Population Genetics and Gene Expression Patterns of *Calanus finmarchicus*

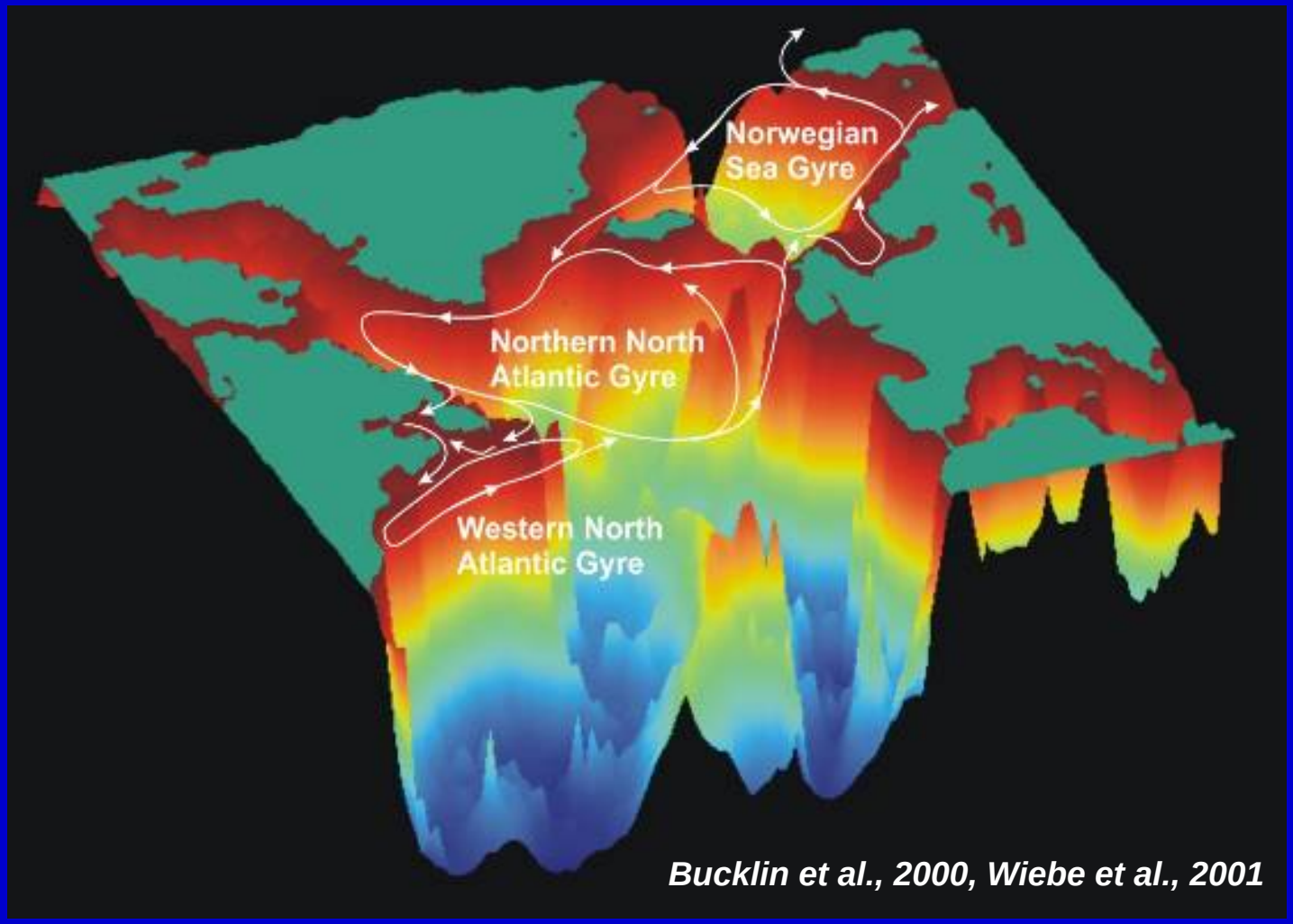
Ebru Unal and Ann Bucklin
Department of Marine Sciences
University of Connecticut, CT

GLOBEC Synthesis PI Meeting, Woods Hole, 28-29 July 2011

University of Connecticut

Three-Gyre Hypothesis

There is limited exchange (gene flow) among *Calanus finmarchicus* populations inhabiting three North Atlantic gyres, resulting in basin-scale population genetic differentiation.

