
17 anova(rda.phy2)
18
19 ## b+ c: All ecology
20 rda.eco2<-
  rda(Aph.dat$KO_temp_worker~biol1+bio5+habitat_v2,data=Aph.dat)
21 anova(rda.eco2)

```

Variance partitioning: summary table

	Species			Colony		
Variance Component	Variance	F-statistic	P-value	Variance	F-statistic	P-value
Independent Phylogeny	0.11	F2,3= 2.5	0.22	0.04	F3,87=1.84	0.61
Independent Environment	0.25	F2,3= 4.5	0.11	0.00	F9,87=1.84	0.09
Phylogeny	0.57	F2,5= 5.65	0.08	0.47	F9,90=10.83	< 0.005
Environment	0.71	F2,5= 9.67	< 0.05	0.42	F3,96=25.03	< 0.005
Phylogeny and Environment	0.46	NA		0.43	NA	
Full model	0.82	F4,3=9.02	0.068	0.47	F12,87=8.14	< 0.001
Residuals	0.18			0.53		

Restructuring results:

1. summary stats (range of ctmax range, phylogeny, Tmax range)
2. species level analyses of CTmax- var part, PGLS
3. colony level random effects model of CTmax
4. Ctmax and hsp params

New body figure layout:

1. Predictions figure: Fold induction vs temperature for each

- parameter
2. Non-phylogenetic controlled analysis: CTmax vs environment
 3. Phylogenetic analysis: ancestral trait reconstruction and variance partitioning
 4. CTmax vs hsp params
 5. hsp params vs habitat type
 6. Methods figure: example of fitting boltzmann to hsp gxp
-

Page 89: 2017-06-20. Project idea: Viability selection on ant founding queens

Natural selection can act upon the individuals at the beginning of their lifestages prior to sexual maturity, also known as viability selection. This is known as the "invisible fraction" of selection because people often times don't measure this (Mojica and Kelly 2010). The greatest opportunity for viability selection in ants are when queens mate and found new colonies. We assume 99% of these queens die! Furthermore, for *Aphaenogaster picea* nad *rudis*, there seems to be no population structure, suggesting that they are panmictic or they are a 1 randomly mating population. So how should colonies invest in queens? In warmer climates, queens should be smaller (temperature size rule) than cooler climates. But queen size is important for surviving colony founding. So, queen size could positively match the local thermal environment due to spatially variable selection. Meaning, all colonies may produce smaller-larger queens, but the ones that are mismatched don't "make" it.

I should read more in plants bc the life history is the same.

Mojica and Kelly 2010 found evidence for viability selection in mimulus. Large flowers had lower viability but higher fecundity than small flowers nad was associated with a QTL.

ref: Mojica JP, Kelly JK. 2010. Viability selection prior to trait expression is an essential component of natural selection. Proceedings of the Royal Society B: Biological Sciences 277:2945–2950.

In ants, you expect something else: larger ants, more fecund and viable.

In ants, a cool experiment would be:

1. Collect 150 new mated queens from 10 sites from whole range. Clip leg to genotype?
2. Divy up 50 queens into 3-6 temperature treatments per site
3. Measure initial size, weight?
4. Track survival, measure dry weight, morphometrics.
5. Do a GWAS for viability selection.
6. assay foragers

Some questions to think about:

1. What is the seasonal timing of nuptial flights in ants? Take a site and measure how different species mate throughout the growing season?
 - 2.
-

Page 90: 2017-07-10. Meeting with Dan Hahn

A few notes:

1. Meeting this Friday to chat about papers and come up with a to do

list.

2. Shadow Chinwin and Chao to see how to work with Eastern Cornborer/ Rhagoletis. I need to gain hands on skills with working with the organisms
3. Check out the science individual development plan---DH wants done by August 1st. (<http://myidp.sciencecareers.org/>)
4. Make a mendeley reading group that anybody can access. Nice thing is we can add our own notes
5. Go over data management practices
6. Lab Safety training? Chao showed me around lab, I have to sign something

For field sampling:

- Plan to go out into the field August-September; depends on fruiting time
- Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois
 - Late in season, collect around Florida because there are many Hawthorne varieties

Page 91: 2017-07-11. Reading Rhagoletis papers:

refs:

1. Feder, J. L., Stolz, U., Lewis, K. M., Perry, W., Roethel, J. B., & Rogers, A. (1997). THE EFFECTS OF WINTER LENGTH ON THE GENETICS OF APPLE AND HAWTHORN RACES OF RHAGOLETIS POMONELLA (DIPTERA: TEPHRITIDAE). *Evolution*, 51(6), 1862–1876.

<http://doi.org/10.1111/j.1558-5646.1997.tb05109.x>

2. Feder, J. L., Roethel, J. B., Wlazlo, B., & Berlocher, S. H. (1997). Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly (sympatric speciation—fitness trade-offs—host plant phenology—*Rhagoletis pomonella*—host x environment interactions). *Evolution*, 94, 11417–11421. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC23485/pdf/pq011417.pdf>
3. Filchak, K. E., Feder, J. L., Roethel, J. B., Stolz, U., & Roethel, J. B. (1999). A Field Test for Host-Plant Dependent Selection on Larvae of the Apple Maggot Fly, A FIELD TEST FOR HOST-PLANT DEPENDENT SELECTION ON LARVAE OF THE APPLE MAGGOT FLY, *RHAGOLETIS POMONELLA*. Source: *Evolution Evolution*, 53(531), 187–200. Retrieved from <http://www.jstor.org/stable/2640931>
4. Feder, J. L., & Filchak, K. E. (1999). It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. In *Proceedings of the 10th International Symposium on Insect-Plant Relationships* (pp. 211–225). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-017-1890-5_27

Feder, J. L., Stolz, U., Lewis, K. M., Perry, W., Roethel, J. B., & Rogers, A. (1997). THE EFFECTS OF WINTER LENGTH ON THE GENETICS OF APPLE AND HAWTHORN RACES OF *RHAGOLETIS POMONELLA* (DIPTERA: TEPHRITIDAE). *Evolution*, 51(6), 1862–1876.

<http://doi.org/10.1111/j.1558-5646.1997.tb05109.x>

- Host plant associated fitness trade-offs --compound noun that is difficult to understand

Goal: determine effects of winter length on the genetics of apple maggot flies

How this relates to trade-offs, not sure and not clear.

Their observation: apple trees fruit earlier and for apple maggot flies to exploit this fruit, they must survive extended winter conditions

Apple maggots are facultative diapausers

But longer winters should select for slower metabolism and longer diapause duration.

Experimental design: Took fruit from apples and hawthorne and incubated at 4C from 1 week to 2 years. They measured survival and # of days to eclosion

Predictions: longer cold periods will select against allozyme markers associated with faster dev rates (typically found in hawthorne)

Result: longer cold periods retained allele frequencies that were more apple like(longer dev time)

Conclusion: Before and during winter selects against certain host races of apple maggots.

background:

Models of sympatric speciation for plant feeding insects:

1. mating on that specific plant -this can act as a pre-zygotic barrier
2. performance on the plant--this point can act as a post zygotic barrier (Dethier 1954 citation)

Trade-offs in plant performance can manifest in the plant phenology and insect development.

Insects must match their life history timing with plant phenology, otherwise they can starve.

Rhagoletis is a good system to understand trade-offs in plant performance

Apple and Hawthorne races are genetically differentiated at 6 allozymes that map to 3 genomic regions.

linkage group:

1. Aat-2, Dia-2
2. Acon-2, Mpi
3. Had

Linkage disequilibrium between loci, but not among loci.

Problem: fitness trade offs is a missing piece of the story. For reciprocal transplants, there are no differences in feeding specialization.

Natural history notes: Apple fruits peak 3 weeks earlier than hawthorne. Eclosion is 10 days earlier. Apple flies overwinter ~1 week longer than hawthorn flies.

Diapause trade-off hypothesis: differential selection such that hawthorne race develops faster (shorter diapause) and apple flies develop slower (longer diapause).

Diapause duration is shaped by plant phenology.

Predictions:

1. there should be an optimal overwintering period
 - long enough to allow for diapause to end

- short enough to prevent energy loss as pupae
2. apple flies should survive longer winters and eclose later than hawthorn flies under common conditions
 3. earlier eclosing adults should have higher allele freq of:
 1. Me 100
 2. Acon-2 95
 3. Mpi 37
 4. Aat-2 +75
 5. Dia-2 100
 6. Had 100
 7. IDH should have no pattern
 4. short overwintering periods should favor these alleles and should be associated with earlier eclosion.

Results

Figure 1: Prob could be represented as a cumulative density function.

In panel A, Apple race eclose at a higher proportion across the extreme ends of time in cold treatment.

In panel B, apple race take longer to eclose than hawthorn race.

Apple races survive over a longer range of cold storage, but it takes longer for them to eclose

Fig 2. almost uninterpretable. But it appears that longer heating periods changes the distribution from shorter to longer days to eclosion

Fig3, allele freq associated with earlier eclosion times as a function of days to eclosion

They dont show IDH.....

The frequency of these alleles decrease through time, which is what you'd expect if they're alleles associated with earlier days to eclosion.

Discussion

diapause trade off hypothesis more laid out: earlier apple fruits in the season selects for slower development rates in apple flies. Apple flies should have lower metabolic rates and have a higher propensity to undergo diapause. Length of winter is potentially an important agent of selection then.

All 4 predictions supported:

1. optimal survivorship was cold treatment over 15-19 weeks. Pupae may not be able to terminate diapause in < 15 weeks and they may not have enough energy stores > 19 weeks.
2. Apple flies had higher survivorship than hawthorn flies under longer cold treatments
3. allozyme patterns matched adult eclosion patterns. Earlier associated alleles were higher than resembled hawthorn race.

Thoughts:

Paper is not really testing a trade off, but more about differences in life histories between host plants. To truly evaluate a trade off, there needs to be a comparison between each race on each host plant.

2017-07-12; cont'd

Feder, J. L., & Filchak, K. E. (1999). It's about time: the evidence for host plant-mediated selection in the apple maggot fly, *Rhagoletis pomonella*, and its implications for fitness trade-offs in phytophagous insects. In Proceedings of the 10th International Symposium on Insect-Plant Relationships (pp. 211–225). Dordrecht: Springer Netherlands.

http://doi.org/10.1007/978-94-017-1890-5_27

Benjamin Walsh in 1860's made the first observation of flies shifting from hawthorn to apple.

3 major conclusions to date:

1. The difference in developmental timing has facilitated the shift between hawthorn to apple for rhagoletis
2. host-specific tradeoffs between occupying different fruits is shaped by selection at different life stages (larvae, pupae)
3. There is a lot of genetic diversity for development, which may allow Rhagoletis to host shift easily

When shifted from cold to warm, hawthorn flies alleles change as a function of days to eclosion. For example, Dia-2 is high early on in freq, but declines with longer eclosion times.

This allele also exhibits clinal variation where it is high in the north and low in the south. The north has later springs, which could mean that these flies need to eclose more quickly to maximize activity/performance/growth over the limited or short growing season.

In these same alleles, they differ in allele freq between host fruit. For example, the earlier eclosion associated allele, Dia-2 100, is overall higher in hawthorn than apple.

Natural history question: How far down in the soil do these rhagoletis go? Looking at figure 6.

Feder, J. L., Roethle, J. B., Wlazlo, B., & Berlocher, S. H. (1997). Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly (sympatric speciation–fitness trade-offs–host plant phenology–*Rhagoletis pomonella*–host x environment interactions). *Evolution*, 94, 11417–11421. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC23485/pdf/pq011417.pdf>

This paper was interested in testing the effect of prewinter temperatures on the eclosion outcomes of maggot flies on hawthorn. There are known loci that differ between apple and hawthorn. Here they focused on prewinter treatments of haw flies to see if they resemble apple like in their eclosion times and allele frequencies.

Apple flies have longer pre winter period than haw flies.

fig 1. The amount of pupae eclosing decreases with longer periods of heat before they overwinter. They prewintered from 2-32 days.

fig 2. allele frequencies for 3 loci decrease through time (number of days pupae heated before overwintering) ; similar story for figure 3.

fig 4. Negative relationship between allel freq and growing degree days. Acon-2 95 is high when there is a shorter growing season. This is important because haw race have higher allele freq than apple race. SO the previous patterns show that the allele freqs become more apple like (low allele freq).

These alleles are associated with how pre-diapause temperatures influence eclosion.

2017-07-13 reading cont'd

Filchak, K. E., Feder, J. L., Roethel, J. B., Stolz, U., & Roethele, J. B. (1999). A Field Test for Host-Plant Dependent Selection on Larvae of the Apple Maggot Fly, A FIELD TEST FOR HOST-PLANT DEPENDENT SELECTION ON LARVAE OF THE APPLE MAGGOT FLY, RHAGOLETIS POMONELLA.

Source: Evolution Evolution, 53(531), 187–200. Retrieved from
<http://www.jstor.org/stable/2640931>

Previous studies show that feeding habit does not differ between races on apple or hawthorn. One possibility that previous studies missed is that rotting fruit can be a selective force on larvae. They wanted to test the effect of fruit rot on larval eclosion

Experimental set up: they varied fruit rot by rearing fruit in the field and in a garage. The garage is suppose to rot the fruit less than the field. And they wanted to associate whether timing of exit of hawthorns/apple is related to linked loci (3)

first two linked groups had SNPs related to lower number of days in the haws. More alleles of loci related to high frequency that eclose earlier were also related to less time in the haws. So these loci are related to faster development.

summary table

Linkage_Group	Allele	race	eclosion_pattern_time	pupae.dev	during.winter.treatment	pre.winter.treatment	GDD
1	Aat-2 75+	higher hawthorne	high earlier on	shorter	but higher in hawthorn	higher earlier	
1	Dia-2 100	higher hawthorne	high earlier on	shorter	higher in hawthorne	higher earlier	
2	Me100	higher hawthorne	high earlier on	shorter	higher earlier on and hawthorn	higher earlier	
2	Acon-2 95	higher hawthorne	high earlier on	shorter	higher earlier on and hawthorn	higher earlier	higher hawthorn and with lower GDD
2	Mpi 37	higher hawthorne	high earlier on	shorter	higher earlier on	higher earlier	
3	Had100	higher hawthorne	high earlier on	shorter	higher in hawthron	Ushaped	

Rnandom thought:

Entering a state of developmental arrest, known as diapause, represents a life-history strategy to cope/survive with unfavorable environmental conditions.

Page 92: 2017-07-12. Meeting prep with Dan: To do/progress list

Set meeting in stone? Fridays 2PM?

1. Paper discussion on Rhagoletis
 - o Thoughts on mendeley?
 - o Future papers to read
 - o Dynmaic QTL idea to integrate genetic basis of traits over life histories

2. Meeting with PIs July 19th-- set time?
3. Which meeting to go to?
 - Evolution 2018
 - Montpellier, France Aug 18-22
 - Entomological Society of America
 - November 5-8 Denver Colorado
 - meet Rhagoletis family
 - Pricing:
 - Registration- \$475 before Sept 13th 2017
 - I need to sign up to be a member; costs covered?
 - Hotel: \$159-201/night
 - SHould I present?
 - SICB 2018
 - San Fran, Ca; Jan3-7
4. When are we collecting?
 - Plan to go out into the field August-September; depends on fruiting time
 - Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois
5. Complete Science IDP by August 1st.
6. Shadow Qinwin and Chao
 - Meeting with Qinwin July 18th, 10 AM
7. Go over data management practices?
 - Get NIH guidelines?
 - Go over data and directories?

8. Which fellowships should I be thinking about apply to?

- NSF?
 - USDA?
-

Page 93: 2017-07-17. Project status; dissertation; climate cascade

1. Hsp rxn norm paper:

- Rearranged figures; so results need to be rewritten; introduction needs tweaking; discussion needs to be finished ; goal mid july send out
- Plan to work on this once I send out range limits paper to AEllison
- Submission goal: December 2017

2. Range limits paper:

- Need to rewrite introduction, tweak results, finish discussion
- Send a draft to AEllison by end of July 2017
- Submission goal: December 2017

3. Proteome stability project:

- Still in data gathering + analysis step
- Wai needs to HPLC fractionate tryptic peptides to ID more of them
- ANBE needs gene annotations for aphaenogaster

4. Thermal niche paper

- In Ichick's hands
- Nsanders wants to look at it.

- Submission goal: yesterday
 - 5. Aphaeno genomes project:
 - Mlau needs to annotate-get GO's and KOGs
 - 6. Stressed in nature project:
 - IN SHC's hands; goal to have draft by end of summer
-

Page 94: 2017-07-14. Meeting with Dan

Set meeting in stone? Fridays 2PM?

1. Paper discussion on Rhagoletis
 - Thoughts on mendeley?
 - Future papers to read
 - Dan gave me 3
 - Dynmaic QTL idea to integrate genetic basis of traits over life histories
2. Meeting with PIs July 19th-- set time? 2 or 3pm
3. Which meeting to go to? I can go to both Evolution and ESA.
 - Evolution 2018
 - Montpellier, France Aug 18-22
 - Entomological Society of America
 - November 5-8 Denver Colorado
 - meet Rhagoletis family
 - Symposium:

- Genomics of Adaptation: Linking the Next Generation of Genome-Wide Analysis to Understand and Manage Complex Traits
Organizers: Gregory Ragland, Glen Hood, Scott Egan, Daniel Hahn, and Meredith Doellman
 - Pricing:
 - Registration- \$475 before Sept 13th 2017
 - I need to sign up to be a member; costs covered?
 - Hotel: \$159-201/night
 - SHould I present?
 - Late breaking abstract
 - I can present or not
- SICB 2018
 - San Fran, Ca; Jan3-7
4. When are we collecting?
- Plan to go out into the field August-September; depends on fruiting time
 - Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois
5. Complete Science IDP by August 1st.
6. Shadow Qinwin and Chao
- Meeting with Qinwin July 18th, 10 AM
7. Go over data management practices?
- Get NIH guidelines?
 - Go over data and directories?
8. Which fellowships should I be thinking about apply to?

- NSF?
- USDA?

Funding opportunities:

Blair(department head) brought up to Dan.

- Sloan INformatics fellowships
- USDA post docs
- NIH Rick Kirkeinstein?
 - somebody teaching how to apply for this

Goal by year 2: Have a killer job talk;

Hvae componenents

1. old work
2. post doc work
3. future work

Rhagoletis- polygenic system European COrnborer- 1 major effect loci

Papers to journal club:

Roethel, J. B., Romero-Severson, J., & Feder, J. L. (2001). Evidence for Broad-Scale Conservation of Linkage Map Relationships Between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, 94(6), 936–947. [https://doi.org/10.1603/0013-8746\(2001\)094\[0936:EFBS\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0936:EFBS]2.0.CO;2)

Filchak, K. E., Roethel, J. B., & Feder, J. L. (2001). Effects of photoperiod and light intensity on the genetics of diapause in the apple maggot (Diptera : Tephritidae). *Annals of the Entomological Society of America*, 94(6), 902–908. <https://doi.org/10.1603/0013->

[8746](#)(2001)094{[]0902:EOPALI]2.0.CO;2

Filchak, K. E., Roethel, J. B., & Feder, J. L. (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, 407(6805), 739–742. <https://doi.org/10.1038/35037578>

Next time, chat about trade-offs.

Page 95: 2017-07-17. Rhagoletis reading batch 2

Filchak, K. E., Roethel, J. B., & Feder, J. L. (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, 407(6805), 739–742. <https://doi.org/10.1038/35037578>

Nature paper showing gene-environment interaction

Figure 1: Seasonal distribution of flies on respective hosts. Panel A shows the distribution of infested larvae on host fruit through July to October. Apple is earlier. Panel B shows the larvae emerging from fruit as a function of the same time frame. Apple is earlier than hawk.

Then they compared the allele frequencies throughout different durations of a wintering lengths, reared at a low(22) and high(26C) temp. For figure 2, for all 3 alleles, there was a significant interaction between temperature and overwintering length on allele frequencies.

Under higher temp(26C), alleles started at a higher frequency, but were lower throughout longer overwintering. What does this mean? These alleles are referenced such that they are higher for shorter eclosion periods. Under low temperatures these alleles are lower in frequency, presumably because lower temperatures hinder growth and earlier

eclosion. So an allele's performance shifts depending on rearing temperature.

Experiments only focused on hawk-moth-reared fruit.

Take home: The performance of an allele depends on the rearing temperature. For eclosion under different overwintering lengths, early eclosion alleles can increase in frequency under high temperatures.

Roethel, J. B., Romero-Severson, J., & Feder, J. L. (2001). Evidence for Broad-Scale Conservation of Linkage Map Relationships Between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, 94(6), 936–947. [https://doi.org/10.1603/0013-8746\(2001\)094\[0936:EFBSCO\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0936:EFBSCO]2.0.CO;2)

Page 96: UF communication tips:

<http://postdoc.aa.ufl.edu/programs/postdoc-research-symposium/communication-tips/>

From Keynote Workshop Presenter: Kevin Folta A Few Good Tips in Science Communication

- The best science is arrested if not communicated. Make every effort to share your discoveries, make them understandable to public audiences. Think about how you'd describe your findings to a non-scientist in an elevator, and practice that message.
- Write every day. Start a blog and paraphrase the scientific work in your field for a common audience. Write about the scientific facets of a controversial topic. Practice generating clever prose.
- Cool it, scientist. Yes, we know you're smart. Dial it down a notch and talk to us. Communicate. The best talks, grant proposals and

papers make the problem understandable and urgent, the effects clear and personal, and the solutions realistic and immediate.

- Tell a story-- don't crush the reader or listener in jargon and statistics.
Remember your audience. The opportunity to share your science might be rare. When you are asked to write, speak, or report about your work think about your audience, spend time considering how to make the message most appropriate and effective. Always prepare well and leave them wanting more.
- Understand rhetoric. Revisit the ancient art of how to frame a persuasive campaign. Use it. Read about how advertisers reach an audience. Use these devices in your presentations, proposals and papers.
- Don't forget about "priming". Psychology says that we make decisions in our brains before we know we made a decision. Make your work beautiful. Only submit it if it has a halo—if it screams about your care, attention to detail, and professionalism.

I need to write a teaching and research statement; both 1-2 pages

Watching this: <https://www.youtube.com/watch?v=DbRdTXDkE7s>

Make it easier for the reader. Make it organized (your application).
Capture interest in first page.

Panel usually approaches application...

1. Cover letter
2. glance at cv
3. research statement

"Have to make a case for yourself"

Detailed parts is important for comparing top 10. For initial screen, don't get discarded.

Look for background (in CV) and what are you going to do

IN teaching statement, state "I can teach these courses"

I could teach:

1. Evolution
2. Insect Physiology
3. General biology
4. Evolutionary Physiology
5. Macroecology?
6. Cell biology?

Advanced stuff...

1. Intro to biostats
 2. Data science (R, open science)
-

Page 97: 2017-07-18. Qinwin training me on european cornborer (ECB)

Rearing room for Adults , pupae, larvae: Enter 3252 and the two strains are located in 2 separate rooms:

- Left side (3251): BE, bivoltine strain (use yellow tape)
- Right side (3253): UZ, univoltine strain (pink tape)

Rearing:

1. Tray with vermiculite layer with a cage on top.
2. Sponge that is watered everyday to maintain humidity
3. wax paper on top of cage so that this allows adults to lay eggs
4. Conditions: 25 C; 14L:10D photoperiod ; also good for Rhagoletis.

Eggs

To extract eggs, take the wax paper and cut into little pieces. Place the wax + egg paper into a clear container with a lid with NO holes (Keeps humidity). Label container with respective colored tape with Name and Date.

It takes 5 days to hatch. Before they do, the eggs turn black (day 4). Add diet to the bottom of the clear container(with wax paper on top) and use a new lid that has a mesh hole (small). Too large of a hole will dry out the diet.

Diet

ECB diet is ready made. Just add powder and boiling water. Then it solidifies. It is made in room 3143. Takes around 4 minutes.

Larval care

Larvae are stored in a 26 C incubator(blue in the hallway) with 16L:8D photoperiod.

Same incubator also used for diapause termination or from switching short to long photoperiod.

It takes 20-21 days to pupate.

Qinwin's project:

Ecdysteroid terminates diapause by promoting growth. IN fact, larvae without this hormone don't pupate. TO compare the effects of ecdysone between strains, she dosed (0, 0.1, 0.4, 0.7, and 1 ug) ECB larvae with ecdysone and measured pupation day. Bivoltine (BE strain) pupated faster as a result(compared to UZ strain).

Qinwin's pathway dissection:

Photoperiod --> PDF --> PDFGR in prostaglandin --> ecdysone in prostoglandin

Diapause induction

Room is 3140, which has many incubators that can differ in photoperiod. The conditions are 12L:12D 23 C.

Page 98: 2017-07-19. UF computing center, tutorial/workshop

1 Navigating UF's High-Performance Computing Environment
2 with
3 Ying Zhang, Applications Specialist
4 University of Florida Informatics Institute and
Information Technology

Ying; started out as high performance computing center. This is what it started out as. There is a huge computing cluster 10,000's computing cores. For large scale computing; parallel, run thousands of nodes. This is used for physics/chem/engineering/math.

But in the last 5 years, more and more biologists need computing power!

Fastest growing user body. Accomodating everybody! Because of sequencing

Support ~500 software installed bio

Cluster called: HiPerGator

Let you run 100s of jobs at the same time. Data within a day instead of a month!

website: <https://www.rc.ufl.edu/>

outlines services; there is a wikipage

(https://wiki.rc.ufl.edu/doc/UFRC_Help_and_Documentation)

[https://wiki.rc.ufl.edu/doc/UFRC_Help_and_Documentation]

Do not provide license. If usergroup has it, then only access to user group. There are 2 bioinformatics specialists has office hours in genetics buildings (ICBR).

Tuesday and Thursday mornings.

Oleksandr Moskalenko Bioinformatics Specialist om@rc.ufl.edu

In the florida museum; need to email him

Matt Gitzendanner Bioinformatics Specialist magitz@hpc.ufl.edu

(works for a quarter of the time); matt does some training; he's out for summer; but in the fall there are training sessions every 2 weeks in the physics building.

Signed up for hipergator: <https://www.rc.ufl.edu/access/request-account/>

Dan needs to approve.

Getting started: https://wiki.rc.ufl.edu/doc/Getting_Started

```
1 | ssh <username>@hpg2.rc.ufl.edu
```

2 policies for use:

1. Don't run programs on login node; to see if script is working, try a test node
 - acc can get suspended
2. Do not run programs from home directory. Your job will get killed; limited storage

Sample scripts for submitting jobs:

(https://wiki.rc.ufl.edu/doc/Sample_SLURM_Scripts)

[https://wiki.rc.ufl.edu/doc/Sample_SLURM_Scripts]

Try interactive jobs, not on home directory, but in working directory.

If you have a problem, don't email, submit a ticket!

Set meeting in stone? Fridays 2PM?

1. Paper discussion on Rhagoletis

- Roethel, J. B., Romero-Severson, J., & Feder, J. L. (2001). Evidence for Broad-Scale Conservation of Linkage Map Relationships Between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, 94(6), 936–947. [https://doi.org/10.1603/0013-8746\(2001\)094\[0936:EFBSCO\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0936:EFBSCO]2.0.CO;2)
- Filchak, K. E., Roethel, J. B., & Feder, J. L. (2001). Effects of photoperiod and light intensity on the genetics of diapause in the apple maggot (Diptera : Tephritidae). *Annals of the Entomological Society of America*, 94(6), 902–908. [https://doi.org/10.1603/0013-8746\(2001\)094\[0902:EOPALI\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0902:EOPALI]2.0.CO;2)
- Filchak, K. E., Roethel, J. B., & Feder, J. L. (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, 407(6805), 739–742. <https://doi.org/10.1038/35037578>

2. Meetings to attend:

- Entomological Society of America
 - November 5-8 Denver Colorado
 - meet Rhagoletis family
 - Symposium:
 - Genomics of Adaptation: Linking the Next Generation of Genome-Wide Analysis to Understand and Manage Complex Traits

Daniel Hahn, and Meredith Doellman

- Pricing:
 - Registration- \$475 before Sept 13th 2017
 - I need to sign up to be a member; costs covered?
 - Hotel: \$159-201/night
 - Evolution 2018
 - Montpellier, France Aug 18-22
3. When are we collecting?
- Plan to go out into the field August-September; depends on fruiting time
 - Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois
4. Complete Science IDP by August 1st. Done.

other notes:

1. Oriented with hipergator computing, met with Ying 2017-07-19
2. switch out desk

Future stuff:

Funding opportunities:

Blair(department head) brought up to Dan.

- Sloan INformatics fellowships
- USDA post docs
- NIH Rick Kirkeinstein?
 - somebody teaching how to apply for this

Goal by year 2: Have a killer job talk;

Hvae componenents

1. old work
2. post doc work
3. future work

Rhagoletis- polygenic system European COrnborer- 1 major effect loci

Page 100: 2017-07-19. Skype meeting with TPowell, GRagland.

Next week is earliest to collect apples.

GRagland: qPCR samples needed

Grant, MI; good for pop gen sample - no wasps

Tpowell , what are the sites?

Lancing looks "good". Haw trees, apartments good.

Urbana, IL has the greatest phenological separation. Apples earlier, hawthorn is later in the south. It is also the easiest site to collect at. Closest to gainesville.

Dont get anything, yet.

AUGust 12, dan's bday. we could collect then.

Are we doing latitudinal cline?

Me: Do Circadian Rythm clocks.

Non-diapause class making it through the winter! Need to explore this.

2 goals for our discussion today:

1. when and where are labs collecting?
2. How do we address diapause issue? 1 lab, or a couple of labs doing it? Have stuff standardized.

Tpowell doing warming experiments. Control, climate change condition and measuring phenology in flies and wasps. Use data to design warming based selection experiments

Alicia started noaa station data and compiling it.

Gainesville greenhaw flies; 30 tubes in a box. Tpowell wants that.

Greg needs urbana for qPCR validation. Number? 4-5 reps /time point/hose race. couple of hundred pupae per host race

Tpowell needs to heavily sample 1 site.

wrap up

1. tom watns urbana, but send greg some samples for qPCR
2. some point, go into -80 to figure out greek stuff, and green haw stuff
3. Greg has a freezer item:

Who is doing diapause classing?

Greg is going to do developmental biology. Take haw flies, class pre-winter and then track cell proliferation and brain morphogenesis. Track individuals. Continuing undergrad.

Pete dataset? What do we want to know?

Is it regressing? Cell prolif?

Write up experimental goals and plans. Need action item completed.
Deadline = Wednesday.

Page 101: 2017-07-21. Meeting with Dan friday 2pm

1. Paper discussion on Rhagoletis

- Roethel, J. B., Romero-Severson, J., & Feder, J. L. (2001). Evidence for Broad-Scale Conservation of Linkage Map Relationships Between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, 94(6), 936–947. [https://doi.org/10.1603/0013-8746\(2001\)094\[0936:EFBSCO\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0936:EFBSCO]2.0.CO;2)
- Filchak, K. E., Roethel, J. B., & Feder, J. L. (2001). Effects of photoperiod and light intensity on the genetics of diapause in the apple maggot (Diptera : Tephritidae). *Annals of the Entomological Society of America*, 94(6), 902–908. [https://doi.org/10.1603/0013-8746\(2001\)094\[0902:EOPALI\]2.0.CO;2](https://doi.org/10.1603/0013-8746(2001)094[0902:EOPALI]2.0.CO;2)
- Filchak, K. E., Roethel, J. B., & Feder, J. L. (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, 407(6805), 739–742. <https://doi.org/10.1038/35037578>

2. Meetings to attend:

- Entomological Society of America

- November 5-8 Denver Colorado
- meet Rhagoletis family
- Symposium:
 - Genomics of Adaptation: Linking the Next Generation of Genome-Wide Analysis to Understand and Manage Complex Traits
Organizers: Gregory Ragland, Glen Hood, Scott Egan, Daniel Hahn, and Meredith Doellman
- Pricing:
 - Registration- \$475 before Sept 13th 2017
 - I need to sign up to be a member; costs covered?
 - Hotel: \$159-201/night

- Evolution 2018
 - Montpellier, France Aug 18-22

3. When are we collecting?

- Plan to go out into the field August-September; depends on fruiting time
- Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois

4. Complete Science IDP by August 1st. Done.

- Send to Dan

5. I need a new desk: Monday, pick up truck?

other notes:

1. Oriented with hipergator computing, met with Ying 2017-07-19
2. met with Qinwin or rearing ECB
3. Q: What are the general rules for co-authorship in the lab?

Future stuff:

Funding opportunities:

Blair(department head) brought up to Dan. Where to find this stuff?

- Sloan INformatics fellowships
 - using existing datasets in the lab; complex genomic architecture problem (diffuse, inversions)
 - charlie baer; *c elegans*; genomic architecture
- USDA post docs
- NIH Rick Kirkeinstein?
 - somebody teaching how to apply for this

Goal by year 2: Have a killer job talk;

Hvae componenents

1. old work
2. post doc work
 - Rhagoletis- polygenic system
 - European COrnborer- 1 major effect loci
3. future work

Prof dev:

Seminar series: There are 2 for biology

1. pop bio group, friday
2. actual seminar series

Dan will show me a writing book.

going over IDP:

Important- teaching in a classroom setting. Boost yourself so that you're marketable. Show that you can teach a full course. norm- 1.5 to 2 courses a year

1. 5 week block of teaching intro bio classes
2. Upper level- grad level course

What kind of PI do I want to be?

Dan is 13 years- transition to doing research to manager. Greg is in the middle, hands on

College of Ag, there is a teacher's college, held in our building. Post docs might be able to attend, 1 day /week course that helps you teach. Faculty level certification course. Pair that with mentored-teaching experience. Take over 1-2 weeks Dan's physiology course. Dan and I could teach a graduate seminar. I'll be an instructor of record. I'll get an official evaluation.

There is a course for mol techniques in the department; and Leigh was a co-instructor of record.

Try to do this in your 2nd year. I could teach a data science.

Rules for Co-authorship?

more generous than less generous

Make a substantial contribution

1. collecting data
2. data analysis
3. data collecting

Revisit IDP in 6 months

Page 102: 2017-07-21. notes on *Rhagoletis* paper; genomic mapping

ref: Roethel, J. B., Romero-Severson, J., & Feder, J. L. (2001). Evidence for Broad-Scale Conservation of Linkage Map Relationships Between *Rhagoletis pomonella* (Diptera: Tephritidae) and *Drosophila melanogaster* (Diptera: Drosophilidae). Annals of the Entomological Society of America, 94(6), 936–947.

intro

First, not all of the *rhagoletis* markers map to fruit fly chromosomes. In fruit flies, chromosome 4 is not mapped at all.

1. *Rhagoletis* has 6 chromosomes (linkage groups)
 - How do they know it is 6 chromosomes?
 - linkage group 1 and 2 are target loci responsible for divergent diapause timing in apple maggots
2. fruit flies have 3 or 4 chromosomes

Some terminology

Need to look up definition for synteny:

- 1 In classical genetics, synteny describes the physical co-localization of genetic loci on the same chromosome within an individual or species. Today, however, biologists usually refer to synteny as the conservation of blocks of order within two sets of chromosomes that are being compared with each other.

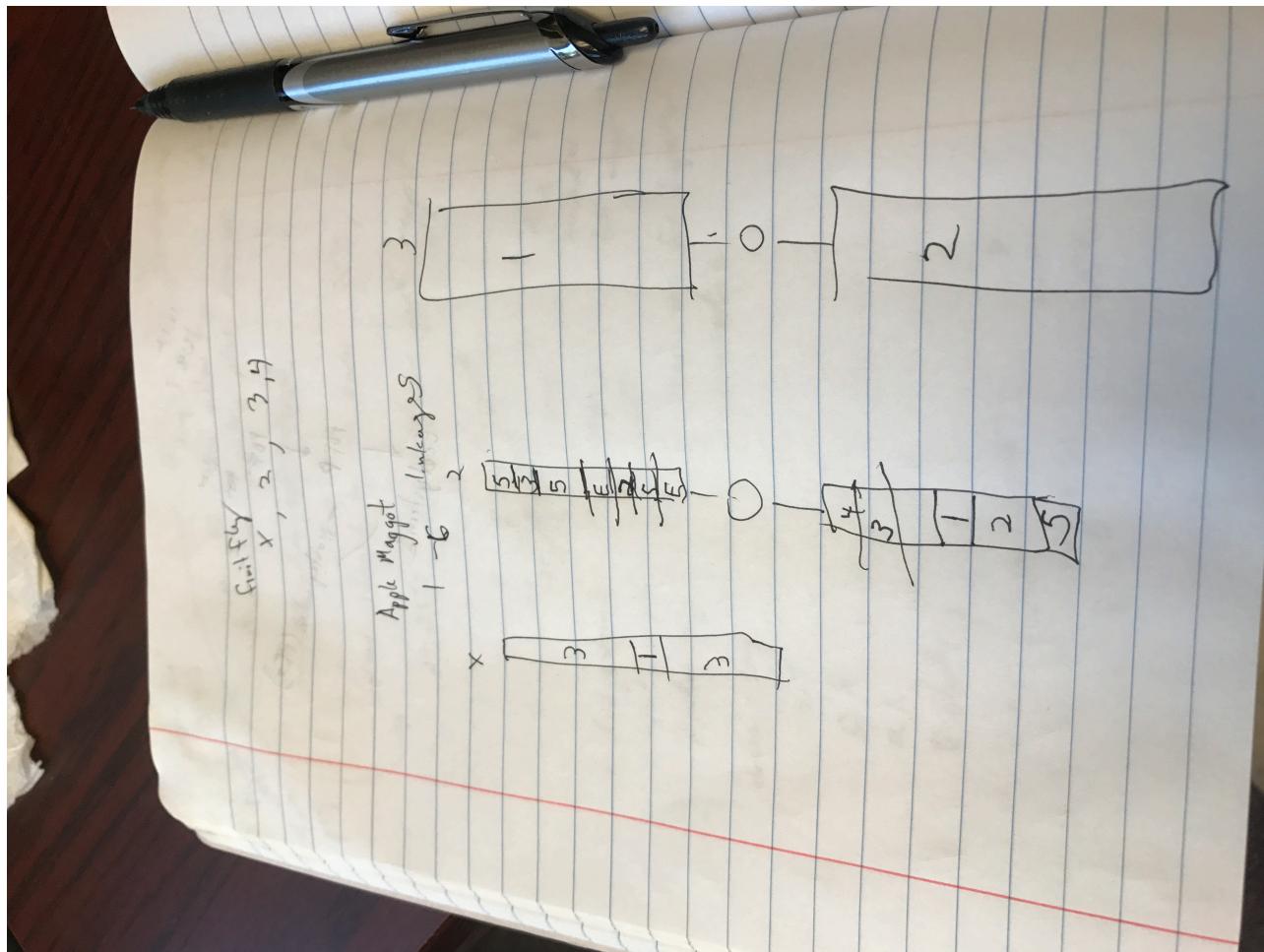
Two diff def, confusing as hell.

Also need to look up types of inversion. pericentric vs paracentric:

1. **Paracentric** does not include the centromere and the breaks in the chromosome occur on 1 arm.
2. **Pericentric** involves a inversion when there are breaks on each chromosome arm, so the inversion includes the centromere.

This paper argues that there are PARacentric inversions. Authors don't really have a clue because there could be pericentric inversions too, at least for chromosome 2.

Some of my hand written notes. I copied figure 4 but placed the rhagoletis linkage groups on the fruit fly chromosomes.



From the paper, it says the chromosome 2 and 3 are a product of chromosome fusings for each arm. Looks like this is true because for chromosome 3, one arm is linkage 1 and the other is linkage 2.

Authors think the common ancestor of rhagoletis and fruit fly is 6. But you cant tell from these data because there is no outgroup! The reverse could be true: ancestral state is 3/4 and rhagoletis expanded it's chromosome numbers.

Structural differences:

1. chromosome 2 is most likely a product of a bunch of translocations
2. chromosome 3 is most likely a product of chromosomal fusions
3. the sex chromosome switched?

So, chromosome 3 should be target for selection in fruit flies for modulating differences in diapause timing. But....

Many of the loci responsible for seasonal timing are mapped to different parts of fruit fly genome. Period is on the x chromosome. So seasonal timing should be a polygenic trait in fruit flies too.

Page 103: Diapause reading papers

1. Koštál, V. (2006). Eco-physiological phases of insect diapause. *Journal of Insect Physiology*, 52(2), 113–127.
<https://doi.org/10.1016/j.jinsphys.2005.09.008>
2. Dambroski, H. R., & Feder, J. L. (2007). Host plant and latitude-related diapause variation in *Rhagoletis pomonella*: A test for multifaceted life history adaptation on different stages of diapause development. *Journal of Evolutionary Biology*, 20(6), 2101–2112.
<https://doi.org/10.1111/j.1420-9101.2007.01435.x>

Koštál, V. (2006). Eco-physiological phases of insect diapause. Journal of Insect Physiology, 52(2), 113–127.
<https://doi.org/10.1016/j.jinsphys.2005.09.008>

Paper focuses on synthesizing all of the previous ways scientists use diapause and arriving at a consensus with words.

Defining diapause

In defining diapause, author argues that it is a specific term for dormancy. Here is how I'd interpret his definitions and my thoughts:

1. Dormancy: any state of suppressed metabolism and development. He argues that it is adaptive-no, it doesn't have anything to do with adaptation, so just leave that part out. Just define the term and then overlay what the term could mean after the fact.
2. Quiescence: a short term suppression of metabolism and development which can be resumed given favorable environmental conditions.
3. Diapause: long term suppression of metabolism and development which significantly shifts away from direct development.

Author's definitions are so obscure, imprecise, and non-direct. The way I've described dormancy is basically short vs long term. I don't lump dormancy depth into the def, but maybe I should because that is part of the process?

stages of diapause

Author subdivides diapause into 3 main stages:

1. pre-diapause
2. diapause
3. post-diapause

Pre-diapause

Diapause is initiated in advance to unfavorable conditions by cues such as photoperiod and temperature. The process might be fixed and completely genetically determined without any external cues, known as obligatory diapause.

1. induction phase - when organisms receive cues to start diapause

2. preparation phase -fixed into entering in diapause by shifting direct development into storage potentially

Diapause: direction development is arrested

1. initiation - direct development completely stops
2. Maintenance - utilizing stored reserves for somatic maintenance in the face of unfavorable environments
3. Termination - when the enviornment becomes more favorable and the organism begins direct development again
 - direct development could begin
 - direct development begins again

Pause-diapause:

Author defines it as environmentally driven inhibition or lowering of development following termination of diapause.

summary

Diapause is a dynamic and gradual process by which insects enter in a state of dormancy when they encounter unfavorable environmental conditions. These conditions could be related to food availability, photoperiod, and temperature. This allows organisms in thermally extreme cold environments to match their life history timing when the environment is favorable. Although the author realizes that diapause is continuous, he attempted to define critical moments in diapause to provide a framework for studying it.

There are a lot of good references to old papers and this paper is a good resource for them!

Page 104: 2017-07-24. Dambroski and Feder 2012; Journal of evo bio paper

Dambroski, H. R., & Feder, J. L. (2007). Host plant and latitude-related diapause variation in *Rhagoletis pomonella*: A test for multifaceted life history adaptation on different stages of diapause development. *Journal of Evolutionary Biology*, 20(6), 2101–2112.

<https://doi.org/10.1111/j.1420-9101.2007.01435.x>

It turns out that the *Rhagoletis* system is more complex than described because there are facultative diapausers. Meaning, not all of these flies diapause at a deep level and they actually vary in this.

Recap: (flies) Apples have longer development than haws, presumably because apple flies are in diapause for a longer duration. The alleles in the papers I've been reading are associated with earlier eclosion, or faster development and are more "haw" like. In haws, these alleles can shift in their allele frequency over overwintering duration depending on the pre-winter temperature conditions.

Work through figures in paper first:

Fig 1. The illustration of the life history timing of the *Rhagoletis* system. In panel A, they show the progression of life stages as adults, on fruit, and diapausing pupae. Panel B shows how apple flies have a longer pre-winter period than hawthorn flies.

Fig 2. The distribution of the different host races. It looks like Haws have the widest distribution.

Fig 3. This shows the # of flies eclosing against the # of days eclosion for haw and apple races. There are "3" populations.

1. *ND flies*- flies that don't diapause or have a shallow diapause and go

into direct development.

2. *SD flies*- shallow diapausers, ones that can exit diapause faster.
3. *CD flies*- chill-dependent diapausers- ones that need to be in cold conditions to diapause.

Fig4- survivorship plotted against lat (N-S). Overall, apples have higher survival at higher latitudes, but converge in survivorship at the lowest latitude.

Fig 5-

- Panel a: Relative %CD flies against latitude (N-S). There is a positive relationship. Looks like an interaction between lat and host race. the % of CD(chill dependent diapausers) is higher as you move south. And the CD flies are overall, higher in apple than hawthorn. This suggests that southern populations need to be primed to diapause by cold temperatures than northern populations. Also, the apple race is more primed to diapause by cold temperatures than hawthorn flies.
- Panel b: Relative ND(non-diapausers) against latitude. There is a negative relationship. There is more of a % of flies that don't diapause in the north than in the south. There is a higher % of ND in haws than apple. Doesn't make sense because you'd expect the opposite pattern.
- panel c: SD+ND vs lat (N-S); positive relationship for apple, but not haws. This panel c looks like or resembles a.

Thoughts:

1. It looks like for haws, the % into CD,ND, and SD are uniformly distributed, which could suggest taht ti bet hedges equally into these different strageties. In contrast, the apple race looks like it invests more in CD flies. so they only invest in CD if they're chilled.

Figure 6:

- Panel A: Survival vs apple, hawthorn, and their reciprocal crosses. Apple had the highest surv, followed by hybrids. Haws had lower survival.
- Panel B: % abs of chill depenndent CD flies vs apple, hawthorn, and their reciprocal crosses. Same story as panel A.
- Panel C: Relative %CD diapause flies vs apple, hawthorn, and their reciprocal crosses. Applies and A x H cross had higher relative %CD flies than H x A cross and hawthorn.
- Panel D: % ND vs apple, hawthorn, and their reciprocal crosses. % of ND flies are higher for hawthorn. The corsses are intermediate.

Thoughts: What does this mean?

Note: In reading the methods, the % is expressed within the class that survived the overwintering treatmnt.

So in fig 5. Applies have higher %CD and SD survivorship, but lower %ND survivorship. Lat patterns still the same: positive relationshp between %CD/%SD with lat. Neg relationship between %ND and lat.

Fr figure 6. this oculd mean that apples have higher survival in CD and overall, but low survival for ND.

Figure 7. shows mean \$ of days to eclosion vs lat. Overall, for the 5 month overwintering treatment, hawthorn have longer developmental times than apple. This is consistent with the paper we prev read:

Feder, J. L6., Stoltz, U., Lewis, K. M., Perry, W., Roethel, J. B., & Rogers, A. (1997). THE EFFECTS OF WINTER LENGTH ON THE GENETICS OF APPLE AND HAWTHORN RACES OF RHAGOLETIS POMONELLA (DIPTERA: TEPHRITIDAE). *Evolution*, 51(6), 1862–1876.

<https://doi.org/10.1111/j.1558-5646.1997.tb05109.x>

2017-07-26, revisited paper:

In fig 5 b, the prop of non-diapausing flies is higher in the haw race than apple race. Why is that? It seems like overall, Apple pupae are higher quality, which may mean low quality flies (that don't have enough energy stores to survive the winter) may not opt into a diapausing strategy.

Page 105: 2017-07-26. diapause paper by Wilsterman et al.

Different systems have different language for how they describe dormancy (diapause, hibernation, torpor). This paper tries to integrate different fields with a common set of words that describe dormancy.

They break down dormancy into a few stages:

1. **Preparation:** remodeling morphology and physiology
 - super non-specific description
2. **Initiation:** decline in sensitivity to environmental conditions that promote growth, development. Stopping activity, cell cycle(growth), and suppression of development.
 - How this is different than preparation, I have no clue

3. **Maintenance**: refractoriness to environmental conditions that will later trigger the completion of dormancy
 - hard to agree with this definition. It is so abstract that I barely know what it means. They are defining this stage by the resistance to warm temperatures, but the emphasis should really be how this stage is characterized by hardiness to winter conditions. Organisms achieve hardiness by....
4. **Potentiation** : sensitivity to environmental conditions that trigger development increases, but not yet because conditions are still unfavorable.
 - Definition could benefit from more directness.
 - Increased sensitivity to environmental conditions that promote development
5. **Activation** : growth resumes because favorable conditions return.

Important distinction: Is dormancy obligate or facultative? Obligate is when species are hard-wired to enter dormancy, while facultative requires environmental cues to enter dormancy. It'd be interesting to know, what is the relative proportions of both of these strategies. It feels like obligate dormancy is rare.

The order of the stages in dormancy are off. Induction written first, then preparation but the authors led with preparation.

**Page 106: 2017-07-27. Meeting with Dan;
progress and to do list**

1. Paper discussion on Rhagoletis

- DAMBROSKI, H. R., & FEDER, J. L. (2007). Host plant and latitude-related diapause variation in *Rhagoletis pomonella*: a test for multifaceted life history adaptation on different stages of diapause development. *Journal of Evolutionary Biology*, 20(6), 2101–2112. <https://doi.org/10.1111/j.1420-9101.2007.01435.x>
- Unpublished diapause paper
- Next set of papers?
- look up papers on pre-winter conditions. Do apple flies enter diapause earlier than haws? Nature paper; 2000s

2. Meetings to attend:

- Entomological Society of America
 - November 5-8 Denver Colorado
 - Symposium:
 - Genomics of Adaptation: Linking the Next Generation of Genome-Wide Analysis to Understand and Manage Complex Traits
Organizers: Gregory Ragland, Glen Hood, Scott Egan, Daniel Hahn, and Meredith Doellman
 - Pricing:
 - Registration- \$475 before Sept 13th 2017
- Evolution 2018
 - Montpellier, France Aug 18-22

3. When are we collecting?

- Dan is away July 31-Aug 10
- Plan to go out into the field August-September; depends on fruiting time

- Collect larvae that has infested fruits (Apple and Hawthorne) around Michigan and Illinois
 - experimental plan? Greg sent but not Tom
4. Dan will show me books on writing.
 - The elements of style by William strunk jr and EB white
 - Style: lessons in clarity and grace by Joe Williams and Joe Bizup
 5. Review a manuscript while Dan is away?
 6. set up meeting with Chao to get training on respirometry
 - for rhagoletis check day 8 and 10 or 11). You can differentiate diapause and non-diapause classed
 7. Dan is buying set up for measuring circadian rythms:
 - Set up from trikinetic: <http://www.trikinetics.com/>
 - data can be analyzed in R : <https://github.com/jstaf/actmon>
 8. update 2017-08-01; get travel pcard. Completed online training and signed agreement form. Linda says ~1 month to get the card.

Future stuff:

Revisit IDP in 6 months

In december

Funding opportunities:

Blair(department head) brought up to Dan.

- Sloan INformatics fellowships
 - using existing datasets in the lab; complex genomic architecture

- problem (diffuse, inversions)
- charlie baer; c elegans; genomic architecture
- USDA post docs
- NIH Rick Kirkeinstein?
 - somebody teaching how to apply for this

Goal by year 2: Have a killer job talk;

Have components

1. old work
2. post doc work
 - Rhagoletis- polygenic system; diffuse genomic architecture
 - European COrnborer- 1 major effect loci
3. future work

Prof dev:

Seminar series: There are 2 for biology where I can present my work

1. pop bio group, friday
2. actual seminar series

PI development

Need to get there first, so I need to write a:

1. Research statement
2. Teaching statement

Teaching

Important- teaching in a classroom setting. Boost yourself so that you're marketable. Show that you can teach a full course. norm- 1.5 to 2 courses a year

1. 5 week block of teaching intro bio classes
2. Upper level- grad level course

College of Ag, there is a teacher's college, held in our building. Post docs might be able to attend, 1 day /week course that helps you teach. Faculty level certification course. Pair that with mentored-teaching experience. Take over 1-2 weeks Dan's physiology course. Dan and I could teach a graduate seminar. I'll be an instructor of record. I'll get an official evaluation.

Try to do this in your 2nd year. I could teach a data science.

Grant related; writing/reviewing

Dan mentioned reviewing grants together; need permission from program officer

What kind of PI do I want to be?

Dan is 13 years- transition to doing research to manager. Greg is in the middle, hands on

I want a set of documents for my lab that will help members get situated.

1. Ray Huey's guide to being a successful grad student/scientist
 2. Data science practices (data entry, storage)
 3. Writing?
 4. Expectations, rules?(co-authorship) , mentoring philosophy
-

Page 107: 2017-07-28. Dissertation work progress/ list

1. *Hsp rxn norm paper:*

- Rearranged figures; so results need to be rewritten; introduction needs tweaking; discussion needs to be finished ; goal mid july send out
- Working on this next!!!
- Submission goal: December 2017

2. **Range limits paper:**

- Send a draft to co-authors by end of July 2017
- Submission goal: December 2017

3. Proteome stability project:

- Still in data gathering + analysis step
- Wai needs to HPLC fractionate tryptic peptides to ID more of them (emailed him 2017-07-27)
- ANBE needs gene annotations for aphaenogaster

4. Thermal niche paper

- In Ichick's hands
- Nsanders wants to look at it.
- Submission goal: yesterday

5. Aphaeno genomes project:

- Mlau needs to annotate-get GO's and KOGs

6. Stressed in nature project:

- IN SHC's hands; goal to have draft by end of summer
-

Page 108: 2017-07-31 Dissertation work progress/ list.

1. *Hsp rxn norm paper:*

- Rearranged figures; so results need to be rewritten; introduction needs tweaking; discussion needs to be finished
 - Figure layout:
 - Predictions figure: Fold induction vs temperature for each parameter
 - Phylogenetic analysis: ancestral trait reconstruction
 - Non-phylogenetic controlled analysis: CTmax vs environment
 - CTmax vs hsp params
 - hsp params vs habitat type
 - Methods figure: example of fitting boltzmann to hsp gxp
 - Working on this next!!!
 - Send out Mid August
 - Submission goal: December 2017

2. **Range limits paper:**

- Sent draft to co-authors 2017-07-28
 - AEllison wants to try overleaf, so far so good
 - included figures , but will take out at submission
- Submission goal: December 2017

3. Proteome stability project:

- Still in data gathering + analysis step

- Wai needs to HPLC fractionate tryptic peptides to ID more of them (emailed him 2017-07-27)
 - He won't get to this for a few months
 - I have been swamped with requests from investigators since Bethany left. I should be able to start working on it in the next couple of months.
 - ANBE needs gene annotations for aphaenogaster-tied to genome project
-

4. Thermal niche paper

- In Ichick's hands
- Nsanders wants to look at it.
- Submission goal: yesterday

5. Aphaeno genomes project:

- SHC needs to annotate-get GO's and KOGs
 - has for pogos, but need aphaeno

6. Stressed in nature project:

- IN SHC's hands; goal to have draft by end of summer
-

**Page 109: 2017-08-01; Paper notes:
Hormonal control of diapause by Denlinger,
Yocum, and Rinehart. Book chapter 10; 2012;
and Thoughts**

Diapause is life history strategy where species opt to enter a state of developmental arrest by slowing metabolism to offset the costs of persisting in unfavorable conditions.

Photoperiod and temperature are main cues to induce diapause and they can covary or be decoupled. Diapause can happen at many life stages from egg to adult. Diapause can be induced transgenerationally when the mother has information about unfavorable conditions and promotes her offspring to diapause (*Bombyx mori* is an example).

What is the molecular basis of diapause?

Photoreceptors appear to be important in receiving daylength information. Generally, shorter daylength induces diapause. There have been a bunch of brain transplant experiments to understand whether different stages still respond to light.

Circadian clock genes seem important:

molecular basis of circadian timekeeping in fruit flies. Stanewsky 2002; Price 2004; Helfrich-forster 20009.

In fruit flies,

1. Period and Timeless heterodimerize and enter the nucleus
2. Per/Tim complex represses the transcriptional activity of Clk/Cyc (clock and cycle)
3. Clk/Cyc transcribe Period and Timeless

This pathway represents an oscillating feedback network that produces temporal variation. So either the pathway gets targeted by selection to shorten or lengthen circadian timing, and/or there is a modifier gene that tunes the circadian timing. This pathway is integrated into the photoperiod response via cry(photoreceptor cryptochrome (CRY)which acts as a pacemaker.

Experimental evidence

1. RNAi knockdown of period caused animals to not enter diapause (*R. pedestris*).
2. Rnai knockdown of cycle induced diapause under non-diapause photoperiod conditions.

(look up ref: Ikeno et al. 2010)

Given this info, what might we expect for their role in rhagoletis and ECB?

1. For Rhagoletis, I have to double check the ref, but if it only pertains to photoperiod then we might just lat clines in period, but not between races. However, there is supposedly segregating variation for period between races.
 - Hypothesis: Period may modulate the tmepo of diapause timing in rhagoletis
 - Apple: longer dev, longer cycles (at 5 month overwinter duration)
 - Haw:
2. ECB,
 - E strain is bivoltine and Z is univoltine (generally)
 - E has shorter diapause, but Z has longer diapause (Wadsworth et al. 2013, Journ of evo bio)
 - E should have shorter period cycling.

But what is period doing? Knocking it down should reduce heterodimerization of per/tim and lead to increased trxn of per/tim. What is the net effect? Circadian timing should extend. How does this cause animals to not opt into diapause?

Random thought which is a fantasized job talk title: **The physiology behind speciation: circadian clocks run seasonal timing to host fruiting phenology**

Wow! Sounds like a "nature" title

Page 110: 2017-08-07. Reading notes

As far as reading, please check out some papers on clocks. Do you still have the general review I sent you a while ago? If not, I'll try to find it and resend. Also look up papers by David Doelzel and/or Shin Goto as authors. They have written some good stuff in recent years and may have useful reviews

Keeping time without a spine in review; Denlinger et al.

One interesting point that is brought up is that organisms can move/migrate such that they avoid unfavorable conditions. Definitely one strategy to escape harsh winters.

Daylength predicts the onset of unfavorable winter conditions.

- short day length = signal for diapause
- long day length = continuous growth

Big question: What causes animals to opt into diapause strategies?

- Evolutionarily, it is to optimize activity and reproduction with favorable conditions
- Mechanistically, it is probably due to hormone control and circadian clocks.

Approaches for studying circadian timing in animals:

1. Nanda-Hamner protocol- different groups of organisms are exposed to different light-day lengths
2. Bunsow night interruption: night time is interrupted with light pulses. .

good reference for both methods

One interesting hypothesis: Circadian clocks and photoperiodic timers are integrated and linked. One problem tackling this is that circadian clocks are pleiotropic and affect many different phenotypes, often simultaneously. For example, circadian genes are transcription factors, so their downstream effects can be really large.

Cool paper on the interaction between day length and cold tolerance:

[http://journals.plos.org/plosgenetics/article?](http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004603)

[id=10.1371/journal.pgen.1004603](http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004603) ; Pegoraro et al. 2014

- done in *frtui* flies
 - measured ccrt (chill coma recovery time) between short day (winter) and long day length(summer) and different sexes.
 - Sig diff between day length for females but not males.
 - These differences went away in mutants of *timless*, *clock* and *period*.
 - So is the stress response tuned by clocks? Or are mutants just overall in poor quality?
 - Comparisons between short and long associated allelic variants show that:
 - for *period*, 2 allelic variants differ in CCRT. Genotype B has greater cold tolerance than genotype A (lower recovery time).
-

Page 111: 2017-08-07. Getting oriented in respirometry with Chao