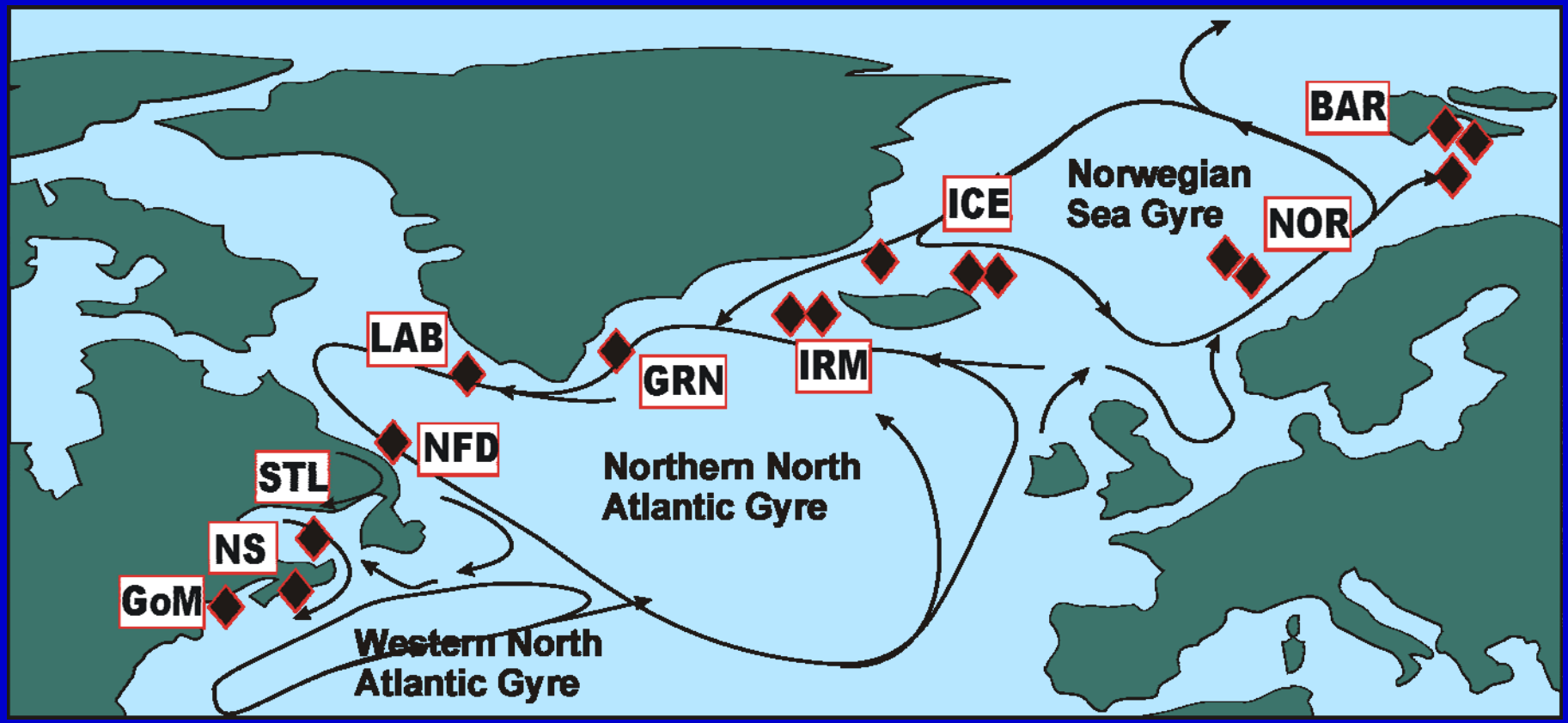


Bucklin et al., 2000, Wiebe et al., 2001

Basin-scale population genetic structure of *Calanus finmarchicus*

24 single nucleotide polymorphisms (SNPs) in 3 nuclear genes were used as population genetic markers to analyze ~350 individuals in 17 samples from 10 regions in 3 gyres (plus Iceland at an interface).

Unal and Bucklin (2010) Prog. Oceanogr.