The main goal of respirometry is to estimate metabolic rate by measuring the release of CO₂ from the animal (in this case, ectotherms). The measurements take place over a time interval to basically get CO₂/second.

Set up:

- 1 * Li-cor (LI-6262) - instrument that measures CO₂
- 2 * MFC - Mass flow controller - controls the flow rate (mL/min)
- 3 * Purge system - getting rid of water CO₂ from existing tubes/line
 - 4 * Dririte - removes water
 - 5 * Ascarite - remove CO₂
- 6 * Expedata - computer program that displays CO₂ acquisition

General steps:

1. Purge the system - air moves through the Dririte, then through Ascarite, then water to rehydrate the air (which is more consistent than ambient air).
2. Turn on Expedata and acquire CO₂. There are a set of markers that you can place in order to separate out sample readouts if you're doing continuous measurements.
3. The gas that goes through the detector is detected as a voltage. Detector goes from 0-5 volts. So set up standard curve that bounds voltage limits.

4. Ok, for each sample, purge, to clear the air and establish baseline for read outs. Keep track of time because you'll need to divide the read out by it (time) to get co2 rate.

5.

- Chao will show me to set up a separate prog set up on expedata when I have my own pupae.
 - For rhagoletis, I'll need to purge and then let them sit over night before I sample them.
-

Page 112: 2017-07-08. Reading on circadian clocks

Looked up David [Dolezel](#) as Dan suggested and works on chronobiology

Reading: [link](#) or [here](#)

Title: Period gene expression in relation to seasonality and circadian rhythms in the linden bug, *Pyrrhocoris apterus* (Heteroptera)

Have to look up stuff I don't know.

Tau- 1 circadian period length. It is usually in the unit of hours. But what is 1 period length? Period is the distance between the peak of activity in a wavey function (focus on x axis; makes sense because time is on the x axis). Amplitude is the absolute max height at peak activity (focus on y axis).

This study was interested in how day length influenced activity throughout the day and a critical circadian clock gene, period. Their experimental approach was to expose animals (linden bugs) to 3 treatments: constant darkness after long day entrainment; short day, and constant darkness after short day. Barely have a clue for the justification of these treatments....

I'm guessig that they want to compare short vs long daays, but the added darkness after long day entrainment treatment is to observe the clock wihtout any light stimulation.

They have two strain:

1. Photoperiod sensitive: wild type, maintained under diapause preventing long days - these go into diapause
2. Photoperiod insensitive: maintained in short days - these don't go itno dipaause

Next they compared mRNA levels of period. Short days looked less cyclical than long days (figure 2). Cant really tell at all. **H**OWever, when comparing overall expression levels, longer days had overall lower expression than short days (fig 3). - main result

It looks liek the magnitude of period is related to the signal transdusction of photoperiod.

So what does this mean? Recap

Short light days promote diapause, longer light days promote reproduction. So higher expression of per may promote diapause. This makes sense in light of Denlinger reviewer, wehre RNai of per disrupted diapause programming. Therefore absolute expression matters, not necessarily the ossication.

Reading review Dolezel, D. (2015).
Photoperiodic time measurement in insects.
Current Opinion in Insect Science, 7, 98–103.
<https://doi.org/10.1016/j.cois.2014.12.002>

Proposed architecture of photoperiodic clock:

1. receptors sense light
2. measurement of that light (photoperiodic clock)
3. internal counter of successive days
4. output

What is output and why does it have no description?

For 1, light enters and receptors sense it. For 2 and 3, clocks measure night length. For seasonal variation, something turns on though an internal counter. Nobody knows how this is achieved. For the output; it is diapause. Hormones are important, but depends on stage that they opt into diapause.

Establishment between photoperiod and circ clocks was suggested by Bunning 1936, wow! How did researchers find this out? They manipulated photoperiod and measured circadian clock genes.

Meuti, M. E., & Denlinger, D. L. (2013). Evolutionary Links Between Circadian Clocks and Photoperiodic Diapause in Insects. *Integrative and Comparative Biology*, 53(1), 131–143. <https://doi.org/10.1093/icb/ict023>

This gives a good run down of how circ clocks are linked to activity to photoperiod.

1. period and timeless proteins heterodimerize and can translocate into nucleus

2. it inhibits clock and cycle
3. but clcok and cycle can increase the gxp o fperiod and timeless
4. this creates a feedback loop

This whole cycling process is integrated with photoperiod with cryptochromes1 (cry1).

5. Cry1 binds to timeless and in the PRESENCE of light, degrades both timeless and cry1!

an additional complication: in lepidopterans, there is an paralogous cryptochrome (cry2) which is not responsive to light, yet it inhibits the expression of clock and cycle.

Therefore, cry2 may be an important negative regulator of circadian clocks.

The review doesn't say, but what does this mean for activity throughout the day? What isn't very clear is the relationship between activity and the cycling of these genes. Gotta look this up.

Cool paragraph on the physiological responses during diapause:

The ability to enter diapause has contributed to the evolutionary success of insects (Denlinger 2008). Insects in diapause are resistant to a range of environmental stresses, and these responses are mediated through physiological mechanisms, such as the generation of polyols and heat shock proteins that enhance survival at low temperatures; elevation of cuticular hydrocarbons that protect against desiccation; increased lipid stores and suppressed metabolic rates that enable diapausing insects to survive long periods without food; and boosts in immune responses to combat increased attacks by pathogens (Denlinger 2002, 2008; Hahn and Denlinger 2011).

Current hypotheses:

Bunning hypothesis (Erwin Bunning 1936) or internal coincidence model: light entrains circ clocks which is set to a light sensitive phase.

One way to test this hypothesis is to interrupt night times. First use short photoperiods followed by brief pulses of light at different times throughout the extended night. "If light falls on the sensitive phase, a long day response is elicited". Wtf....

Nanda-Hamner resonance experiments. Subject short period of light followed by periods of darkness to create Light:Dark cycles that range from 24072 hours. (Nanda and Hamner 1958). Light responses should follow day light length.

Page 113: 2017-08-14. Two new papers on niche breadth, related to range limits, thermal tolerance breadth projects

Both refs have Jason Sexton, a cool guy!

refs:

1. Hirst, M. J., Griffin, P. C., Sexton, J. P., & Hoffmann, A. A. (2017). Testing the niche breadth-range size hypothesis: Habitat specialization versus performance in Australian alpine daisies. *Ecology*. <https://doi.org/10.1002/ecy.1964>
2. Sexton, J. P., Montiel, J., Shay, J. E., Stephens, M. R., & Slatyer, R. A. (2017). Evolution of Ecological Niche Breadth. *Annu. Rev. Ecol. Evol.*

Skimming, but Sexton et al. 2017 has some nice definitions:

- Niche breadth - the variety of resources, habitats, or env used by a given species
 - I think this is a good def, but I don't think resources necessarily needs to be in there
 - my wording: the range of environments a species inhabits
- Niche dimensions- by which the niche is quantified or measured; often interchangeable with the term niche axis.
 - I guess the word niche itself is a nebulous term, but niche dimensions grounds us in reality by focusing on what we can measure.
 - my wording: the measurable aspects of the environment that species inhabit
- Fundamental niche: the range of environmental conditions and resources in which a species can survive and reproduce without the influence of predation or competition.
- Realized niche: the range of environmental conditions and resources in which a species can survive and reproduce in the presence of species interactions (predation and competition)

These different definitions of niches are functional, but I think it is good to realize that it is hard to tease apart +/- species interactions. I need to look up more specific cases.

my wording: The environment in which species inhabit. Or, the environment that promotes survival and reproduction for a species.

Let's keep things simple and concise.....

- Local adaptation- a process of nat sel whereby resident pops evolve higher relative fitness in their local habitat than populations elsewhere
 - I think this is a good def, how would I word it?
 - Well I might want to separate out the process that generated local adaptation and what it actually is
 - When local populations have more optimal fitness than surrounding populations
 - Local adaptation can be brought about by natural selection where local populations have higher relative fitness than surrounding pops
 - The expectation or signature of local adaptation, for example, is clinal variation in a particular trait.
 - hmm, the defs focus on fitness and the sig of local adaptation is usually measured as a trait; leap from trait to fitness, but one could argue that traits are a proxy of fitness
- Niche filling: expansion of the realized niche to fill a greater proportion of the fundamental niche
 - wait what...why
- Niche unfilling: contraction of the realized niche to fill a smaller proportion of the fundamental niche..

These defs focus on how range limit dynamics in the +/- of species interactions. I dont havnt really thought about species distr this way.

Notes:

- Confusing disucssion about phenotypic plasticity vs niche breadth. If phen plast is defined as *slope of the reaction norm*, then where does that leave niche breadth? Well, the niche breadth for a perf curve is the distance between the lower and upper limits of

performance. This is not inherently phen plast. To assess whether limits are phen plast, you'd have to generate a perf curve of thermal limits vs env. This way, you can get the slope of the reaction norm.
Not clearly laid out in this paper

- their def of phen plas is focused on individuals, I think the focus should be on genotypes

Brown (1984) envisaged species arising in localized regions that differ in several ecological respects from those occupied by parent species, thus promoting multidimensional specialization through speciation, but to-date, few direct tests of this idea have been made (but see Bonetti & Wiens 2014 for one example).

Dont think this is true if im reading multidim specialization through speciation correctly...Apple maggot flies?

- I really like this part of the sentence: *Environmental heterogeneity—in space, time, or both—is ubiquitous yet uneven in nature...*
- *Levins (1968) distinguished between fine- and coarse-grained scales of environmental heterogeneity, which generally refer to variation experienced within and between individual lifetimes, respectively* I should reread this paper because I didnt grab this from taht paper

Hirst et al. paper, figures are impenetrable. They could have potted survival vs elevation with species as different lines....you know, show the range of elevation , assuming this is the axis of variation they care about.

But from abstract they wanted to test the niche breadth- range size hypothesis: More speciose clades have more restricted distributions. The expectation is that more common species within a clade should have wider range sizes compared to relative taxa or sister taxa.

Page 114: 2017-07-14. Meeting with Dan 2pm

1. Paper readings:

- Focusing on circadian clocks genes and rhythms
 - Tau, per/tim and cyc/clk daily rhythms, Bunning hypothesis, Nanda Hamner experiments
 - dan says to look up carpet beetles,

2. Meetings to attend:

- Entomological Society of America
 - November 5-8 Denver Colorado
 - Symposium:
 - Genomics of Adaptation: Linking the Next Generation of Genome-Wide Analysis to Understand and Manage Complex Traits
Organizers: Gregory Ragland, Glen Hood, Scott Egan, Daniel Hahn, and Meredith Doellman
 - Pricing:
 - **Registration- \$475 before Sept 13th 2017 before September 13**

- Evolution 2018

- Montpellier, France Aug 18-22

3. When are we collecting?

4. Review a manuscript while Dan is away? Yep, insect sociaux

5. set up meeting with Chao to get training on respirometry.
 - Chao showed me general set up, will need to do it myself to get the hang of it
6. update 2017-08-01; get travel pcard. Completed online training and signed agreement form. Linda says ~1 month to get the card.
7. Undergrad volunteer
 - Free Aug 17th-21st to meet.
8. Two ms to send out to the group for review: September?

Future stuff:

Revisit IDP in 6 months

In december

Funding opportunities:

Blair(department head) brought up to Dan.

- Sloan INformatics fellowships
 - using existing datasets in the lab; complex genomic architecture problem (diffuse, inversions)
 - charlie baer; c elegans; genomic architecture
- USDA post docs
- NIH Rick Kirkeinstein?
 - somebody teaching how to apply for this; emailed Blockwood for his proposal

Goal by year 2: Have a killer job talk;

Have components

1. old work
2. post doc work
 - Rhagoletis- polygenic system; diffuse genomic architecture
 - European Cornborer- 1 major effect loci
3. future work

Prof dev:

Seminar series: There are 2 for biology where I can present my work

1. pop bio group, friday
2. actual seminar series

PI development

Need to get there first, so I need to write a:

1. Research statement
2. Teaching statement

Teaching

Important- teaching in a classroom setting. Boost yourself so that you're marketable. Show that you can teach a full course. norm- 1.5 to 2 courses a year

1. 5 week block of teaching intro bio classes
2. Upper level- grad level course

College of Ag, there is a teacher's college, held in our building. Post docs might be able to attend, 1 day /week course that helps you teach. Faculty level certification course. Pair that with mentored-teaching experience. Take over 1-2 weeks Dan's physiology course. Dan and I could teach a graduate seminar. I'll be an instructor of record. I'll get an official evaluation.

Try to do this in your 2nd year. I could teach a data science.

Grant related; writing/reviewing

Dan mentioned reviewing grants together; need permission from program officer

What kind of PI do I want to be?

Dan is 13 years- transition to doing research to manager. Greg is in the middle, hands on

I want a set of documents for my lab that will help members get situated.

1. Ray Huey's guide to being a successful grad student/scientist
2. Data science practices (data entry, storage)
3. Writing?
4. Expectations, rules?(co-authorship) , mentoring philosophy

dissertation progress

1. *Hsp rxn norm paper:*
 - **Sent out 2017-08-14!!!**
 - Submission goal: December 2017
2. **Range limits paper:**

- Sent draft to co-authors 2017-07-28
 - AEllison wants to try overleaf, so far so good
- Submission goal: December 2017

3. Proteome stability project:

- Still in data gathering + analysis step
 - Wai needs to HPLC fractionate tryptic peptides to ID more of them (emailed him 2017-07-27)
 - He won't get to this for a few months
 - I have been swamped with requests from investigators since Bethany left. I should be able to start working on it in the next couple of months.
 - ANBE needs gene
-

Page 115: 2017-08-17. meeting with Dan.

1. Field collecting:

- pick up mini van tomorrow at airport; state contract rates are etter at enterprise;
- grab mini van tomorrow; 2PM, i'll take it home and pick dan up saturday morning ; 7:45AM.
- eat breakfast and pack a lunch
- get to nashville 5PM; stay somewhere north of nashville ;
- address : 1630 northwest 89th terrace , gainesvile fl 32606; eagle point development

2. meeting with undergrad tomorrow 10AM

3. traveling to ESA arrangements; Nov 5-8

- pcard activated, special procedures?
 - start with travel auth form
 - Hotel covered
 - Airfare ; stay from 4-9?
 - Registration fee
 - Member fee
-

Page 116: 2017-08-24. Apple maggot collecting trip log

- We drove for 2 days 2017-08-19 to 2017-08-20 from fl to michigan
- Then we collected from 3 sites for Apple maggots on 2017-08-20
 - MSU - college campus; located in lansing
 - Close to campus we collected 1 site with Haws on the floor, somebody's front yard
 - Old grant (OG);
 - parked on somebody's property and asked them to sample, across the street deep in the woods and also an apple tree by the road
 - Ferris road (grant);
 - dirt road, microsympatry between haws and apples
- Drove back from michigan to florida for 2 days (Aug 22-23)

- On Aug 23, we put apples into bins with wiring to let the larvae fall to the bottom of the tub, which will make for easier collecting

MSU site pictures

Haw tree that has fruit down!



Where we actually collected apples



transfer bins to the lab:



apples in set up to collect larvae:



Summary of collection:

Number of tubs we've laid out in the lab. We poured apples into a tub with mesh so that larvae can be collected at the bottom. Apples were spread out 1-2 layers thick.

Site	Site Name	tubs
Grant	Old Grant	11
Grant	Ferris road	18
Lansing	MSU	25

Chao's protocol for handling Rhagoletis

Preparation of tub for holding fruits with maggot!

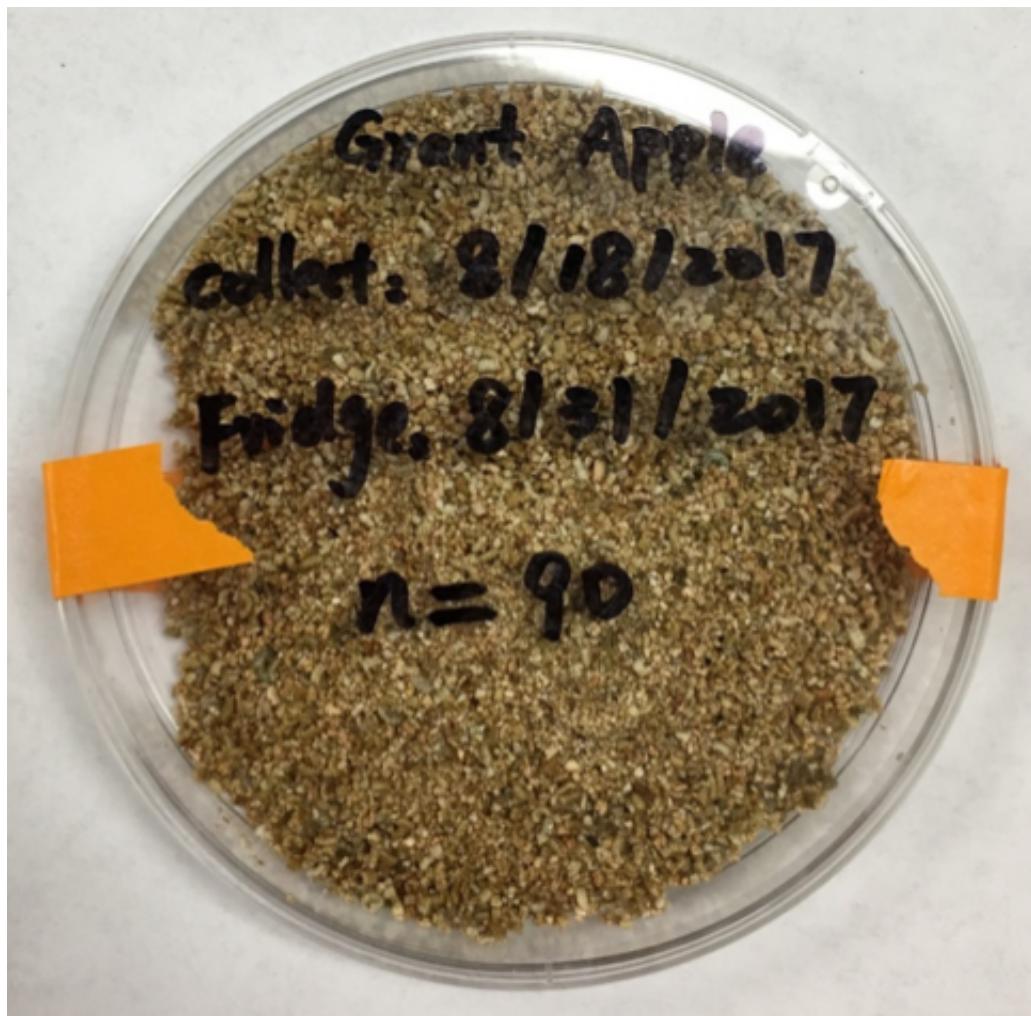
1. Put paper towels into the bottom of tubs to absorb the fruit juice. (As fruits keep rotting, you may need to increase the usage of paper towels)
2. Put the metal mesh into the tubs, and pour the fruits onto the metal mesh (don't put too many apples in each tub).
3. Label all the tubs clearly especially when the fruits are collected from different spots.

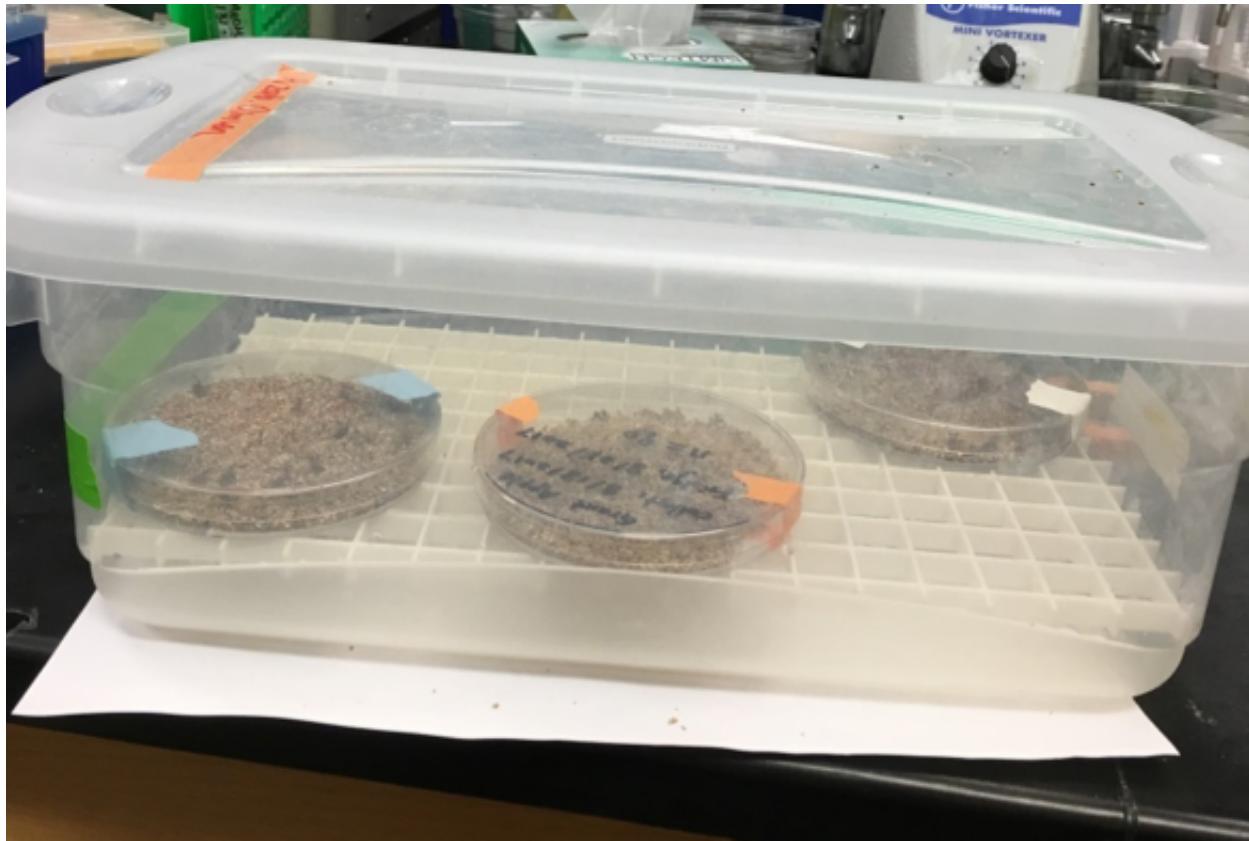
Collection of wandering larvae and puape

1. The wandering larvae will come out from the apple and drop into the tub. Take the metal mesh out so you can see the wandering

larvae or pupae at the bottom of the tub.

2. Pick the larvae or pupae out from the tub very gently and put them into one or more petri dish for now depends on the number.
3. Clean the tub with paper towels and put the metal mesh with apples back.
4. Split the petri dish with all of your collections into several petri dishes with wet (not too wet) vermiculite. Put no more than 100 pupae or larvae into each petri dish.
5. Label the petri dishes with collection spot, race, collection date, fridge date (we transfer the insect into fridge 13 days after collection), and number of samples.
6. Seal each petri dish with tape, and place the taped petri dishes into a container with saturated potassium chloride solution to keep the humidity (85%).





- Note: At the last few days of collection, the fruits will become very juicy so we sometimes clean the insects with water before moving into the vermiculite. Try to collect the insect at the same schedule every day

Respirometry:

1. Take the pupae out from the vermiculite at the day of respirometry. Normally we do purge in the afternoon and injection in the next morning. I can show you the details of respirometry later.

Notes: I need to autoclave vermiculite to provide a sterile environment for the larvae.

Sorted flies today, summary day 1

Table summary with how I think the data table can be laid out:

Site	Site_name	Host	Collect_larvae	Date_fridge	Petri_dish	Petri_density	Resp_day	Resp_date	respirometry	individ
Grant	Ferris	Apple	2017-08-24	2017-09-06	1	26	NA	NA	NA	NA
Grant	Ferris	Apple	2017-08-24	2017-09-06	2	46	NA	NA	NA	NA
Grant	OG	Apple	2017-08-24	2017-09-06	1	26	NA	NA	NA	NA
Grant	OG	Apple	2017-08-24	2017-09-06	2	80	NA	NA	NA	NA
Lansing	MSU	Apple	2017-08-24	2017-09-06	0	0	NA	NA	NA	NA

- Site is the city
 - Site name is the micro site ID
 - Host is apple or haw
 - When we collected larvae, date
 - Date when we should stick em in fridge
 - Petri dish block
 - petri dish density = number of individuals in dish
 - Resp day (7, 10, 11?)
 - REsp date- for blocking potentially
 - Respirometry - reading in C02
 - Individual for when we measure each pupae; can prob get a sense of what lives from beginning to when we sample for resp
-

Page 117: 2017-08-25. Apple sampling day 2

summary table

Site	Site_name	Host	Collect_fruit	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	2017-08-25	2017-09-07	1	AN	23
Grant	Ferris	Apple	2017-08-21	2017-08-25	2017-09-07	2	PE	21
Grant	OG	Apple	2017-08-21	2017-08-25	2017-09-07	1	AN	16
Grant	OG	Apple	2017-08-21	2017-08-25	2017-09-07	2	PE	45
Lansing	MSU	Apple	2017-08-21	2017-08-25	2017-09-07	0		0
Lansing	Coll Halls	Haw	2017-08-21	2017-08-25	2017-09-07	0		0
Lansing	MilfordST	Haw	2017-08-21	2017-08-25	2017-09-07	1	AN	2

When the apples turn too rotten (> ~75% brown), then put them in a separate tub to see what comes out of them before discarding.

Today we had the following "rotting tubs"

1. Grant - 1
 2. Ferris - 2
-

Page 118: 2017-08-26 Sampling apple maggots day 3

summary of collecting

Site	Site_name	Host	Collect_fruit	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	2017-08-26	2017-09-08	1	AN	49
Grant	OG	Apple	2017-08-21	2017-08-26	2017-09-08	1	AN	88
Lansing	MSU	Apple	2017-08-21	2017-08-26	2017-09-08	0		NA
Lansing	Coll Halls	Haw	2017-08-21	2017-08-26	2017-09-08	1	AN	3
Lansing	MilfordST	Haw	2017-08-21	2017-08-26	2017-09-08	1	AN	2

Notes:

The "rotting tubs" tend to have more maggots emerging. They probably have to get out before facing bacterial/fungal infection from rotting fruit.

Things for me to order:

- 5 mL syringe Norm-Ject Luer lock
- table?
- soft forceps
- lower syringe thingy
- potassium chloride (done through myufl and mymarket. assign person to pamela)

Page 119: 2017-08-27. Sampling apple maggots day 4

summary table

Site	Site_name	Host	Collect_fruit	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	2017-08-27	2017-09-09	1	AN	81
Grant	OG	Apple	2017-08-21	2017-08-27	2017-09-09	1	AN	96
Lansing	MSU	Apple	2017-08-21	2017-08-27	2017-09-09	0		0
Lansing	Coll Halls	Haw	2017-08-21	2017-08-27	2017-09-09	1	AN	2
Lansing	MilfordST	Haw	2017-08-21	2017-08-27	2017-09-09	1	AN	3

Page 120: 2017-08-28. Calibrating Li-cors (measures co2)

The li-cors need to be calibrated to make sure the readings for co2 are right. The hahn lab has 2, which are stacked together (6262 and 7000). The 6262 cant be hooked up to the computer, but the 7000 one can. Calibrate once a month usually.



references for the tubing and samples--

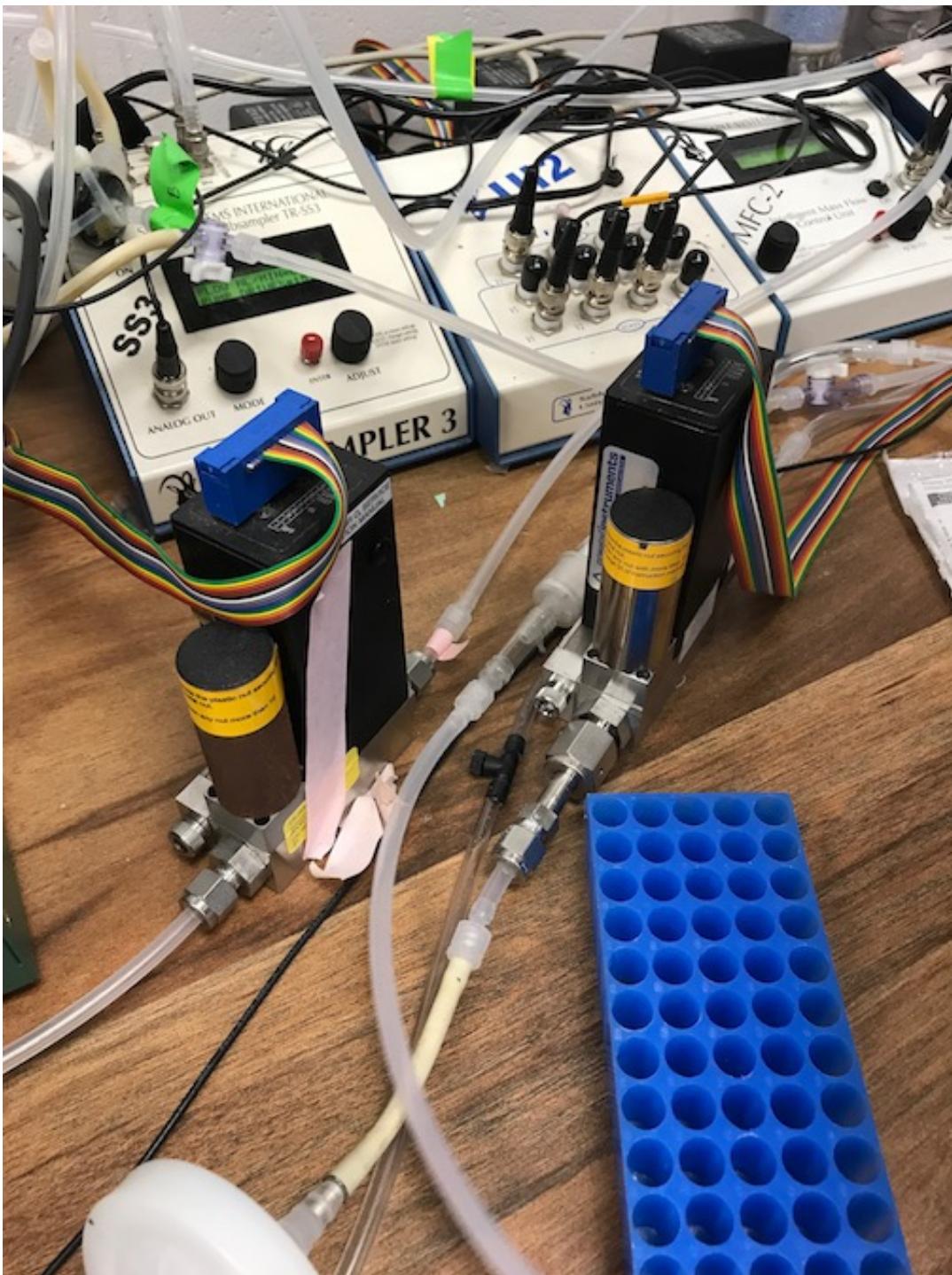
1. A is the reference cell
2. B is the sample cell



For calibration there are 3 general steps:

1. calibrate the reference(A) and sample cells(B) with zero co2.
2. calibrate sample cell(B) with co2
3. calibrate both cells with co2

Flow rate hook ups

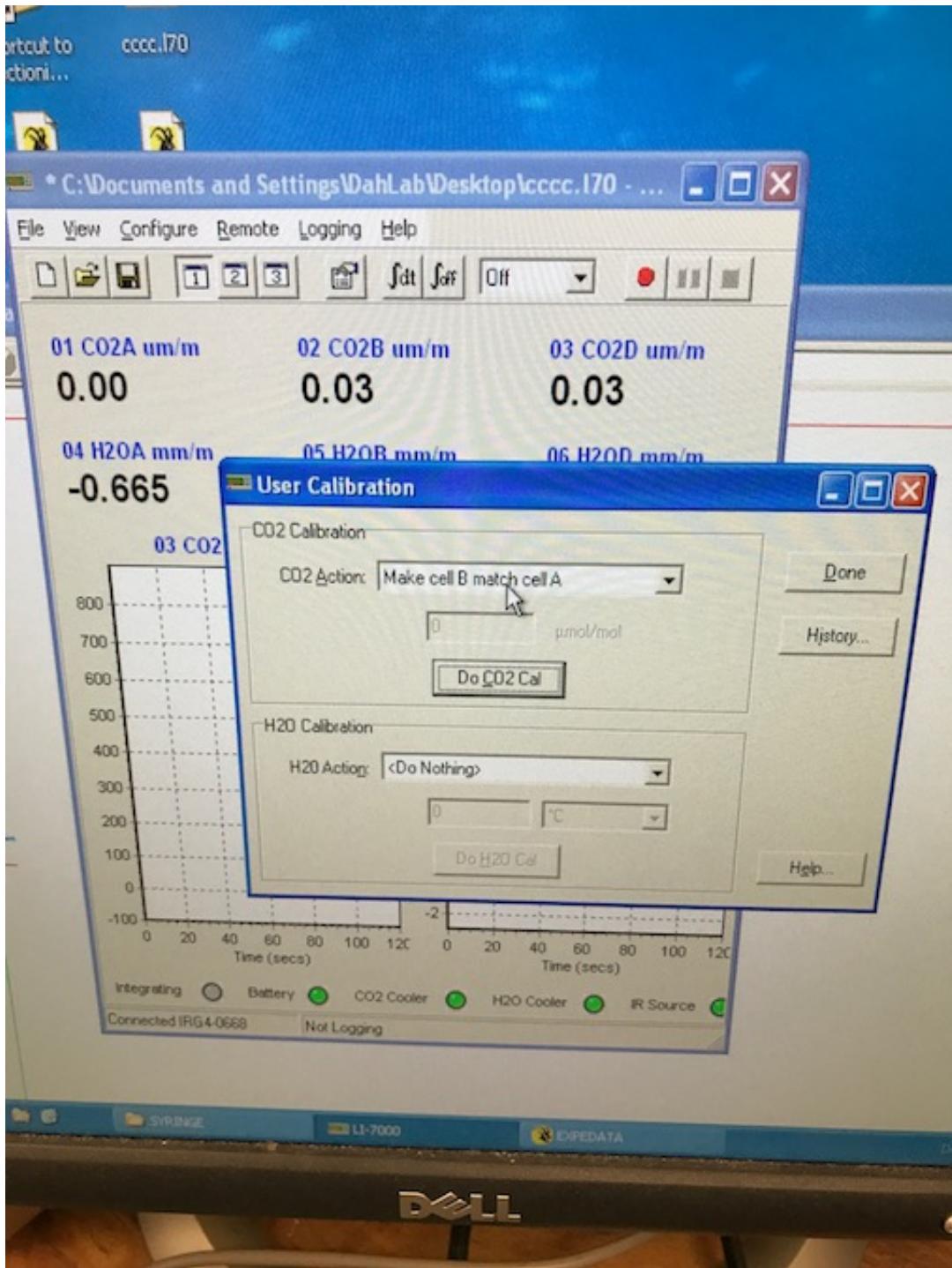


more specific steps:

**You need two gas cylinders:

- N2 for control
- co2 for calibration

1. Connect N2 to both cells (B in and A in). The flow rate will be 150 mL/min
 - Nitrogen will serve as the "zero"
 - Make sure A out and B out are not connected to anything (connected to air).
 - The flow controllers go from 0-200 mL/min, so setting the interfaced control thingy to 75% will result in 150ml/min
 - make sure flow rate is the same for both cells
 - set up is in my written notes below
2. On the computer, go to "Remote", then "User Calibration"
3. Under co2 action: click "Make cell A read 0". And click "Do co2 cal".
4. Then click "Make cell B read 0". and click "Do co2 cal"



5. Calibrate sample cells with co2. Disconnect sample cell b, then connect co2 line.
6. On the computer, go to "Remote", then "User Calibration"
7. "Make cell B read", spent gas which is 465.8 for the tank. **This will be tank specific, so you have to look at the tag**
8. Connect co2 to both A in and Bin.
8. On computer, "Make cell A" read the spent gas (465.8-again this is

tank specific)

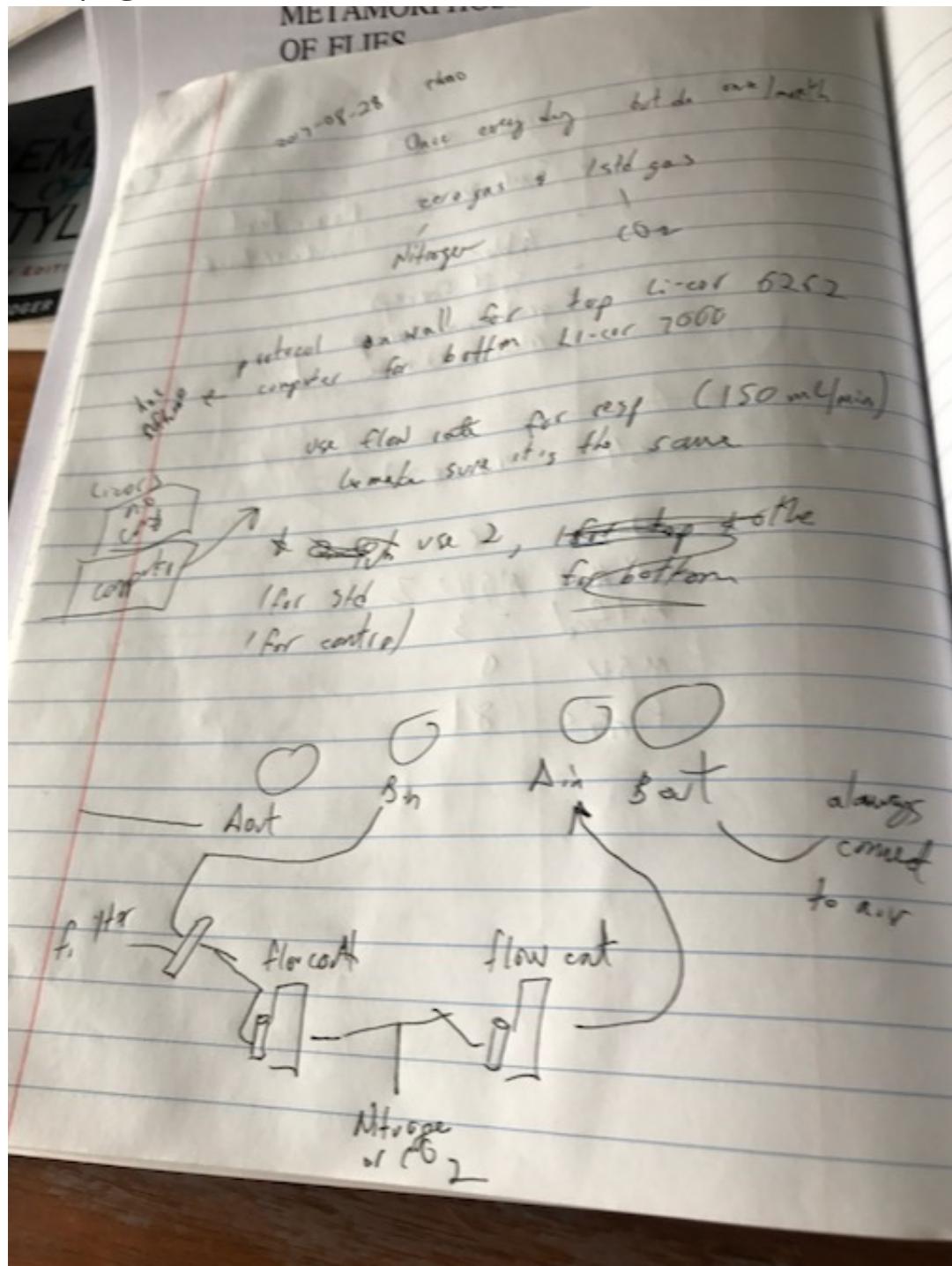
For the licor 6262, do the same, but you have to interface the machine.

zero'ing manually the 6262 licor



my written notes

first page



second page

A is return
 B is sample cell
 controlling flow rate
 0-200 mL/min
 set to 75% to get 150 mL/min

on half
 last

- 1) connect
- 2) do O₂ to both cells
- 3) disconnect sample cell, B, then
connect std(CO₂)
to check ppm (465.8) in cylinder
- 3) connect CO₂ (465.8) to both cells

make cell A read
 465.8

Make cell B read
 "spent gas"
 465.8

- Dr CO₂ out

1) rezero
 2) vsr calibration

- CO₂ active make cell A read 0
 Do CO₂ cal

- Make cell B read 0
 Do CO₂ cal

Page 121: 2017-08-28. Sampling apple maggots day 5

Site	Site_name	Host	Collect_fruit	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	2017-08-28	2017-09-10	1	PE	102
Grant	OG	Apple	2017-08-21	2017-08-28	2017-09-10	1	AN	94
Lansing	MSU	Apple	2017-08-21	2017-08-28	2017-09-10	0		0
Lansing	Coll Halls	Haw	2017-08-21	2017-08-28	2017-09-10	0		0
Lansing	MilfordST	Haw	2017-08-21	2017-08-28	2017-09-10	1	AN	11

I should think about the fridge date. I want to measure wet mass and co2 production at 2 time points. Which ones? day 10 and 15? Then weigh mass day 9 and 14. Fridge a subset day 15.

written workflow:

I also need to get a physical notebook.

Page 122: 2017-08-29. Apple maggot sampling - day6

summary table:

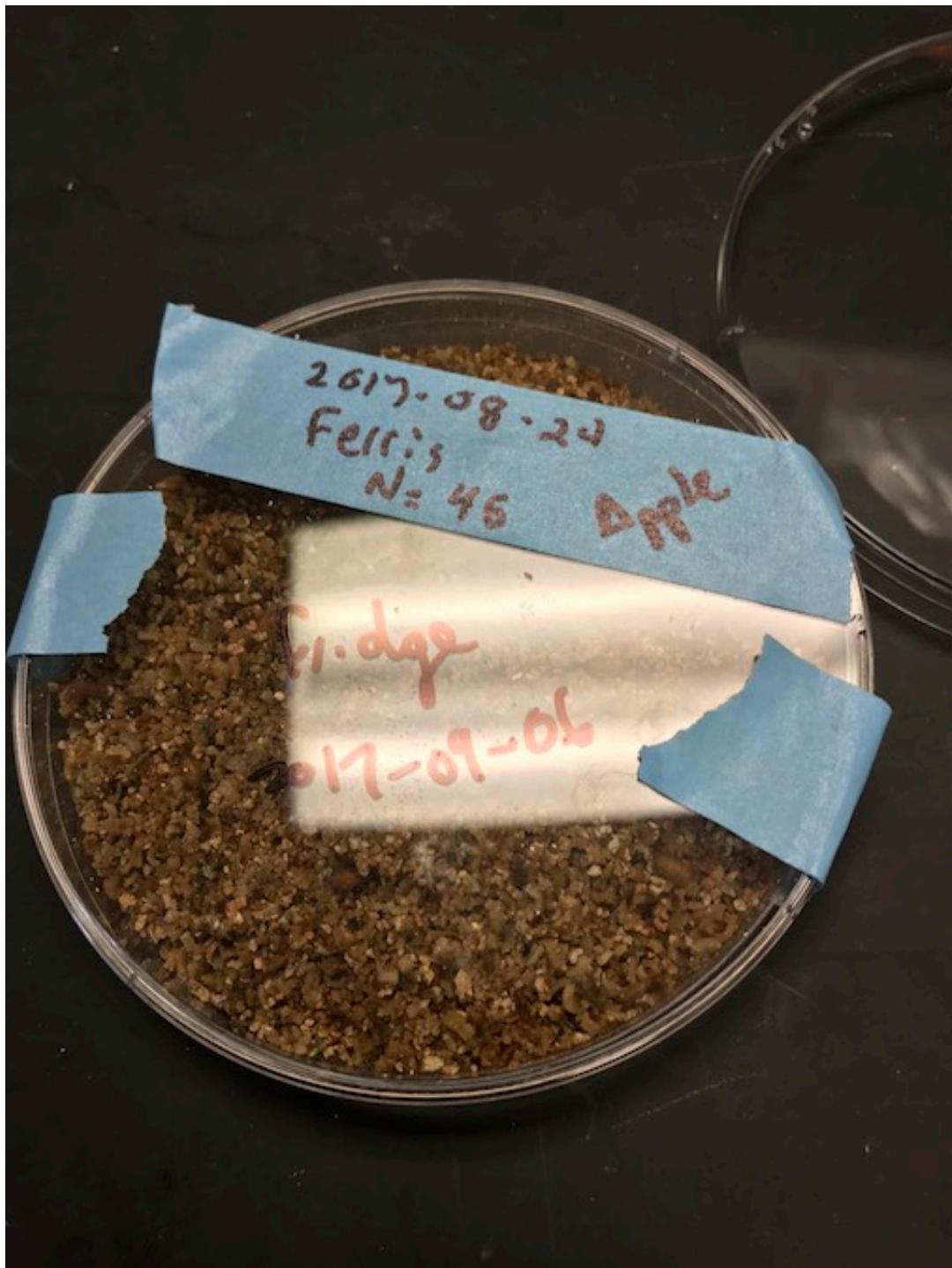
Site	Site_name	Host	Collect_fruit	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	2017-08-29	2017-09-11	1	PE	88
Grant	OG	Apple	2017-08-21	2017-08-28	2017-09-10	1	AN	111
Lansing	MSU	Apple	2017-08-21	2017-08-28	2017-09-10	1	AN	2
Lansing	Coll Halls	Haw	2017-08-21	2017-08-28	2017-09-10	0		0
Lansing	MilfordST	Haw	2017-08-21	2017-08-28	2017-09-10	1	AN	8

Notes: The rotting tubs still have a lot of maggots coming out. PE dropped ferris.

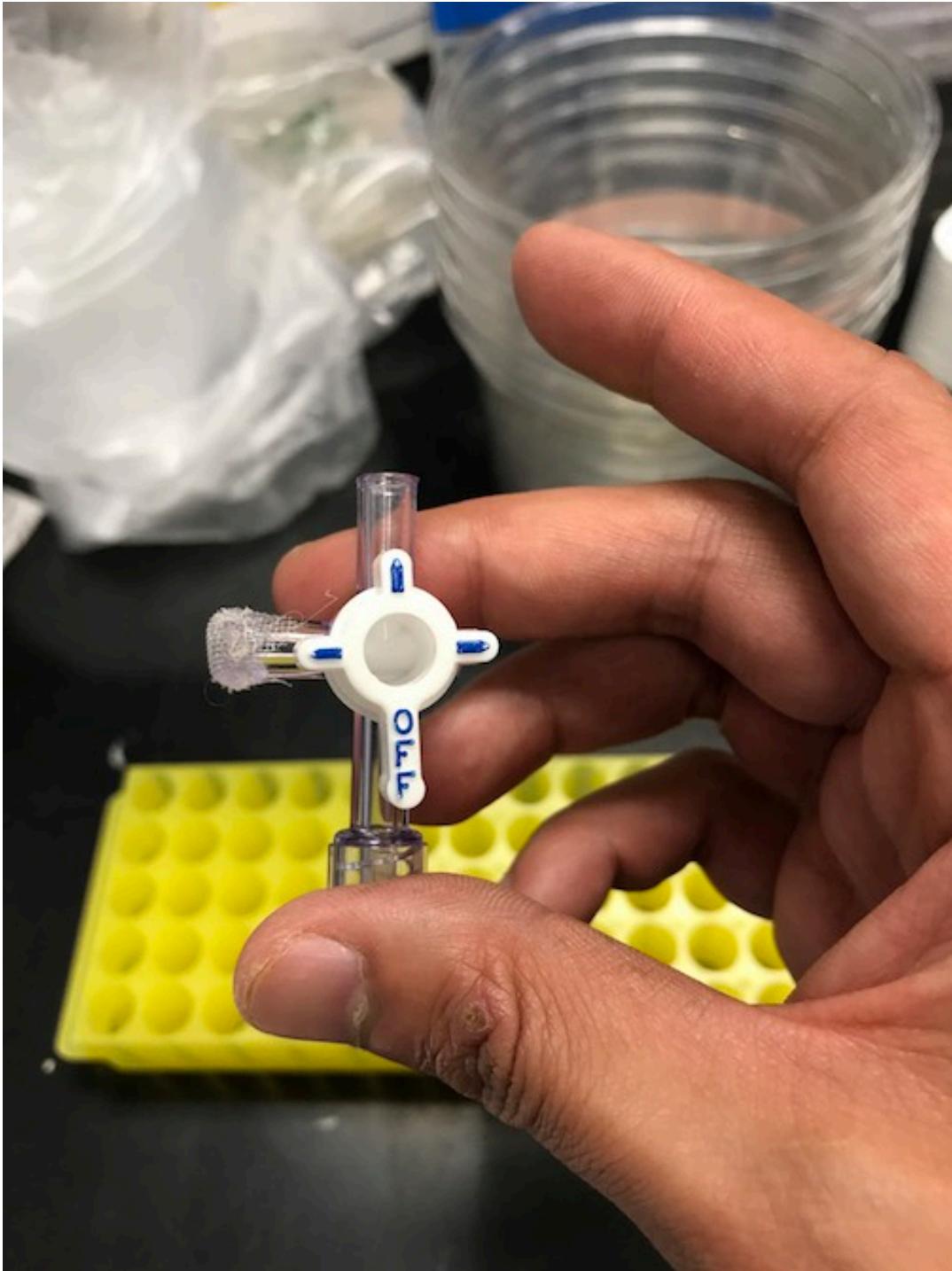
Page 123: 2017-08-29. Rhatoletis respirometry- purging

1. We took the first day set of pupae:

- 2017-08-24, Ferris, apple, n=46
- took 9 pupae out
- we included two blanks per block (there were two blocks)



2. Placed them in syringe nozzle with



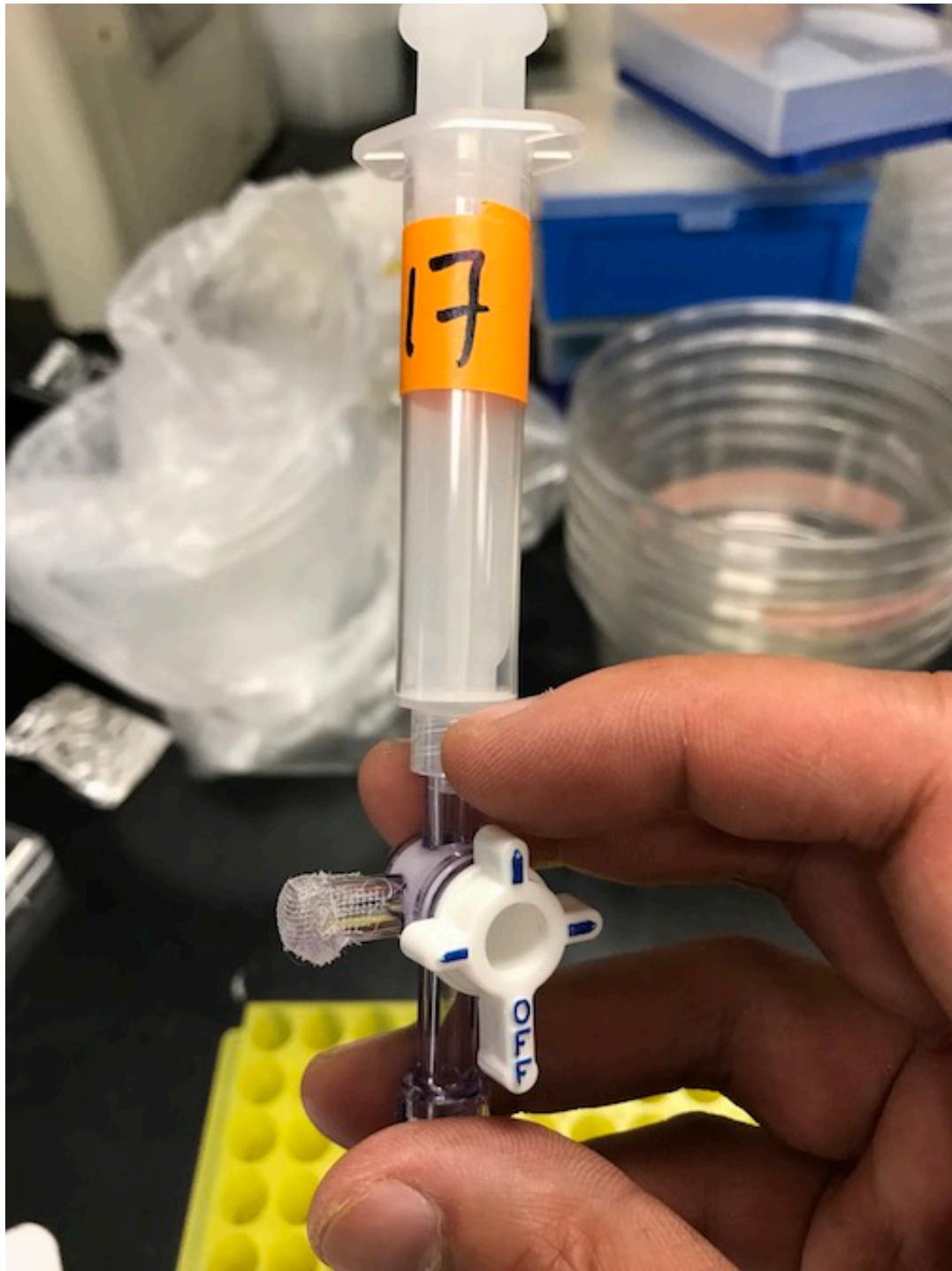
1 | * do this for a set of blocks

3. Take the syringe and hook it up to the nozzle

syringe



syringe + nozzle



4. Purging setup:



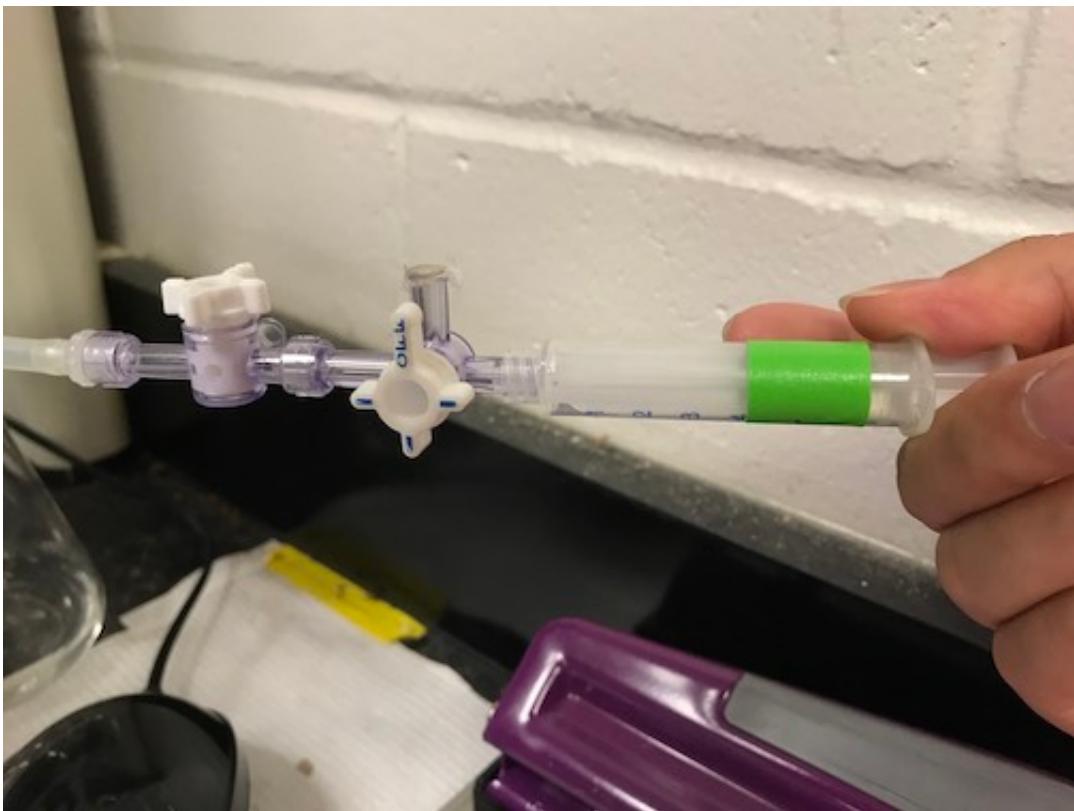
It looks like air moves through the dririte and then through the water for hydration.

5. Allow air flow for both tubes of dririte

picture of one



6. The pupae in the tubes are ready to be purged. Meaning we're removing the existing air and putting in air without co₂ to standardize future co₂ measurements.
7. First, twist white part until off is facing syringe to remove air in the connection. Then twist off toward you so that the off is pointed to you and pull air for 1 mL. Then twist the off towards the mesh. ***need to double check the steps!!****



Throughout process, document the date, start time, finish time, and any notes. Keep syringe set ups on rack and place back into the rearing room. For rhagoletis, we need to let them go over night so that enough co2 is accumulated for measurements.

Stuff I need to order, do, and think about:

1. Luer stockcocks -[3 way](#)
2. order tubs with white grates
3. find a suitable fridge to overwinter these animals
4. order 5 mL [syringe](#)
5. get a notebook

2017-08-30 Purging revisited:

When purging, it is important to get rid of the previous air. Here are the steps for purging when everything is turned on .

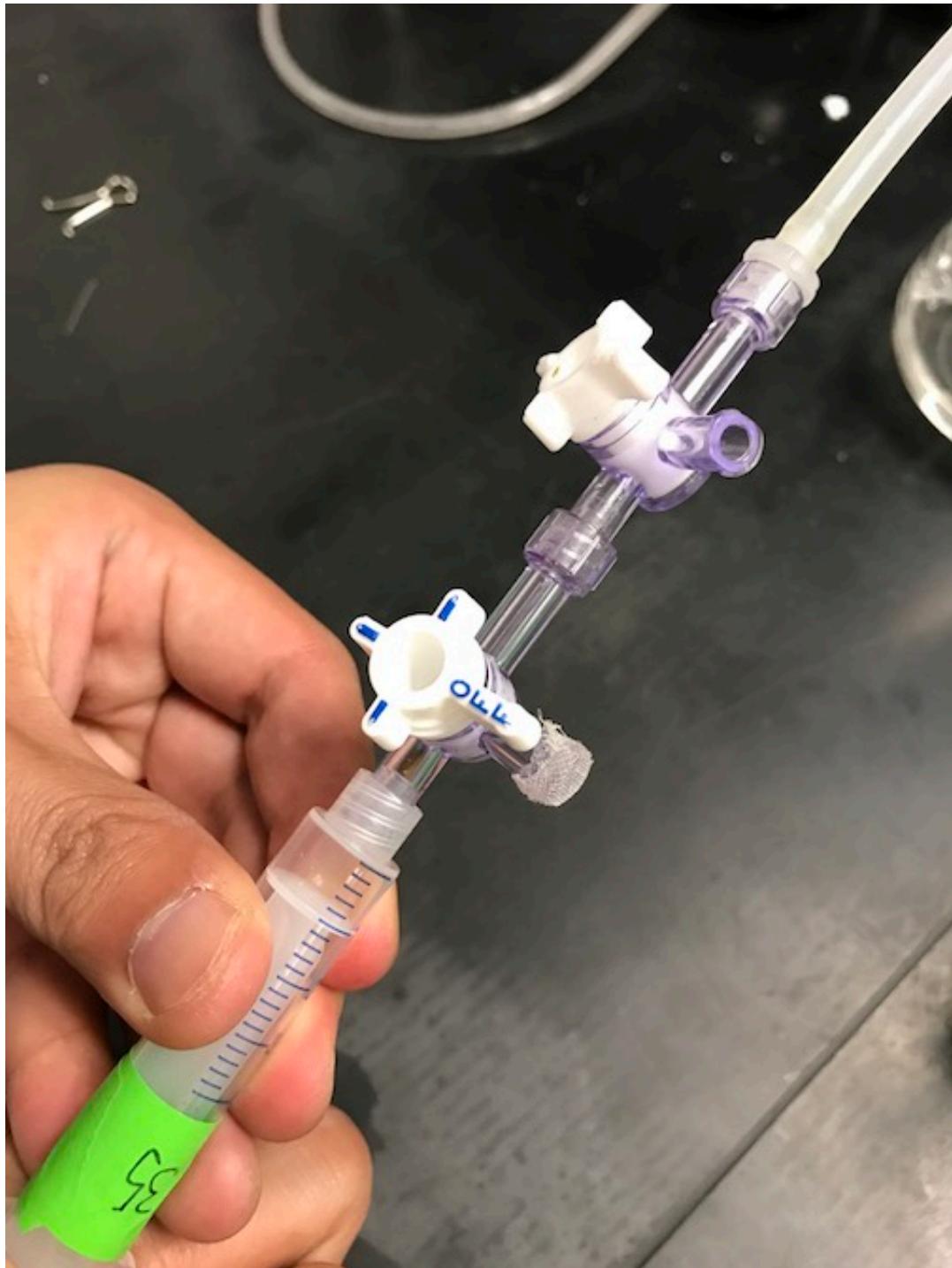
1. Turn the off button left to let air flow out from the connection and syringe.



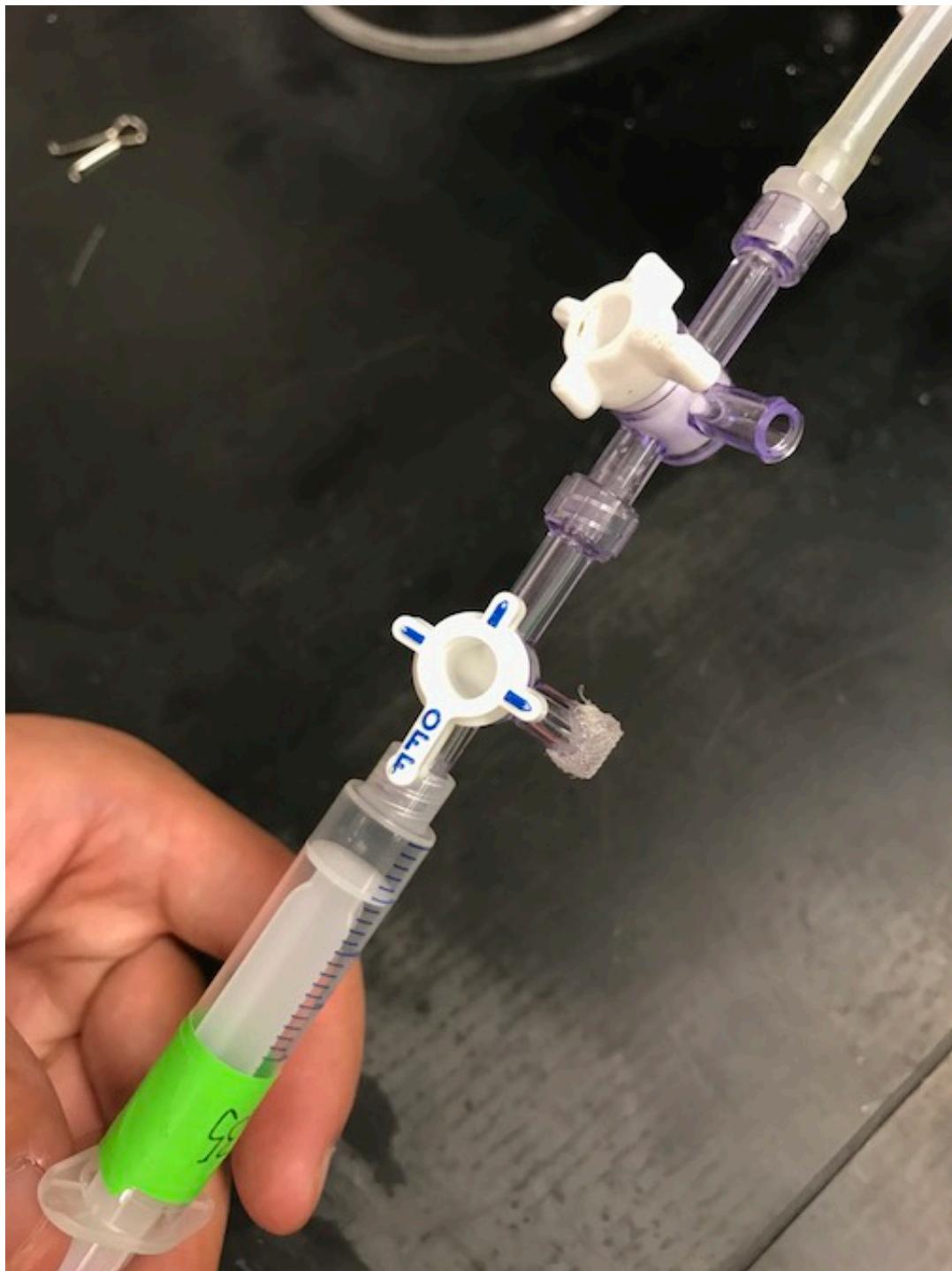
2. Turn 180 degrees counter clockwise. This lets air from purger to syringe only and covers the outlet with the mesh



3. For rhagoletis, pull 1 mL of air



4. Turn nozzle clockwise (towards you) to shut off air going into syringe



5. unplug! :-)

Page 124: Apple maggot sampling day 7

summary table

Site	Site_name	Host	Collect_fruit	Collect_day	Collect_larvae	Date_fridge	Petri_dish	Collector	Petri_density
Grant	Ferris	Apple	2017-08-21	7	2017-08-30	2017-09-12	1	PE	87
Grant	OG	Apple	2017-08-21	7	2017-08-30	2017-09-12	1	AN	122
Lansing	MSU	Apple	2017-08-21	7	2017-08-30	2017-09-12	0		0
Lansing	Coll Halls	Haw	2017-08-21	7	2017-08-30	2017-09-12	1	AN	2
Lansing	MilfordST	Haw	2017-08-21	7	2017-08-30	2017-09-12	1	AN	13
				NA			NA	total	1249

with total so far.... .

Page 125: 2017-09-18. Rhagoletis project update:

No time to enter everything I'm doing everyday on electronic notebook, but it is in a physical notebook.

Here is our workflow(I'll add a picture later):

Here is a calendar of what our group has been doing:

Today		September 2017						
Sun	Mon	Tue	Wed	Thu	Fri	Sat		
27	28	29	30 10 Tatiana hours	31	Sep 1	2		
3	4 2017-08-25 cohort day 10 weights	5 2017-08-26 cohort day 10 weights	6 2017-08-26 cohort day 11 resp 2017-08-27 cohort day 10 weights 2p Chelsea hours	7 2017-08-27 cohort day 11 resp 2017-08-28 cohort day 10 weights 10 Tatiana hours	8 2017-08-25 cohort day 14 weights 2017-08-28 cohort resp day 11 10 Tatiana hours	9 2017-08-25 cohort day 15 resp 2017-08-26 cohort day 14 weights 9 Chelsea hours		
10 2017-08-26 cohort day 15 resp 2017-08-27 cohort day 14 weights	11 2017-08-27 cohort resp day 15 2017-08-28 cohort day 14 weights	12 2017-08-28 cohort resp day 15 2017-09-02 cohort day 10 weights 10 Tatiana hours	13 2017-09-02 cohort resp day 11 2017-09-03 day 10 weights 10 Tatiana hours	14 2017-09-03 day 11 resp 2017-09-04 cohort day 10 weights 10 Tatiana hours	15 2017-09-04 cohort day 11 resp 2017-09-05 day 10 weights 10 Tatiana hours	16 2017-09-02 cohort weights day 14 2017-09-05 day 11 resp 2017-09-06 cohort day 10 weights 9 Chelsea hours		
17 2017-09-02 cohort resp day 15 2017-09-03 day 14 weights 2017-09-06 cohort day 11 resp 9 Chelsea hours	18 Collect Haws 2017-09-03 day 15 resp 2017-09-04 cohort day 14 weights 2017-09-08 cohort(d16) day10 weigl	19 2017-09-04 cohort day 15 resp 2017-09-05 day 14 weights 2017-09-08 cohort(d16) day10 resp	20 2017-09-05 day 15 resp 2017-09-06 day 14 weight 2017-09-10 cohort(d18) day10 weigl	21 2017-09-06 day 15 resp 2017-09-07 cohort(d18) day10 resp 2017-09-11 cohort(d19) day10 weigl	22 2017-09-08 cohort(d16) day14 weigl 2017-09-11 cohort(d18) day11 resp skipped cohort day 15 2017-09-07 10 Tatiana hours	23 2017-09-08 cohort(d16) day15 resp 9 Chelsea hours		
24 2017-09-10 cohort(d18) day14 weig 9 Chelsea hours	25 2017-09-10 cohort(d18) day15 resp 2017-09-11 cohort(d19) day14 weig 10 Tatiana hours	26 2017-09-11 cohort(d19) day15 resp 9 Kylie hours 2p Chelsea hours	27 9 Kylie hours 10 Tatiana hours	28 9 Kylie Hours 10 Tatiana hours	29 10 Tatiana hours	30 9 Chelsea hours		

Page 126: 2017-09-22; field notes from 2017-09-19 hawthorne collection

Collected hawthorne fruit from lansing and grant michigan with Greg and Mac.

Lansing

this is the raintree site, trees are next to apartment complex. Typically, we collect haws/apple off of tree by laying a tarp and shaking branches so the fruit falls off.



Grant

We only collected on ferris road.



Collected 5-6 hours



Page 127: 2017-11-10. Overall project status, planning, to do 's

Dissertation

1. Hsp rxn norm paper:

- SHC comments: mainly intro , results focus
- Plan to work on this once I send out range limits paper to AEllison
- Submission goal: December 2017

2. Range limits paper:
 - Address SHC and Hahn lab comments
 - Send out to collabs end of Nov
 - Submission goal: December 2017
3. Proteome stability project:
 - Still in data gathering + analysis step
 - Wai needs to HPLC fractionate tryptic peptides to ID more of them
 - ANBE needs gene annotations for aphaenogaster
4. Thermal niche paper
 - In Ichick's hands
 - Nsanders wants to look at it.
 - Submission goal: yesterday
5. Aphaeno genomes project:
 - Mlau needs to annotate-get GO's and KOGs
6. Stressed in nature project:
 - IN SHC's hands; goal to have draft by end of summer

Post doc

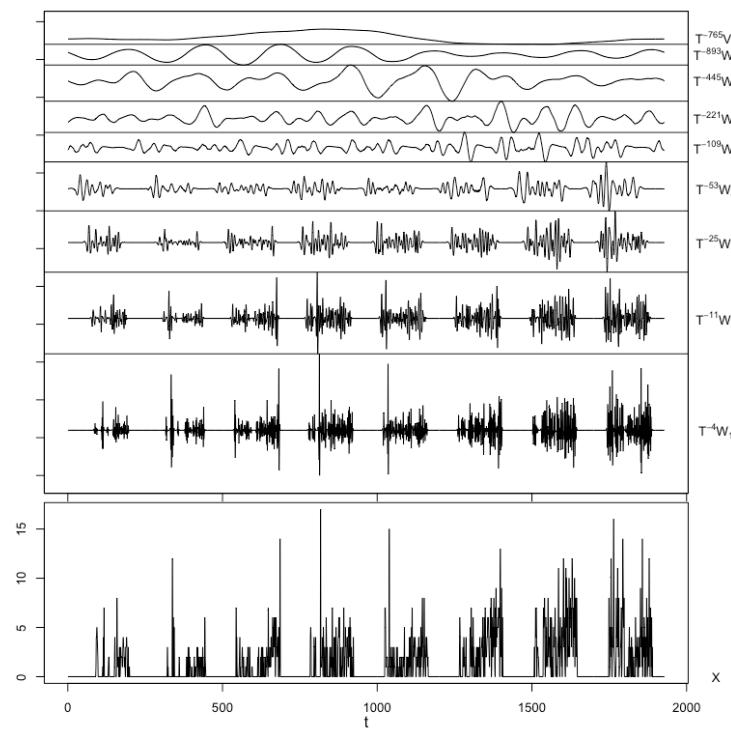
1. Circadian Rhythms
 - Need to input all of the resp and weight data for day 11 and day 15
 - Plot and show data (mass specific metabolic rate), consistent with Greg's work?
 - Analyze Trikinetics data with wavelet analysis
 - Discrete wavelet analysis to give interval bouts of activity
2. Meta-analysis of Circ Rhythms

- read more papers and make an excel spreadsheet
 - What shapes variation in biological rythms? day-season ?
-

Page 128: 2017-11-14. Prepping meeting with Dan ; Thursday 2017-11-16

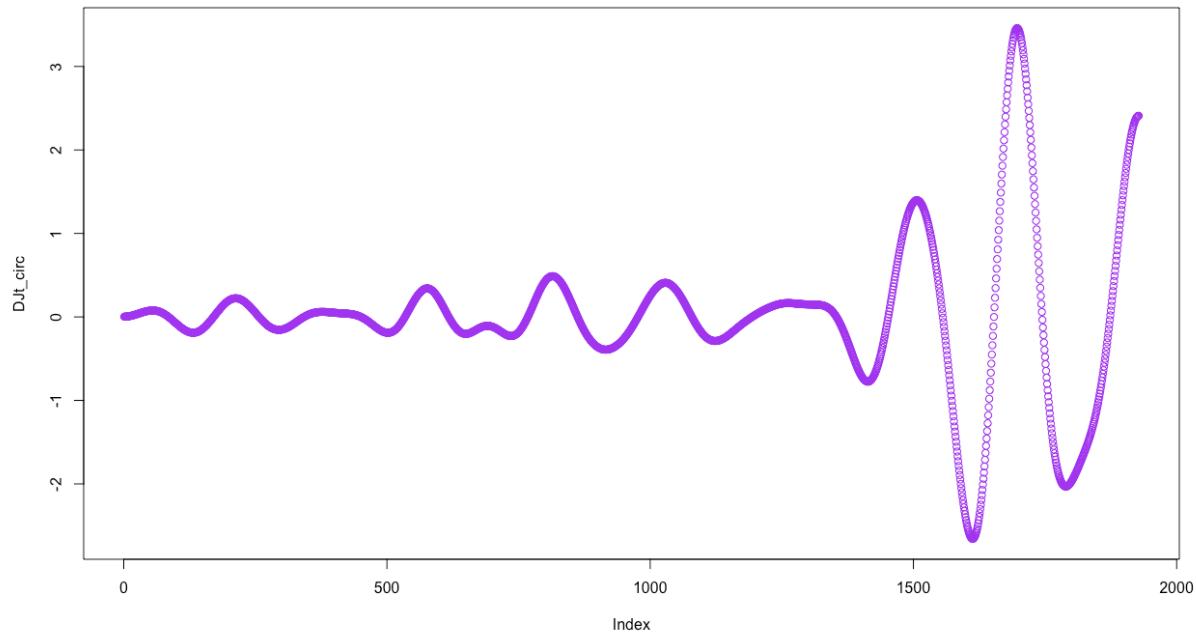
1. Show some biological rhythm data
 - Show interval bout data vs time of day

Discrete Wavelet Analysis

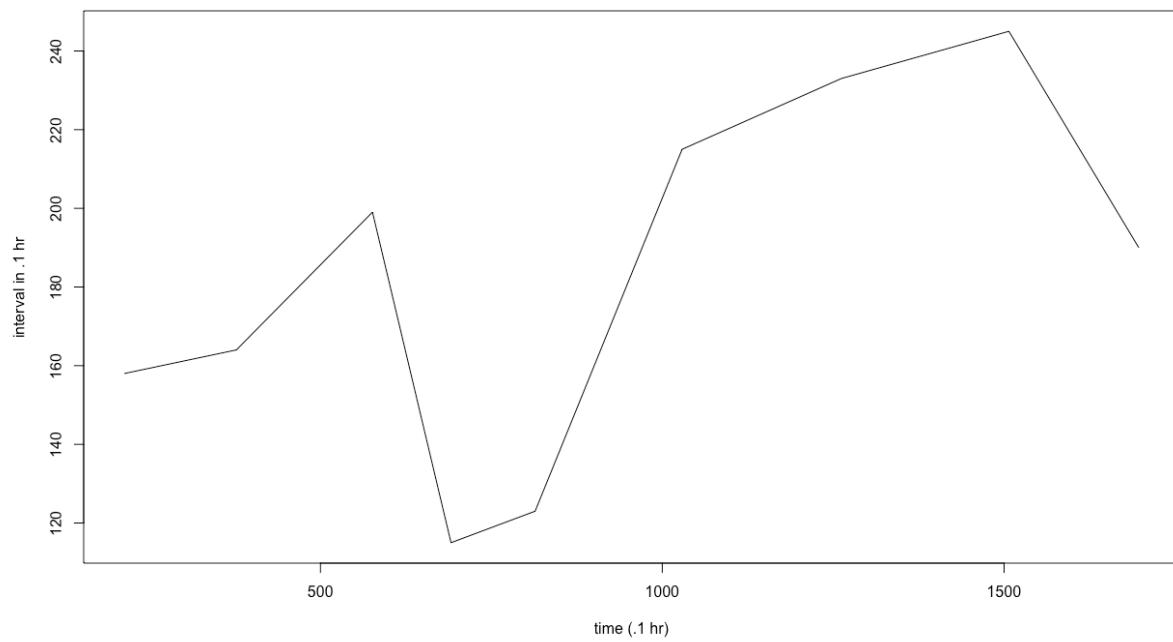


this shows the coefficients that break down the biological rythms into different period slices

This shows that when you add all the details



Interval bout figure



Code dump

```
1  ### Tanya Leise suggestion
2  Jcirc <- floor(log2(round(24/.1)))
```

```

3 DJt_circ <-
4   wavMRDSum(col.3.1$count, levels=Jcirc, keep.smooth=FALSE,
5   keep.details=TRUE, reflect=TRUE, wavelet="s12", xform="mod
6   wt")
7 DJt_circ
8
9
10 library(pracma)
11 IBL<-data.frame(findpeaks(DJt_circ))
12 names(IBL)<-
13   c("Height", "mid_time", "initial_time", "final_time")
14 #plot(IBL)
15 diff(IBL$mid_time)
16
17 plot(IBL$mid_time[-1],diff(IBL$mid_time),xlab="time (.1
18 hr)",ylab="interval in .1 hr",type="l")
19 lines(loess(diff(IBL$mid_time)~IBL$mid_time[-1],span=1)
20 )
21
22 plot(IBL[,2:1])
23 plot(diff(IBL$mid_time),IBL[-1,1])
24 lines(loess(IBL[-1,1]~diff(IBL$mid_time)))

```

```

1 * Status of the rigs:
2   * rigs 1-2 = 16L:8D cycle
3   * rigs 3-6 = free run (complete darkness)

```

2. Undergrad proposal writing for SURF

- she knows R and wants to learn more
-

3. Greg looking for samples:

- *GJR drawer 1, box/slot 2, and the box should say "Mexican pupae" on them.*
- 4. Read more papers for meta-analysis
 - What exact papers should I target?
 - **before doing this, phil trans paper on field/wild clocks before deciding what to do**

2017-11-16: Meeting notes

Mike wade suggestion on on setting up behavioral assays with volatiles; talk to Tom, Shannon, Cheyenne, and read more

for #2,made 3rd round of edits and sent to out

Page 129: 2017-11-17: Rhago notes , looked at Feder papers again; refs- [Page 91: 2017-07-11](#). Reading rhagoletis papers

Feder et al. 1997; PNAS

2017-11-17

paper readings

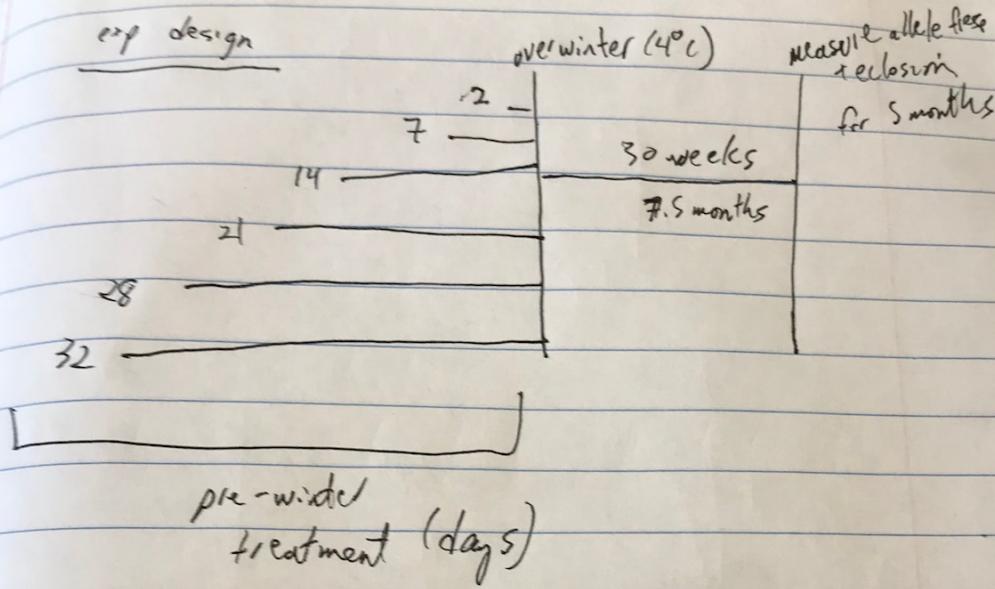
Giant size

PNAS

* exp on
Haws

Fedor et al. 1997 ^ → selective maintenance of
allozyme shift among host races

exp design



Loci - high in haws

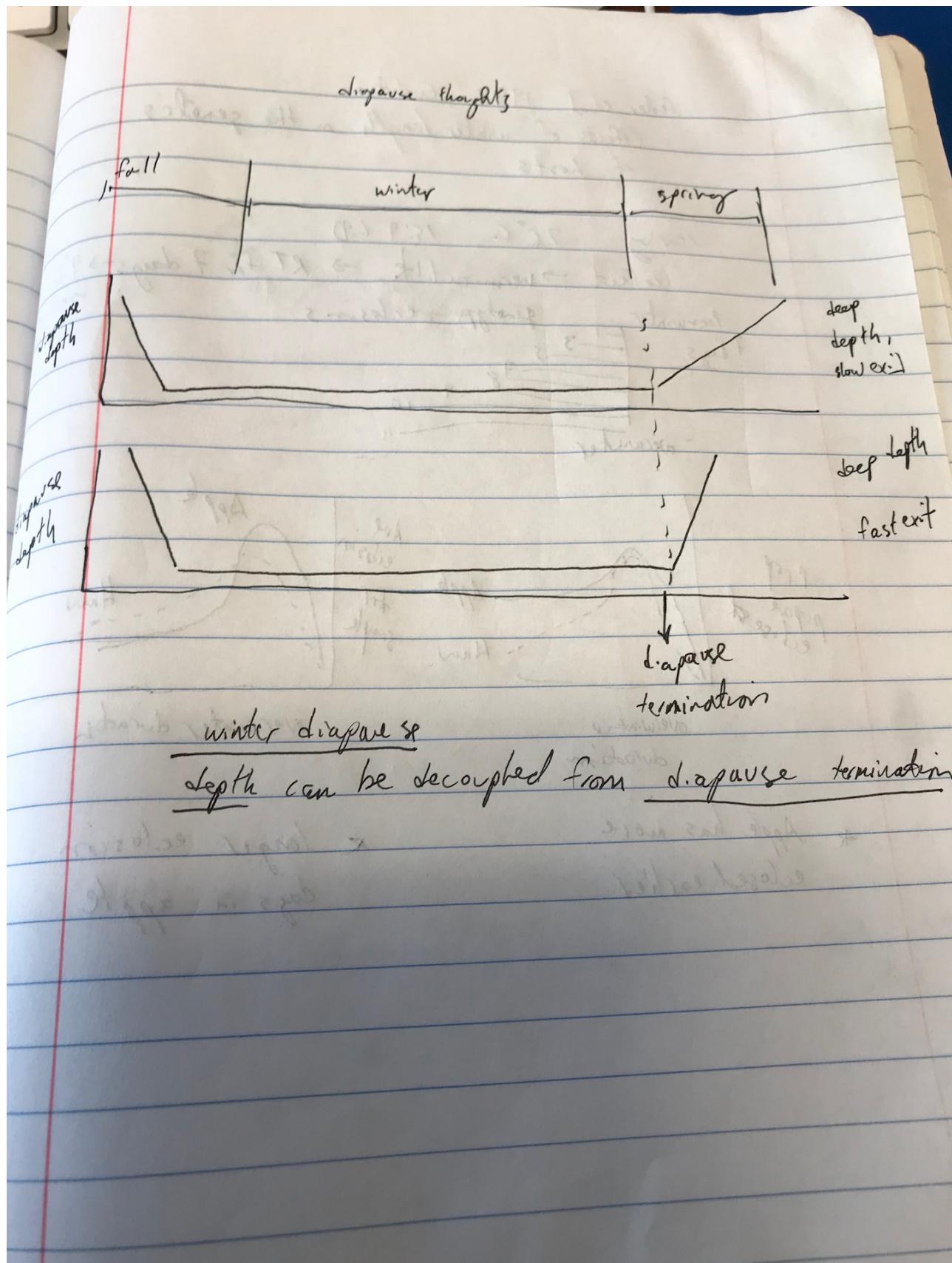
1) Me100 - high in haws

2) Acon-295

3) Mpi37

* allele freq shift from Haw to apple-like
occurred at longer prewinter lengths
(in Haw population)

Is it possible for diapause depth to differ or become decoupled from diapause exit? Yes, if the rate of diapause termination is more directly related to eclosion timing. Ragland has a phys paper on how termination is biphasic, it'd be interesting to test this hypothesis by determining whether the first initial increase in metabolic rate or just overall metabolic rate in diapaused individuals better explains eclosion timing.



Feder et al. 1997; Evolution

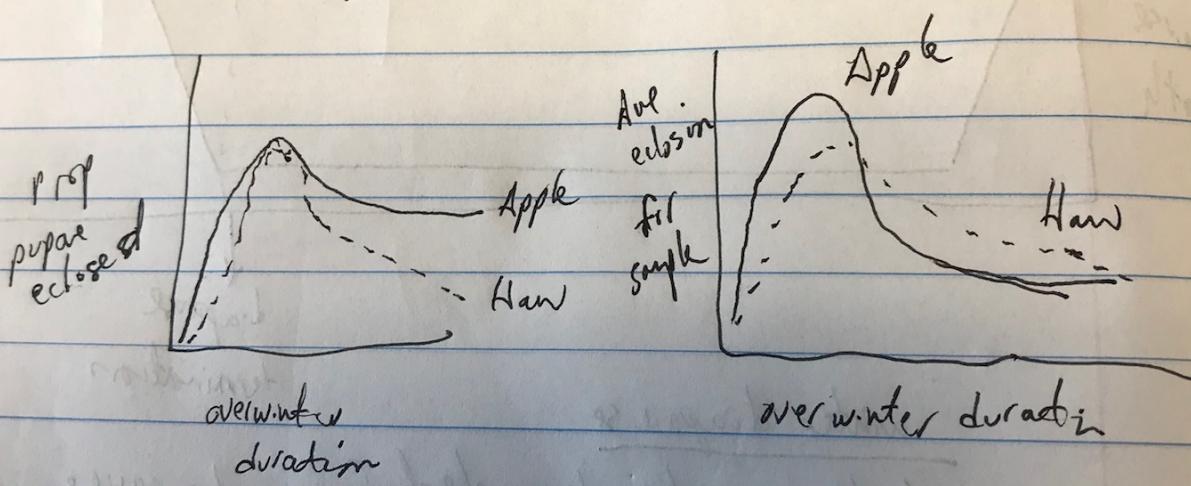
Feulner et al. 1997; Evolution
Effects of winter length on the genetics
of hosts

1 early 26°C 15:9 L:D
1b →

collect → vermiculite → RT for 7 days → 4°C winter

Prewinter genotype + eclosions

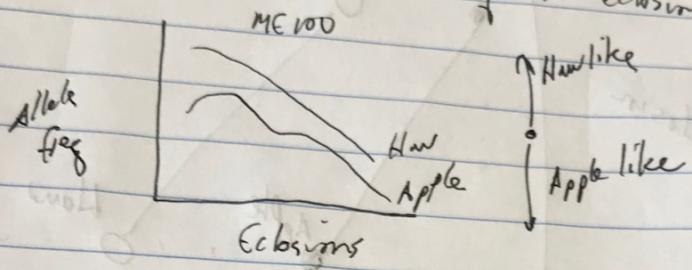
Pre-winter $\frac{7 \text{ days}}{3}$



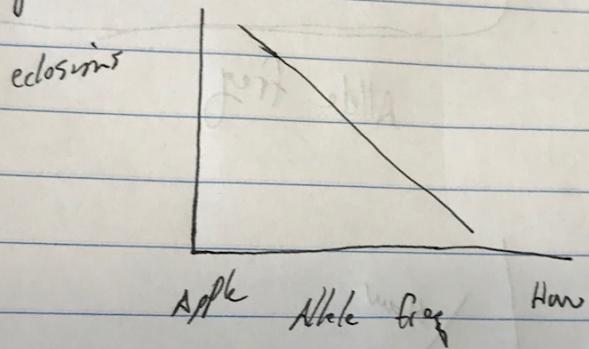
- * Apple has more eclosed earlier
 - * longer eclosion days in apple

cont'd

paper plots Allele freq vs. eclosion ...



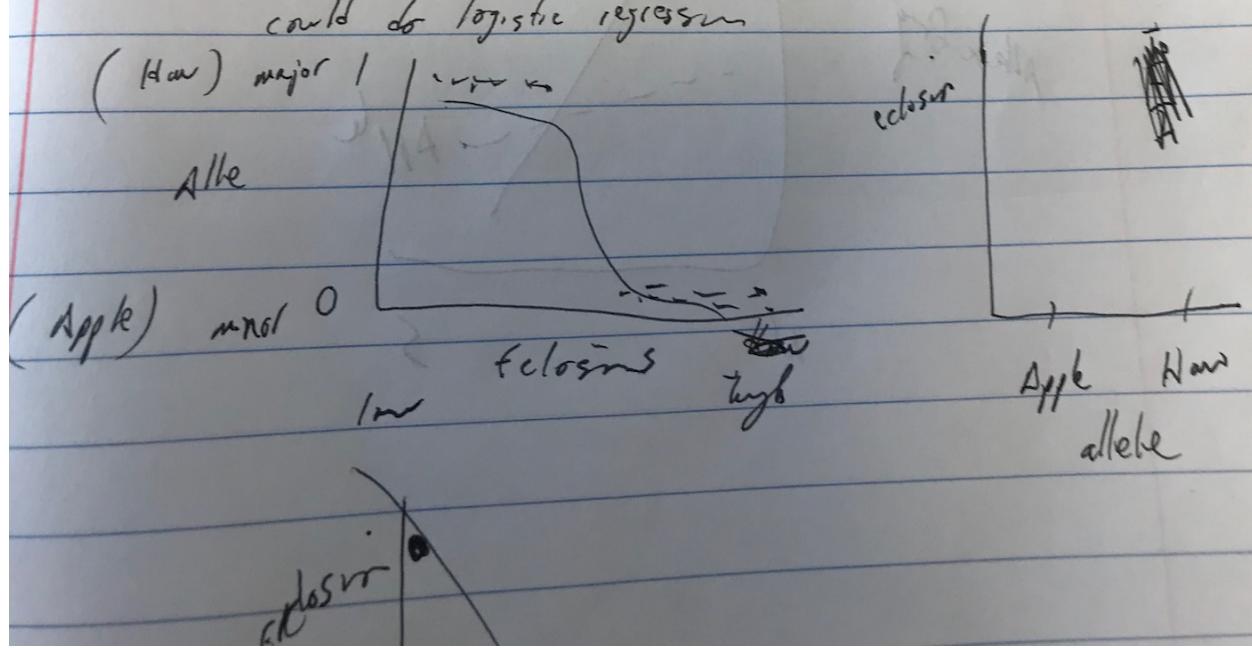
try reverse?



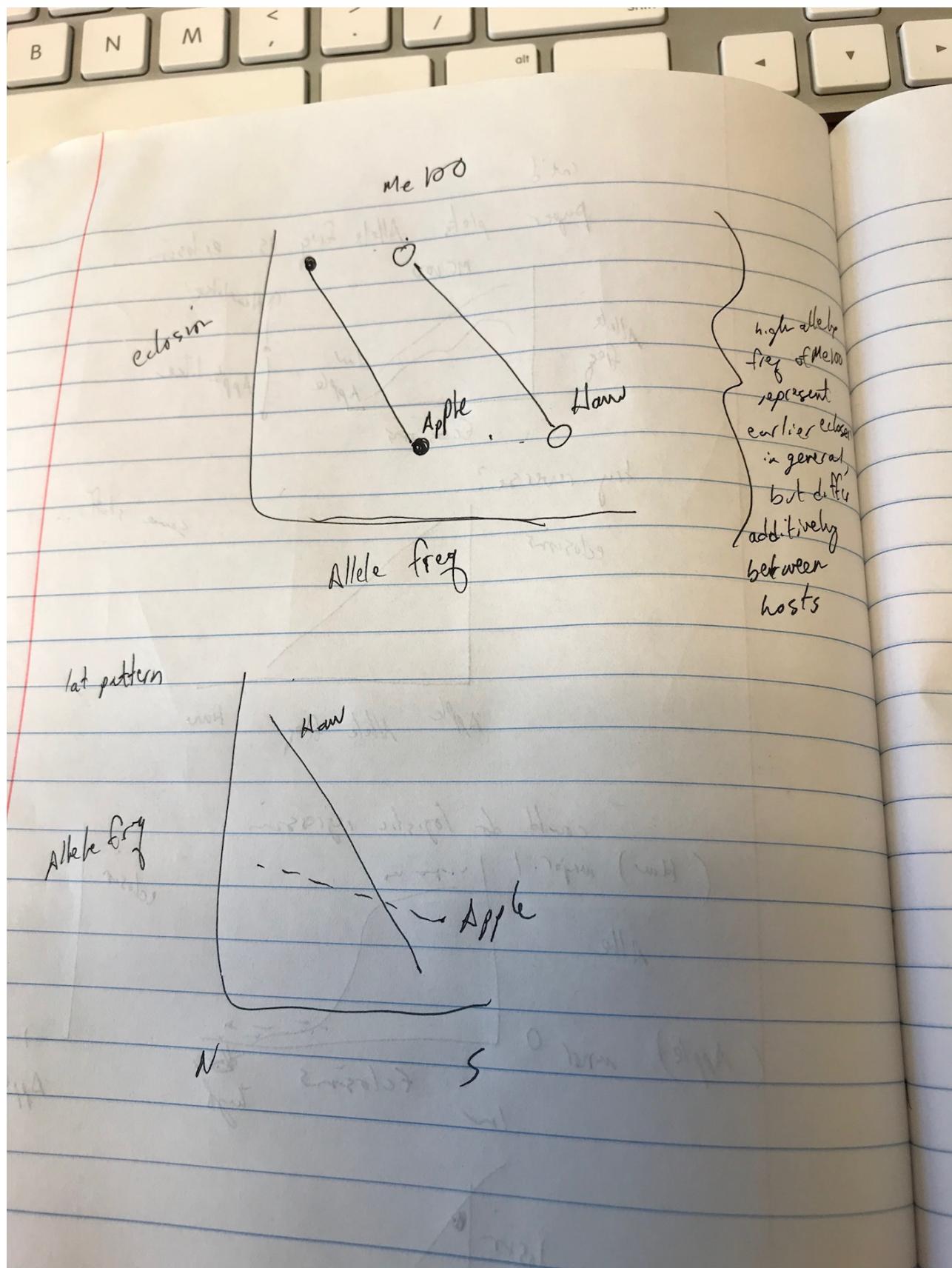
some stuff...

could do logistic regression

(How) major / minor



Different way of thinking about "haw" alleles. Higher freq in me100 for example is overall higher in earlier eclosers, but there is a main effect of host. It is important to remember that hosts are different environments! So differences are not purely due to genetics. There really needs to be a reciprocal transplant of maggots on host fruit to determine if the performance under different genetic backgrounds is real.



Page 130: 2017-11-19 Initial reading of Philosophical Transactions B review articles on wild clocks

(<http://rstb.royalsocietypublishing.org/content/372/1734>)

Preface: Wild Clocks- preface and glossary

Chronobiologist; programmes....annoying term.

Outlines, compares, contrasts chronobiologist(lab) and ecologist; and approach to studying time---basically lab vs ecological setting.

They want to unite these two approaches and thinking.

Summary of papers in the whole issue

1. Describe diff timing phenotypes as chronotypes and how they can be phenotypically plastic ; a new area for integration between fields
2. next two articles talk about new tools that advance both feilds.
3. NExt 3 describe both chronobio and eco approaches in seasonal biology
4. 3 final ones consider interspecific interactions in setting biological timing (plant-pollinator interactions)

First paper in issue- Two sides of a coin: eco and chrono perspectives of timing in the wild

ref- Review article: Two sides of a coin: ecological and chronobiological perspectives of timing in the wild Barbara Helm, Marcel E. Visser, William Schwartz, Noga Kronfeld-Schor, Menno Gerkema, Theunis Piersma, Guy Bloch Phil. Trans. R. Soc. B 2017 372 20160246; DOI: 10.1098/rstb.2016.0246. Published 9 October 2017

Key questions:

Chronobiologist

- How do organisms time biological processes?

Ecologist

- Why do they do it?

Timing can be both abiotic (light intensity) and biotic (veg cover). Organisms can respond to whatever timing cue by altering behaviors (ie. migration). So the combination of external and internal timing influence the ultimate expression of when organisms do things.

Focus on chronobiologists:

It identifies effects of components of time that act to modify internal clock time through entrainment of the clock (synchronization of internal clock time by environmental cues or 'zeitgebers' [25])

What the...

Zeitgeber is the signal/cue that organisms synchronize their internal clock with? Don't know why chronobiologists would care about this per se. Its an abstraction that means very little on its own. I really have no clue what this thing is and why it matters.

I'm assuming that chronobiologists just care about the internal clock? And how organisms can respond to different cues and match their clocks accordingly.

Thus, chronobiologists increasingly conclude that pure consideration of internal clocks and photic entrainment, although perfected in laboratory settings, has insufficient explanatory power in the real world.

No citation with this claim...needs a meta analysis to support.

Paragraph focusing on ecological perspective doesn't link back to "Why do they do it" question. And the examples are focused on interspecific interactions. What they needed was to link back was to relate timing with fitness.

It seems like the big discrepancy between small scale and large scale temporal biologist is asking questions in terms of proximate or ultimate mechanisms.

Importantly, therefore, an organism can be thought of as consisting of many clocks (millions in complex multi-cellular organisms; [45]) that are co-ordinated within the body in various ways (the 'second revolution' in chronobiology; [40]).

What....the emphasis should be on traits/characteristics which occur at certain times...not on "timing" in general.

Introduces chronotype without actually explaining what it is....

2017-11-20 cont'd

Focus on Ecology:

history- ecology minded chronobiologists ; Pat DeCoursey, Eberhard Gwinner, Serge Daan

thoughts: ecology inherently considers timing in its very definition-- study of how organisms interact with their environment over space and time

temporal niche- temporal segregation of resource use among potentially competing species or individuals sharing the same habitat

a good example would be the timing of bees and flowers whereby each party shifts their foraging or flowering time, respectively.

The things that shape timing

Macro-ecologists have recently identified large-scale patterns in daily timing of mammals across the globe [68,71]. The proportion of nocturnal species is highest in arid regions and lowest at extremely high latitudes, while crepuscularity (activity during dawn and dusk) is correlated with longer twilight durations.

Cathemerality (activity that is spread across day and night) is also more common in cold habitats and under long hours of daylight and twilight in the northern Holarctic region [68,72]

- Nocturnal species highest in arid regions and lowest at the poles
- dawn/dusk activity correlated with longer twilight durations
- activity over night and day is more common in cold habitats and under long daylight hours

Converging key concepts of both fields

One problem they mention: plasticity

Many studies focus on the rigidity of rhythms, but there is a lot of plasticity. (I have no clue in what way they mean, they should define plasticity explicitly.). This could be a source for the differences between the field and lab. (no duh, they are diff environments....). There is "less" plasticity in the lab....but during an entrainment or free-run experiment,

animals may not respond consistently if you're analyzing at the unit of the day.

I think they mean plasticity as the ability to change timing in response to different environments.

but, then this:

To some authors, a biological rhythm in a behaviour or physiological process in itself is seen as 'plasticity' ('endogenous' plasticity [102]; table 1a).

These authors would benefit from a more quantitative way of thinking about plasticity. Yes, behavior vs time can be imagined as plasticity, for a genotype!!!!, but how much this "matters" in an evolutionary framework is comparing among plasticity among genotypes. And plasticity can be imagined as the slope of the reaction norm

And the link between the lab and field biologist could be this: Go out in the field and identify critical factors shaping variation in the timing of a trait; then go into the lab and dissect out that timing trait, even the molecular basis of it, by manipulating the critical factors identified from ecology. If we know the key genetic constituents and the important environmental factors, we can predict(with a predictive model) the timing of that trait out in nature....functional outcomes: inform species responses to climate change; predict matchiness between biocontrol agent and pest species; They mention this later on in the article!

Introducing more jargon: malleable temporal programme; seriously wtf...

Binary distinction between plasticity through the clock (entrainment) versus plasticity outside the clock (masking) falls short of capturing the ability of animals to adjust their rhythms to the environment

What...I didnt even know entrainment and masking was related to plasticity.....

Such plasticity is enhanced by its mechanistic complexity, whose many parts and multiple oscillators offer numerous ways of adjusting internal clock time, and thereby making the system less rigid. Therefore, complexity could be the key to addressing the apparent paradox of 'How can a biological system be rigid and conserved, but at the same time plastic?' Resolving this paradox requires us to understand how the system works—what are the gears, how do they work together and how do they respond to ecologically relevant factors, such that the timing system as a whole generates plasticity.

Reading this: Rattenborg, N. C., de la Iglesia, H. O., Kempenaers, B., Lesku, J. A., Meerlo, P., & Scriba, M. F. (2017). Sleep research goes wild: new methods and approaches to investigate the ecology, evolution and functions of sleep. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1734), 20160251.

<https://doi.org/10.1098/rstb.2016.0251>

The temporal programme of behaviours and physiology expressed by an organism is driven by a vast network of clocks and rhythms distributed across tissues throughout the body.

somebody tell me what "the temporal programme" is?

Organisms comprise of characteristics whose functional consequences are largely determined by the timing in their expression. This timing usually varies among different characteristics themselves.....

Hypothalamus in mammals is the place where decisions about timing are made, specifically, the suprachiasmatic nucleus(SCN). It allows for organisms to take in light cues and entrain to it. It also transmits this information to other parts of the body. All non-SCN is known as the 'peripheral timing system'.

Why not just describe the process of how organisms receive input from external signals and respond to them....? That's basically figure 1. Because many external forces are cyclical in nature, organisms can respond cyclically as well in order to maximize survival and reproduction...

All the things that can influence the cyclical behavior of animals (Zeitgebers):

- light
- predators
- food availability

cool walk through of how SCN ablated individuals(assuming mice) are arrhythmic under controlled conditions; under natural conditions(light, social interactions, social cues, ambient temp, fear), SCN ablated individuals regain rhythmicity.

this suggests that there is timing independent of SCN. So there should be a role for peripheral tissues in contributing to the timing of behaviors/phys

Interesting point: blood glucose can be cyclical even if it is constant in the blood, but the cells in different tissues cycle their cell surface receptors..

Glucocorticoids can modulate input from Zeitgebers for peripheral tissues.

Page 131: 2017-11-21. Figured out code to subset out a dataframe based on a time interval

This'll be useful for parsing out individual behavioral count data (from trikinetics rigs) based on monitor, position in monitor, then date+time.

```
1 #specify time interval:  
2  
3 t<-interval("2017-11-01 00:01:00","2017-11-13  
23:00:00", tzonе = tz(start))  
4 t  
5  
6 ##  
7 read in monitor data and change month into a number  
8 x<-read.table("Monitor6.txt")  
9 x$V3<-match(x$V3,month.abb)  
10
```

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19										
V20	1	46586	12	11	17	20:41:00	1	0	0	0	0	0	0	0	0	0	0	2	46587	12	11							
17	20:42:09	1	0	0	0	0	0	0	0	0	0	0	0	3	46588	12	11	17	20:44:50	1	0	0						
0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	46589	12	11	17	20:47:18	1	0	0	0	0	0	0	0	0
0	0	5	46590	12	11	17	20:52:13	1	0	0	0	0	0	0	0	0	0	0	6	46591	12	11						
17	20:57:19	1	0	0	0	0	0	0	0	0	0	0	0	V21	V22	V23	V24	V25	V26	V27								
V28	V29	V30	V31	V32	V33	V34	V35	V36	V37	V38	V39	V40	1	0	0	0	0	0	0									
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0		
0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0		
V41	V42	V43	V44	1	0	0	0	2	0	0	0	3	0	0	0	4	0	0	0	5	0	0	0	6	0	0	0	

```
1 #convert to time vector
2 test<-data.frame(time=strptime(time,"%Y-%m-%d
%H:%M:%S",tz=tz(start)))
3
4 #subset
5 head(subset(test,time %within% t))
```

Page 132: 2017-11-21. Field clocks reviews - methods in field chronobiology

ref: Methods in field chronobiology Davide M. Dominoni, Susanne Åkesson, Raymond Klaassen, Kamiel Spoelstra, Martin Bulla Phil. Trans. R. Soc. B 2017 372 20160247; DOI: 10.1098/rstb.2016.0247. Published 9 October 2017

Critical questions:

1. How variable are rhythms within and between, taxa, species, populations, individuals?
2. What drives such variation? (env , internal state, sociality, anthropogenic disturbance, climate change)
3. Are rhythms of captive animals comparable to rhythms of animals in the wild?
4. How can we disentangle the relative contribution fo genes vs environmental drivers of rhythmicity in wild species?
5. Is variation in rhythms associated to fitness?

case study: incubation rhythms of biparental care in shorebirds. They incubate at different times, this varies among species and individuals. There is high phylo signal and the lenght of encubation is related to latitude, and some proxy for predation risk.

a diff approach:

insect migration- lidar tech; lazer beam detection

Another case study targeting whether we can relate circ to fitness:
Measuring fitness consequences of circ organization in the wild. They created lesions in SCN in antelope squirrels and chipmunks and tracked their rhythms. Related to survival but not fitness?

Release of CK1etau tau allele reduces endogenous circ period; so they released control mice vs tau mutants and tau mutants reduced in frequency.

needs statistical methods

Page 133: 2017-11-24. Prep for meeting with Dan, 2017-11-28 + thoughts on diapause + to do list

Main stuff to talk about:

1. Reading 3 papers on odors in Rhago system

- Linn, C. E., Dambroski, H. R., Feder, J. L., Berlocher, S. H., Nojima, S., & Roelofs, W. L. (2004). Postzygotic isolating factor in sympatric speciation in *Rhagoletis* flies: Reduced response of hybrids to parental host-fruit odors. Retrieved from <http://www.pnas.org/content/101/51/17753.full.pdf>
- Linn, C., Feder, J. L., Nojima, S., Dambroski, H. R., Berlocher, S. H., & Roelofs, W. (n.d.). Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. Retrieved from <http://www.pnas.org/content/100/20/11490.full.pdf>
- Dambroski, H. R., Linn, C., Berlocher, S. H., Forbes, A. A., Roelofs, W., & Feder, J. L. (n.d.). THE GENETIC BASIS FOR FRUIT ODOR DISCRIMINATION IN RHAGOLETIS FLIES AND ITS SIGNIFICANCE FOR SYMPATRIC HOST SHIFTS. Source: Evolution, 59(9), 1953-1964. <https://doi.org/10.1554/05-133.1>

Stuff on my mind that would be nice to discuss

1. Diapause phases:

Ragland 2009 JIP paper on biphasic MR. First logarithmic increase is termination? Then, the following exponential increase is post diapause?

ref:

Biphasic metabolic rate trajectory of pupal diapause termination and post-diapause development in a tephritid fly. (2009). Journal of Insect Physiology, 55(4), 344–350.

<https://doi.org/10.1016/J.JINSPHYS.2008.12.013>

Has Dan Read the Kostal et al. 2017, PNAS paper?

refs:

Koštál, V., St Etina, T., Poupartin, R., Korbelová, J., & Bruce, A. W. (n.d.). Conceptual framework of the eco-physiological phases of insect diapause development justified by transcriptomic profiling.

<https://doi.org/10.1073/pnas.1707281114>

Koštál, V. (2006). Eco-physiological phases of insect diapause. Journal of Insect Physiology, 52(2), 113–127.

<https://doi.org/10.1016/j.jinsphys.2005.09.008>

Argues for distinct phases of diapause, but not very clear that they are separated --. They should have used a discriminant analysis, which would show better separation, suggesting that certain aspects of the transcriptome are associated with transitions in diapause among phases.

2. "Early eclosing alleles"
3. Data analysis
 - Show some figures
 - for mass specific metabolic rate, some values are negative

to do:

1. check undergrad schedules
2. get back to writing manuscripts

3. generate spreadsheet for samples that need to be transferred to free run
 - update dataset
4. set up animals in trikinetics
5. write up methods so it is less painful later on

2017-11-28 meeting with Dan

Fellowship opportunity with [Mozilla](#)

Supplement request REU program, 6K ; asking

-10 hour/week, spring, 30hr week summer -1 paragraph highlighting my experience with REU(commitment to mentoring) and Tat's traits and she'll benefit from this experience-excellent career biologist

Kylie for 3 credit , 10 hour /week

- have to write a report at the end; showing sufficient understanding about project and outcome ; 4-6 pages ; 10 references

Dan looking at data,

- What predicts non-diapause/diapause , using eclosions and metabolic rate, and weights ; discriminant analysis ; classificaiton tree
-

Going to biological rhythms [conference](#):

-Look into sable training for respirometry

Page 134: 2017-11-28. Reading paper for lab meeting Overgaard and MacMillan 2017; Ann Rev Phys

ref: Overgaard, J., & Macmillan, H. A. (2017). The Integrative Physiology of Insect Chill Tolerance. *Annu. Rev. Physiol.*, 79, 187–208.

<https://doi.org/10.1146/annurev-physiol-022516-034142>

Key negative impact of cold temperature is disruption of behaviors/motor function because cells are unable to maintain ion flux sufficient enough to fire action potentials.

Cold tolerance defined by freeze-tolerant (ice in the extracellular fluid) or avoidance (lowering freezing point of extracellular fluid). What....the heck?

SCP = super cooling point

Chill susceptible - dying to cold above freezing points or sensitive to low temperatures

Chill tolerant - withstanding low temperatures to some degree but still die at temperatures above the SCP

Measurements of cold susceptibility

CTmin - temperature at which motor function is lost

CCO - chill coma onset temperture ; temperature that induces complete paralysis

CCRT - chill coma recovery time ; time required to recover the ability to stand following cold exposure

LT 50 - temperature at which 50% of individuals survive a given exposure duration

Lt50 ; time at a given temp that causes 50% mortality

Survival

Ion homeostasis

Main cause of cold damage: loss of motor function due to disruption of membrane function. One example is loss of membrane potential. Resting is usually -70mV but the onset of cold coma is when the membrane potential is -40 mV . This is called membrane depolarization, which can be simulated in locusts by adding extracellular K⁺ and reduces muscle force production in vitro.

1 | * This is cool, because it'd be interesting to determine what other threats to cold damage there are by comparing to a simulated K⁺ in the extracellular space

Cold damage can cause more extracellular K⁺ surrounding neurons, impairing the nervous system. K⁺ might be adaptive because it shuts down activity in the face of cold.

Disruption of ion and water balance linked to chill injury

Two reasons:

1. cooling causes phase change in cell membranes and they lose their selective permeability (refs 103; 104; look up)
2. ion balance is lost more gradually during cold exposures because

thermodynamics of passive diffusion are only slightly affected by temperature

Page 135: 2017-11-29. Notes on how to analyze biological rhythm data

Behaviors are cyclical and can be captured as a cosine function. Behaviors are time series essentially because measurements depend on the preceding measurement. A periodogram is used to ID the dominant frequencies /periods/cycles(properties of a cosine function) in a time series.

Period (T) is the number of time steps it takes to complete a single cycle of cosine function

Frequency is omega = $1/T$; fraction of the cycle that's completed in a single time period

converting freq to period:

The dominant peak area occurs somewhere around a frequency of 0.05. Investigation of the periodogram values indicates that the peak occurs at nearly exactly this frequency. This corresponds to a period of about $1/0.05 = 20$ time periods. That's 10 years, since this is semi-annual data. Thus there appears to be a dominant periodicity of about 10 years in sunspot activity.

so confused about the units of [frequency](#)

The frequency is measured in cycles per unit time where unit time is defined to be the distance between adjacent points. A frequency of 0 corresponds to an infinite cycle while a frequency of 0.5 corresponds to a cycle of 2 data points. Equi-spaced time series are inherently limited to detecting frequencies between 0 and 0.5.

Discrete Fourier Transform, (synonymous with periodogram)

one of the major goals is to estimate the dominant frequencies that occur in a stationary time series

a good [resource](#):

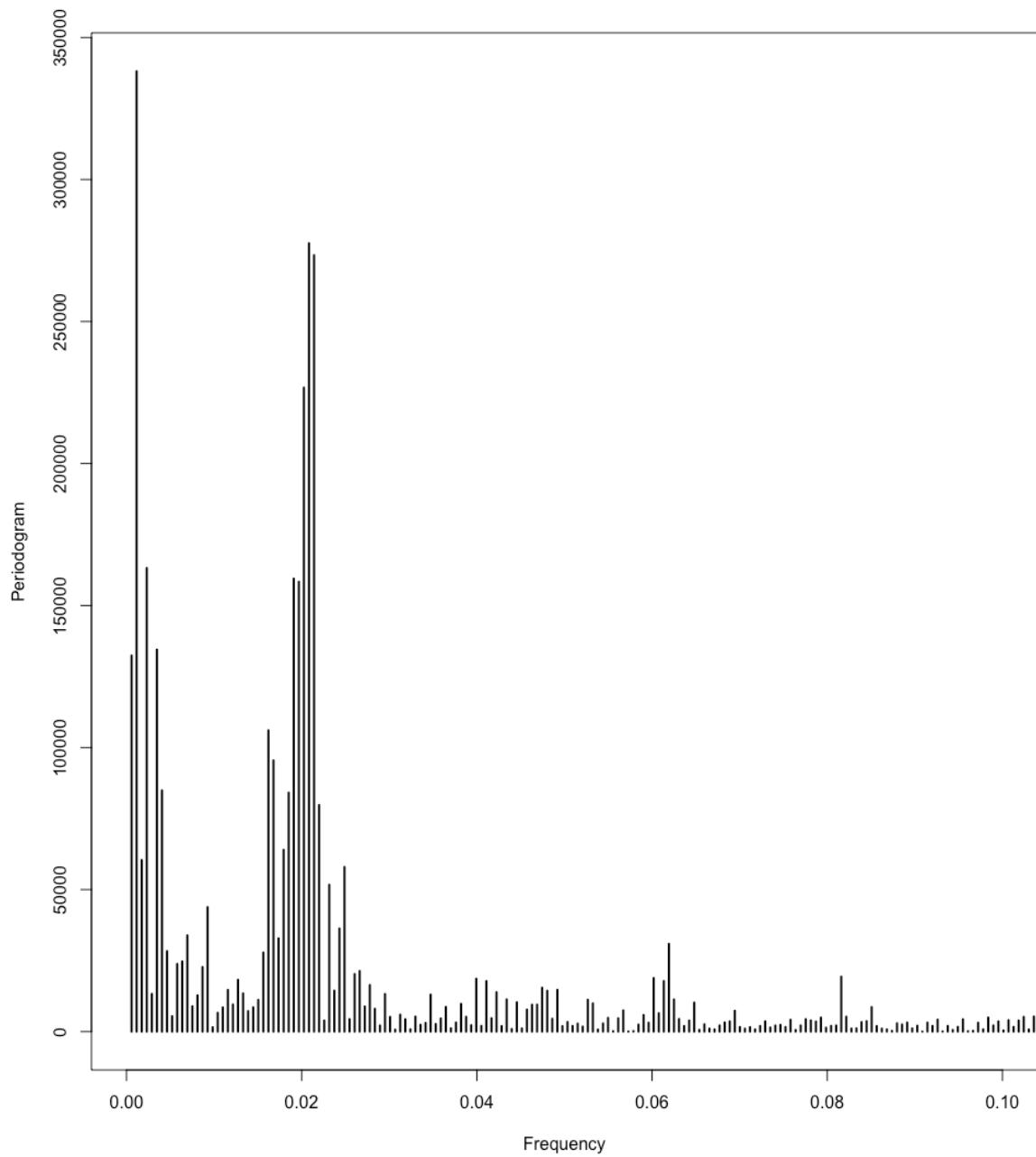
Interpretation and Use: A relatively large value of $P(j/n)$ indicates relatively more importance for the frequency j/n (or near j/n) in explaining the oscillation in the observed series. $P(j/n)$ is proportional to the squared correlation between the observed series and a cosine wave with frequency j/n . The dominant frequencies might be used to fit cosine (or sine) waves to the data, or might be used simply to describe the important periodicities in the series.

Fourier transform explained <https://www.r-bloggers.com/the-fourier-transform-explained-in-one-sentence/>

Trying different time intervals: 6 min (.1hr), 15 min(.25 hr), 30 min (0.5 hr) on h4o4

some notes: h4o4, 35 days, first 8 days are entrainment (16L:8D), remaining days are free-run (dark)

30 min

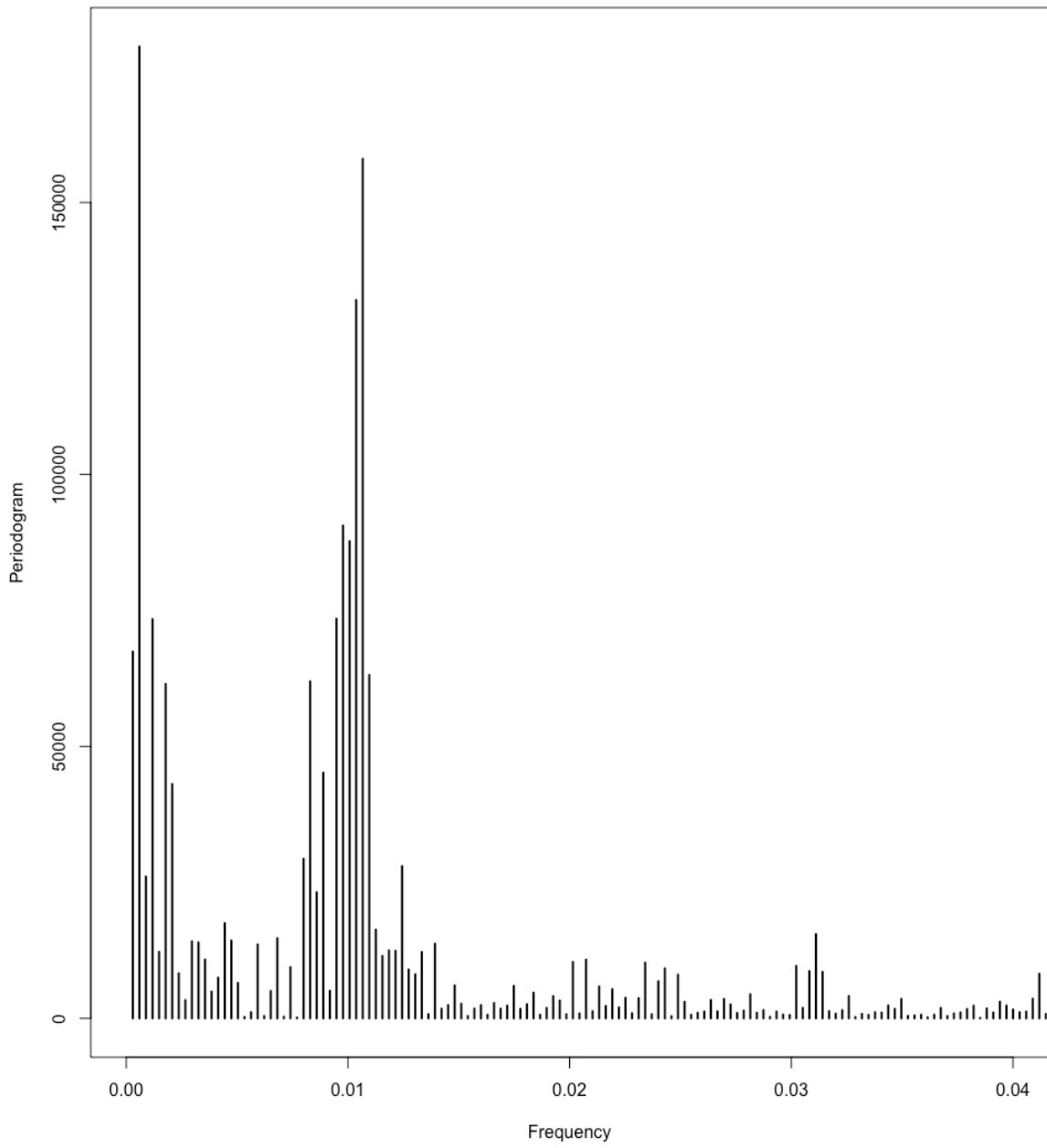


```
1 library(TSA)
2
3 #30 min
4 p0<-periodogram(counts30$counts30)
5 p0
6
7 dd0<-data.frame(freq=p0$freq,p0$spec)
8 order0<-dd0[order(-dd0$p0.spec),]
9 top0<-head(order0,10);top0
10 1/top0$freq/2/24 # convert frequency into period and in
units in days
```

Vector of days of most important periods

```
[1] 18.000000 1.000000 0.972973 1.028571 9.000000 1.090909
1.058824 6.000000 36.000000 1.285714
```

15 min



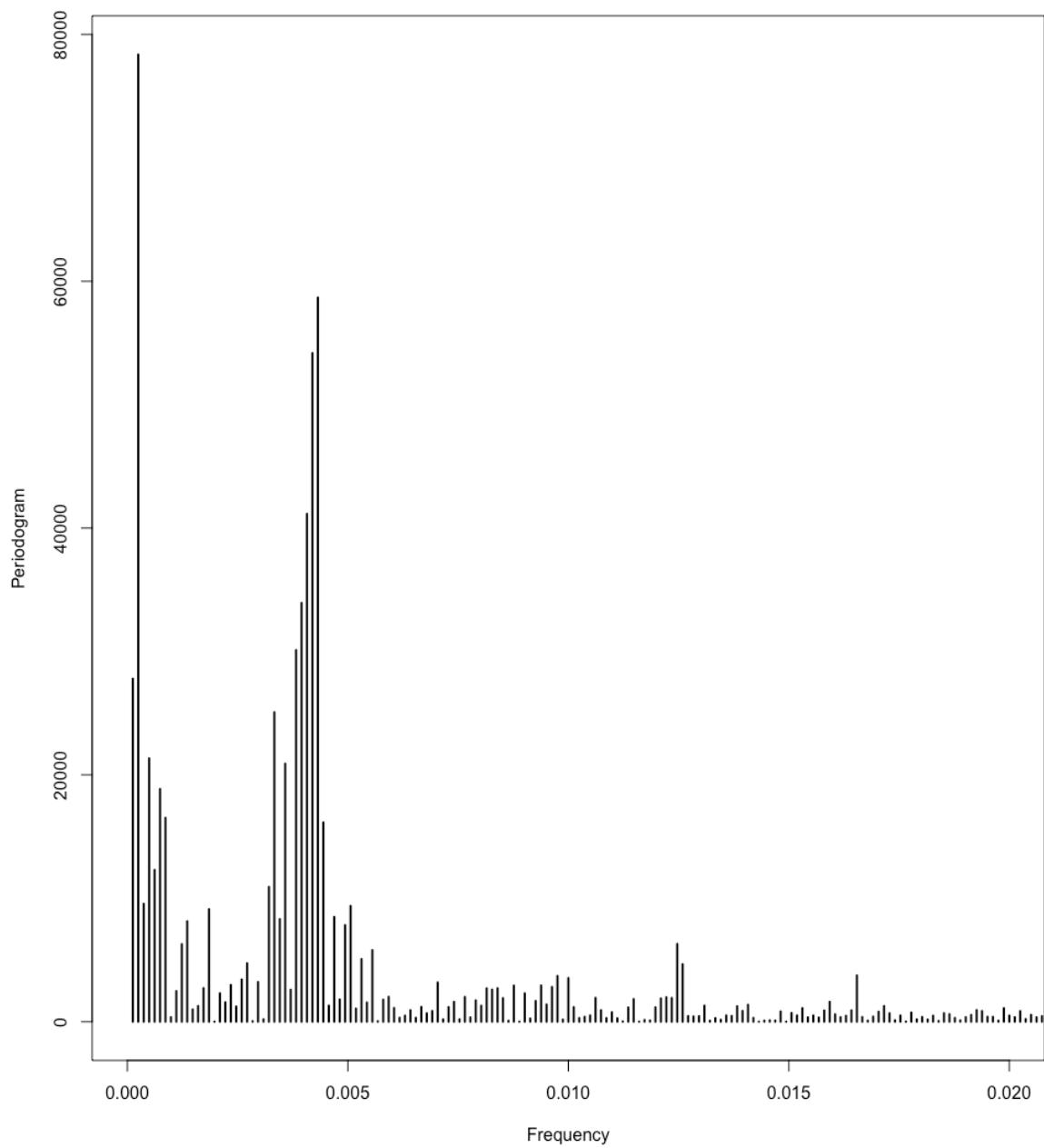
```
1 # 15 min
2 p<-periodogram(counts15$counts15)
3 p
4
5 dd<-data.frame(freq=p$freq,p$spec)
6 order<-dd[order(-dd$p.spec),]
7 top2<-head(order,10);top2
8
```

```
9  1 / top2$freq  
10  
11 (1 / top2$freq) / 4 # to get hours  
12 (1 / top2$freq) / 4 / 24 # days  
13  
14
```

Vector of days of most important periods

```
[1] 17.5781250 0.9765625 1.0044643 1.0653409 1.0340074 1.0986328  
8.7890625 35.1562500 0.9501689 1.2555804
```

6 min



```

1
2 ### 6 min
3 p2<-
4 periodogram(counts06$counts06,plot=TRUE,xlim=c(0,.02))
5 dd2<-data.frame(freq=p2$freq,p2$spec)
6 order2<-dd2[order(-dd2$p2.spec),]
7 top10<-head(order2,10);top2
8 1/top10$freq
9 (1/top10$freq)/10 # to get hours
10 (1/top10$freq)/10/24 # days
11

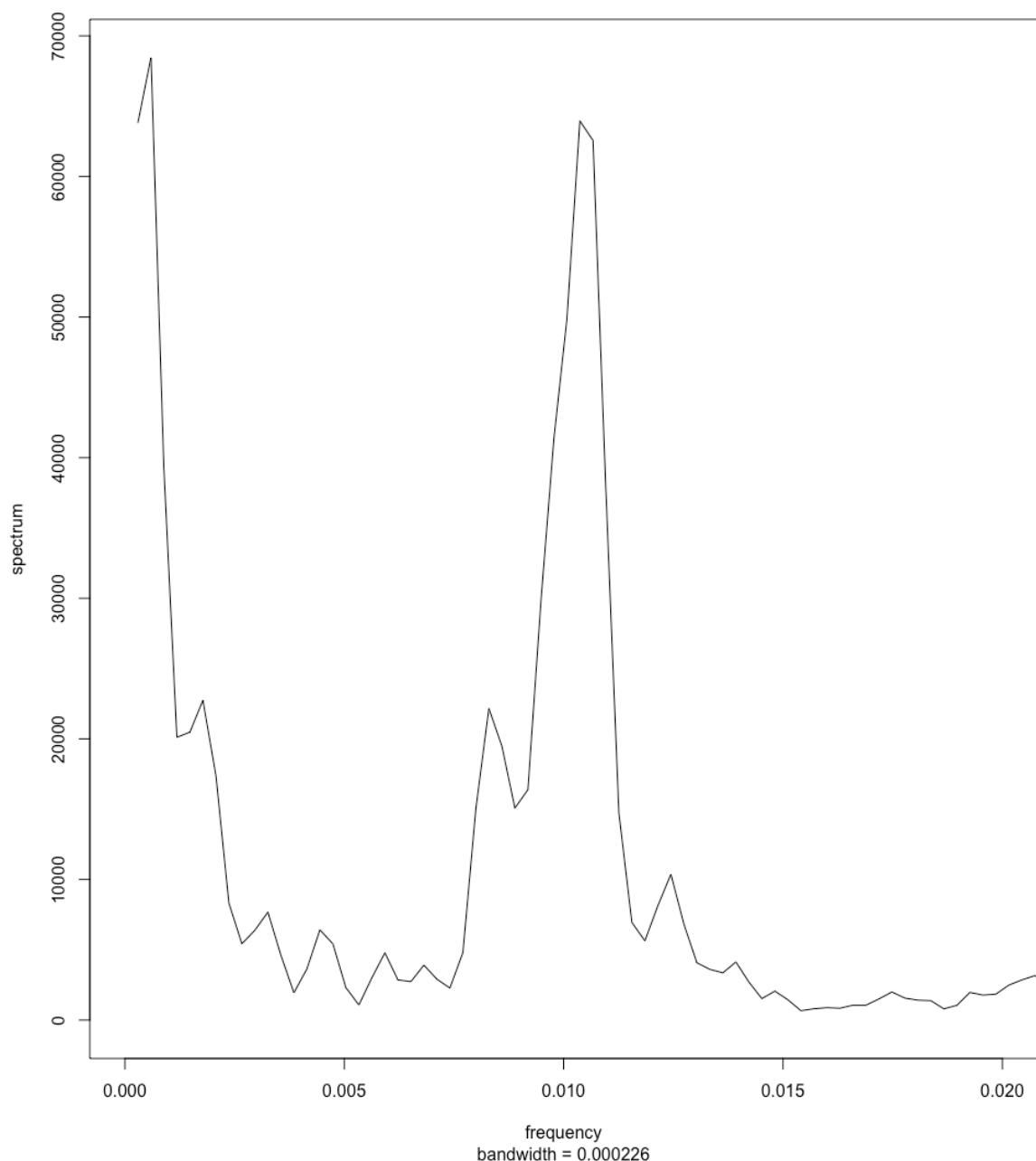
```

Vector of days of most important periods

```
[1] 16.8750000 0.9642857 0.9926471 1.0227273 1.0546875 1.0887097
33.7500000 1.2500000 8.4375000 1.1637931
```

Diff way to do spectral analysis/periodogram; 15 min bins

Series: counts15\$counts15
Smoothed Periodogram



```

1 ##15 min bins
2 raw.spec1<-
  spec.pgram(counts15$counts15,taper=0,kernel=c(1,1,1),spa
  ns=3,detrend=FALSE)
3 plot(raw.spec1,log="no",xlim=c(0,.02))
4 1/.01*4/24
5
6 rs2<-data.frame(f=raw.spec1$freq,s=raw.spec1$spec)
7 1/head(rs2[order(rs2$s,decreasing=TRUE),],10)
8
9 1/head(rs2[order(rs2$s,decreasing=TRUE),],10)[,1]/4/24

```

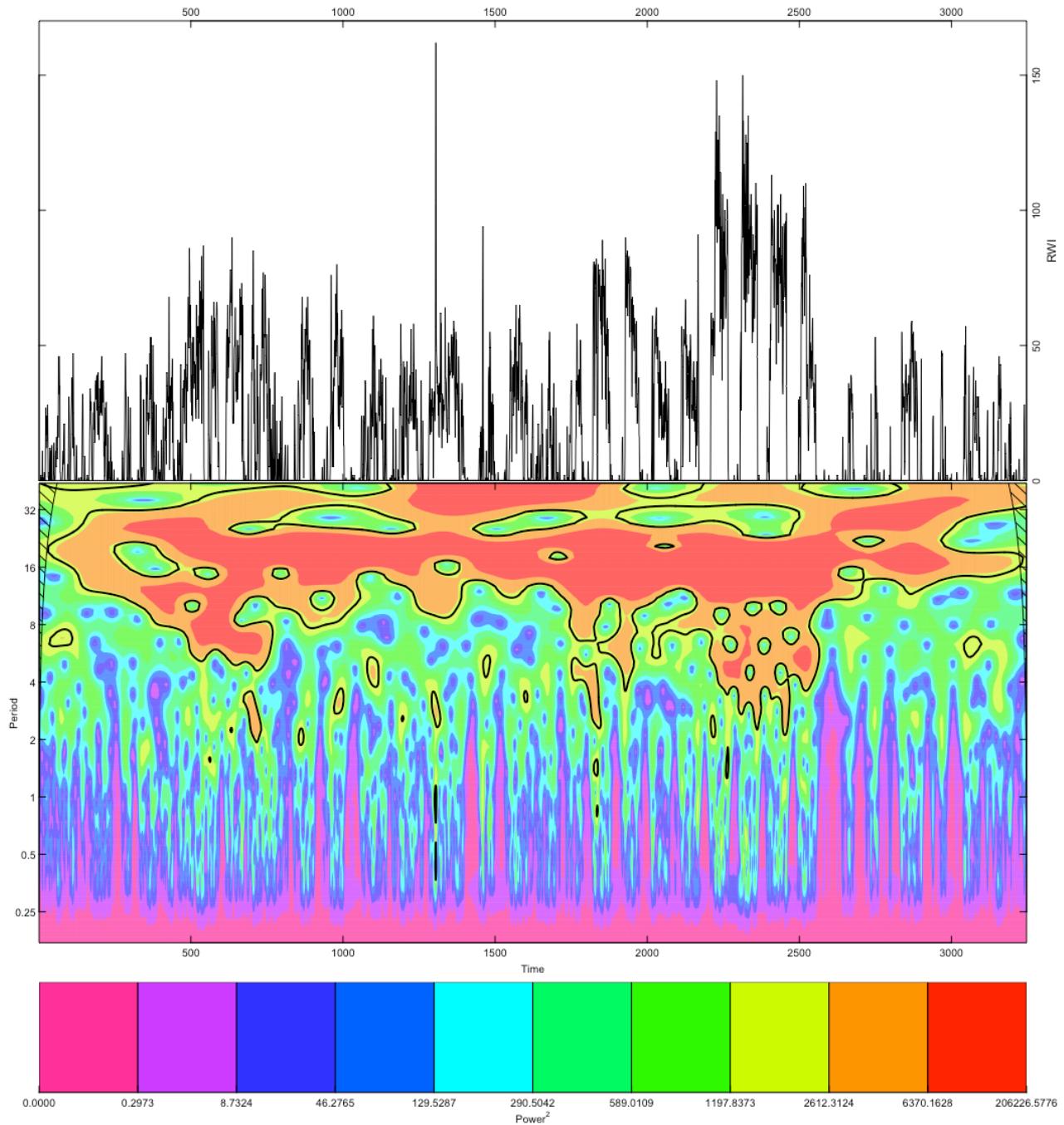
Vector of days of most important periods

```
[1] 17.5781250 0.9765625 1.0044643 1.0653409 1.0340074 1.0986328
8.7890625 35.1562500 0.9501689 1.2555804
```

Biological rhythms may not be stationary (non-stationary) and it is good to check this assumption with wavelet analyses. Wavelets transform the data into a set of small wavelets that are bounded which can be shrunk or expanded to fit the data.

Doing a continuous wavelet analysis on h4o4

We'll try to find the dominant period signal in the data (although we can get the instantaneous peak along each time point and the amplitude can vary).