Population Genetic Sampling and Analysis

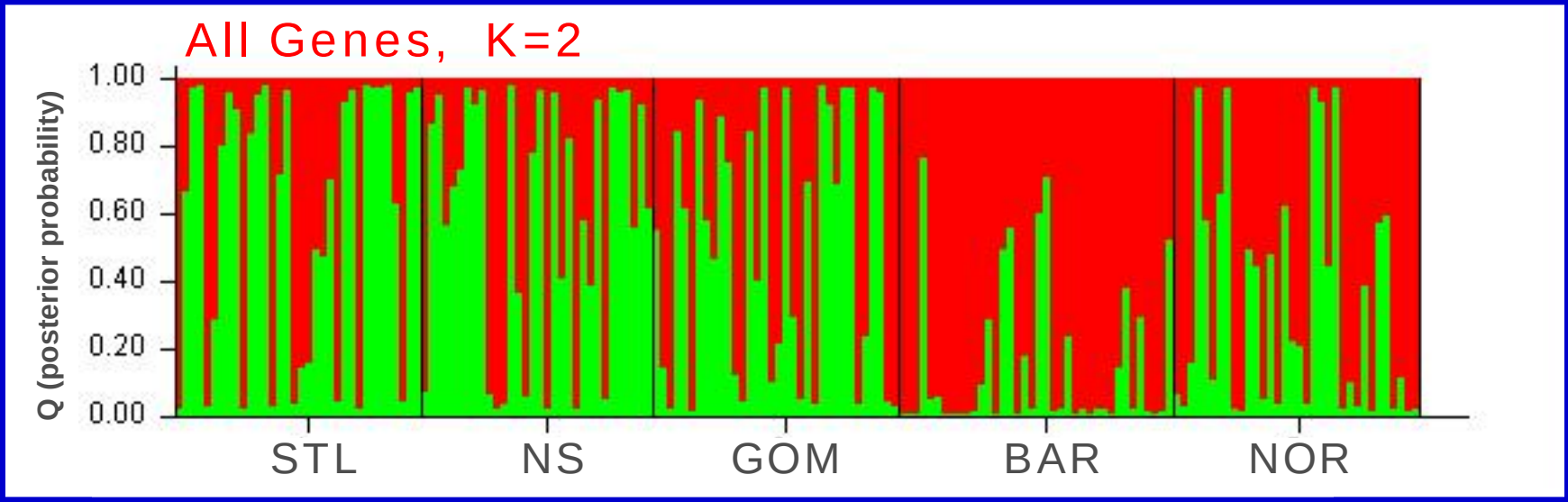
Analysis of SNPs in three genes encoding metabolic proteins:

- ***AMP-Activated Protein Kinase (AMPK)*** **– cellular energy homeostasis**
- ***Citrate synthase (CS)*** **– metabolic enzyme of the citric acid cycle**
- ***Heat shock protein (HSP-70)*** **– molecular chaperones / damage protection**

Sample	Region	AMPK (7)	CS (9)	HSP-70 (8)	All (24)
Gulf of Maine	NW	34	34	34	34
Nova Scotia	NW	32	32	32	32
St Lawrence	NW	33	33	34	34
Newfoundland	NW	32	32	31	32
Labrador	Central	34	33	30	34
Greenland	Central	24	23	24	24
Irminger	Central	34	34	27	34
Iceland	Central	58	45	47	58
Norway	NE	34	34	30	34
Barents	Barents	36	35	27	38
TOTAL		351	335	316	354