```

1 library(dplR)
2 ##15 min
3 wave.out <- morlet(y1 = counts15$counts15, p2 = 8, dj =
4   0.1, siglvl = 0.95)
5
6 levs <- quantile(wave.out$Power, c(0, 0.25, 0.5, 0.75,
7   0.95, 1))

```

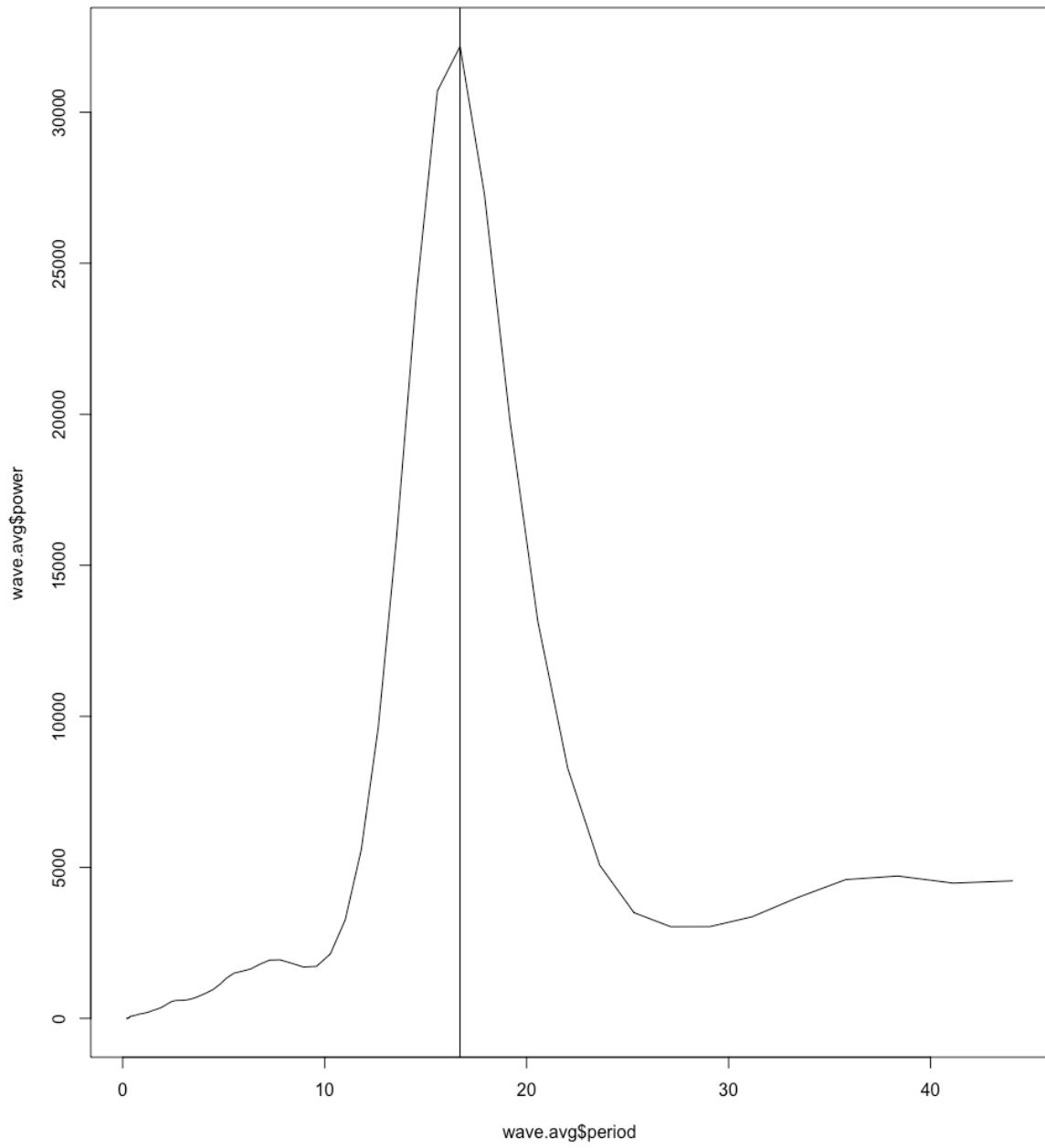
```
7 #wavelet.plot(wave.out, wavelet.levels = levs)
8 wavelet.plot(wave.out)
9
10
11 wave.avg <- data.frame(power = apply(wave.out$Power, 2,
12 mean), period = (wave.out$period))
13
14 findpeaks(wave.avg$power)
15 wave.avg[findpeaks(wave.avg$power)[2,2],] #period in
16 days
17 plot(wave.avg$period, wave.avg$power, type = "l")
18 abline(v=wave.avg[findpeaks(wave.avg$power)[2,2],])
```

The period is very similar to previous estimates with periodograms

wave.avg[findpeaks(wave.avg\$power)[2,2],] #period in days

power period

32174.79 16.70188



**2017-11-30 update: yeah,
mistake with the data itself,
needed to convert date into**

a numeric and order the dates

test code to show how to order dates:

```
1 > tt<-unique(counts15.1$date)
2 > tt
3 [1] 2017-10-16 2017-10-17 2017-10-18 2017-10-19 2017-
10-20 2017-10-21 2017-10-22 2017-10-23 2017-10-24 2017-
10-25
4 [11] 2017-10-26 2017-10-27 2017-10-28 2017-10-29 2017-
10-30 2017-10-31 2017-11-1 2017-11-10 2017-11-11 2017-
11-12
5 [21] 2017-11-13 2017-11-14 2017-11-15 2017-11-16 2017-
11-17 2017-11-18 2017-11-19 2017-11-2 2017-11-3 2017-
11-4
6 [31] 2017-11-5 2017-11-6 2017-11-7 2017-11-8 2017-
11-9
7 35 Levels: 2017-10-16 2017-10-17 2017-10-18 2017-10-19
2017-10-20 2017-10-21 2017-10-22 2017-10-23 ... 2017-
11-9
8 > tt<-as.Date(unique(counts15.1$date))
9 > tt
10 [1] "2017-10-16" "2017-10-17" "2017-10-18" "2017-10-
19" "2017-10-20" "2017-10-21" "2017-10-22" "2017-10-23"
"2017-10-24"
11 [10] "2017-10-25" "2017-10-26" "2017-10-27" "2017-10-
28" "2017-10-29" "2017-10-30" "2017-10-31" "2017-11-01"
"2017-11-10"
12 [19] "2017-11-11" "2017-11-12" "2017-11-13" "2017-11-
14" "2017-11-15" "2017-11-16" "2017-11-17" "2017-11-18"
"2017-11-19"
```

```

13 [28] "2017-11-02" "2017-11-03" "2017-11-04" "2017-11-
      05" "2017-11-06" "2017-11-07" "2017-11-08" "2017-11-09"
14 > tt[order(as.Date(tt,format = "%Y%m%d"))]
15 [1] "2017-10-16" "2017-10-17" "2017-10-18" "2017-10-
      19" "2017-10-20" "2017-10-21" "2017-10-22" "2017-10-23"
      "2017-10-24"
16 [10] "2017-10-25" "2017-10-26" "2017-10-27" "2017-10-
      28" "2017-10-29" "2017-10-30" "2017-10-31" "2017-11-01"
      "2017-11-02"
17 [19] "2017-11-03" "2017-11-04" "2017-11-05" "2017-11-
      06" "2017-11-07" "2017-11-08" "2017-11-09" "2017-11-10"
      "2017-11-11"
18 [28] "2017-11-12" "2017-11-13" "2017-11-14" "2017-11-
      15" "2017-11-16" "2017-11-17" "2017-11-18" "2017-11-19"

```

results dont change much for the spectral analysis, but it does for the wavelet analyses

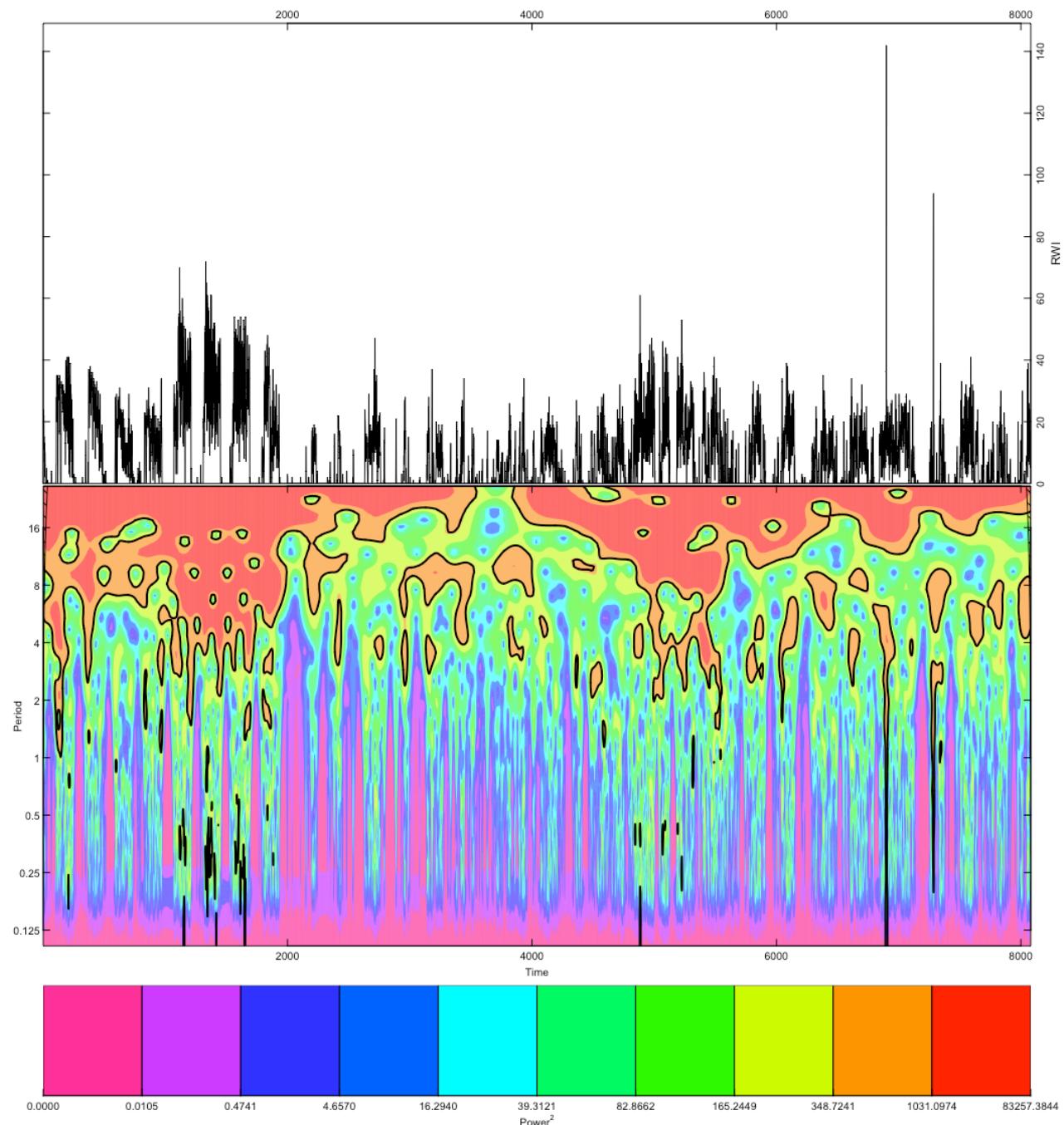
Page 136: 2017-11-30. Continuous Wavelet analyses for different time bins: 6 min, 15 min, 30 min, 1 hr

Take home/thoughts:

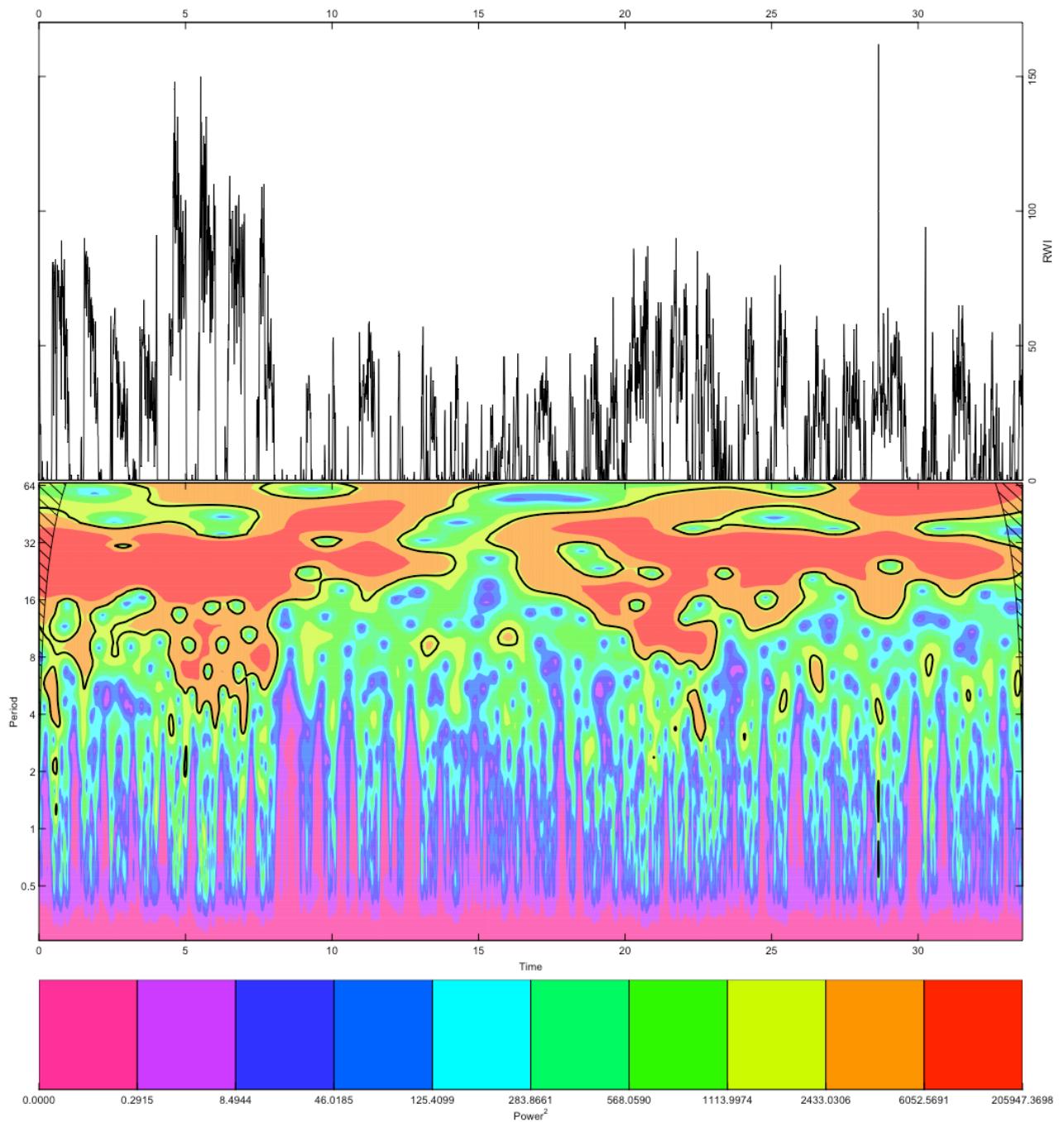
- smaller time scales lets you detect daily fluctuations better and larger time scales lets you detect long term fluctuations better.
- All time bins can detect 24 hour periods
- Prob want to play with 1 hour time bin to detect period changes on the order of days

- 6 min bins good for detecting period within days
- See below to basically take the dominant period in the transformation
- **Remember power is equivalent to the amplitude squared!!!**

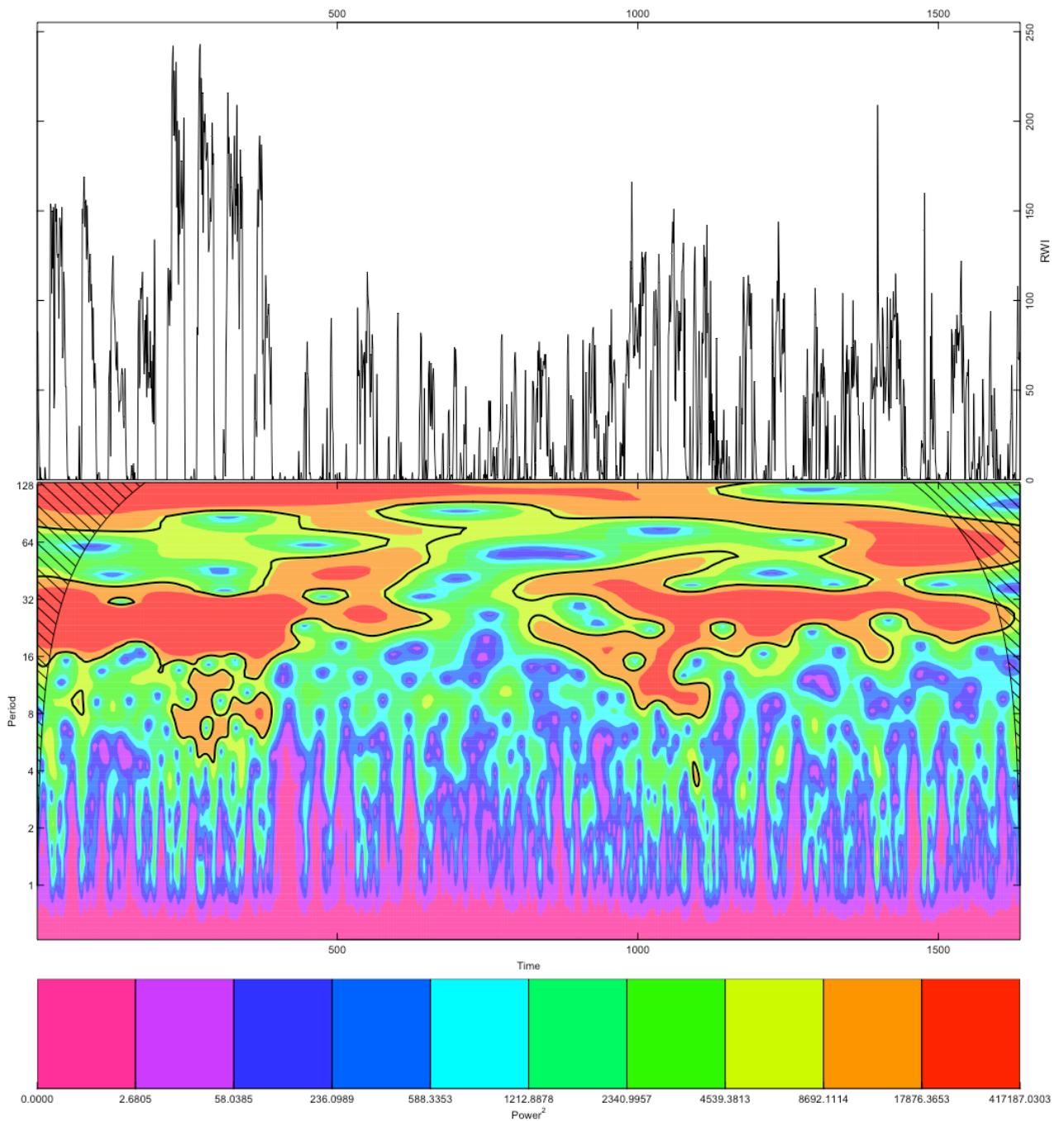
6 min time bins



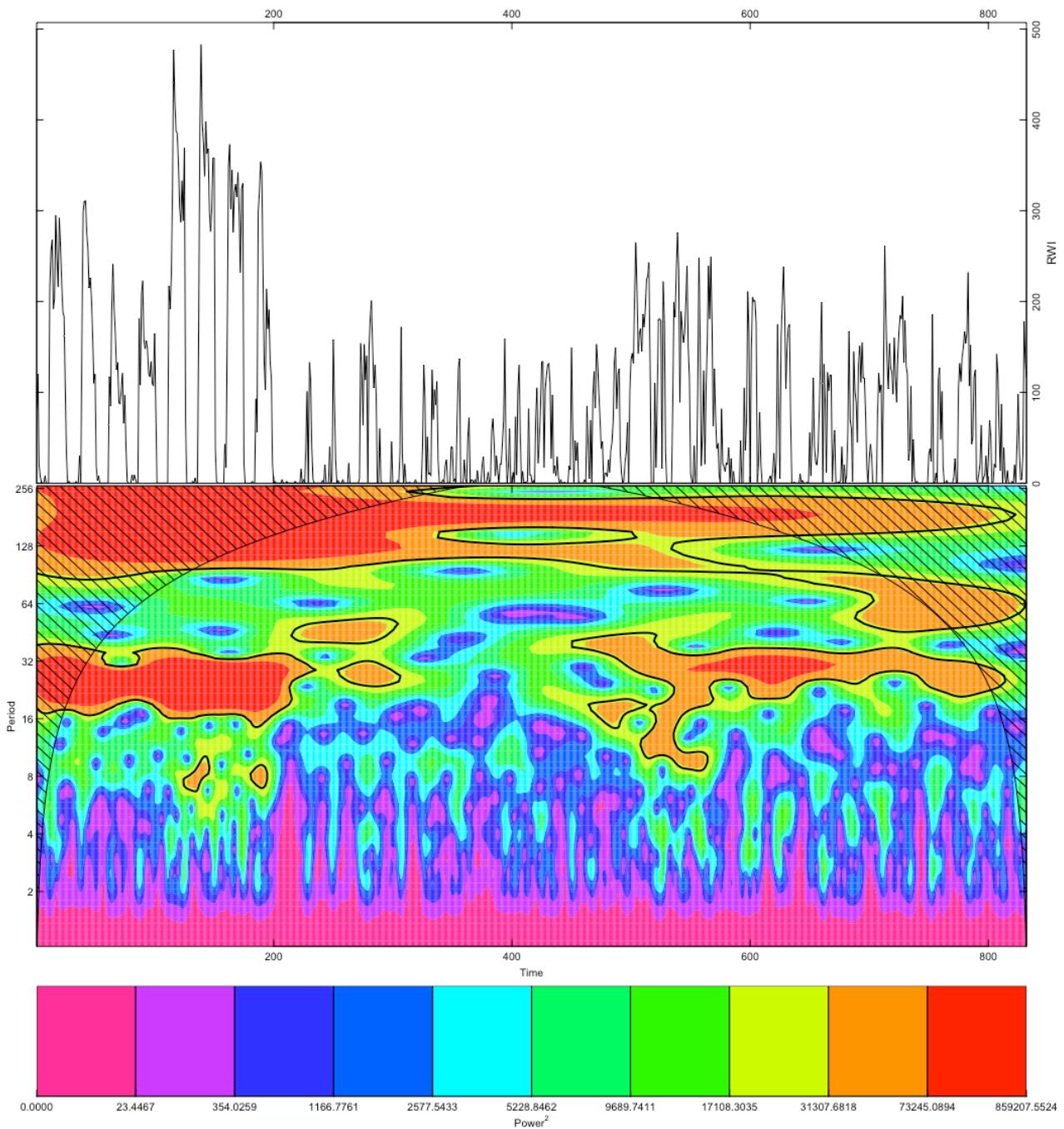
15 min time bins



30 min time bins



1 hr time bins



Code

```

1 library(dplR)
2
3 # try 1 hour interval
4
5 wave.out00 <- morlet(y1 = counts60$counts60, p2 = 8, dj
  = 0.1, siglvl = 0.95)

```

```

6 #wave.out0$period <- wave.out0$period
7 wavelet.plot(wave.out0)
8
9 ### 30 min
10 wave.out0 <- morlet(y1 = counts30$counts30, p2 = 8, dj
11   = 0.1, siglvl = 0.95)
12 #wave.out0$period <- wave.out0$period/2/24
13 wave.out0$period <- wave.out0$period/2
14 wavelet.plot(wave.out0)
15
16 ##15 min
17 wave.out <- morlet(y1 = counts15$counts15,counts15$day,
18   p2 = 8, dj = 0.1, siglvl = 0.95)
19 #wave.out$period <- wave.out$period/4/24
20 wave.out$period <- wave.out$period/4
21
22 #levs <- quantile(wave.out$Power, c(0, 0.25, 0.5, 0.75,
23   0.95, 1))
24 #wavelet.plot(wave.out, wavelet.levels = levs)
25 wavelet.plot(wave.out)
26
27
28 wave.avg <- data.frame(power = apply(wave.out$Power, 2,
29   mean), period = (wave.out$period))
30
31 findpeaks(wave.avg$power)
32 wave.avg[findpeaks(wave.avg$power)[2,2],] #period in
33 hours
34
35 plot(wave.avg$period, wave.avg$power, type = "l")
36 abline(v=wave.avg[findpeaks(wave.avg$power)[2,2],])
37
38 ###6 min
39 wave.out2 <- morlet(y1 = counts06$counts06, p2 = 8, dj
40   = 0.1, siglvl = 0.95)

```

```

35 wave.out2$period <- wave.out2$period/10
36
37 levs <- quantile(wave.out$Power, c(0, 0.25, 0.5, 0.75,
38 0.95, 1))
39 #wavelet.plot(wave.out, wavelet.levels = levs)
40 wavelet.plot(wave.out2)

```

Finding the dominant period from the wavelet analysis

general code:

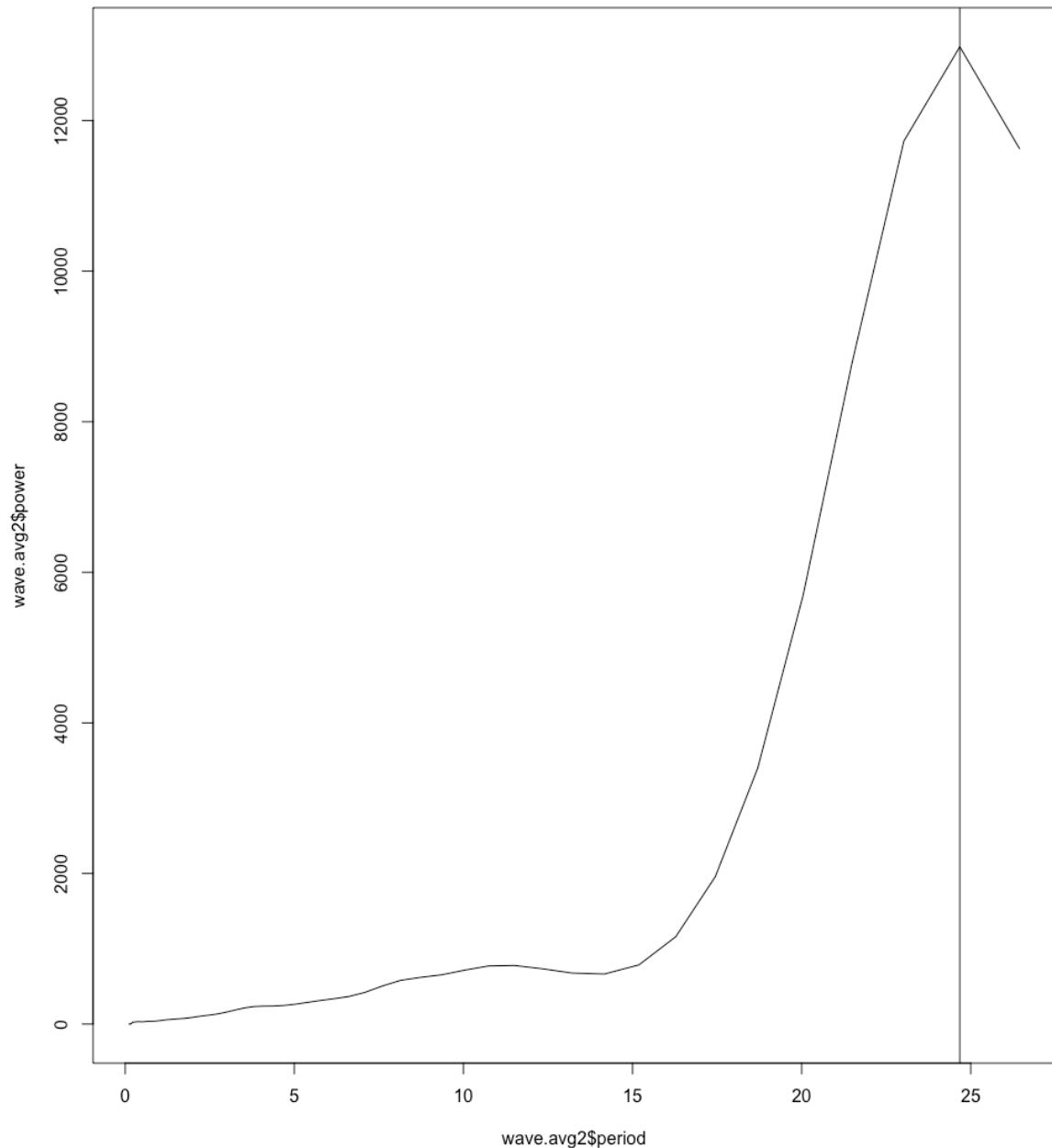
6 min time bin example

```

1 wave.avg2 <- data.frame(power = apply(wave.out2$Power,
2 , mean), period = (wave.out2$period))
3
3 findpeaks(wave.avg2$power)
4 wave.avg2[findpeaks(wave.avg2$power)[3,2],] #period in
hours
5   power    period
6 80 12979.82 24.67491
7
8 plot(wave.avg2$period, wave.avg2$power, type = "l")
9 abline(v=wave.avg2[findpeaks(wave.avg2$power)[3,2],])

```

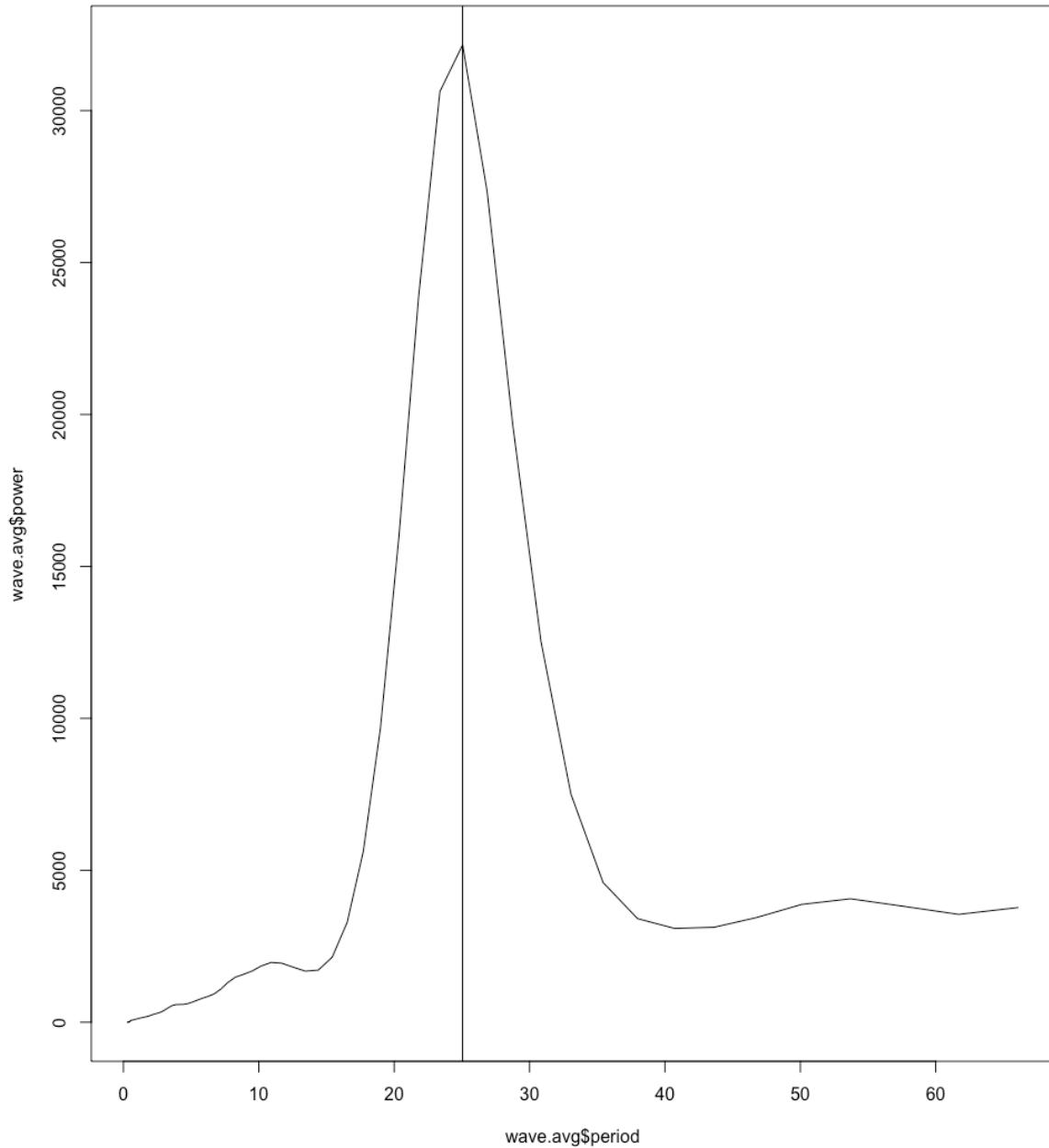
Fig showing power vs period



15/30/60 min have same values

```
1 wave.avg[findpeaks(wave.avg$power)[2,2],] #period in  
hours  
2 power    period  
3 67 32160.23 25.05282
```

Fig showing power vs period



A cool thing to explore is taking a slice of the period and plotting power vs time

this'll let you explore the amplitude at a given period across time

to my understanding, this is very similar to discrete wavelet analysis

trying it out for 6 min time bins

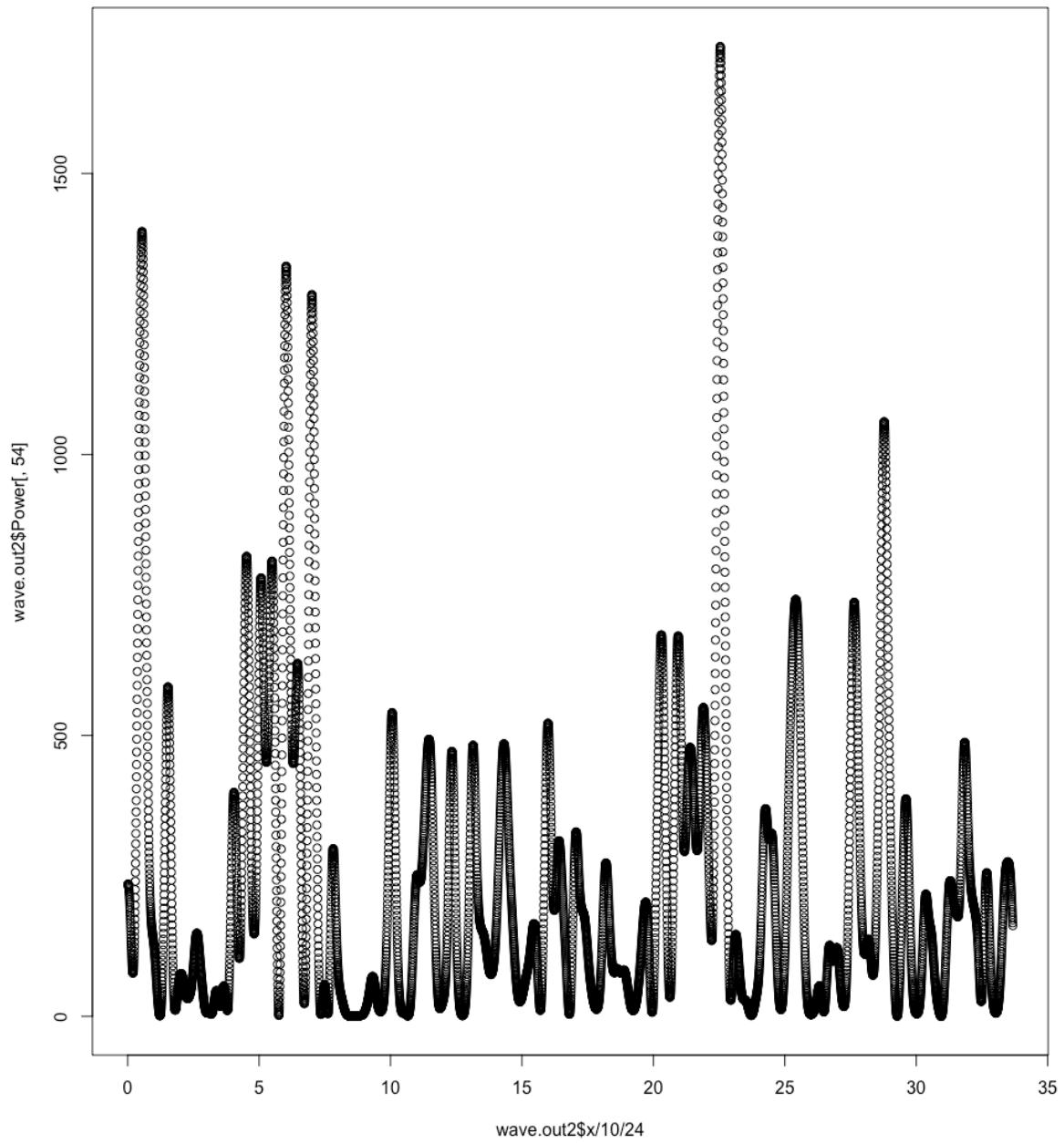
identify the periods I want:

```
1 wave.out2$period
2 [1] 0.1033044 0.1107189 0.1186656 0.1271826
     0.1363109 0.1460944 0.1565801 0.1678184 0.1798633
     0.1927728
3 [11] 0.2066087 0.2214378 0.2373311 0.2543652
     0.2726219 0.2921889 0.3131603 0.3356369 0.3597267
     0.3855455
4 [21] 0.4132175 0.4428755 0.4746622 0.5087304
     0.5452437 0.5843777 0.6263205 0.6712737 0.7194534
     0.7710910
5 [31] 0.8264349 0.8857510 0.9493244 1.0174607
     1.0904874 1.1687555 1.2526411 1.3425475 1.4389068
     1.5421821
6 [41] 1.6528698 1.7715020 1.8986489 2.0349215
     2.1809748 2.3375109 2.5052822 2.6850950 2.8778135
     3.0843642
7 [51] 3.3057397 3.5430041 3.7972977 4.0698429
     4.3619496 4.6750219 5.0105644 5.3701899 5.7556271
     6.1687284
8 [61] 6.6114793 7.0860081 7.5945954 8.1396859
     8.7238993 9.3500438 10.0211288 10.7403799 11.5112541
     12.3374567
9 [71] 13.2229587 14.1720162 15.1891909 16.2793717
     17.4477986 18.7000875 20.0422575 21.4807598 23.0225083
     24.6749134
10 [81] 26.4459174
```

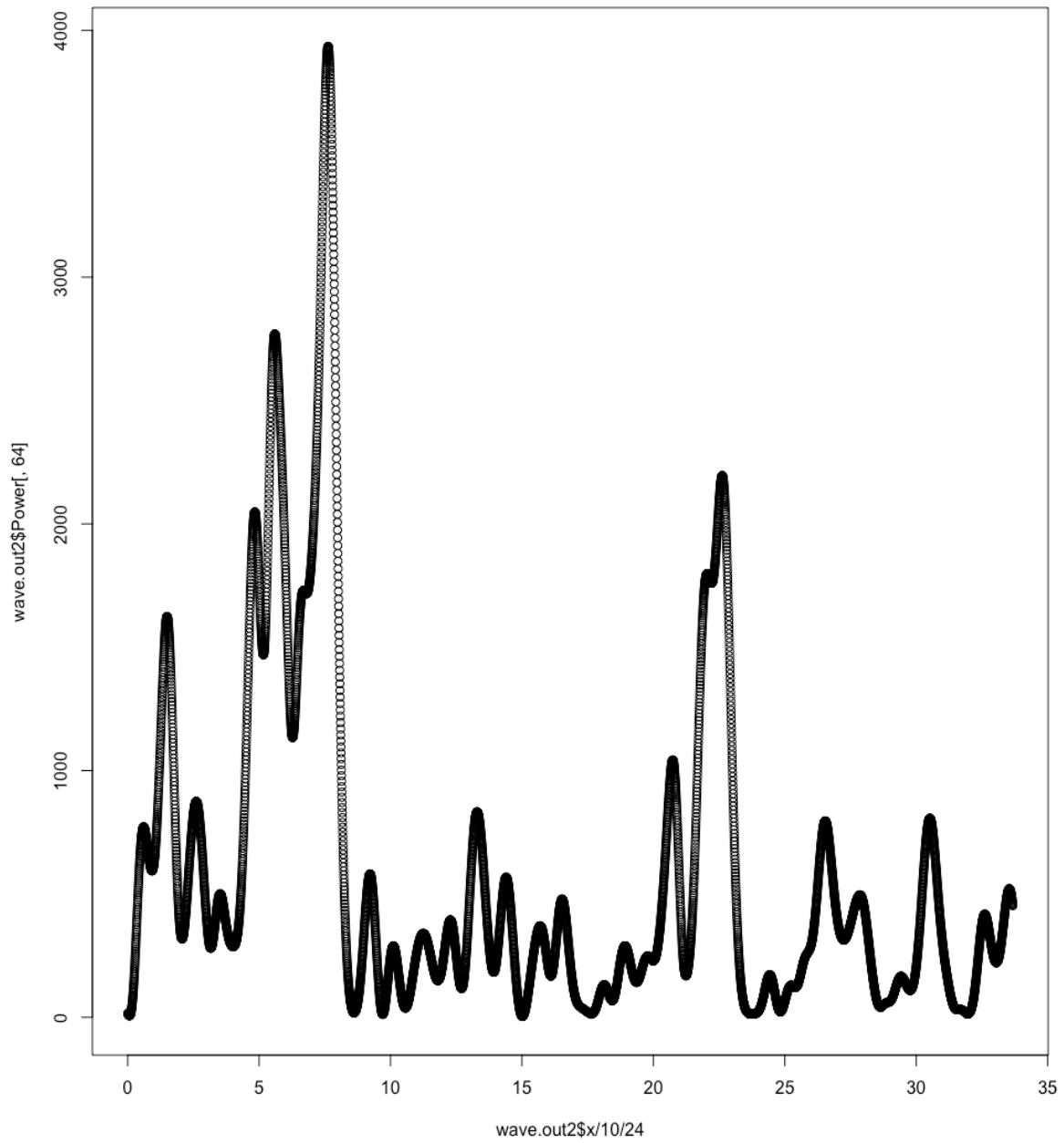
Grab the 4 and 8 hour periods I care about (54 and 64)

```
1 #4 hour rhythms
2 plot(wave.out2$x/10/24, wave.out2$Power[, 54])
3 findpeaks(wave.out2$Power[, 54])
4
5 #8 hour rhythms
6 plot(wave.out2$x/10/24, wave.out2$Power[, 64])
7 findpeaks(wave.out2$Power[, 64])
```

4 hour



8 hour



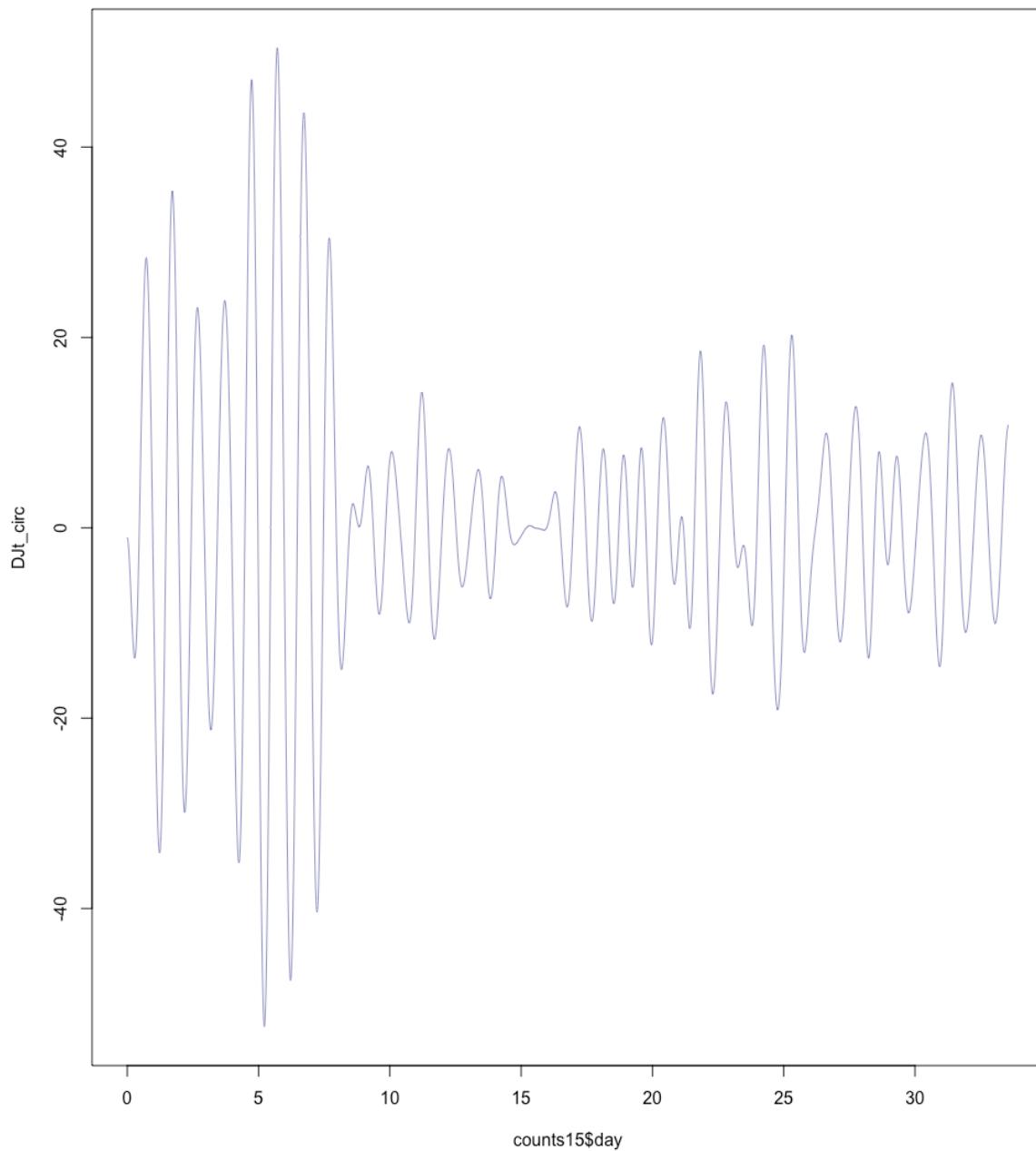
Page 137: 2017-12-03. Discrete wavelet analysis on h404; 15 min bins

The discrete wavelet analysis lets you look at the power for a given period window.

Ok, we can set the window, looking at the circadian window (detail)

```
1 #circadian detail
2 Jcirc <- floor(log2(round(24/.25)))
3 DJt_circ <- wavMRDSum(counts15$counts15,levels=Jcirc
,keep.smooth=FALSE,
keep.details=TRUE,reflect=TRUE,wavelet="s12",xform="modwt")
4
5 plot(counts15$day,DJt_circ,col=rgb(0,0,.5,.5),type="l")
```

Plotting the circadian detail itself

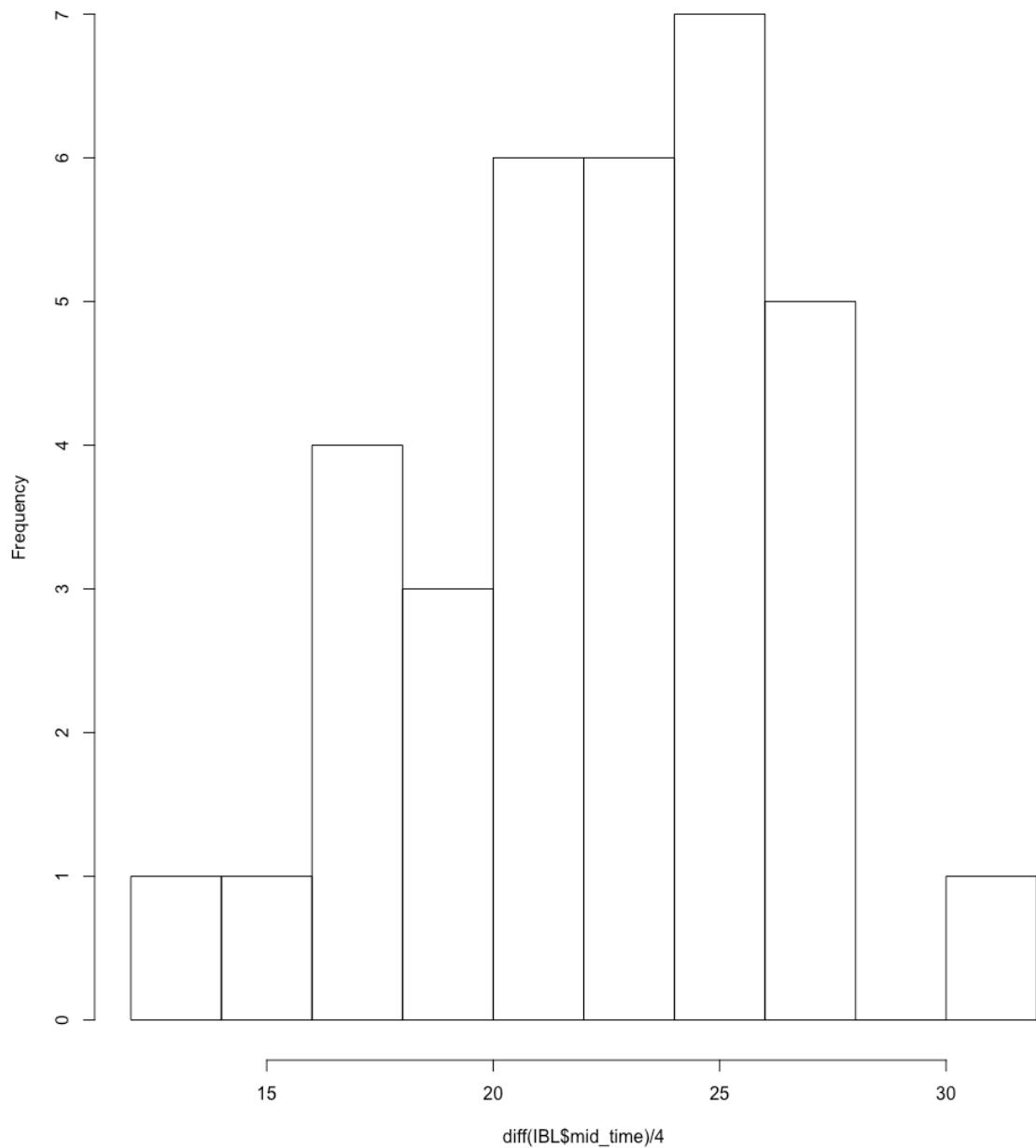


Calculating bout intervals (I'm thinking about it as instantaneous period), taking mid points of peak amplitudes and subtracting them from the next peak midpoint (units are in time then).