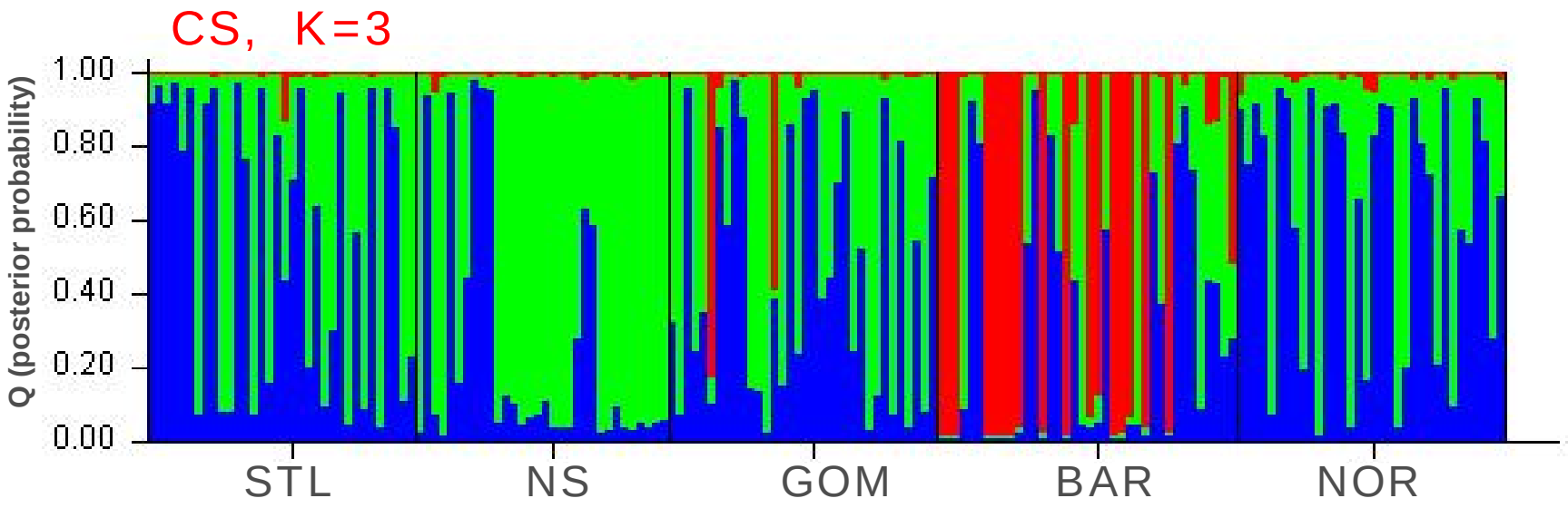
Large-Scale Analysis using Bar Plots

- **STRUCTURE 2.2 (Pritchard, et al., 2000) uses a posterior likelihood analysis to determine number of groups (population clusters = K) based on best fit to data.**
- **Bar plot shows Q (estimated membership coefficients by posterior probability) for each individual; K = clusters derived from genetic similarities; POP = pre-defined geographic populations**
- **In the bar plot, each individual is represented by a single vertical line broken into K colored segments with lengths proportional to each of the K inferred clusters; individuals are sorted by predefined populations.**



Large-Scale Analysis using Bar Plots

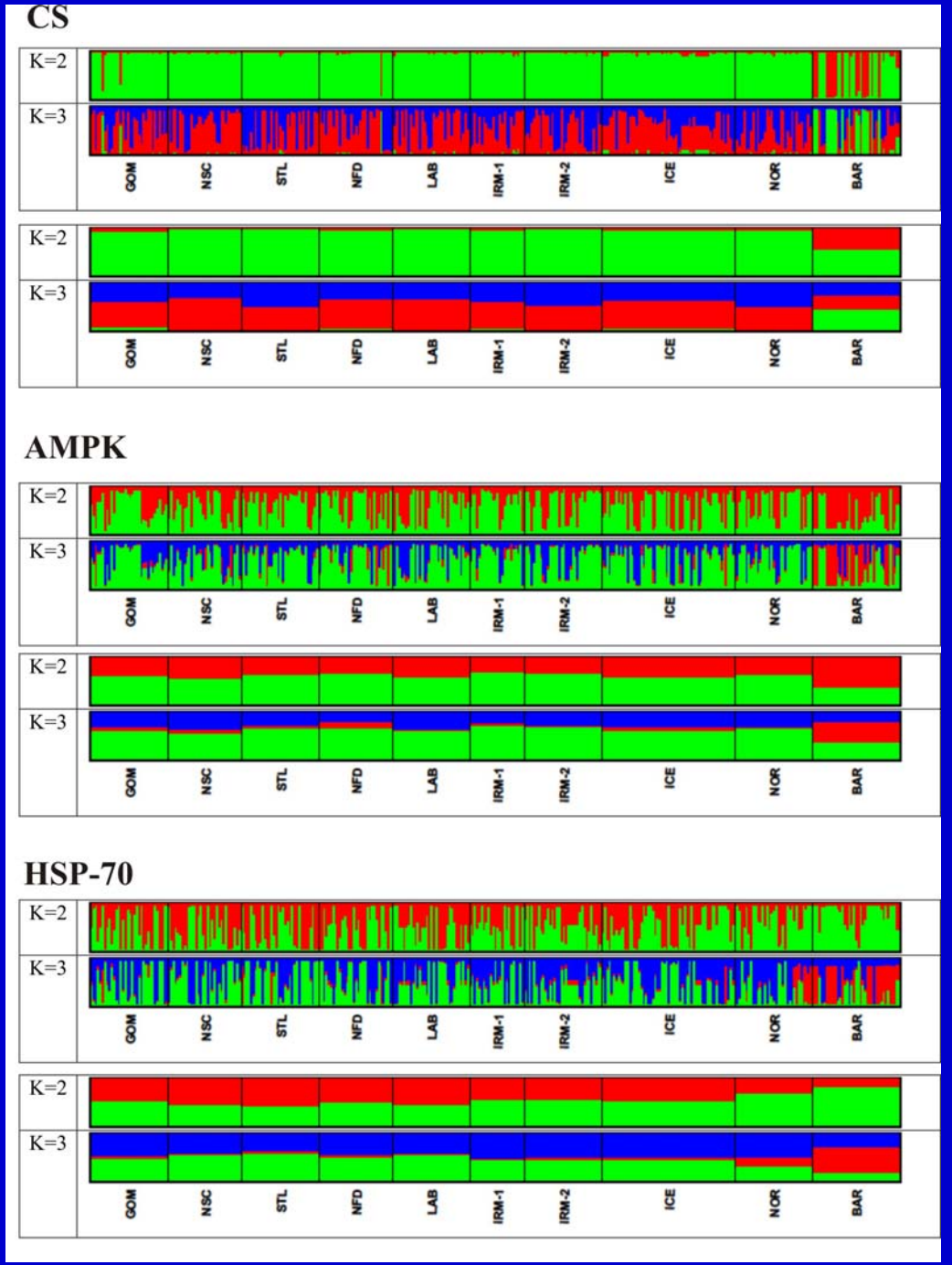
- **SNP allele frequencies for 3 genes showed different patterns in analysis with STRUCTURE 2.2.**
- **For citrate synthase (CS), there was a good fit to K = 3; both BAR and NS are distinctive.**
- **Q = estimated membership coefficients by posterior probability for each individual; K = clusters derived from genetic similarities; POP = pre-defined geographic populations.**



Analysis by Gene

Separate analysis of SNP allele frequencies for each of the 3 genes revealed different patterns among genes.

Additional genes and genetic characters should be analyzed to examine effects of selection on estimations of small- to large-scale structure and exchange.



Interannual Variation in Wilkinson Basin

Allele frequencies at SNP sites within citrate synthase are shown in the pie diagrams (with colors indicating the four nucleotides, A,G,T,C) were determined for 1997, 1998, 1999, 2003, 2004 and 2005 samples. Numbers indicate SNP sites, based on the 294 base-pair region of CS sequenced. 9 SNPs with statistically significant (p<0.05) differences in allele frequencies among the six years are indicated by red colored font above. The ones that show highly significant (p<0.01).differences are also indicated by asterisk (*) below

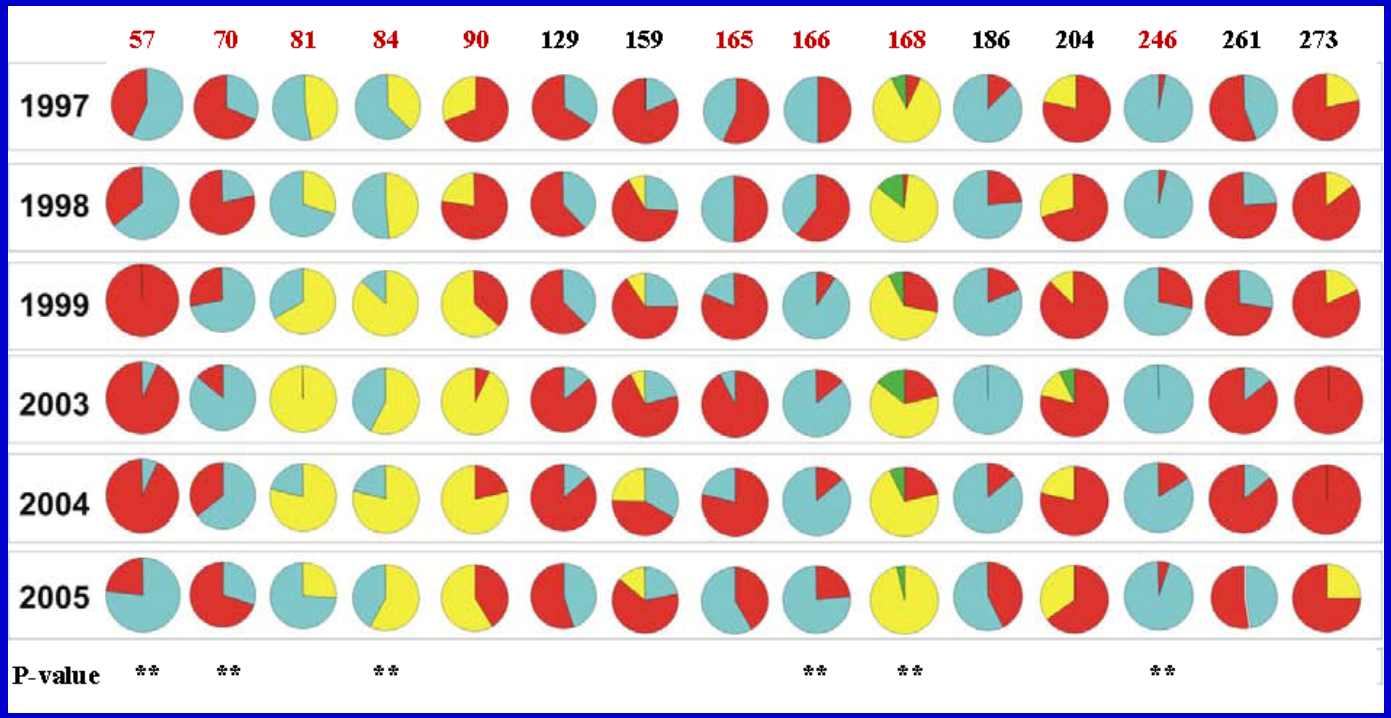
Date
10/16/1997

10/22/1998
12/05/1998

10/17/1999
12/05/1999
04/20/2003

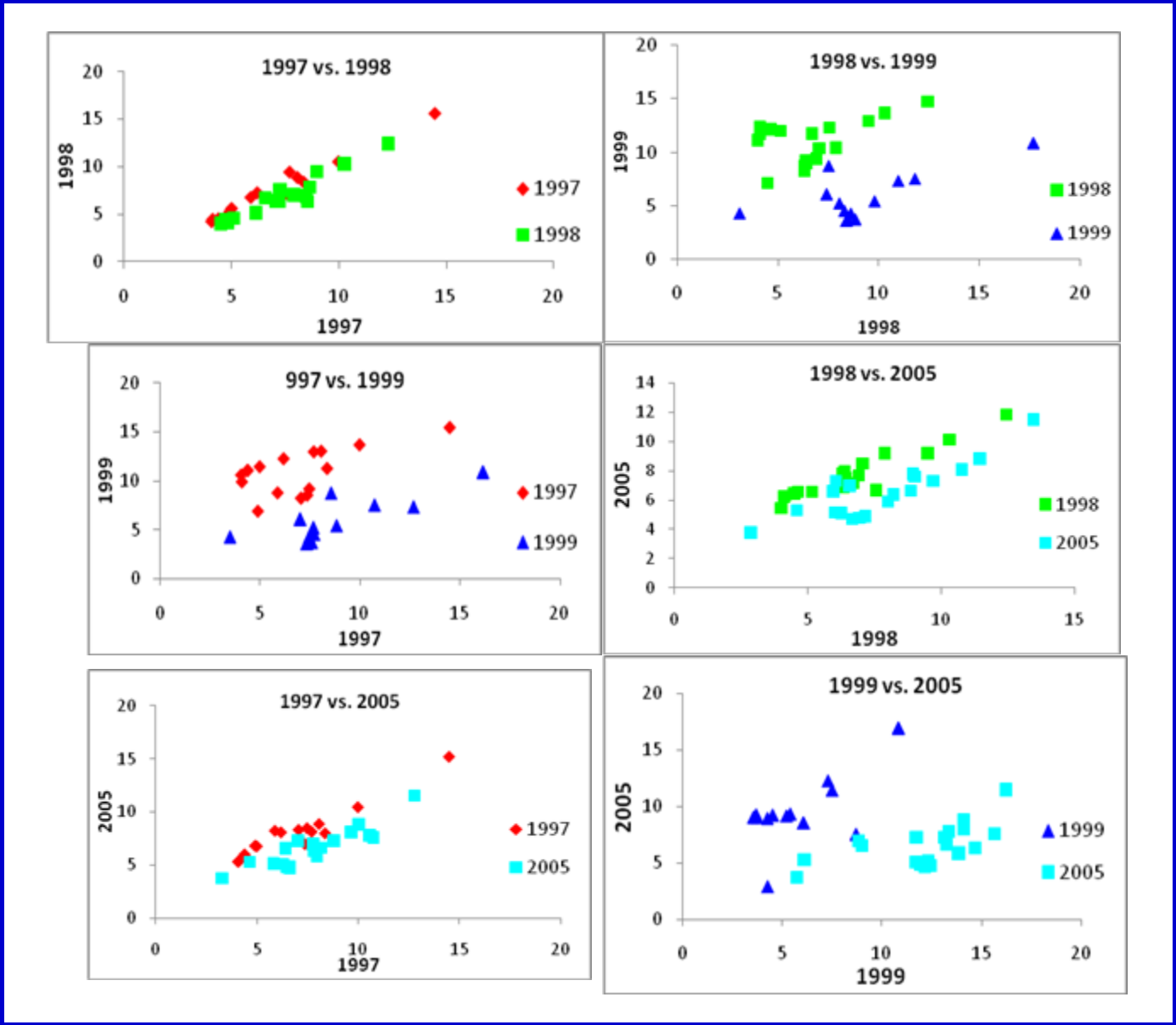
05/03/2004

03/18/2005
03/18/2005

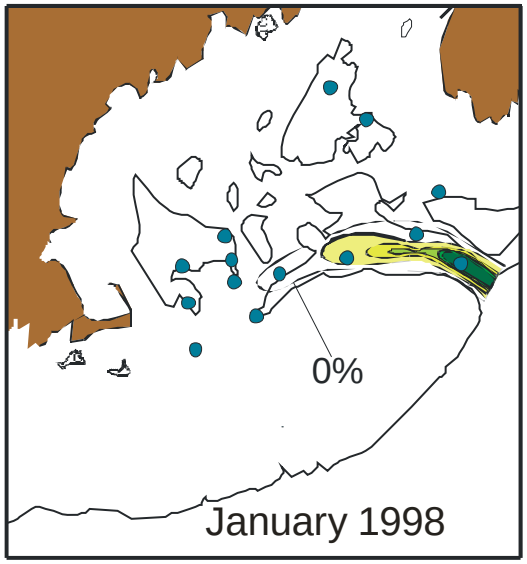


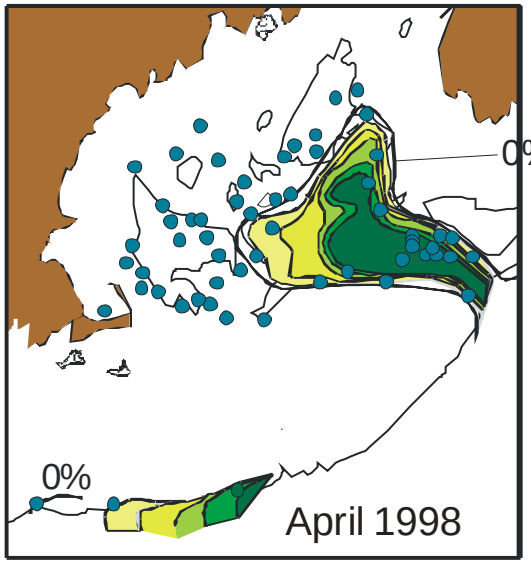
Interannual Variation in Wilkinson Basin