```
1 library(pracma)
2 IBL<-data.frame(findpeaks(DJt_circ))
3 names(IBL)<-
4   c("Height", "mid_time", "initial_time", "final_time")
5 IBL
6 diff(IBL$mid_time) / 4
7 hist(diff(IBL$mid_time) / 4)
```

Plotting histogram to see the range of values (x axis is in days)

Histogram of $\text{diff}(\text{IBL}\$mid_time)/4$

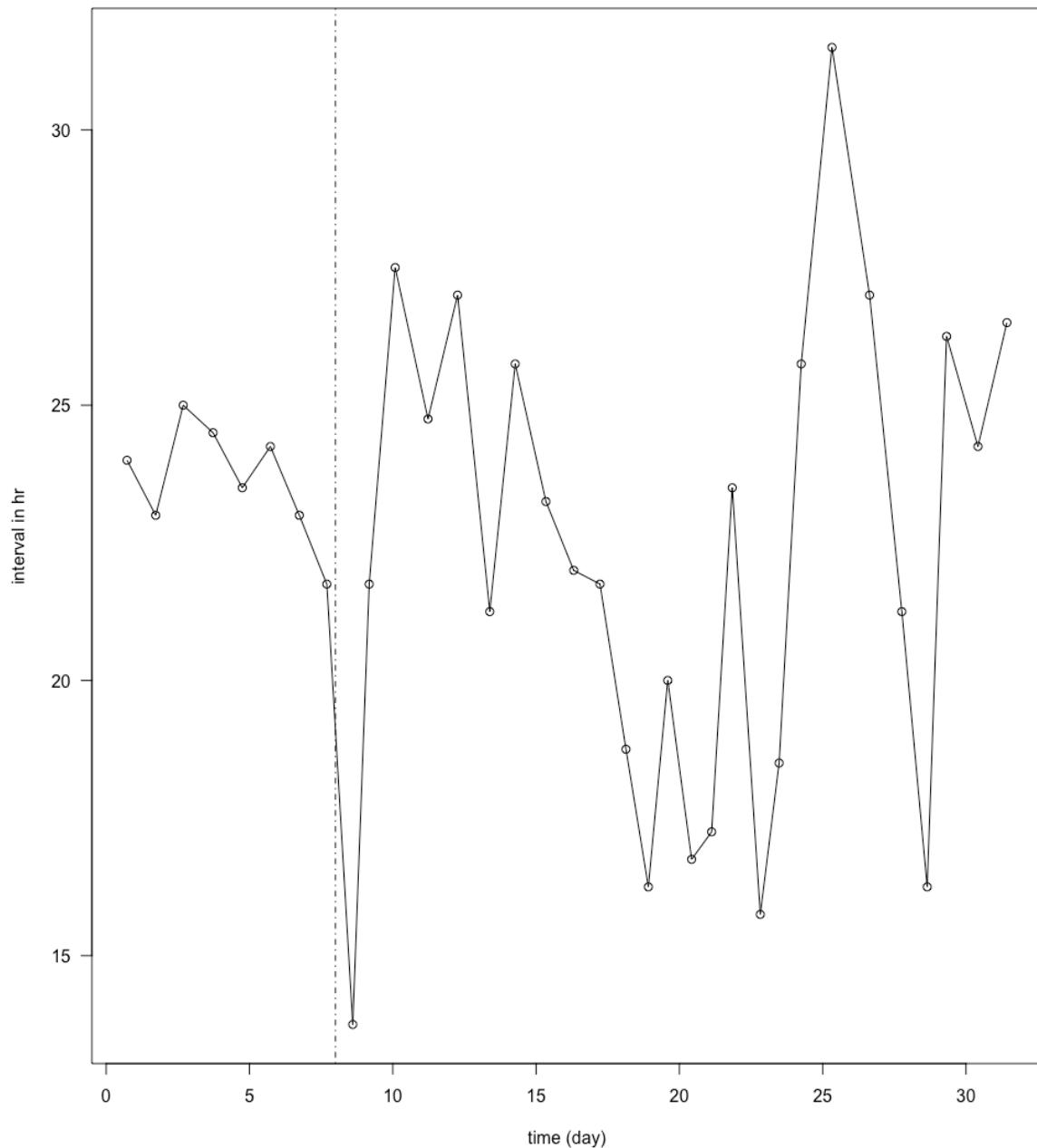


Ok, plotting bout intervals vs time

```

1 plot(IBL$mid_time[-35]/4/24,diff(IBL$mid_time)/4,xlab="time (day)",ylab="interval in hr",las=1)
2 #lines(loess(diff(IBL$mid_time)~IBL$mid_time[-31],span=1))
3 s<-(-IBL$mid_time[-35]/4/24)
4 lines(loess(diff(IBL$mid_time)/4~s,span=1))
5 abline(v=8,lty="dotdash")

```

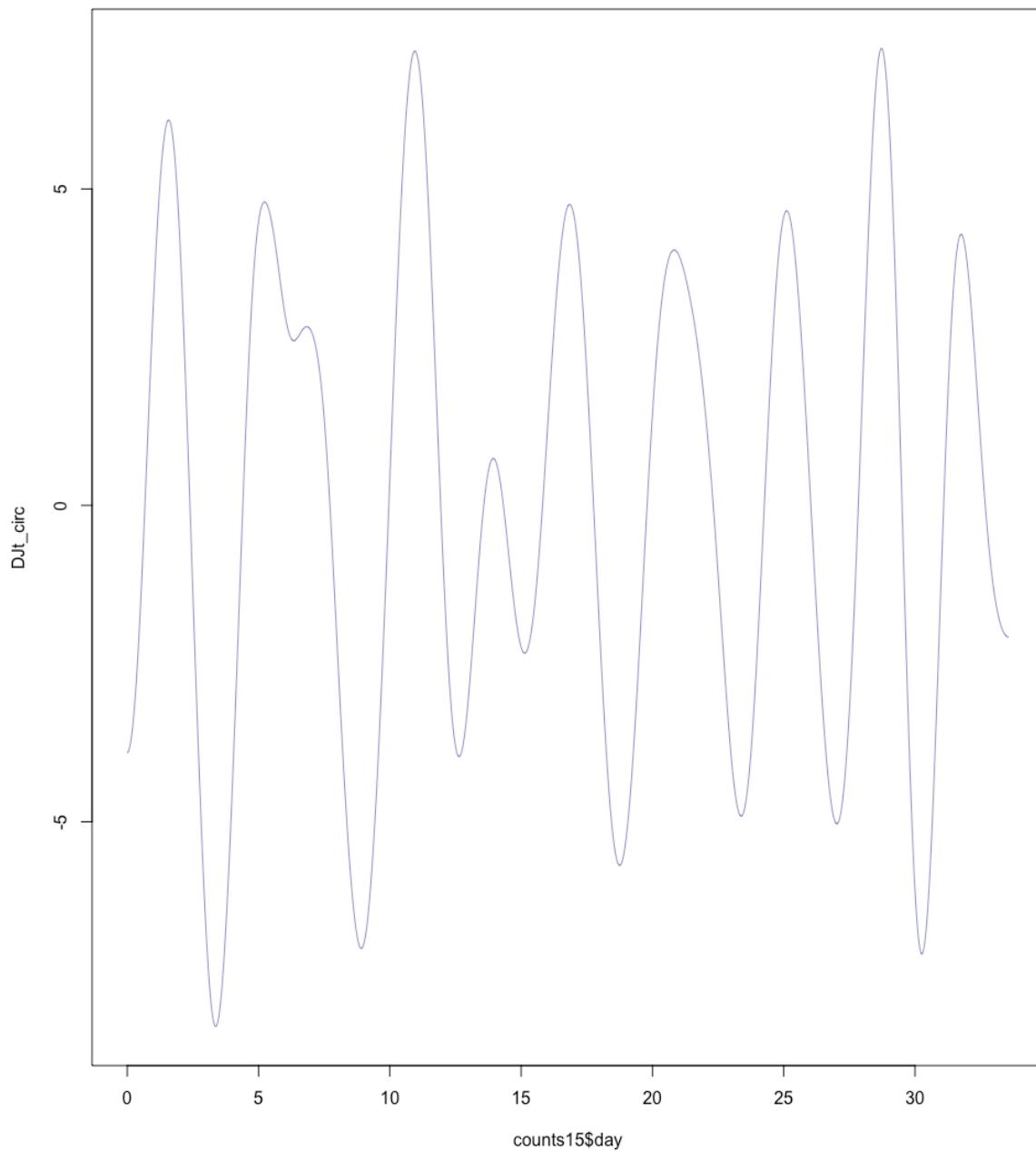


Notes: I'm analyzing both entrainment and free run data together. I could separate.

Looking at highest detail to explore if we want to see if there are more long term patterns

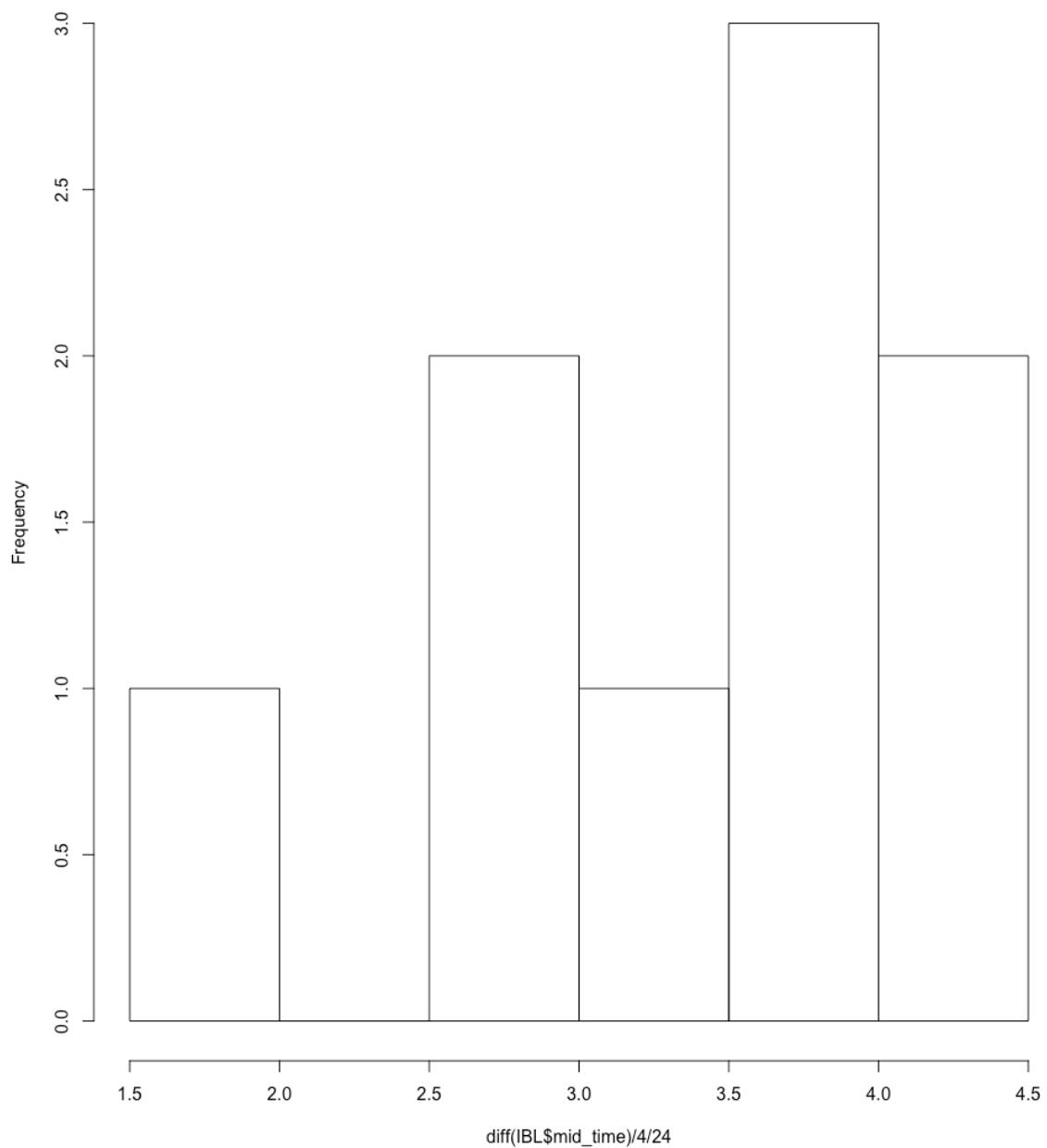
```
1 #highest detail
2 DJt_circ <- wavMRDSum(counts15$counts15, levels=8
, keep.smooth=FALSE,
keep.details=TRUE, reflect=TRUE, wavelet="s12", xform="modwt")
```

wavelet detail

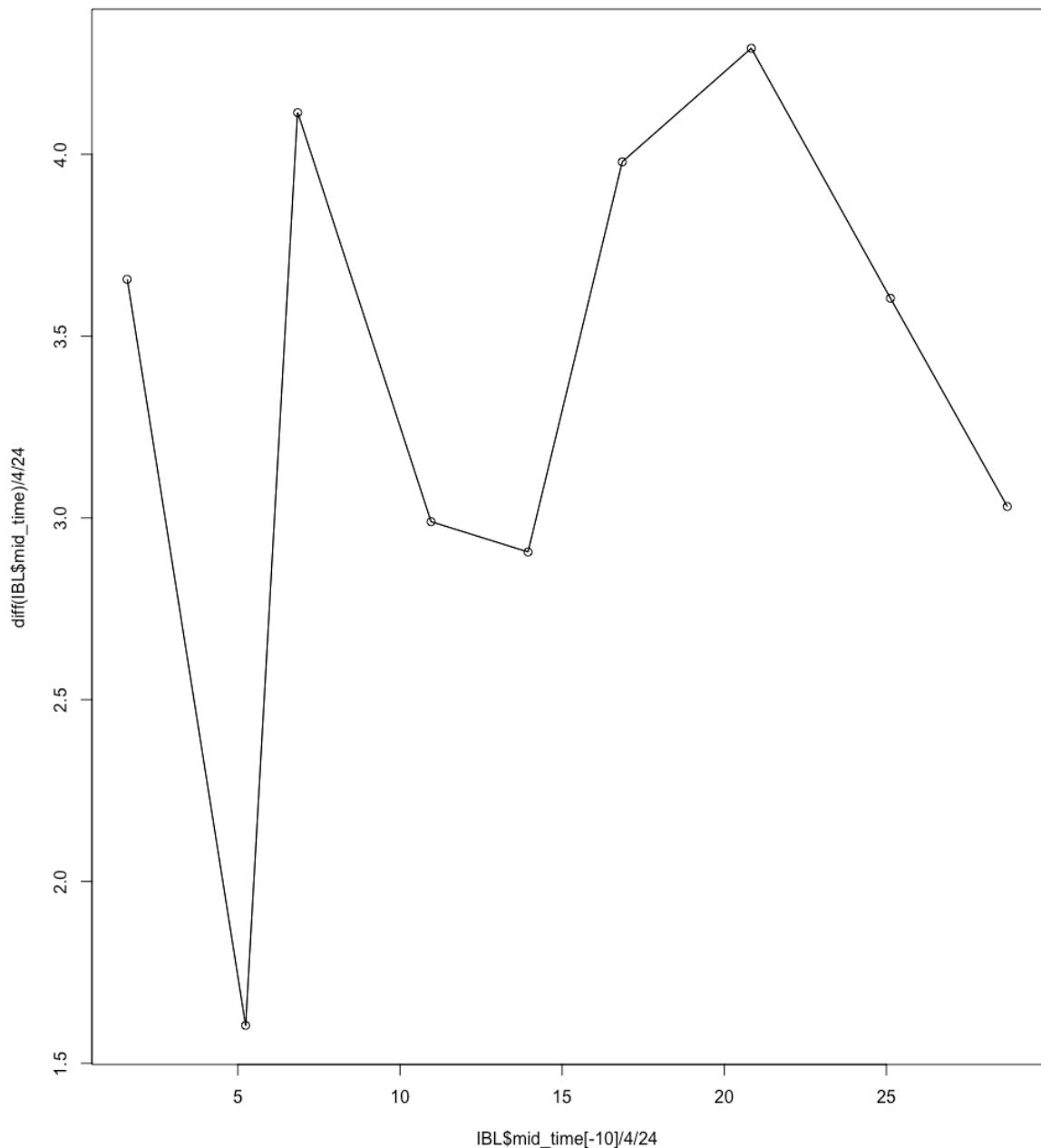


histogram

Histogram of $\text{diff}(\text{IBL}\$\text{mid_time})/4/24$



bout intervals



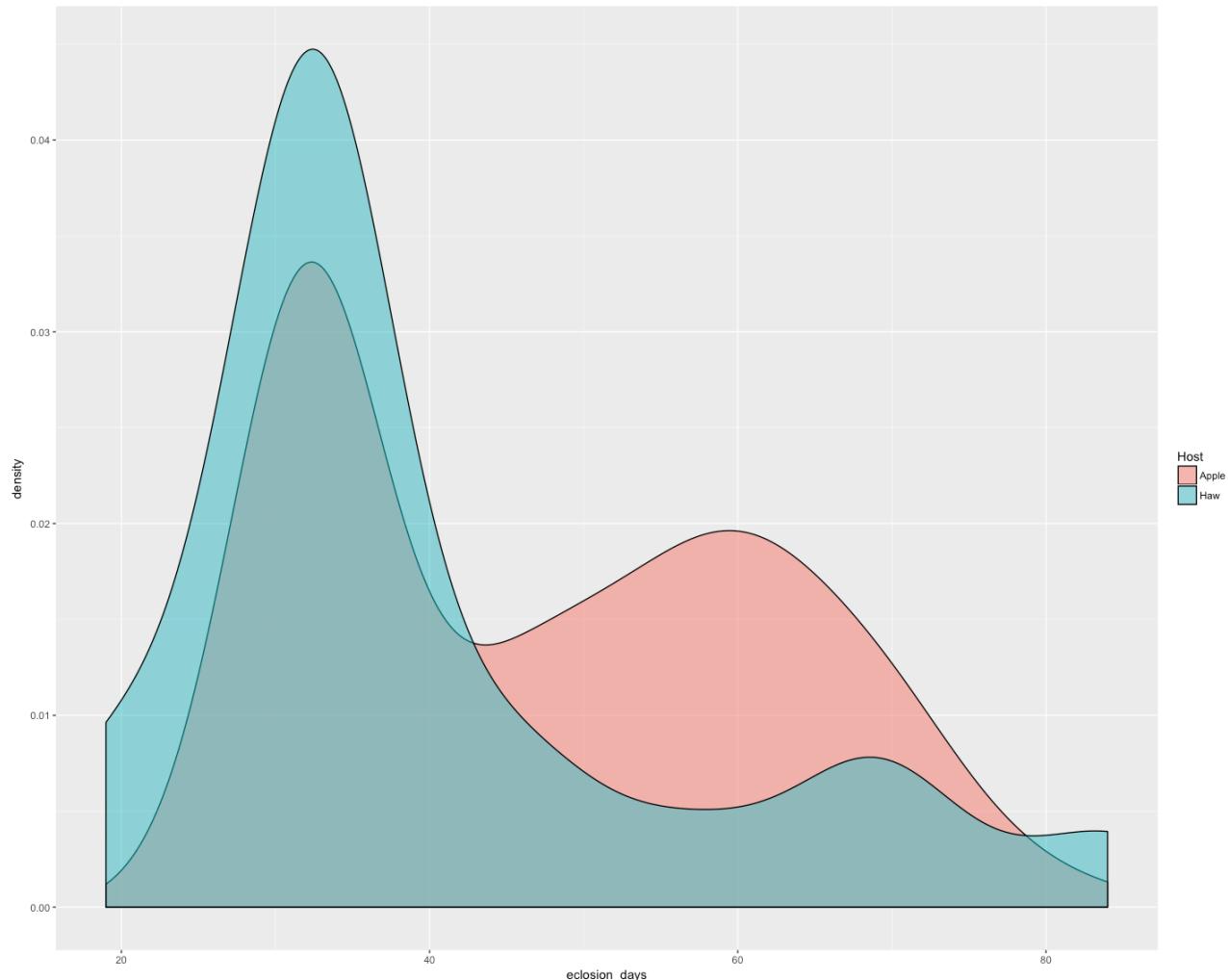
**Page 138: 2017-12-05. Prepping meeting with
Dan on 2017-12-07, 10:30AM**

Stuff to go over:

1. Go over data analysis

- Discriminant analysis - difficult to separate non-diapause(high and low MR), diapause

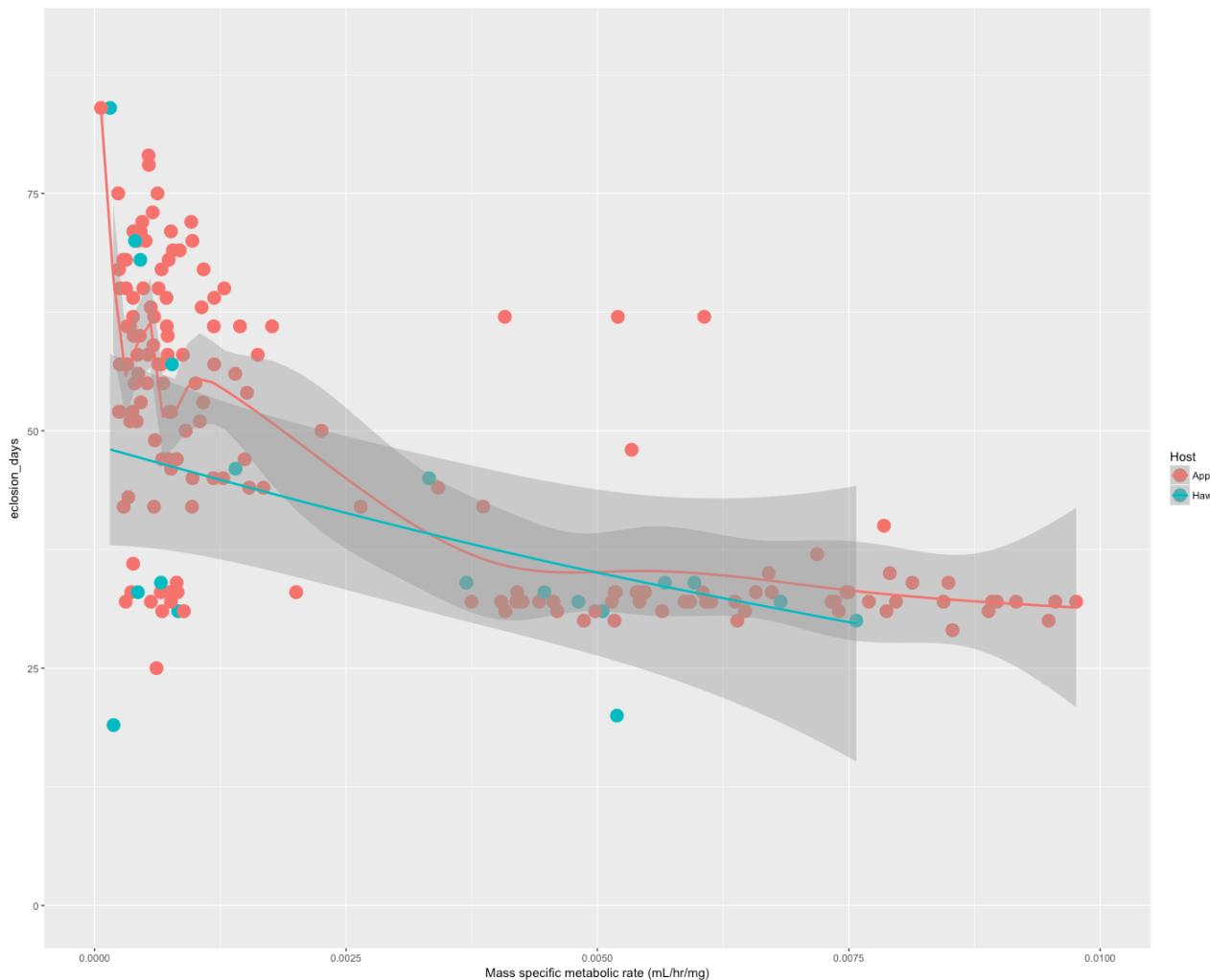
just to recap, here is eclosion data:



pre-designated groups based on cut off:

```
1 trait2$groups<-ifelse(trait2$MR15<  
0.00375, ifelse(trait2$eclosion_days<35, "non-  
diapause_lowMR", "diapause"), "non-diapause_highMR")
```

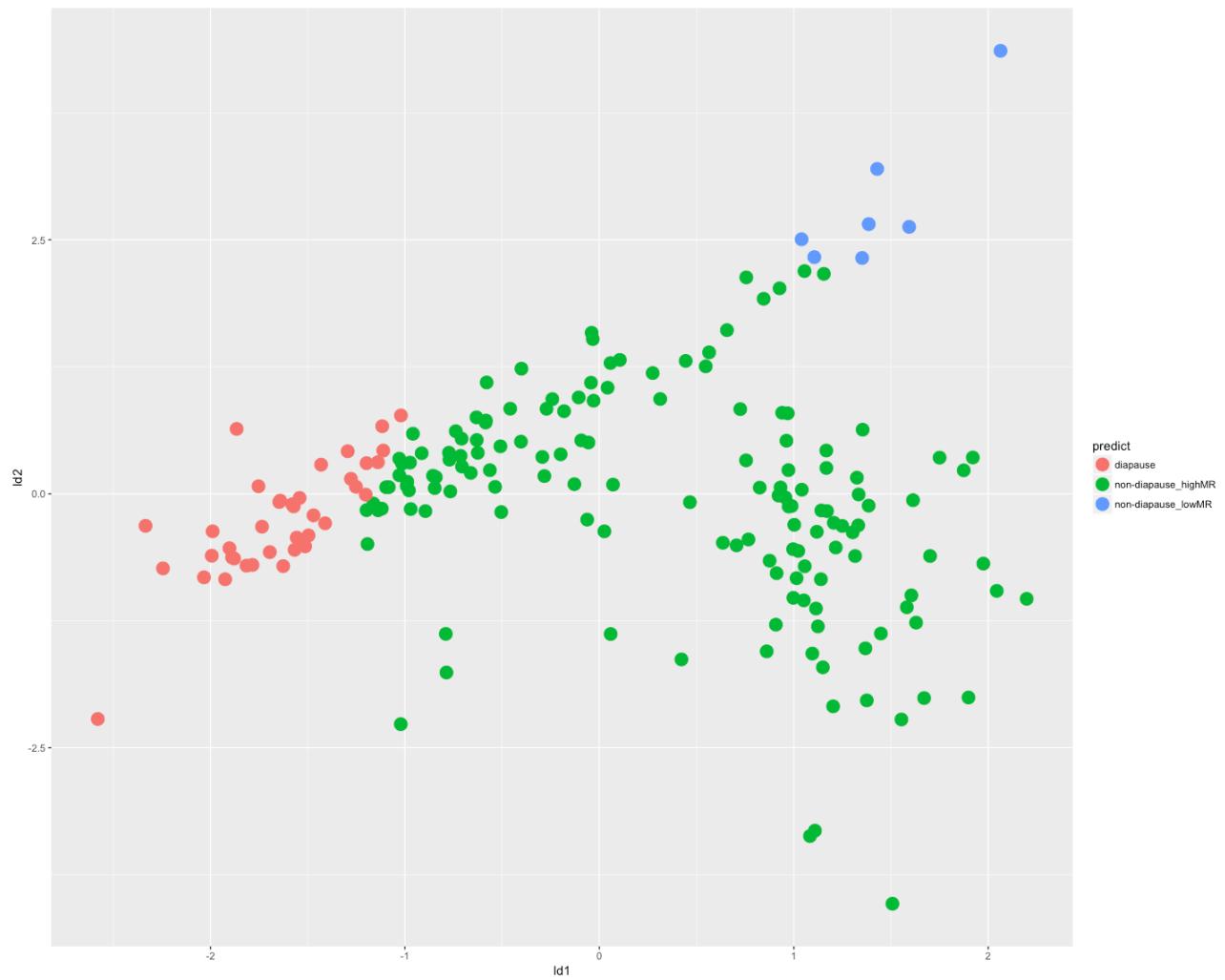
see fig for day 15 mr



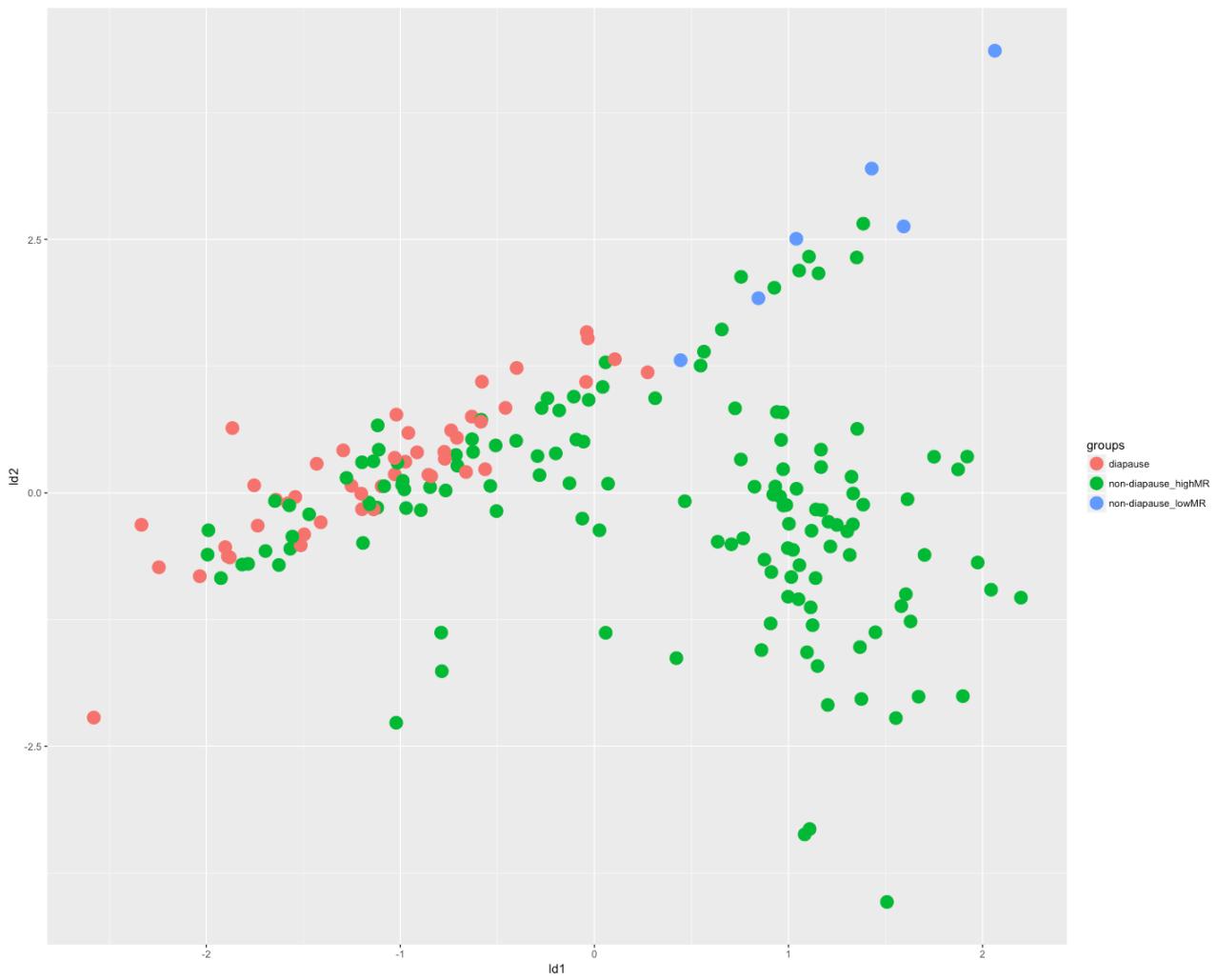
So I'm trying to discriminate diapause classs based on: MR, weight, eclosion day.

```
1 | fit2<-
qda(groups~mass_day10+mass_day14+MR11+MR15+eclosion_days
, data=trait2, CV=FALSE)
```

Discriminant analysis biplots: predicted groups



Discriminant analysis biplots: pre-designated groups



Confusion matrix

	diapause	non-diapause_highMR	non-diapause_lowMR
diapause	21	17	0
non-diapause_highMR	26	120	2
non-diapause_lowMR	0	3	4

75.12% accurate

- Analyzing biological rhythms
 - [Periodograms](#) at different time bins
 - [Continuous wavelet analysis](#) at different time bins -[identifying dominant period](#)
 - [Discrete wavelet analysis](#)

summary table

Time_bin	Method	Frequency_level	Estimate_day	Estimate_hr
1.00	Periodogram	1	18.0000000	432.00000
1.00	Periodogram	2	1.0580000	25.41176
0.50	Periodogram	1	18.0000000	432.00000
0.50	Periodogram	2	0.9729000	23.35135
0.25	Periodogram	1	16.8750000	405.00000
0.25	Periodogram	2	1.0227000	24.54545
0.10	Periodogram	1	16.8750000	405.00000
0.10	Periodogram	2	0.9926000	23.82353
1.00	CWT	2	1.0438670	25.05282
0.50	CWT	2	1.0438670	25.05282
0.25	CWT	2	1.0438670	25.05282
0.10	CWT	2	1.0281213	24.67491
0.25	DWT	2	0.9353554	22.44853

2. Paper discussions

refs:

- Helm, B., Visser, M. E., Schwartz, W., Kronfeld-Schor, N., Gerkema, M., Piersma, T., & Bloch, G. (2017). Two sides of a coin: ecological and chronobiological perspectives of timing in the wild. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1734), 20160246.
<https://doi.org/10.1098/rstb.2016.0246>
- Bulla, M., Valcu, M., Dokter, A. M., Donua, A. G., Kosztolányi, A., Rutten, A. L., ... Kempenaers, B. (2016). Unexpected diversity in socially synchronized rhythms of shorebirds. *Nature*, 540(7631), 109–113. <https://doi.org/10.1038/nature20563>

3. Decide about when to take Rhagos out of fridge

		month (30 day intervals)	
Cohort_day	Cohort_date	4	5
2	2017-08-25	2017-12-23	2018-01-22
3	2017-08-26	2017-12-24	2018-01-23
4	2017-08-27	2017-12-25	2018-01-24
5	2017-08-28	2017-12-26	2018-01-25
10	2017-09-02	2017-12-31	2018-01-30
11	2017-09-03	2018-01-01	2018-01-31
12	2017-09-04	2018-01-02	2018-02-01
13	2017-09-05	2018-01-03	2018-02-02
14	2017-09-06	2018-01-04	2018-02-03
16	2017-09-08	2018-01-06	2018-02-05

18	2017-09-10	2018-01-08	2018-02-07
19	2017-09-11	2018-01-09	2018-02-08
20	2017-09-12	2018-01-10	2018-02-09
1	2017-09-22	2018-01-21	2018-02-20
2	2017-09-23	2018-01-22	2018-02-21
3	2017-09-24	2018-01-23	2018-02-22
4	2017-09-25	2018-01-24	2018-02-23
5	2017-09-26	2018-01-25	2018-02-24
6	2017-09-27	2018-01-26	2018-02-25
7	2017-09-28	2018-01-27	2018-02-26
8	2017-09-29	2018-01-28	2018-02-27
9	2017-09-30	2018-01-29	2018-02-28
			2018-03-

10	2017-10-01	2018-01-30	01
11	2017-10-02	2018-01-31	2018-03-02
12	2017-10-03	2018-02-01	2018-03-03

2017-12-07 meeting with Dan Notes:

For next week:

- scale up to handful of more individuals.
- do at all possible bins
- all analyses

Email Tom's data to Tom, Dan, and Greg to coordinate and decide what time to take samples out of the fridge 4, 4.5 or 5 months.

Go over next week and read 2 more papers, (I choose).

1. Research article: Marine biorhythms: bridging chronobiology and ecology

Martin Bulla, Thomas Oudman, Allert I. Bijleveld, Theunis Piersma, Charalambos P. Kyriacou

Phil. Trans. R. Soc. B 2017 372 20160253; DOI:

10.1098/rstb.2016.0253. Published 9 October 2017

2. Review article: Flexible clock systems: adjusting the temporal programme

Daan R. van der Veen, Sjaak J. Riede, Paul D. Heideman, Michaela Hau, Vincent van der Vinne, Roelof A. Hut

Phil. Trans. R. Soc. B 2017 372 20160254; DOI:

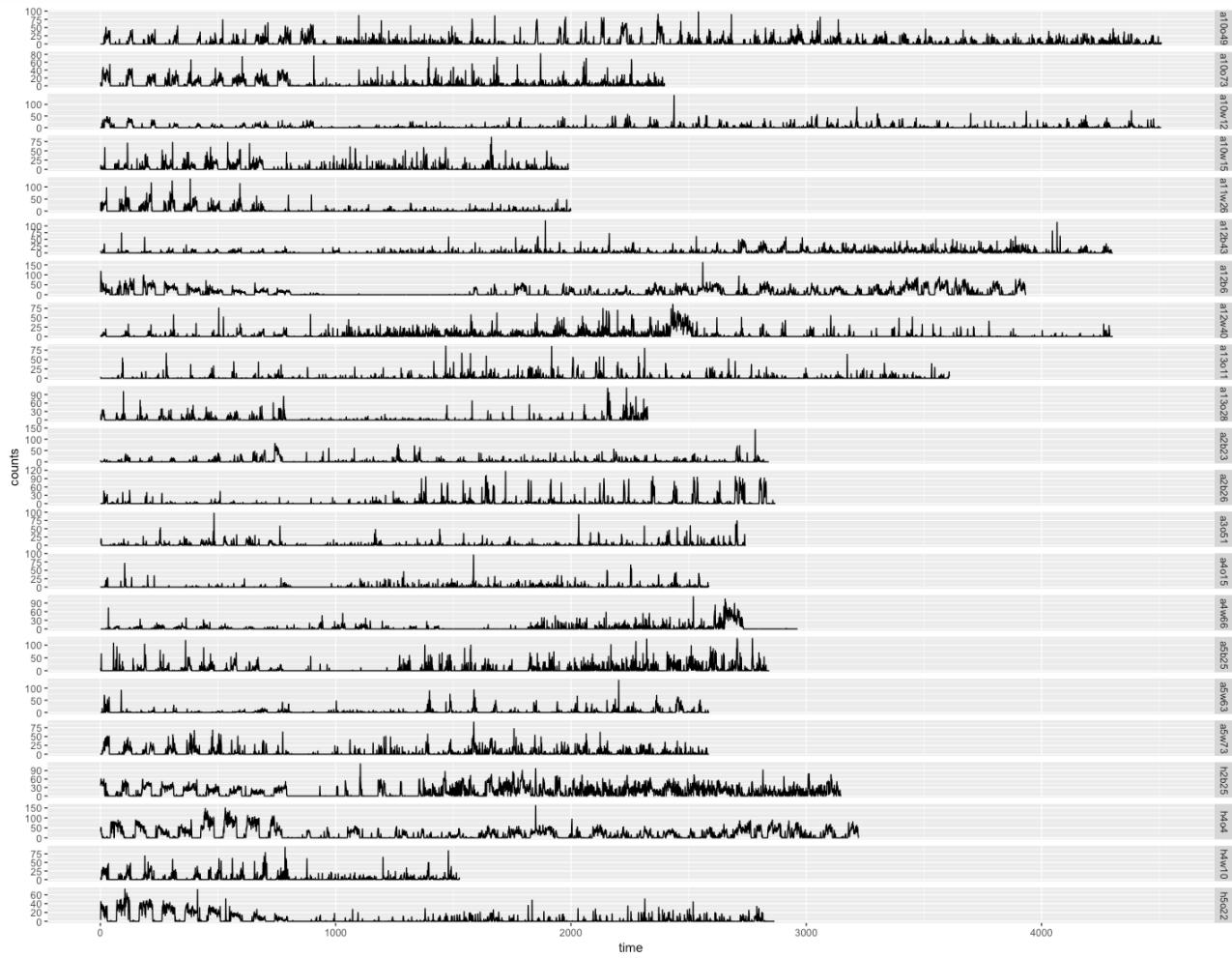
10.1098/rstb.2016.0254. Published 9 October 2017

Monday go to dimensions biodiv data wrangling meeting.

Page 139: 2017-12-08. behavioral data (trikinetics) data parsing update--fixing date order again; update from [2017-11-30 update](#).

Showing code to accurately order dates for 15 min bin:

```
1 bins15=c(paste0(rep(c(paste0(0,0:9),10:23), each=4)," : ",  
c("00",15,30,45))[-1],"24:00")  
2 all.data$bins15=cut(all.data$time2,breaks=seq(0,24,.25),  
labels=bins15)  
3  
4 nall.data15<-ddply(all.data,.  
(uniqueID,date,experiment,bins15),summarize,counts=sum(c  
ounts))  
5 nall.data15_2<-na.omit(ddply(nall.data15,.  
(uniqueID,experiment),transform,order(date)))  
6 nall.data15_3<-ddply(nall.data15_2,.  
(uniqueID),transform,time=seq(1,length(bins15),1))  
7 #write.csv(nall.data15_3,"2017-12-  
08_15min_bins_subset_behavior.csv")  
8 ggplot(nall.data15_3,aes(x=time,y=counts))+geom_line()+f  
acet_grid(uniqueID~,scales="free")
```



These data were manually extracted from the trikinetics monitors and are a represented set of individuals to estimate biological rhythms

ok, i'll estimate the dominant period with periodogram

making a function first: estimate the dominant 3 peaks of the periodogram and convert to days

```

1 ##making a function to grab top period values
2 per.fun<-function(ts=nall.data15_3$counts){
3 m<-
4   data.frame(freq=periodogram(ts,plot=FALSE)$freq,spec=periodogram(ts,plot=FALSE)$spec)
5   return(1/(m[order(m$spec,decreasing=TRUE),]
6   [1:3,1])/4/24)
7 }
```

Using function in ddply, so estimate dominant 3 peaks for each unique id

```

1 ### using function in ddply
2 ## for each unique ID
3 test<-ddply(nall.data15_3,.(uniqueID),function(sub)
4   per.fun(sub$counts))
5 #summary(nall.data15_3$uniqueID)
6 test
```

uniqueID	V1	V2	V3
a10o49	0.9411765	0.9795918	0.8727273
a10o73	0.9615385	1.0000000	25.0000000
a10w12	1.0000000	0.9600000	24.0000000
a10w15	0.9469697	1.0416667	10.4166667
a11w26	20.8333333	0.9920635	0.9469697
a12b43	45.0000000	0.9782609	0.9574468
a12b6	41.6666667	1.0162602	1.1574074
a12w40	45.0000000	22.5000000	11.2500000
a13o11	0.9991776	37.9687500	0.9492187

a13o28	25.0000000	8.3333333	0.8928571
a2b23	0.8823529	1.0000000	7.5000000
a2b26	1.0000000	0.9375000	30.0000000
a3o51	0.5000000	30.0000000	10.0000000
a4o15	27.0000000	1.0000000	0.9642857
a4w66	31.2500000	7.8125000	15.6250000
a5b25	30.0000000	0.9090909	0.9677419
a5w63	0.9000000	0.9310345	27.0000000
a5w73	1.0000000	0.9642857	13.5000000
h2b25	1.0101010	33.3333333	16.6666667
h4o4	1.0227273	33.7500000	0.9642857
h4w10	1.0000000	0.5000000	0.9411765
h5o22	30.0000000	15.0000000	1.0000000

Using function in ddply, so estimate dominant 3 peaks for each unique id and experiment (entrainment vs free run)

```

1  ### for each unique ID and experiment
2  test2<-ddply(nall.data15_3,.
   (uniqueID,experiment),function(sub) per.fun(sub$counts))
3  test2

```

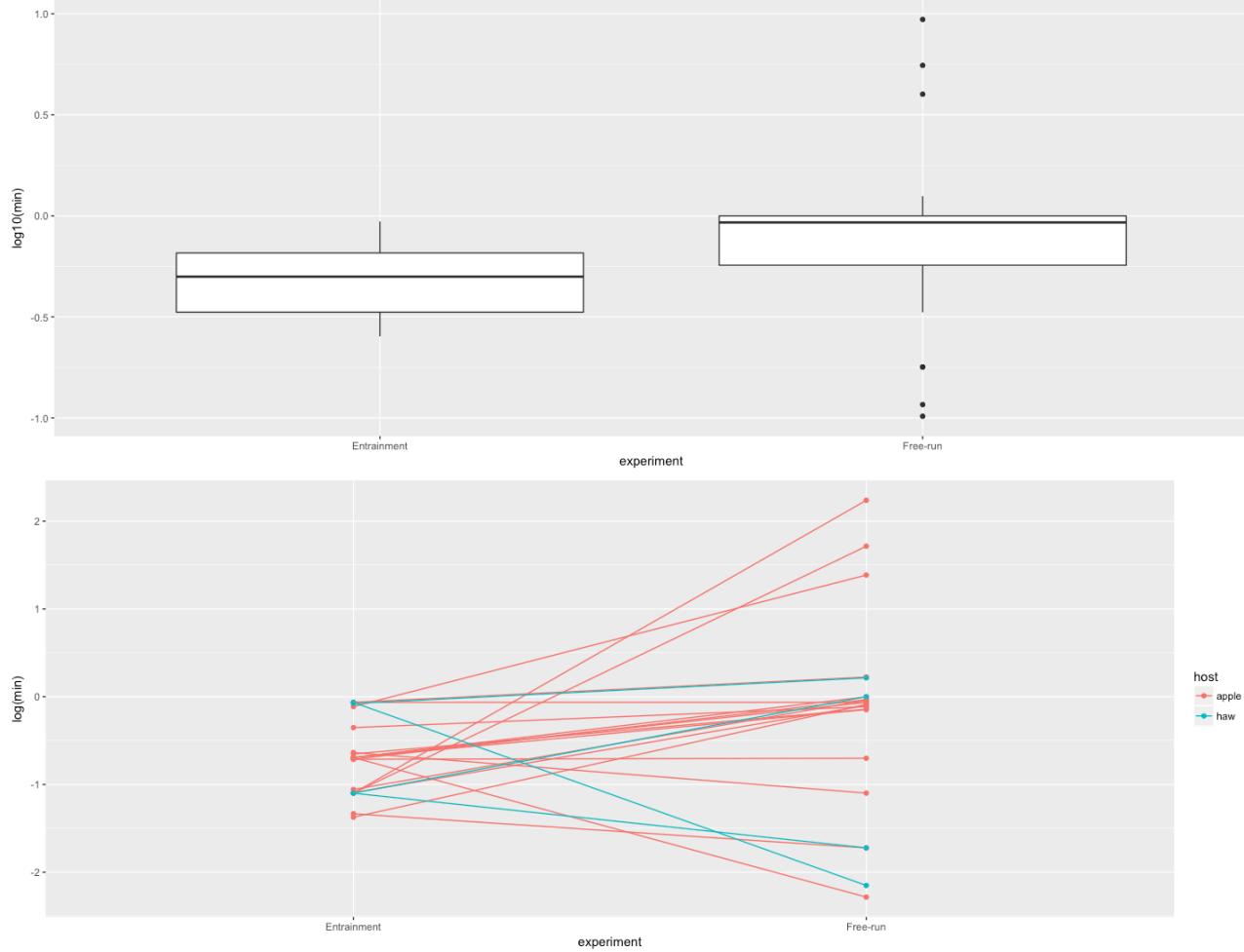
uniqueID	experiment	V1	V2	V3
a10o49	Entrainment	1.0000000	0.5000000	10.0000000

a10o49	Free-run	0.9492187	37.9687500	0.8629261
a10o73	Entrainment	1.0000000	1.1250000	0.5294118
a10o73	Free-run	0.9803922	0.9259259	0.3333333
a10w12	Entrainment	1.0000000	0.5000000	1.1111111
a10w12	Free-run	0.9991776	37.9687500	9.4921875
a10w15	Entrainment	1.0714286	0.9375000	0.5000000
a10w15	Free-run	1.0044643	0.9375000	14.0625000
a11w26	Entrainment	0.9375000	1.0714286	0.5000000
a11w26	Free-run	0.9375000	14.0625000	0.1019022
a12b43	Entrainment	1.0416667	0.5208333	0.3472222
a12b43	Free-run	37.5000000	0.9615385	0.9868421
a12b6	Entrainment	0.9375000	1.0546875	8.4375000
a12b6	Free-run	32.5520833	10.8506944	1.2520032
a12w40	Entrainment	1.0416667	0.4901961	0.3333333
a12w40	Free-run	37.5000000	12.5000000	9.3750000
a13o11	Entrainment	1.0416667	0.4901961	1.6666667
a13o11	Free-run	30.0000000	0.9677419	1.0000000
a13o28	Entrainment	1.0416667	0.3333333	0.4901961
a13o28	Free-run	16.6666667	8.3333333	5.5555556
a2b23	Entrainment	0.9375000	9.3750000	0.5208333
a2b23	Free-run	0.8680556	0.9920635	0.9469697
a2b26	Entrainment	4.6875000	9.3750000	0.2533784
a2b26	Free-run	1.0044643	0.9171196	10.5468750
a3o51	Entrainment	0.4901961	0.9259259	8.3333333
a3o51	Free-run	0.9920635	20.8333333	0.4960317
a4o15	Entrainment	8.4375000	0.2636719	0.3515625
a4o15	Free-run	0.9868421	0.3348214	0.1785714
a4w66	Entrainment	1.0125000	7.5937500	0.8933824
a4w66	Free-run	16.0000000	4.0000000	8.0000000

a5b25	Entrainment	1.0416667	0.4934211	0.3348214
a5b25	Free-run	0.9057971	20.8333333	10.4166667
a5w63	Entrainment	8.4375000	0.7670455	0.7031250
a5w63	Free-run	0.8928571	0.9375000	9.3750000
a5w73	Entrainment	0.9375000	1.0546875	4.2187500
a5w73	Free-run	0.9868421	18.7500000	0.9375000
h2b25	Entrainment	1.0416667	0.9259259	0.3333333
h2b25	Free-run	25.0000000	12.5000000	1.0000000
h4o4	Entrainment	1.0416667	0.9259259	8.3333333
h4o4	Free-run	1.2400794	13.0208333	26.0416667
h4w10	Entrainment	1.0000000	0.5000000	0.3333333
h4w10	Free-run	1.0714286	0.9375000	0.1785714
h5o22	Entrainment	1.0416667	0.9375000	9.3750000
h5o22	Free-run	10.4166667	0.9920635	0.1163873

2017-12-13. Adding figures

Comparing the min value across treatments and hosts



Trying to find dominant peak with continuous wavelet analysis

```

1 #function to find dominant peak with continuous wavelet
  anlaysis (dplr package)
2 cwa<-function(series=counts15$counts15){
3   wave.out <- morlet(y1 = series, p2 = 8, dj = 0.1,
4   siglvl = 0.95)
5   wave.out$period <- wave.out$period/4
6   wave.avg <- data.frame(power = apply(wave.out$Power,
7   2, mean), period = (wave.out$period))
8   aa<-data.frame(findpeaks(wave.avg$power))
9   line<-aa[order(aa,decreasing=TRUE),][1,2]
10  return(wave.avg[line,][2])
11 }
```

```

10
11 cwa()
12
13 #applying function for each unique id and experiment
14 test3<-ddply(nall.data15_3,.
15   (uniqueID,experiment),function(sub) cwa(sub$counts))
16 test3

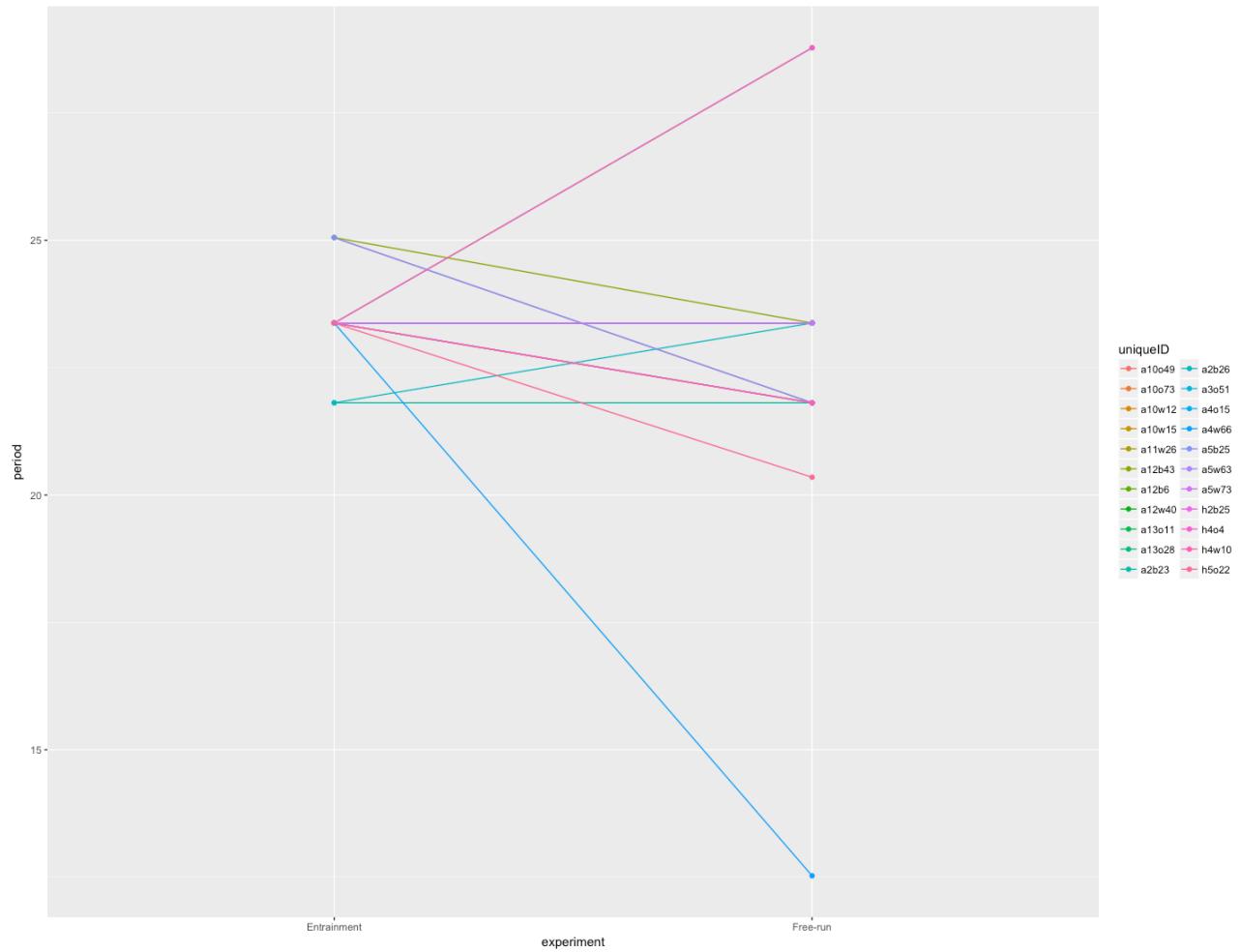
```

uniqueID	experiment	period
a10o49	Entrainment	25.05282
a10o49	Free-run	21.80975
a10o73	Entrainment	23.37511
a10o73	Free-run	23.37511
a10w12	Entrainment	23.37511
a10w12	Free-run	23.37511
a10w15	Entrainment	23.37511
a10w15	Free-run	23.37511
a11w26	Entrainment	23.37511
a11w26	Free-run	21.80975
a12b43	Entrainment	25.05282
a12b43	Free-run	23.37511
a12b6	Entrainment	23.37511
a12b6	Free-run	28.77814
a12w40	Entrainment	23.37511

a12w40	Free-run	23.37511
a13o11	Entrainment	23.37511
a13o11	Free-run	23.37511
a13o28	Entrainment	23.37511
a13o28	Free-run	21.80975
a2b23	Entrainment	21.80975
a2b23	Free-run	21.80975
a2b26	Entrainment	21.80975
a2b26	Free-run	23.37511
a3o51	Entrainment	23.37511
a3o51	Free-run	23.37511
a4o15	Entrainment	23.37511
a4o15	Free-run	23.37511
a4w66	Entrainment	23.37511
a4w66	Free-run	12.52641
a5b25	Entrainment	25.05282
a5b25	Free-run	21.80975
a5w63	Entrainment	23.37511
a5w63	Free-run	23.37511
a5w73	Entrainment	23.37511
a5w73	Free-run	23.37511

h2b25	Entrainment	23.37511
h2b25	Free-run	21.80975
h4o4	Entrainment	23.37511
h4o4	Free-run	28.77814
h4w10	Entrainment	23.37511
h4w10	Free-run	21.80975
h5o22	Entrainment	23.37511
h5o22	Free-run	20.34921

```
1 ggplot(test3,aes(x=experiment,y=period,group=uniqueID,co  
lour=uniqueID))+geom_point() +geom_line()
```



Discrete wavelet analysis

creating the function to id bouts, height, length of activity in time per peak

```

1 ## creating function for discrete wavelet analysis
2 ### getting the circadian detail
3 dwa<-function(time){
4   DJt_circ <- wavMRDSum(time,levels=6
5   ,keep.smooth=FALSE,
6   keep.details=TRUE,reflect=TRUE,wavelet="s12",xform="mod
7   wt")
8   IBL<-data.frame(findpeaks(DJt_circ))
9   names(IBL)<-
10  c("Height","mid_time","initial_time","final_time")
11  #return(diff(IBL$mid_time)/4)
12  return(data.frame(cbind(IBL[,-3:-4],bout=as.numeric(c
13  ("NA",diff(IBL$mid_time)/4))),act_length=IBL[,4]-
14  IBL[,3]))
15 }

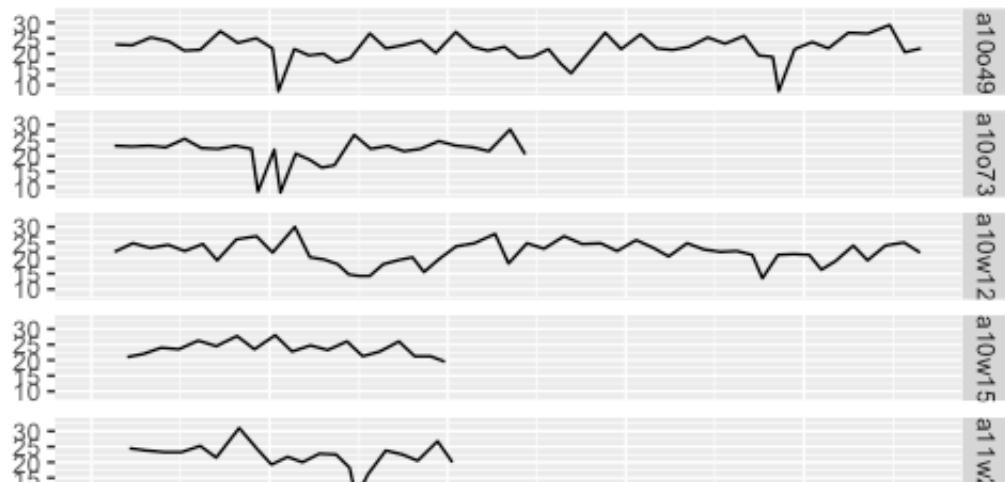
```

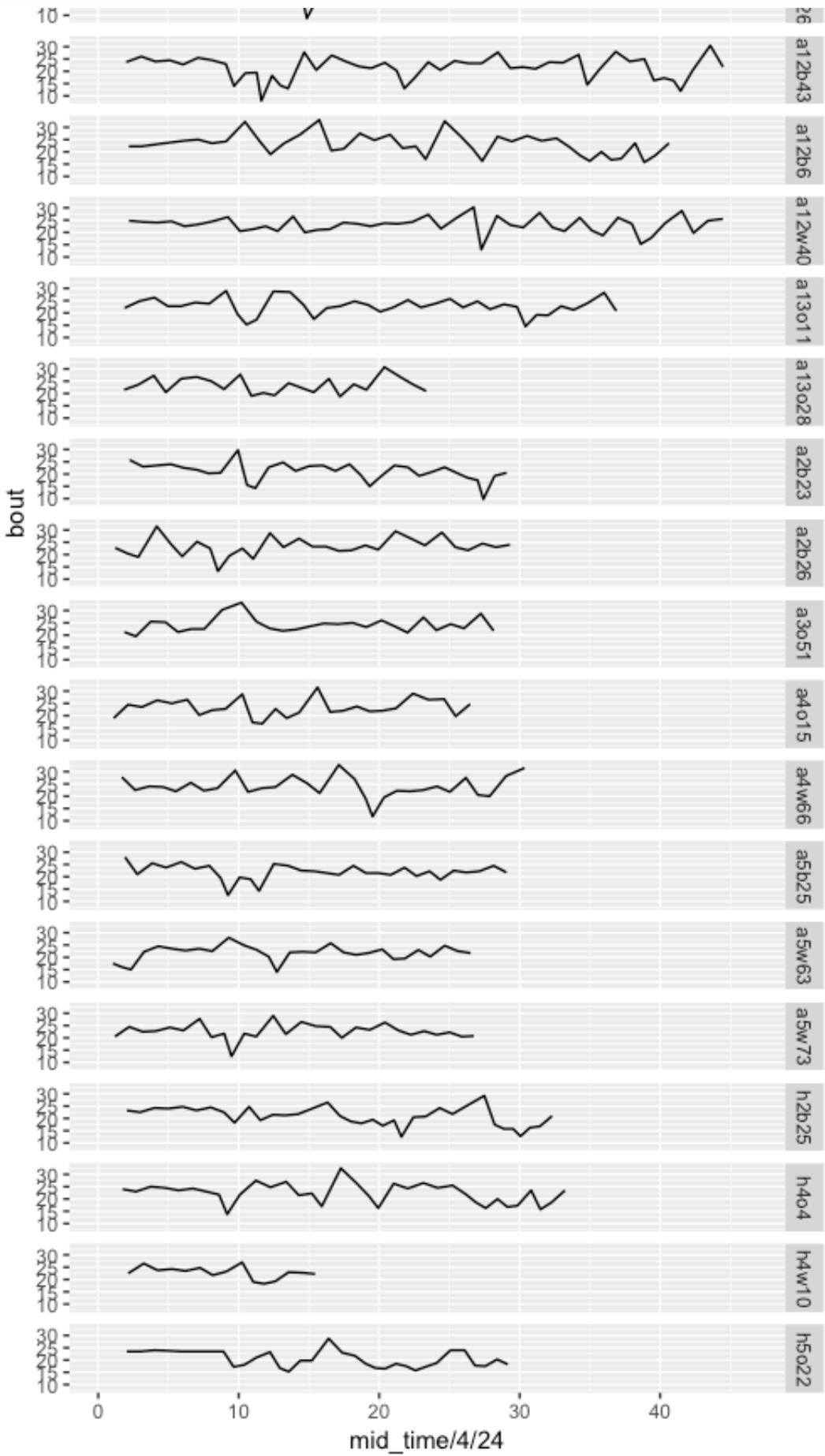
Finding the bouts, height, act time,time ,for each unique ID

```

1 test4<-ddply(nall.data15_3,.(uniqueID),function(sub)
2   dwa(sub$counts))
3
4 ggplot(test4,aes(x=mid_time/4/24,y=bout))+geom_line()+
5   facet_grid(uniqueID~.)

```





Page 140: 2017-12-11. Dimensions of biodiversity workshop: managing big data

Mainly stayed the first part of day where PIs went over their projects and data management challenges

2017-12-11

data science
workshopdata management challenges of large-scale
projects

Intro.

6 projects

Data-host

Aiculms

- * time scales
- * levels of bio org
- * geo regions
- * co-funded

Catherine Malone (DEB → fellow) - presentation

Reed Beaman (Div Bio infrastructure, Program Director)
to Div Bio working group

what's next? what's needed?

goals

- discuss challenges data management, integration
- discuss improvements
- strategize
- common framework for all biodiversity data
- or connect diff platforms (perfect old platforms)
- new de novo platforms

12
12
24
36
48
60
72
84
96
108
120
32
44

2017-12-11

data science
workshop

data management challenges of large-scale projects

Intro

6 projects

Data-host

Aim/Goals

- * time scales
- * levels of bio org
- * geo regions
- * co-funders

Catherine Malone (DEB → fellow) - presentation

Ross Beauman (Div Bio infrastructure, Program Director)

↳ Div Bio working group

what's next? what's needed?

goals

- discuss challenges: data management, integration
- discuss improvements
- strategize
- common framework for all biodiversity data
- or connect diff platforms (perfect old platforms)
new de novo platforms

12
12
24
36
48
60
72
84
96
108
120
32
44

2017-12-11

Pam Soltis

DIV of eastern forests (China vs US)

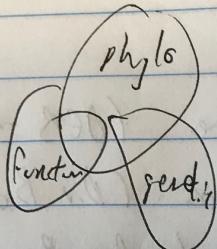
eastern deciduous forests

65 genera are disjunct across continents

dogwood, tulips, opportunity for international collaboration

↳ Chinese collaborating

how diversity is structured b/w those forests?



data types - vouchers, leaves, DNA, RNA, later microbial samples

- community phylogenies, 20 genera, phylogenies
- microbial diversity
- functional trait data
- remote sensing data

NEON sites in US
5 sites in China

data formats

- DNA sequences
- excel sheets
- Images
- diverse software

tolken.org - Tree of Life Knowledge & Information Network

(Cellinese & Beaman 2012)

Ana Carnaval

Fundação
Nasa
FAPESP

Brazilian Atlantic Forest hotspots

Brazil
State
fundação

mode 1. patterns of diversity

use microsensors

ecology → genetic/omics

response to gradients past & future

locality data points across forests

17 labs, 8 institutions

Hav. → data stored → students

types → genomic, sanger, georef, env data, phys
data

1	12
1	12
2	24
3	36
4	48
5	60
6	72
7	84
8	96
9	108
10	120
11	132
12	144

Jun-Yung-Lim
Andy

Hawaii dimensions in biodiversity (NSF)

DFG

Carnegie Institute

macroecology + evolution

understand how communities change over time

Hawaiian chrono sequence

to compare community dynamics vs. age



older

○

microevolution

too

younger

focus → arthropod community (time) sampling

quantitative sampling

lots of material!!

field crew 15 people

Data → sites - events - specimens

relative data

sort down to size

↳ Shiny app

Page 141: 2017-12-12. paper notes- Bulla et al. 2017 on Marine Biorhythms and van der Veen et al. 2017 on flexible clock systems

refs:

- Research article: Marine biorhythms: bridging chronobiology and ecology
Martin Bulla, Thomas Oudman, Allert I. Bijleveld, Theunis Piersma, Charalambos P. Kyriacou
Phil. Trans. R. Soc. B 2017 372 20160253; DOI:
10.1098/rstb.2016.0253. Published 9 October 2017

Rhythms on earth driven by rotations of the Earth on its axis (24 hr), movement of moon + rotation of sun exerting different gravitational pull on the oceans (tidal/lunar cycles- 15 dys), and tilted movement of earth around sun (seasonal).

3 goals:

1. highlight scarcity of data on intertidal "higher" vertebrates in the context of biological rhythms
2. talk about new studies on circadian clocks genes in intertidal behavior/phys of arthropods and worms
3. Discuss value of collaborations between chronobio and ecologists

Shorebirds are a good example to study tidal rhythms. Their period should be reflected by tidal period. So they're showcasing shorebirds and they're using novel automated tracking tech to describe foraging rhythms of red knots at Baie d'Arguin, and their wintering grounds on coastal Mauritanian. In Mauritanian, it is an environment with tidal

rhythms and have diel fluctuations in light intensity.

Natural history: red knots are long distance migratory shorebirds that breed in the high arctic, live in coastal intertidal environments for res of year. They eat hard shelled moluscs.

INdividual rhytms are unknown

At high tides, birds generally close to the roost and as the tide retreats, birds moved away from it. This reflects foraging strategy. Low tide- feeds, high tide- stay at roost.

To see the genetic componenet of this trait, we will need a common garden and free run birds to quantify their endogenous clock.

on page 3, they highlight a potential pitfall of estimating biological rhythms with a periodogram because this analysis picks up the dominant period.

Molecular studies of tidal rhythms

other variables might be important such as species interactions that can alter an entrained biological rhythm.

four orgs that have been used to study tidal rhythms

1. *Eurydice pulchra* - marine isopod
2. *Apteronomobius asahinai* - mangrove cricket
3. *Clunio marinus* - marine midge
4. *Platynereis dumerilii* - bristle worm

marine isopod lives on euro coasts, as tide comes in it swims out of sandy burrows and forages. as the tide goes out, it goes back into the sand so it isnt dragged out to thesea. IT has a 12.4 hr rhythm , set by vibrations and influenced by temp. KO of period gene has no effect on circatidal swimming periodicity of 12.4 hrs. pharmacologically inhibiting

Casein Kinase 1[Math Processing Error] lengthened tidal swimming and circadian chromatophore cycle. CK1[Math Processing Error] phosphorylates period. (look up interaction). Altogether this suggests, unique and common pathways for tidal and circadian rhythms.

marine midge - during full and new moon, millions of males and females emerge from sea in low tide (more habitat exposure), where they develop from eggs to pupae. adults only live for a few hours, so they need to synchronize with themselves and in low tide. Suggest 2 clocks: circalunar and circadian. Think they did some genomics of some sort and found that calcium/calmodulin-dependent kinase II.1 (CaMKII) associated with circadian adaptations. For the lunar related phenotype, and found a QTL it looks like. But in this region, they didn't find any clock genes.

Bristle worm - spawns monthly rhythm - number of worms are sexual peak at new moon and troughs at full moon. It looks like they are entrained by moonlight ; recapitulated in the lab. The strength of rhythm is modulated by phases of the moon, suggests cross talk between 2 oscillators. Treatment of CK1 kinase inhibitor (used in Eurydice) impaired circadian movement/growth, but not circalunar maturation. Molecularly, these things can be independent.

Next big thing: Find tidal/lunar genes. Bridging the gap will simply involve longer time series data.

- Review article: Flexible clock systems: adjusting the temporal programme
Daan R. van der Veen, Sjaak J. Riede, Paul D. Heideman, Michaela Hau, Vincent van der Vinne, Roelof A. Hut
Phil. Trans. R. Soc. B 2017 372 20160254; DOI:

in mammalian systems, the suprachiasmatic nucleus of hypothalamus coordinates biological rhythms. It can be entrained to light and dark cycle. All non-scN clocks are known as PERIPHERAL TIMING SYSTEM and every cell in the body can express a circadian clock. So the peripheral timing system is complex. SCN may act as a conductor to orchestrate peripheral clocks

There are many external inputs from the environment and occur simultaneously in nature.

Temporal niche faces two challenges: how to align circ rhythms of multiple tissues with each other and what is the most beneficial internal alignment pattern for any given environmental temporal niche?

example:

1. liver entrained by food
2. SCN entrained by light/dark

Glucocorticoids are key zeitgebers for peripheral tissues. It can interfere with zeitgebers such as food availability.

This paper is like a big laundry list..wtf.

Page 142: 2017-12-13. Prepping meeting with Dan, 9AM , 2017-12-14

1. Discuss analyses
 - scaled to periodogram, cwt, and dwt to ~ 22 individuals
 - try more bins

2. Paper discussions

3.

Dan's suggestion for circadian rhythm analysis;

- What do other people do to decide which dominant peaks in periodogram to choose from?
- particular in Fourier transform approach; somebody develop stats? what is heuristic for picking out circadian rhythm
- What is going on with continuous wavelet? Binning at lower time give more fine estimates?
- Focus on Fourier transform data.
- look up some general methods papers

paper discussion

- Read the base papers in the review for marine biorhythms. Do the midge papers. Charalambos P. Kyriacou papers
-

metaanalysis ideas given what we've read

- compare different rhythms, circ more or less flexible than other rhythms (short, longer)
- are different components of biological rhythms heritable
- what env factors best explain bio rhythms
 - light
 - temp
 - food
- What's the potential for evolution ?

Organize evolution and biological rhythms meeting (tell Dan about time line)

Page 143: 2017-12-17. Douse 2013; Maximum entropy spectral analysis (MESA) paper

Dowse, H. B. (n.d.). Maximum entropy spectral analysis for circadian rhythms: theory, history and practice. Retrieved from <https://pdfs.semanticscholar.org/d564/6a61fad1e03e61681fc1ec3694c4e36b7456.pdf>

Developed by John Parker Burg in 1960s to address shortcomings of fourier analysis. Go over fourier analysis and how MESA covers it's weaknesses.

Sampling frequency must be no less than twice the frequency of the cyclic process of interest

2 important things to consider:

1. resolution - ability to separate two frequencies as being distinct.
Very similar to optical resolution where two objects in an image want to be separated. Resolution is limited by the number of cycles in the data set, or in optics, diameter of the lens.
2. Discern periodic signal in the presence of noise , aka, sensitivity

Autocovariance and autocorrelation functions

One way to test if the data is rhythmic is to statistically test it with an autocorrelation analysis.

Fourier analysis

Any function can show certain minimal properties called "Dirichlet Conditions" and be approximated by a harmonic series of orthogonal sine and cosine terms.

[Math Processing Error]

$$f(t) \cong a_0/2 + a_1 \sin t + a_2 \sin 2t + \dots + b_1 \cos t + b_2 \cos 2t$$

if the function consists of an ordered set of values $x(t)$, then the "power" in the series is the ensemble average of the squared values. If the mean is 0, this is the variance.

The Fourier transform is an extension of a fit of the Fourier series and has the property that the coefficients approximate the power spectrum. Power at each frequency can be computed for a continuous series as:

[Math Processing Error]

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{(-i\omega t)} dt$$

exponential function consolidates sine and cosine terms. $F(\omega)$ is the spectrum of the function, with [Math Processing Error] as the angular velocity (aka $2\pi f$), where f is the frequency. This is also known as the **periodogram** of the function.

If the spectrum is calculated directly from data sampled at intervals, it is known as the **Discrete Fourier Transform**. This is rare and Fourier spectra are produced from either autocovariance or autocorrelation. One reason to use autocovariance is that the output is equivalent to partitioning the variance in the signal by frequency and the area under the curve is the power.

Compromises in Fourier analysis

~1/3 of the data are used to compute the correlation coefficients because autocov or autocorr functions lose power with each power of points lost. This adversely affects the resolution in the spectrum.

Solution - pad out the rest of the function with 0's. Fast fourier transform constrains input series to [Math Processing Error] points.

Problem- adding 0s adds side lobes

Maximum entropy spectral analysis

gives highest possible resolution while eliminating side lobes. "linchpin" of technique is the stochastic modelling. System moves forward in term as a function of previous values + noise. example is a Markov process:

$$X_t = aX_{t-1} + Z_t$$

- t is time
- a is the coefficient derived from the data
- Zt is white noise

This process can be extended by going backwards earlier in time , with each weighted by a coefficient derived from the known observed data.

$$X_t = aX_{t-1} + bX_{t-2} + \dots + cX_{t-p} + Z_t$$

a, b,c are the model's coefficients and p is the order of the filter

The model can be used to predict future values based on what is known from the past. Entropy = ignorance in information theory, so calculating values that max ignorance means the values are the most honest. The spectrum is calculated from the coefficients as follows:

$$S(\omega) = P / \left| 1 - \sum_{k=1}^p a_k e^{-i\omega k} \right|^2$$

- Again, $S(w)$ is the spectral power as a function of angular velocity
- P is the power passed by the power error filter
- a_k is the set of PEF coefficients

Number of coefficients is important, generally $N/4$ or $N/3$,

Data conditioning

section on how to deal with data prior to analysis

1. detrending step which fits a line by regression and subtracts it, removing linear trend and mean. DC component can obscure rhythmic one ?
2. Removal of non-linear component through filters
 - look into Butterworth filter

examples of analysis at work:

1. fruit fly cryptochrome gene - Krishnan et al. 2001 Nature; Plautz et al. 1997 J Biol Rhythms
2. lunar rhythmicity in marine systems (Hamming 1983) --odd citation
3. Ultradian and circ rhythms in premature infants -- Tenriero et al. 1991
4. Electroencephalography and ultradian period in rats to study models of sleep-wake dynamics - Stephenson et al. 2012 J Biol Rhythms
5. ultradian in mice loco activity -- Dowse et al. 2010 J Exp Biol
6. Vertical migration in antarctic krill - Gaten et al. 2008 J Genetics
7. cardiac pacemaker in fly heart - Dowse et al. 1995 J Neurogenet; Bodmer et al. 2005
8. Mating in drosophila control under period gene -- Alt et al. 1998 Anim Behav

Page 144: 2017-12-17. Time series analysis demo

Following this demo: <http://www.stat.pitt.edu/stoffer/tsa4/tsaEZ.pdf>

Common approaches to time series:

1. time domain approach
 2. frequency domain approach
-

Page 145: 2017-12-18. Costantini et al. 2010; *Ecology Letters*; Ecological processes in a hormetic framework

Paradox: sublethal stressors can confer resistance/tolerance to future stressors

This occurs when stressors are at low levels, hence sublethal. What the stressor does is actually "prime" physiological processes such that they are prepared later in life. Termed "hormesis"

So unclear what hormesis is by the way they describe it...

Problem they outline: even though hormesis happens, its unclear how other use the word and how you can test it.

Most work done on toxicology and human health. Check Calabrese & Blain 2005

Argument: provides conceptual link between prevailing environmental circumstances and organism function. So this is good for eco and evo biologists

They want to frame hormesis in the context of fitness consequences. They say Darwinian fitness, but fitness is just fine guys.

What is hormesis?

mid 1880s German pharmacologist Hugo Shulz found that small doses of disinfectants stimulated yeast growth, but was lethal at high doses. People thought this was cool, but homeopathy ruined everything. These authors like to site Calabrese a lot. Anyway, non-linear dose responses have been established between organisms and chemicals/env stressors.

Hormesis was coined by Chester Southam and John Ehrlich 1943. It is used in two different ways.

1. response to env stressor varies with level of exposure
 1. linear no-threshold model - level of exposure to stressor always negatively affects fitness
 2. threshold model - fitness is maintained in the face of a stressor up until a certain point where higher doses/exposure negatively affects fitness
 3. hormetic model - Fitness is increased at lower doses/exposure but is negatively affected at higher doses/exposure.

good example in fruit flies(Gibbs et al. 2003)

2. "priming" or "conditioning" effect where exposure to a low level of stressor results in the organism being better to cope with exposure to higher levels in the future. AKA Conditioning hormesis

Underlying mechanisms

- changes in genes in pathways that respond to stressors
 - ie hsp, antioxidant enzymes, hormones
 - Early life conditioning may have the strongest hormetic effect
 - may be due to epigenetic modifications

Environmental Agents

Nutrition

- Non-essential dietary compounds can generate hormetic responses
- Dietary restriction
 - hsp upreg in liver of rats ; increasing cryoprotective molecules and reducing levels of oxidative stress
 - mito activity can change
- they argue that a group that was starved earlier in life might fare better (fitness) as adults compared to fully fed group earlier in life (as adults) .

temperature

physical activity

immune system

Hormesis and phenotypic plasticity

Hormesis might increase the stress tolerance and response of an organism ; conditioning hormesis might translate into increase in phenotypic plasticity

hormesis and fitness

They argue there has to be a link. Because there are energetic trade-offs among life history traits, not all traits can exhibit a hormetic response at the same time. ex: fitness trade-off between longevity and reproduction in response to heat stress (Maynard Smith 1958; Le Bourg et al. 200; Le Bourg 2005; Sorensen et al. 2007,2008)

Also should consider some costs associated with mounting a hormetic response. Not only measure ATP_o but also other nutrients. If it is too costly, then organisms won't mount a hormetic response.

Perspective and conclusions

Low hanging fruit: compare hormetic response under early and late life sublethal stress exposure

Need to understand hormesis in the context of multiple stressors

Thoughts:

Overall, interesting paper. I think how they described hormesis was useful in that there were two cases, one being an overall fitness increase at low doses and the other focusing on the timing of exposure that confers fitness advantage. I'm more used to thinking about hormesis in terms of the 2nd case: prior exposure conferring resistance/tolerance to future stressors. They could have done a better job of putting everything in an evolutionary framework. Like...why would this matter in a natural setting.

- They could have laid out some ecology, for example, under which conditions do you expect more or less hormetic response? For temperature, you'd expect more hormesis for cold tolerance in northern, more variable environments. This relates to my range limits ms, and could use this as a citation.
 - ants are long lived, so they experience the environmental

variation throughout(within) their life time, giving the possibility for hormesis

- For example, what is the adaptive potential for a hormetic response? They could have added some quantitative genetics in there.
- It's also fun to think about adding more layers given this framework. For example, basically exploring plasticity of plasticity. Measure fitness along an env gradient(temp for example) and overlay that with another (desiccation for example). And you can see how beta shifts with desiccation.

Page 146: 2017-12-18. Trying spectral analysis with spectrum() function

scaling up spectrum() function

```
1 sa.an<-function(ts=counts15$counts15){  
2   sa1<-  
3   spectrum(ts,method=c("pgram","ar"),plot=FALSE,demean=TRUE,  
4             detrend=TRUE,tape=.2)  
5   spx<-sa1$freq  
6   spy<-2*sa1$spec  
7   pw<-data.frame(spx,spy)  
8   cc<-head(pw[order(pw$spy,decreasing=TRUE),],4)  
9   return(1/cc[,1]/4)  ## hours  
10  #return(1/cc[,1]/4/24)  ## days  
11 }  
12 sa.an()
```

Units are in hours

```
1 # 270.00000 26.12903 810.00000 23.14286  
2
```

Operate function for each unique ID and experiment

```
1 ### for each unique ID and experiment  
2 test10<-ddply(nall.data15_3,.  
  (uniqueID,experiment),function(sub) sa.an(sub$counts))  
3 test10
```

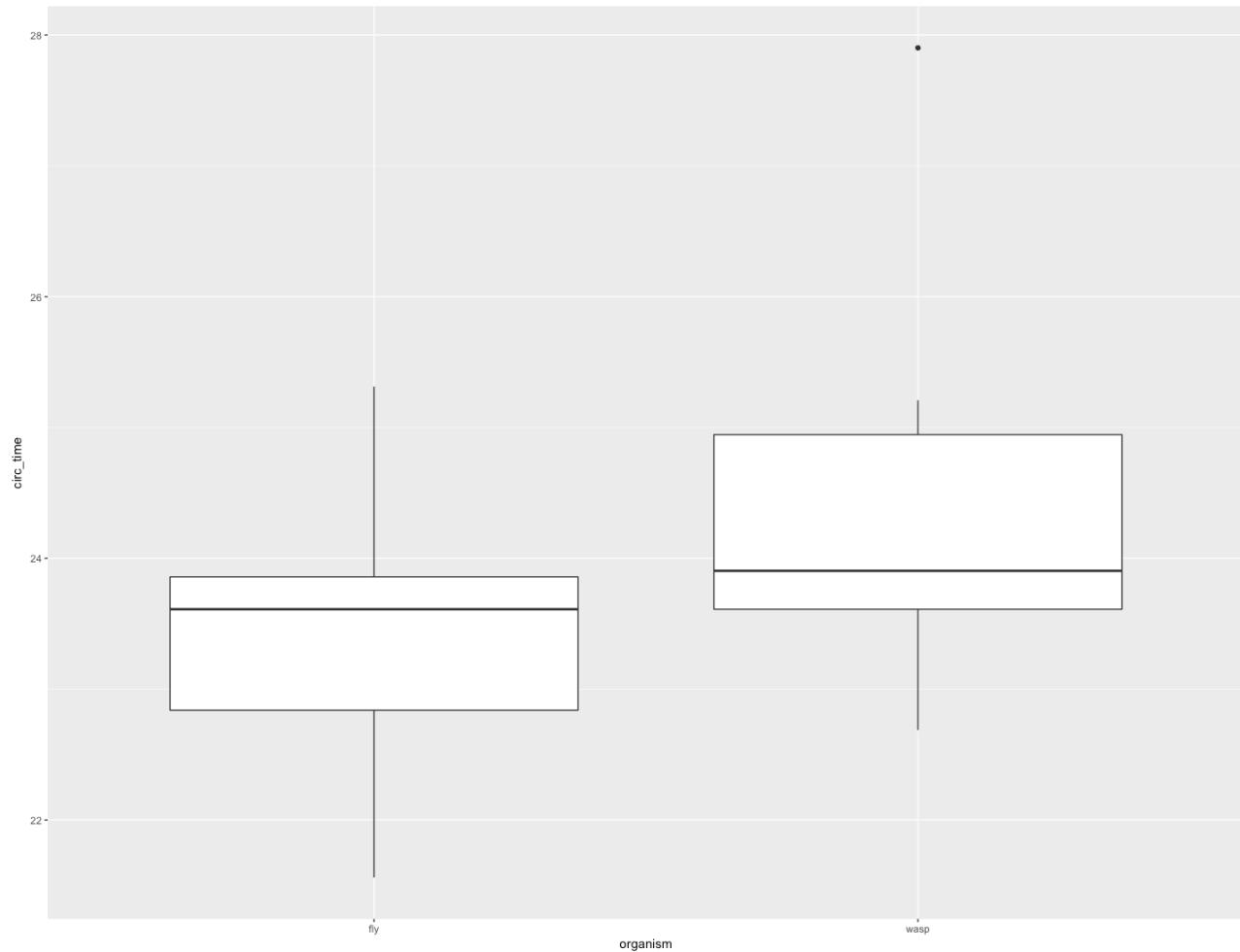
uniqueID	experiment	V1	V2	V3	V4
a10o49	Entrainment	24.000000	12.000000	240.000000	30.000000
a10o49	Free-run	22.781250	20.710227	21.191861	10.013736
a10o73	Entrainment	24.000000	21.600000	19.636364	12.000000
a10o73	Free-run	23.529412	22.222222	12.121212	8.000000
a10w12	Entrainment	24.000000	12.000000	8.000000	21.818182
a10w12	Free-run	23.980263	23.365385	11.682692	24.628378
a10w15	Entrainment	22.500000	25.714286	12.000000	1.764706
a10w15	Free-run	24.107143	22.500000	3.214286	12.053571
a11w26	Entrainment	22.500000	25.714286	12.000000	8.181818
a11w26	Free-run	22.500000	8.035714	21.093750	7.670454
a12b43	Entrainment	25.000000	12.500000	22.222222	8.333333
a12b43	Free-run	23.684210	23.076923	21.951220	22.500000
a12b6	Entrainment	22.500000	25.312500	11.911765	20.250000
a12b6	Free-run	195.312500	23.674242	260.416667	26.041667
a12w40	Entrainment	25.000000	22.222222	11.764706	8.000000
a12w40	Free-run	900.000000	450.000000	23.684210	300.000000
a13o11	Entrainment	25.000000	11.764706	22.222222	2.000000
a13o11	Free-run	23.225807	720.000000	24.000000	22.500000
a13o28	Entrainment	25.000000	8.000000	22.222222	11.764706
a13o28	Free-run	400.000000	21.052632	25.000000	23.529412
a2b23	Entrainment	22.500000	12.500000	13.235294	225.000000

a2b23	Free-run	20.833333	23.809524	22.727273	10.869565
a2b26	Entrainment	6.081081	18.750000	4.326923	22.500000
a2b26	Free-run	24.107143	22.010870	25.312500	11.005435
a3o51	Entrainment	22.222222	11.764706	25.000000	4.347826
a3o51	Free-run	23.809524	11.904762	4.950495	12.195122
a4o15	Entrainment	3.432203	25.312500	5.955882	8.437500
a4o15	Free-run	23.684210	4.285714	8.035714	11.842105
a4w66	Entrainment	24.300000	21.441177	182.250000	28.038462
a4w66	Free-run	96.000000	76.800000	128.000000	64.000000
a5b25	Entrainment	25.000000	11.842105	22.500000	8.035714
a5b25	Free-run	21.739130	23.809524	12.195122	250.000000
a5w63	Entrainment	202.500000	22.500000	25.312500	11.911765
a5w63	Free-run	21.428571	22.500000	25.000000	10.975610
a5w73	Entrainment	22.500000	25.312500	20.250000	101.250000
a5w73	Free-run	23.684210	22.500000	450.000000	112.500000
h2b25	Entrainment	25.000000	22.222222	20.000000	8.000000
h2b25	Free-run	600.000000	200.000000	24.000000	75.000000
h4o4	Entrainment	25.000000	22.222222	20.000000	100.000000
h4o4	Free-run	208.333333	29.761905	26.041667	312.500000
h4w10	Entrainment	24.000000	12.000000	8.000000	12.705882
h4w10	Free-run	4.285714	25.714286	22.500000	20.000000
h5o22	Entrainment	25.000000	22.500000	8.035714	28.125000
h5o22	Free-run	250.000000	23.809524	2.793296	166.666667

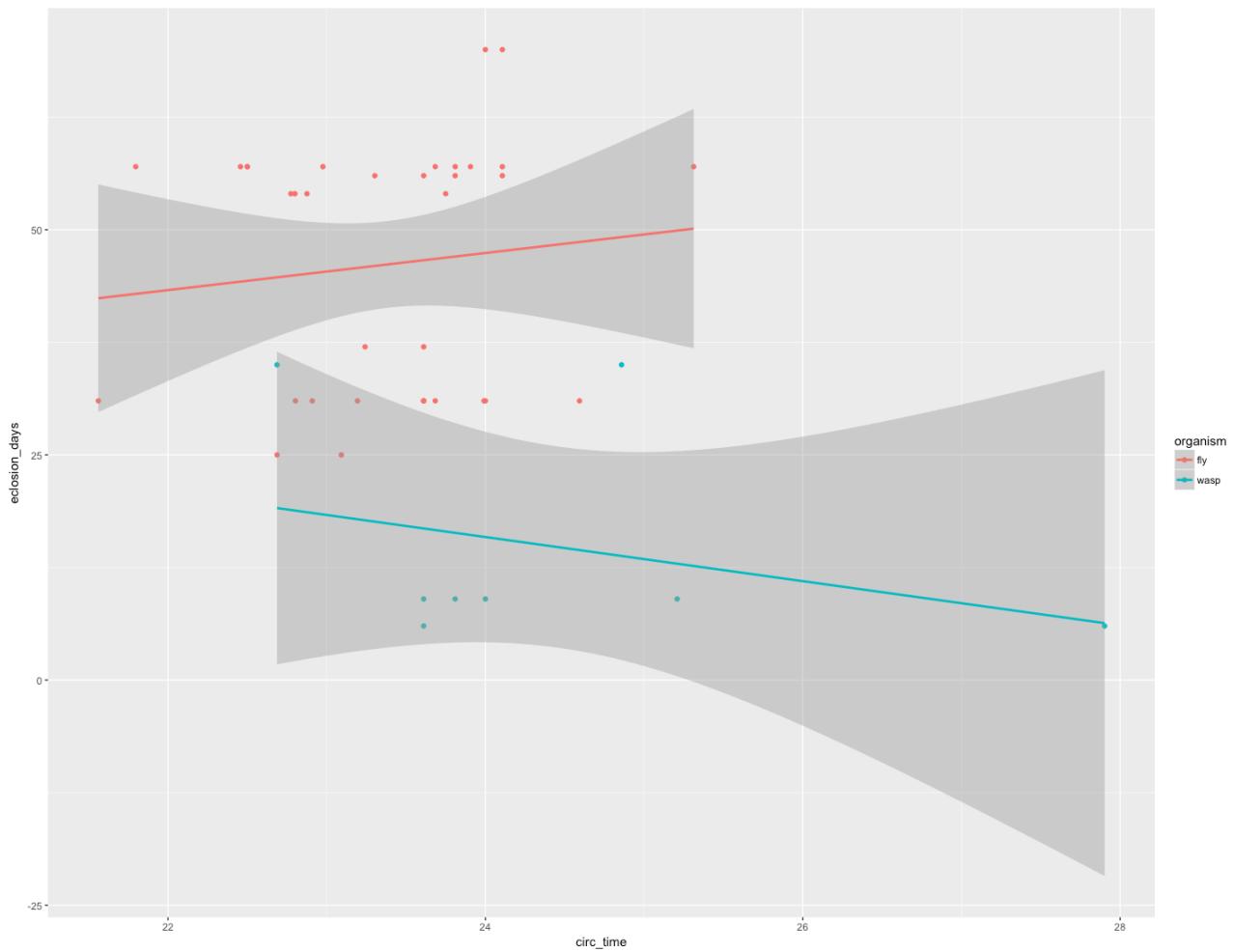
Values are everywhere, trying to take the average circ timing/period by taking average of values between 20 and 30 hours in data set.