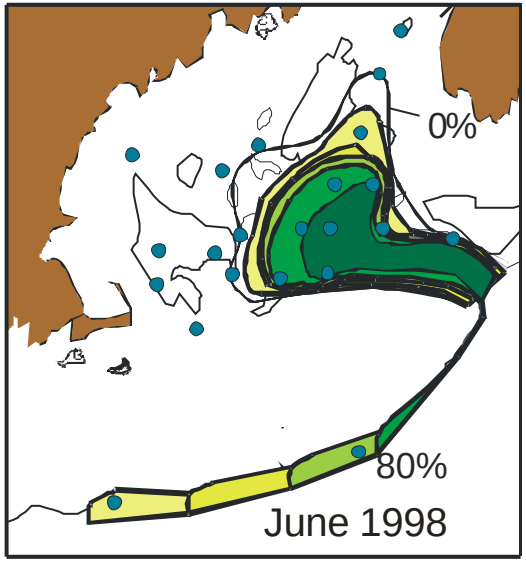
Pairwise population assignment graphs generated by using GENALEX for *C. finmarchicus* individuals collected during 4 years (1997, 1998, 1999, and 2005) and analyzed for 15 SNP loci. The results indicate the distinctive status of 1999 samples as compared to others. Samples from 2003 and 2004 were excluded from this analysis.

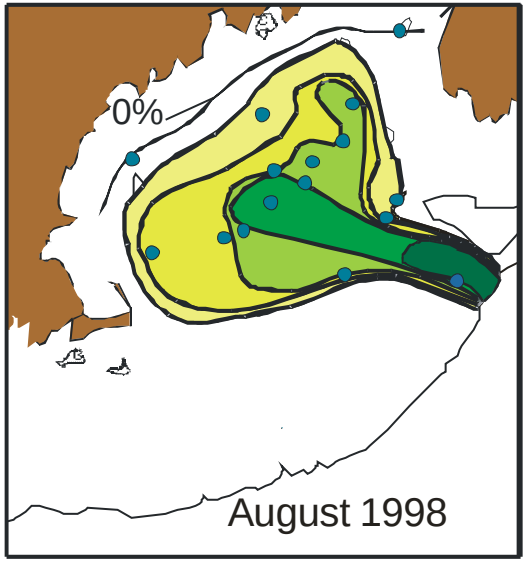


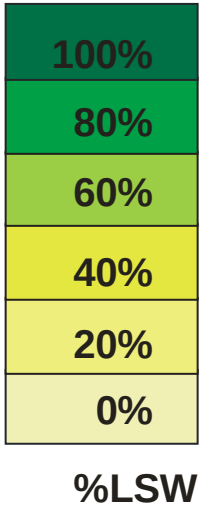
Interannual Variation in Wilkinson Basin