```
1 ##taking circadian value
2
3 test10.merg$circ_time<-
4 apply(test10.merg[,3:6],1,function(x){mean(subset(x,x<30
& x > 20))})
```

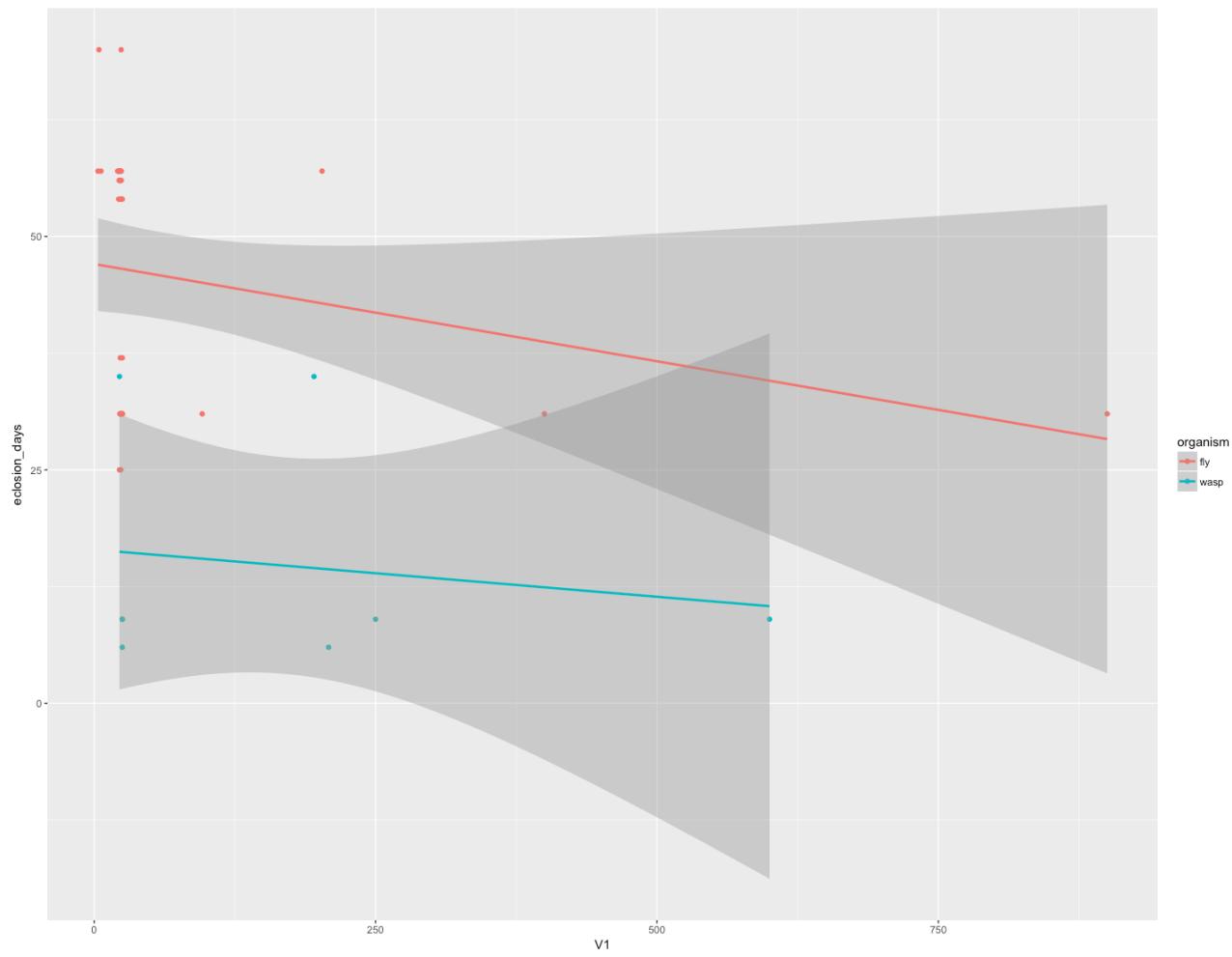
diff between fly and wasp



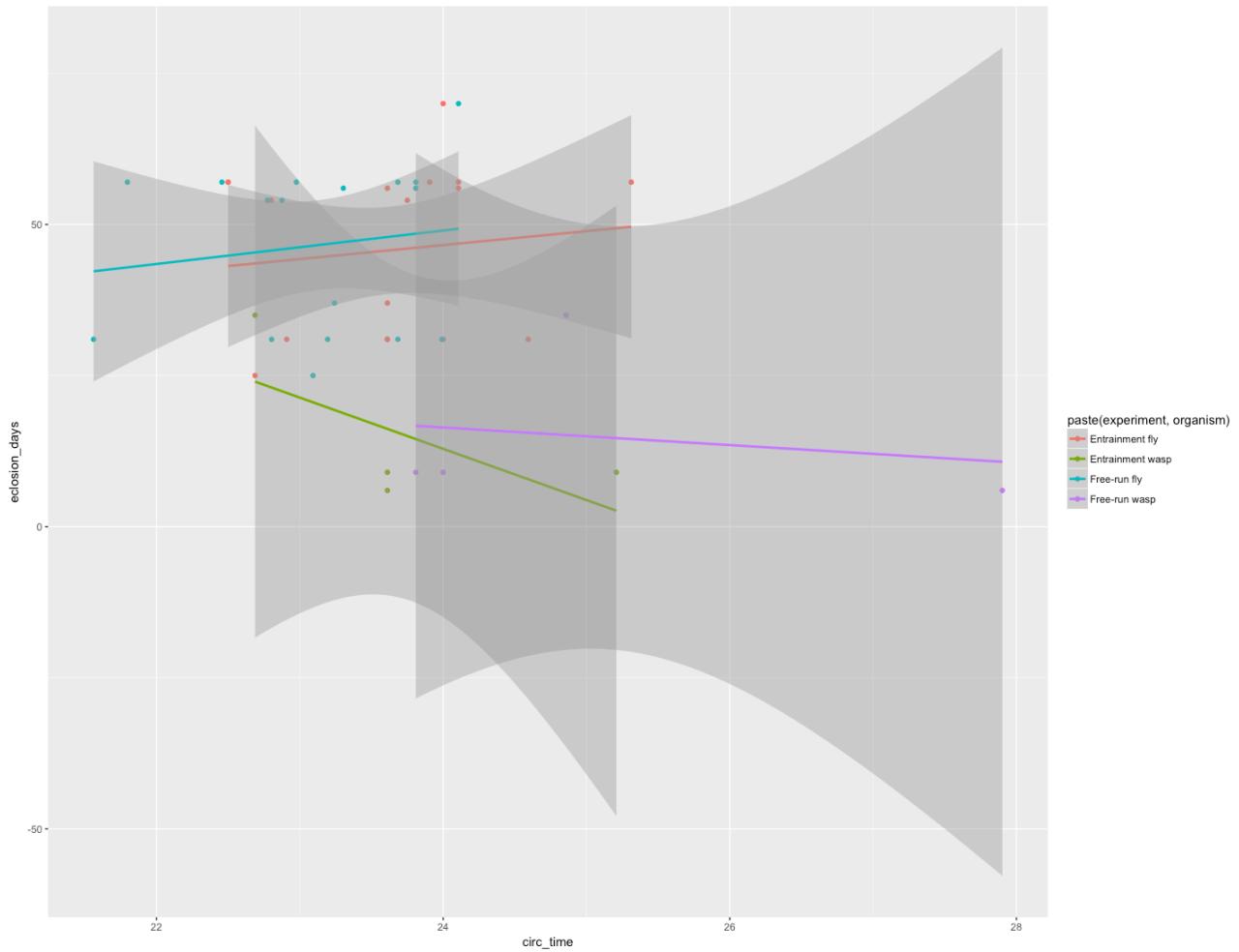
no relationship between eclosion days vs circ time



compare previous fig with just comparing the dominant period



overlays and adding entrainment vs free run



doesn't look like much of a pattern

stats: testing interaction between organism and circ time on eclosion days

not a balanced design, but just seeing how the stats come out

```

1 summary(aov(eclosion_days~organism*circ_time,data=test10
.merg))
2
3   Df Sum Sq Mean Sq F value    Pr(>F)
4
5 organism          1   6417   6417  33.587 9.96e-07
6 ***

7 circ_time         1      0      0  0.001   0.972
8
9 organism:circ_time 1   190    190  0.994   0.325
10
11 Residuals        39   7452   191
12 ---
13 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
14 ' ' 1
15 1 observation deleted due to missingness

```

ok what about diff in circ time between organisms? yes

```

1 summary(aov(circ_time~organism,data=test10.merg))
2
3   Df Sum Sq Mean Sq F value    Pr(>F)
4 organism       1   7.57   7.572  8.286 0.00632 **
5 Residuals     41  37.47   0.914
6 ---
7 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
8 ' ' 1
9 1 observation deleted due to missingness

```

no differences between host races

```
1 tt<-subset(test10.merg,organism!="wasp")
2 summary(aov(circ_time~Host,data=tt))
3   Df Sum Sq Mean Sq F value Pr(>F)
4 Host       1  0.006  0.0064   0.011  0.918
5 Residuals  33 19.677  0.5963
6 1 observation deleted due to missingness
```

Page 147: 2017-12-18. Table of contents of physical notebook: NB#001; Apple Maggots; year 2017; grant # 108398

COMPOSITION BOOK

Lab : Hahn lab NB #001

Project: Apple maggots

Year 2017

Grant # 168398

Wide Rule

100 Sheets • 200 pages

9.75 in x 7.5 in/24.7 cm x 19.0 cm



TOP|FLIGHT

Wide Ruling

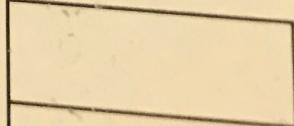


Table of contents

Date	Title	Page
20170830	Summary of Project so far	1
20170830	Respirometry w/ Chao	
20170830	Purging for test icsp	2-4
20170830	Apple maggot sampling day 7	5
20170831	Housekeeping notes, order	6
20170831	Methods for handling apple maggots	7
20170831	Apple maggot sampling day 8	8
20170901	Test respirometry handling, purging, mass	10-12
20170902	Apple maggot sampling day 9	13
20170903	Apple maggot sampling day 10	14
20170904	" " " day 11	15
20170904	" " " day 12	16
20170905	Day 2 cohort weighing & purge day 10 dev ~ 2017-08-25	17
20170905	Apple maggot sampling day 13	18
20170905	Day 3 cohort weighing & purge day 10 dev	19
20170906	Day 4 cohort weighing & purging	20
20170907	Respirometry notes Day 4 cohort	21
20170907	Apple Maggot Sampling Day 15	22
20170907	Day 5 cohort weigh & purge	23
20170908	Apple Maggot Sampling Day 16	24
20170909	Apple Maggot Sampling Day 17	25
20170910	Apple Maggot Sampling Day 18	26
20170911	Apple Maggot Sampling Day 19	27

Table of Contents

Date	Title	Page
2017 09 02	Apple Maggot Sampling Day 20	28
2017 09 13	Respirometry Day 11 cohort	29
2017 09 13	Day 11 cohort weighing & purge	30
2017 09 13	Day 3 cohort Data sheets	31
2017 09 13	Day 3 cohort mass data	32
2017 09 13	Respirometry notes	33
2017 09 13	Day 2 cohort Data sheet	34
2017 09 13	Day 4 cohort organizational sheets	35
2017 09 13	Day 4 cohort data sheet	36
2017 09 13	Day 5 cohort organizational sheets	38
2017 09 13	splitting 2017-08-29 cohort	39
2017 09 13	Day 5 cohort data sheets	40
2017 09 14	Day 11 cohort respirometry	41
2017 09 14	Day 7 cohort treatment	42
2017 09 15	Day 13 cohort weigh & purge	43
2017 09 15	Day 12 cohort respirometry	44
2017 09 16	Day 13 cohort respirometry	45
2017 09 16	Day 14 cohort weigh & purge	46
2017 09 16	splitting skipped cohort D9 into treatments	47
2017 09 17	Day 10 cohort respirometry on day 15	48
2017 09 17	Day 14 cohort respirometry day 11	49
2017 09 17	Day 14 cohort weighing & purge day 14	50
2017 09 17	Day 11 cohort weigh & purge day 14	51
2017 09 17	Day 10 cohort data sheet splitting into groups 51-51	52
2017 09 17	Day 10 cohort mass data	

Table of contents

	Title	Page
31	Day 11 cohort respirometry day 15	53
32	Day 16 cohort merge + purge day 10	54
33	Day 11 cohort data sheet splitting into graphs	55
4-135	Day 11 cohort data sheet mass D10 & 14	56
3-25	Day 12 cohort respirometry day 15	57
32	Day 12 cohort data sheet splitting into graphs	58
9	Day 18 cohort merge & purge day 10	59
-14	Day 18 cohort data sheet mass D10 & 14	60
2	Day 18 & 13 cohort data sheet splitting into graphs	61
3	Day 13 cohort data sheet mass D10 & 14	62
4	Day 14 cohort respirometry d15	63
5	Day 14 cohort data sheet mass D10 & 14	64
6	Day 14 cohort data sheet ^{splitting} mass into graphs	65
7	Day 19 cohort respirometry day 11	66
8	8Kipped Day 15 cohort split into 2 dates	67
9	Day 1 Plantherne collection	68
10	Day 20 cohort respirometry day 11	69
11	Day 16 cohort respirometry day 15	70
12	Day 18 cohort respirometry day 15	71
53	Day 18 cohort respirometry day 15	72
13	Checking eclogites for RT plates	73
14	Day 16 cohort data sheet mass & treatment	74
156	Day 16 cohort data sheet mass & treatment	75
16	Day 18 cohort data sheet mass & treatment	76
59	Day 19 cohort respirometry day 15	77
17	Day 19 cohort respirometry day 15	78
18	Test samples of Rhago for kinetics	79
19	Test samples of Rhago for kinetics	80

Date	Title	Page
2017/09/21	samples that eclosed	79
2017/09/28	Trikinetics set up + eclosion	80
2017/09/28	Day 20 cohort mass + treatment check	81
2017/09/29	Day 19 cohort mass + treatment check	82
2017/09/29	Checking eclosion + putting adults in Trikinetics	83
2017/09/29	Skipped cohort eclosion	84
2017/09/30	eclosion times	85
2017/10/01	eclosion times	86
2017/10/02	collecting pupae + checking eclosion	87
2017/10/03	Checking eclosions	88
2017/10/04	Checking eclosions	89
2017/10/04	Skipped cohort eclosion	90
2017/10/05	Skipped cohort eclosion	91
2017/10/05	Checking eclosions	92
2017/10/06	Day 4 Haw DLL respirometry	93
2017/10/06	Checking eclosions	94
2017/10/07	Checking eclosions	95
2017/10/07	Checking eclosions	96
2017/10/07	Day 1 cohort Haw split into 3 treatments	97
2017/10/08	Checking for eclosions	98
2017/10/09	Haw cohort day 2 data checks	99
2017/10/09	Checking for eclosions	100
2017/10/10	Skipped cohort eclosions	101
2017/10/10	Haw day 3 data check	102
2017/10/10	Haw day 4 data checks	103

Table of contents

Date	Title	Page
2017/10/10	Checking for eclosions	104
2017/10/11	Checking for eclosions	105-106
2017/10/12	Haw cohort 5 data sheets	107
2017/10/12	Checking eclosions	108
2017/10/12	Haw cohort 6 data sheets	109
2017/10/12	Free running files	110
2017/10/13	Checking eclosions	111
2017/10/13	entrainment to freenning	112
2017/10/14	Checking eclosions	113
2017/10/14	Haw cohort 7 data sheets	114
2017/10/14	Haw cohort 8 data sheets	115
2017/10/15	Checking eclosions	116
2017/10/15	Haw cohort 9 data sheets	117
2017/10/15	Freenning cohort transfers	118
2017/10/16	Freenning cohort transfer	119
2017/10/16	Freenning cohort transfer	120
2017/10/16	Putting eclosed skipped cohorts into freezer	121
2017/10/16	Haw cohort 10 data sheets	122
2017/10/16	Checking eclosions Putting skipped cohorts into freezer	123
2017/10/17	Checking eclosions	124
2017/10/16	Checking eclosions	125
2017/10/17	Checking eclosions	126
2017/10/17	Haw cohort 11 data sheets	127
2017/10/17	Tokinetics data sheet	128
2017/10/18	Haw cohort 12 data sheets	129
		130

Table of contents

date	title	page
2017/10/18	checking eclosions	184
2017/10/18	Free nn deaths	181
2017/10/19	checking eclosions	183
2017/10/20	checking eclosions	184-185
2017/10/21	checking eclosions	186-187
2017/10/22	Labs notes	188
2017/10/22	checking eclosions	189
2017/10/23	checking eclosions	190-191
2017/10/23	checking free nn deaths	192
2017/10/24	free running samples dates listed	193
2017/10/24	cheecking eclosions	194
2017/10/25	checking eclosions	195
2017/10/25	Free nn experiment deaths	196
2017/10/26	checking eclosions	197
2017/10/26	changed water	198
2017/10/27	changed water	199
2017/10/27	cheecking eclosions	200
2017/10/28	Free nn deaths	201
2017/10/28	checking eclosions	202-203
2017/10/29	checking eclosions	204
2017/10/30	checking eclosions	205-206
2017/10/30	Free nn deaths	207
2017/10/31	checking eclosions	208-209
2017/10/31	Free nn Deaths	210

Table of Contents

Date	Entry	Page
2017/11/01	checking enclosures for tracks 11/01	161
2017/11/01	change water 11/01	162
2017/11/02	checking enclosures 11/02	163
2017/11/02	free running samples 11/02	164
2017/11/02	free nn deaths 11/02	165
2017/11/03	checking enclosures 11/03	166-167
2017/11/04	checking enclosures 11/04	168-169
2017/11/05	checking enclosures 11/05	170
2017/11/06	checking enclosures 11/06	171
2017/11/06	free nn deaths 11/06	172
2017/11/07	checking enclosures 11/07	173
2017/11/07	free nn deaths 11/07	174
2017/11/08	checking enclosures 11/08	175
2017/11/09	checking enclosures 11/09	176
2017/11/09	free nn deaths 11/09	177
2017/11/10	checking enclosures 11/10	178
2017/11/11	tasks 11/11	179
2017/11/11	checking enclosures 11/11	180
2017/11/12	checking enclosures 11/12	181
2017/11/13	checking enclosures 11/13	182
2017/11/13	free nn deaths 11/13	183
2017/11/13	checking enclosures 11/13	184
2017/11/14	change water 11/14	185
2017/11/14	free nn deaths 11/14	186

Table of contents

Page 148: 2017-12-18. Chao leading lab meeting

2017-12-18

Class reading

In China, you need to always connect to humans
why study diapause, sex?

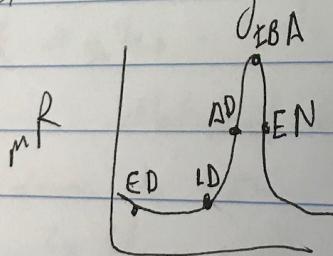
Metabolic cycles

→ focus

- 1) generate energy
- 2) replenish key metabolites
- 3) purge anaerobic products
- 4) boost stress defense
- 5) repair damage

how CO_2 produced? TCA releases CO_2

arousal \rightarrow 2-5 days



Days

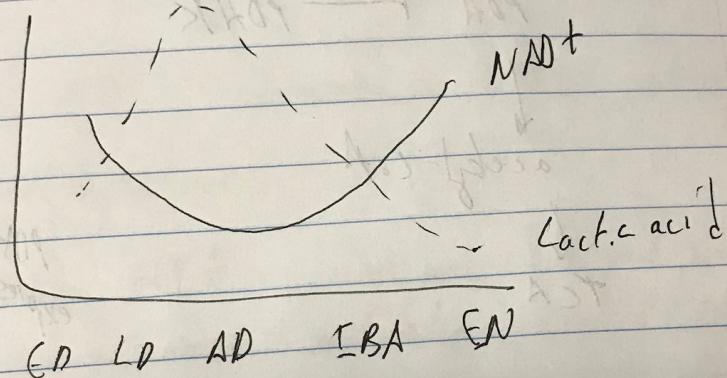
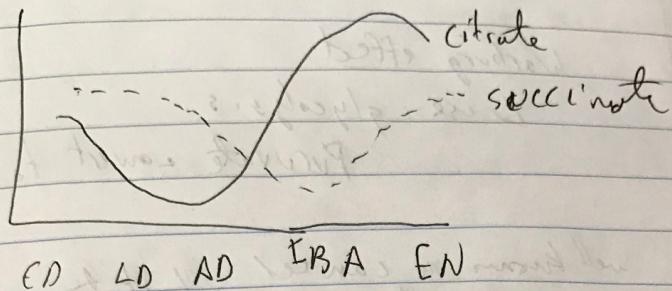
succinate \rightarrow fumarate

regulation step

sim. lat
to respond
to hypoxia

2017-12-18

(heat resists)



needs to know NAD^+/NADH ratio

focus on AMPK → kinase be of ATP resists

AMPK activates glycolysis to make more ATP

* Insect relies on glycolysis in normoxic
In field they burrow, so they're using making CO_2
w/ TCA

2017-12-18

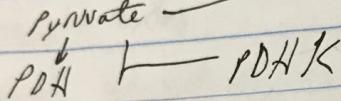
Chao lab meeting

Warburg effect

use glycolysis

Pyruvate convert to lactate

well known we can use lactate

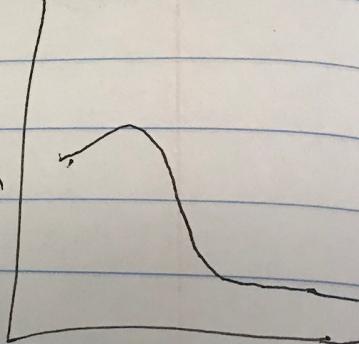


* when PDH is
phosph., it is
inhibited

↓
acetyl-CoA

↓
TCA

PDHK
expression



* people use PDHK inhibitors EP LD AD IBA EN

PDHK antagonizes TCA

* Inject PDHK inhibitor, caused immediate arousal

What regulates metabolic arousal & depression?

PDHK is an important target for TCA regulation

lactate

HIF

PDHK

TCA

inhibit HIF → increase PDHK activity
↓ w/injections

Inhibiting HIF, flies were aroused, but then caused constant metabolic arrest (CO_2)

HIF is involved in arousal

what regulates HIF?

ROS → HIF

NADH/FADH₂

TCA

tested whether ROS was important
ROS scavenger → get rid of ROS (NAC)

Add paraquat to induce ROS

Leyh asks about reductive carboxylation
get pyruvate to fumarate → generates acetyl/CO₂
under Hypoxia

indirectly

not anaerobic

ROS generated by hypoxia & hyper

Page 149: 2017-12-20. continuous wavelet analysis revisited: 6 min bins and exploring power vs period to id multiple dominant peaks

Exploring multiple dominant peaks

reworking function to include time bins that I can divide to calculate period

```
1 cwa<-function(series=counts15$counts15,N=4){  
2   wave.out <- morlet(y1 = series, p2 = 8, dj = 0.1,  
3   siglvl = 0.95)  
4   wave.out$period <- wave.out$period/N  
5   wave.avg <- data.frame(power = apply(wave.out$Power,  
6   2, mean), period = (wave.out$period))  
7   aa<-data.frame(findpeaks(wave.avg$power))  
8   line<-aa[order(aa,decreasing=TRUE),][1,2]  
9   return(wave.avg[line,][2])  
10 }  
11  
12 cwa()
```

estimating period for 6 min bins for each unique ID and experiment

```
1 test3.6<-ddply(nall.data06_3,.  
2   (uniqueID,experiment),function(sub)  
3   cwa(sub$counts,N=10))  
4 knitr::kable(test3.6)
```

uniqueID	experiment	period

a10o49	Entrainment	24.67491
a10o49	Free-run	21.48076
a10o73	Entrainment	23.02251
a10o73	Free-run	23.02251
a10w12	Entrainment	24.67491
a10w12	Free-run	23.02251
a10w15	Entrainment	24.67491
a10w15	Free-run	23.02251
a11w26	Entrainment	23.02251
a11w26	Free-run	NA
a12b43	Entrainment	24.67491
a12b43	Free-run	23.02251
a12b6	Entrainment	23.02251
a12b6	Free-run	10.74038
a12w40	Entrainment	24.67491
a12w40	Free-run	24.67491
a13o11	Entrainment	24.67491
a13o11	Free-run	23.02251
a13o28	Entrainment	24.67491
a13o28	Free-run	21.48076
a2b23	Entrainment	21.48076

a2b23	Free-run	21.48076
a2b26	Entrainment	NA
a2b26	Free-run	23.02251
a3o51	Entrainment	23.02251
a3o51	Free-run	NA
a4o15	Entrainment	NA
a4o15	Free-run	24.67491
a4w66	Entrainment	24.67491
a4w66	Free-run	13.22296
a5b25	Entrainment	24.67491
a5b25	Free-run	23.02251
a5w63	Entrainment	23.02251
a5w63	Free-run	23.02251
a5w73	Entrainment	23.02251
a5w73	Free-run	23.02251
h2b25	Entrainment	23.02251
h2b25	Free-run	23.02251
h4o4	Entrainment	24.67491
h4o4	Free-run	10.74038
h4w10	Entrainment	24.67491
h4w10	Free-run	23.02251

h5o22	Entrainment	24.67491
h5o22	Free-run	NA

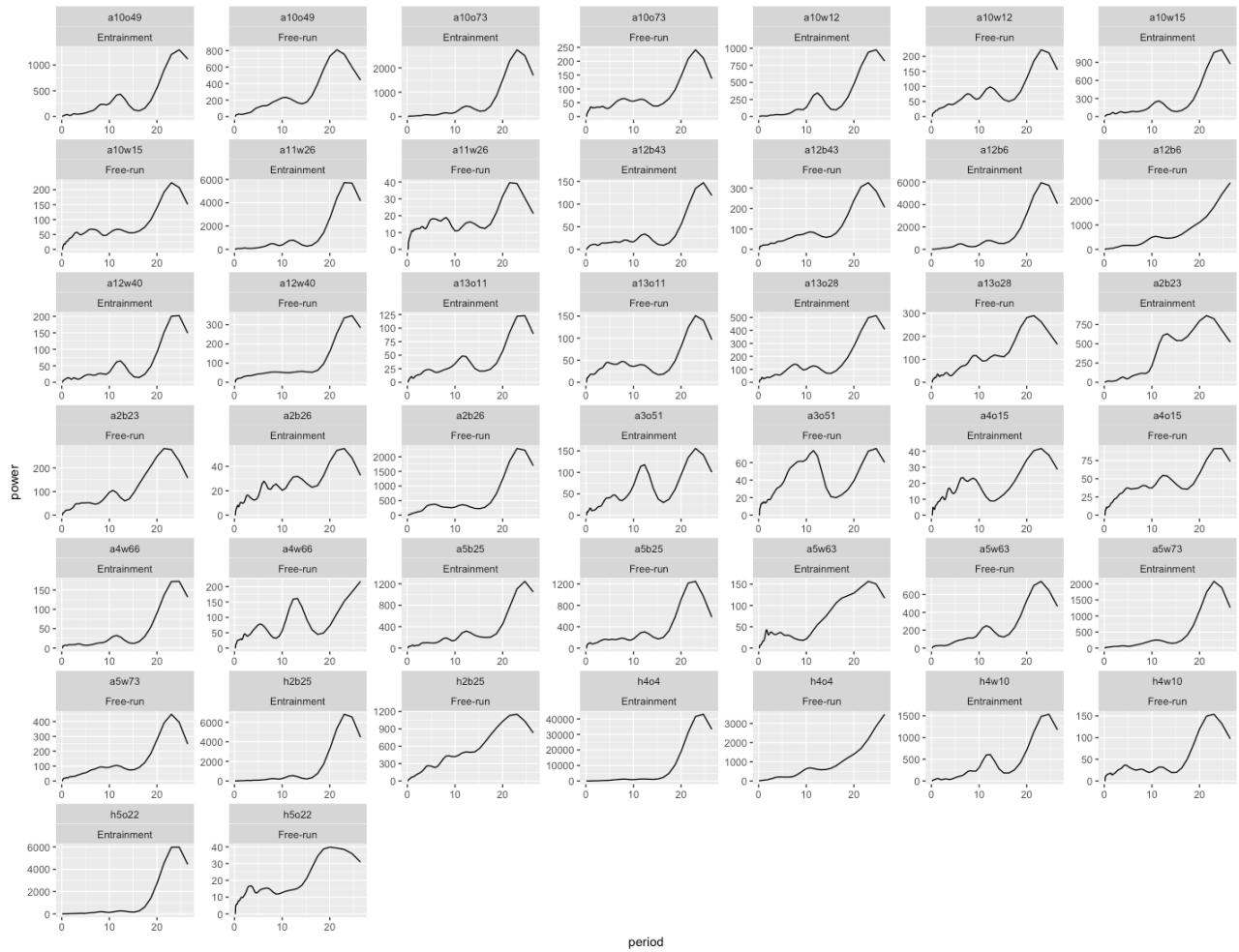
some estimates are NA, which is odd. Why...

I'll explore the power vs period spectrum..

function to grab power vs period and then estimating it for each unique ID and experiment

```

1 cwa.1<-function(series=counts15$counts15,N=4){
2   wave.out <- morlet(y1 = series, p2 = 8, dj = 0.1,
3   siglvl = 0.95)
4   wave.out$period <- wave.out$period/N
5   return(data.frame(power = apply(wave.out$Power, 2,
6   mean), period = (wave.out$period)))
7
8 tt3<-ddply(nall.data06_3,.
9 (uniqueID,experiment),function(sub)
10 cwa.1(sub$counts,N=10))
11 ggplot(tt3,aes(x=period,y=power))+facet_wrap(uniqueID~e
xperiment,scale="free")+geom_line()
```



Page 150: 2017-12-21 Meeting with Dan

Dan could not make meeting , but still prepping stuff to talk about

1. Paper discussion

- Dowse, H. B. (n.d.). Maximum entropy spectral analysis for circadian rhythms: theory, history and practice. Retrieved from <https://pdfs.semanticscholar.org/d564/6a61fad1e03e61681fc1ec36>

[94c4e36b7456.pdf](#)

2. Data discussion

- show spectral analysis

from last time,

Dan's suggestion for circadian rhythm analysis;

- 1 * What do other people do to decide which dominant peaks in periodogram to choose from?
- 2 * particular in fourier transform approach; somebody develop stats? what is heuristic for picking out circadian rhythm
- 3 * What is going on with continuous wavelet? Binning at lower time give more fine estimates?
- 4 * Focus on fourier transform data.
- 5 * look up some general methods papers

Still running into problem with how to estimate peaks.

- use findpeak function

What the dataset looks like altogether

```
1 knitr::kable(ddply(test10.merg, .  
(organism, experiment, Host), summarize, count=length(organ  
ism)))
```

organism	experiment	Host	count
fly	Entrainment	Apple	15
fly	Entrainment	Haw	3
fly	Free-run	Apple	15
fly	Free-run	Haw	3
wasp	Entrainment	Apple	1
wasp	Entrainment	Haw	3
wasp	Free-run	Apple	1
wasp	Free-run	Haw	3

reworking function to use findpeaks() function in pracma package

```

1 sa.an<-function(ts=counts15$counts15){
2   sa1<-
3     spectrum(ts,method=c("pgram", "ar"),plot=FALSE,demean=TRUE,
4       detrend=TRUE,tape=.2)
5   spx<-sa1$freq
6   spy<-2*sa1$spec
7   pw<-data.frame(spx,spy)
8   cc1<-pw[order(pw$spy,decreasing=TRUE),]
9   cc2<-subset(cc1,spx<0.05)
10  cc2$density<-density(cc2$spy,n=length(cc2$spy))$y
11  cc2$t<-density(cc2$spy,n=length(cc2$spy))$x
12  #cc<-findpeaks(cc2[,3],minpeakheight=1E-6)
13  cc<-findpeaks(cc2[,1])
14  cc2[order(cc2$density,decreasing=TRUE),]
15  out<-1/cc2[cc[,2],][,1]/4
16  #out<-1/cc2[cc[,2],][,1]/4/24

```

```

15   return(out[1:4])
16   ## hours
17
18
19 }
20
21 sa.an()
22

```

[1] 26.12903 23.14286 21.89189 21.31579

estimating biological rhythms for each unique ID and experiment

```

1  ### for each unique ID and experiment
2  test10<-ddply(nall.data15_3,.
  (uniqueID,experiment),function(sub) sa.an(sub$counts))
3  knitr::kable(cbind(test10[,1:2],round(test10[,3:6],3)))
4

```

uniqueID	experiment	V1	V2	V3	V4
10o49	Entrainment	12.000	8.000	6.000	7.500
10o49	Free-run	20.710	10.014	18.984	22.226
10o73	Entrainment	8.000	7.714	6.353	6.000
10o73	Free-run	7.843	9.302	5.970	6.557
10w12	Entrainment	8.000	12.632	13.333	8.571
10w12	Free-run	11.683	7.594	7.409	5.457
10w15	Entrainment	6.000	5.455	9.474	5.294
10w15	Free-run	6.027	5.357	5.533	10.227
11w26	Entrainment	6.000	6.667	7.500	7.200
11w26	Free-run	8.036	5.819	5.037	5.444
12b43	Entrainment	12.500	8.333	5.882	7.692

12b43	Free-run	21.951	12.000	21.429	20.000
12b6	Entrainment	11.912	5.956	18.409	6.328
12b6	Free-run	23.674	26.042	27.902	30.048
12w40	Entrainment	8.000	6.061	5.714	5.263
12w40	Free-run	23.684	75.000	90.000	24.324
13o11	Entrainment	11.765	8.000	5.128	5.405
13o11	Free-run	7.500	5.902	5.854	5.714
13o28	Entrainment	8.000	6.667	5.263	6.061
13o28	Free-run	21.053	8.511	9.091	8.696
2b23	Entrainment	12.500	11.842	16.071	11.250
2b23	Free-run	10.870	10.417	16.129	10.638
2b26	Entrainment	8.654	5.625	8.333	5.921
2b26	Free-run	22.011	11.005	5.219	6.027
3o51	Entrainment	11.765	6.250	6.061	5.882
3o51	Free-run	11.905	9.434	8.065	8.333
4o15	Entrainment	5.956	6.328	5.786	8.100
4o15	Free-run	8.036	6.000	5.921	10.976
4w66	Entrainment	21.441	11.391	11.045	5.975
4w66	Free-run	76.800	54.857	42.667	12.000
5b25	Entrainment	11.842	8.036	7.759	6.081
5b25	Free-run	12.195	11.628	11.364	5.208
5w63	Entrainment	22.500	11.912	5.956	5.625
5w63	Free-run	10.976	8.036	5.921	5.422
5w73	Entrainment	20.250	6.532	7.788	10.125
5w73	Free-run	22.500	11.250	7.627	8.333
h2b25	Entrainment	8.000	6.061	5.263	5.882
h2b25	Free-run	24.000	8.000	8.108	9.091

h4o4	Entrainment	20.000	8.000	11.765	6.061
h4o4	Free-run	26.042	52.083	23.148	27.174
h4w10	Entrainment	8.000	18.000	5.838	6.000
h4w10	Free-run	5.143	5.294	6.000	6.923
h5o22	Entrainment	8.036	12.500	6.081	5.921
h5o22	Free-run	23.810	22.727	5.556	5.155

Merge dataset with original one with eclosion data.

```

1  ###match data with eclosion data
2
3  ### have to adjust uniqueID dataframe in test10 to
   match orig.dat
4  test10$uniqueID<-
   as.factor(ifelse(substr(test10$uniqueID,1,1)=="a",subst
   r(test10$uniqueID,2,10),substr(test10$uniqueID,1,10)))
5
6  ##orig.dat
7  orig.dat<-read.csv("../Data/2017-11-
   17_subset_host_comparison_trik_data_extract.csv")
8
9  test10.merg<-inner_join(test10,orig.dat,by="uniqueID")
10

```

Since dominant periods are at different time scales, I explored comparing biological rhythms at the same scale among individuals. Have to extract and summarize that data from the 4 columns

```

1 test10.merg$circ_time<-
  apply(test10.merg[,3:6],1,function(x){mean(subset(x,x<35
  & x > 20))})
2
3 test10.merg$ult_time<-
  apply(test10.merg[,3:6],1,function(x)
  {mean(subset(x,x<20))})
4 test10.merg$day_time<-
  apply(test10.merg[,3:6],1,function(x)
  {mean(subset(x,x>48))})
5

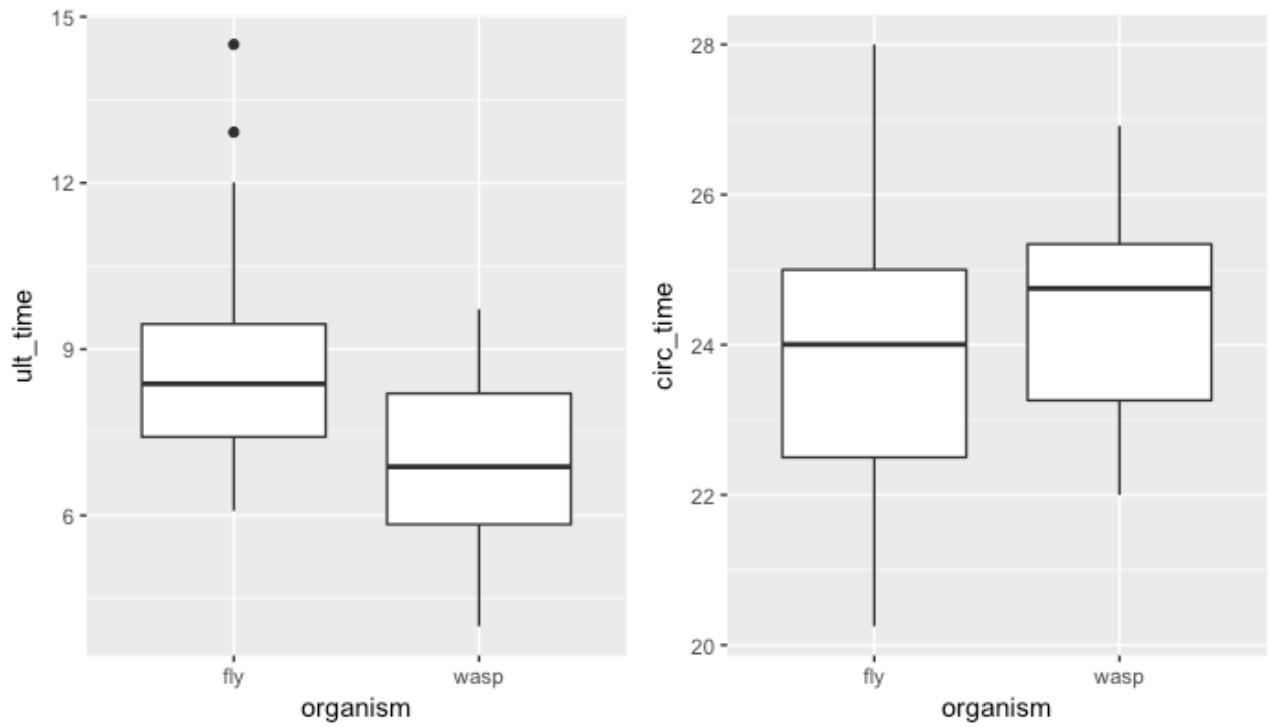
```

Comparisons for each specific biological rhythms among organisms: flies vs wasps

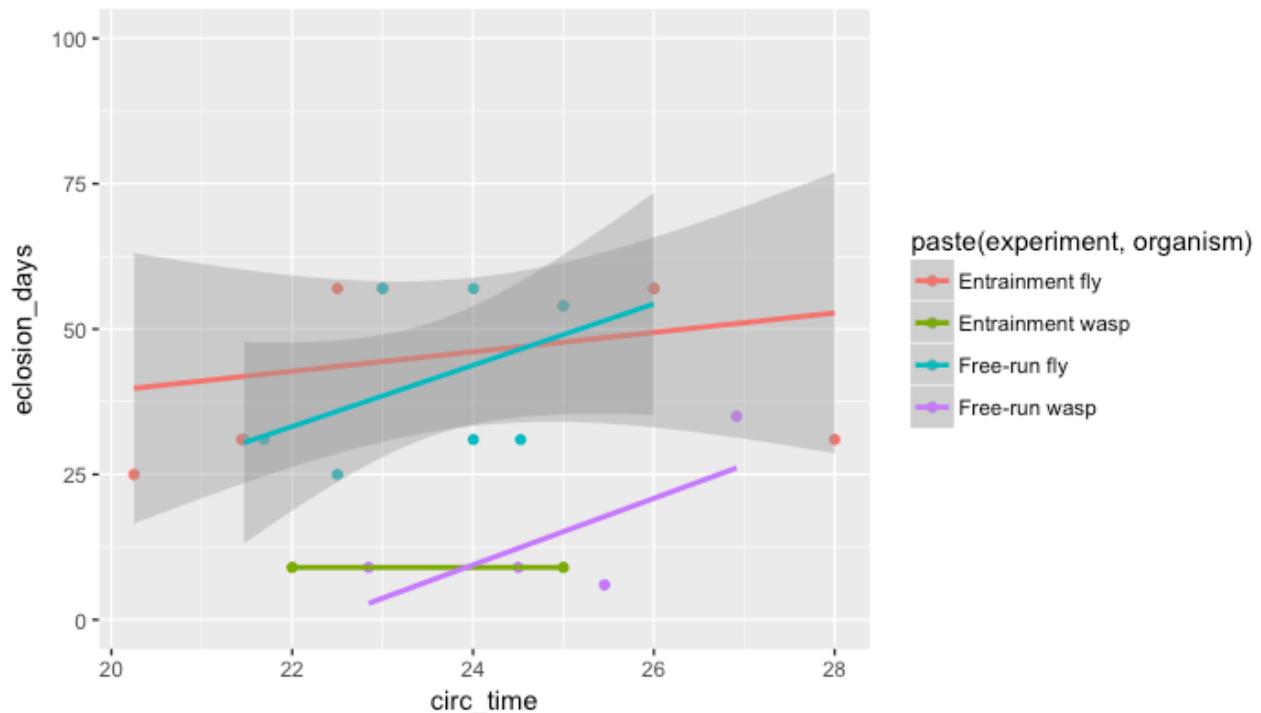
```

1 ab<-
  ggplot(test10.merg,aes(x=organism,y=circ_time))+geom_box
  plot()
2 cd<-
  ggplot(test10.merg,aes(x=organism,y=ult_time))+geom_boxp
  lot()
3 grid.arrange(cd,ab,ncol=2)

```



Relationship to eclosion days, for circ time



stats

```

1 summary(aov(eclosion_days~organism+circ_time,data=test1
0.merg))
2
3      Df Sum Sq Mean Sq F value    Pr(>F)
4 organism     1   4227   4227  28.411 3.23e-05 ***
5 circ_time     1     739     739   4.966  0.0375 *
6 Residuals   20   2976     149
7 ---
8 Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
9   ' '
10 21 observations deleted due to missingness

```

ultradian not significant.

Reminders : from 2017-12-13

metaanalysis ideas given what we've read

compare different rhythms, circ more or less flexible than other rhythms
 (short, longer) are different components of biological rhythms heritable
 what env factors best explain bio rhythms

- light
- temp
- food

What's the potential for evolution ?

Organize evolution and biological rhythms meeting (tell Dan about time line)

- Evolution 2018: Montpellier France;
 - August 19-22
 - Post doc travel grant:
<http://evolutionmontpellier2018.org/travel-grants>

- Deadline is January 17th
- Biological rhythms meeting 2018:
<https://srbr.org/meetings/upcoming-meeting/registration/>
 - May 12-16 2018
 - registration:
 - early(nov 15-march 1) reg non member - **\$725 ; \$425 member**
 - late (march 2 to onsite) reg non member - **\$825; \$525 member**
 - early(nov 15-march 1) post doc non member - **\$475 ; \$425 post doc member**
 - late (march 2 to onsite) post doc non member - **\$575; \$525 post doc member**

2017-12-22. actual meeting

Look into Cold spring harbor protocols

```
1 * set up trikinetics
2 * data analysis
3 * data handling
```

Basically, look at protocols online, actually check papers that have similar data structure like ours. Look up drosophila work.

Analyze existing, publicly available datasets to see if we can recapitulate results.

How do we separate signal from noise?