Interannual Variation in Wilkinson Basin

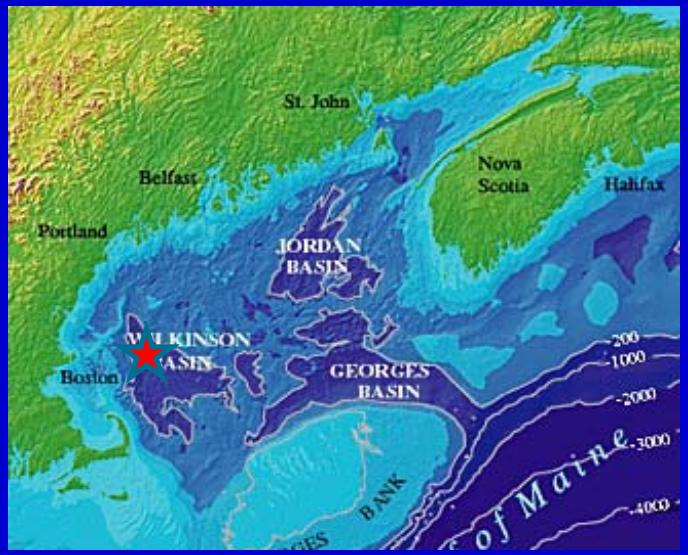
F_{ST} values (below diagonal) and significance levels (above diagonal) for pairwise comparisons between samples of *C. finmarchicus* collected in 4 years, 1997, 1998, 1999, and 2005. Significance levels are indicated by symbols as follows: $p \leq 0.01$ (+), and $p > 0.01$ (?). Samples collected in 2003 and 2004 were not used for analysis due to small sample sizes.

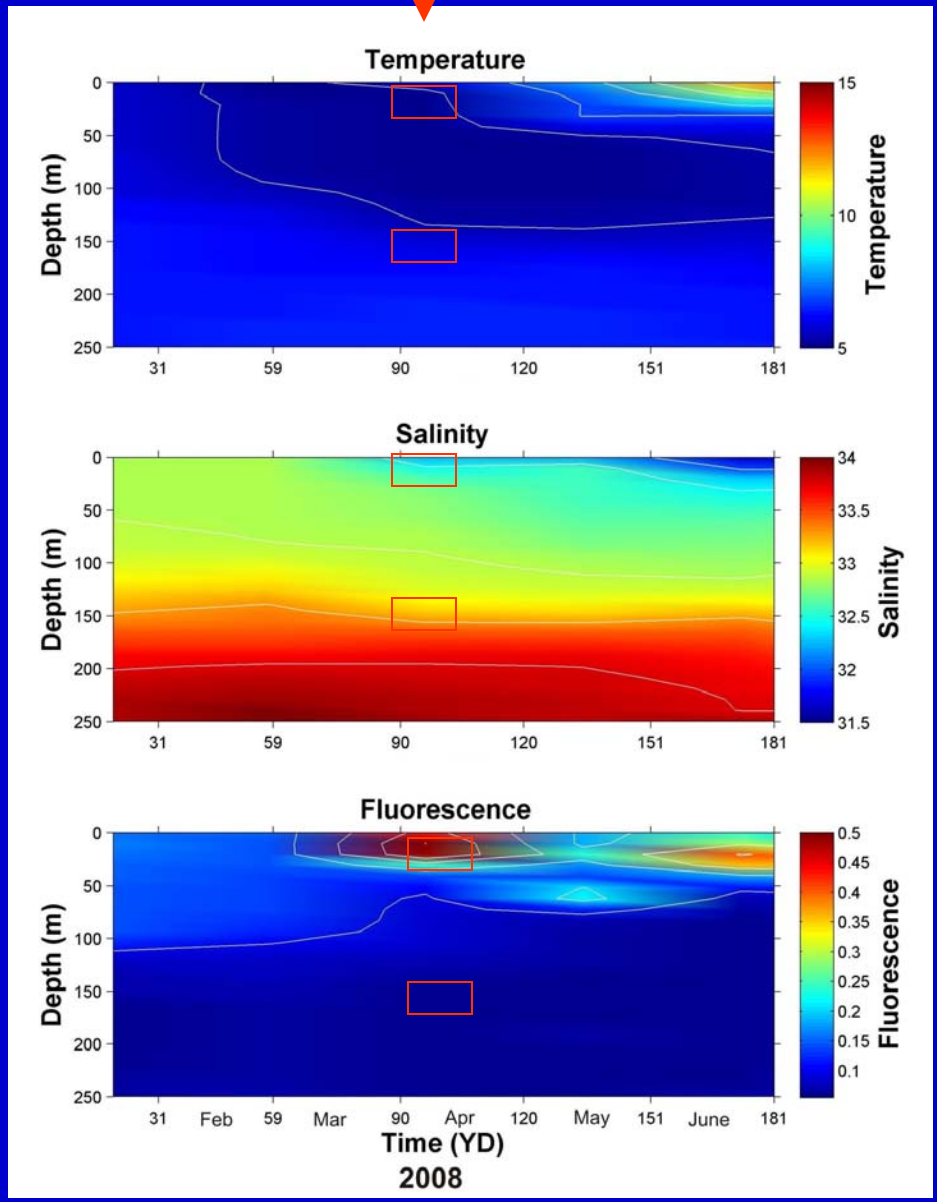
	1997	1998	1999	2005
1997		?	+	?
1998	-0.00107		+	?
1999	0.15353	0.19815		+
2005	0.03631	0.04484	0.19096	

Differential Gene Expression on Small Scales

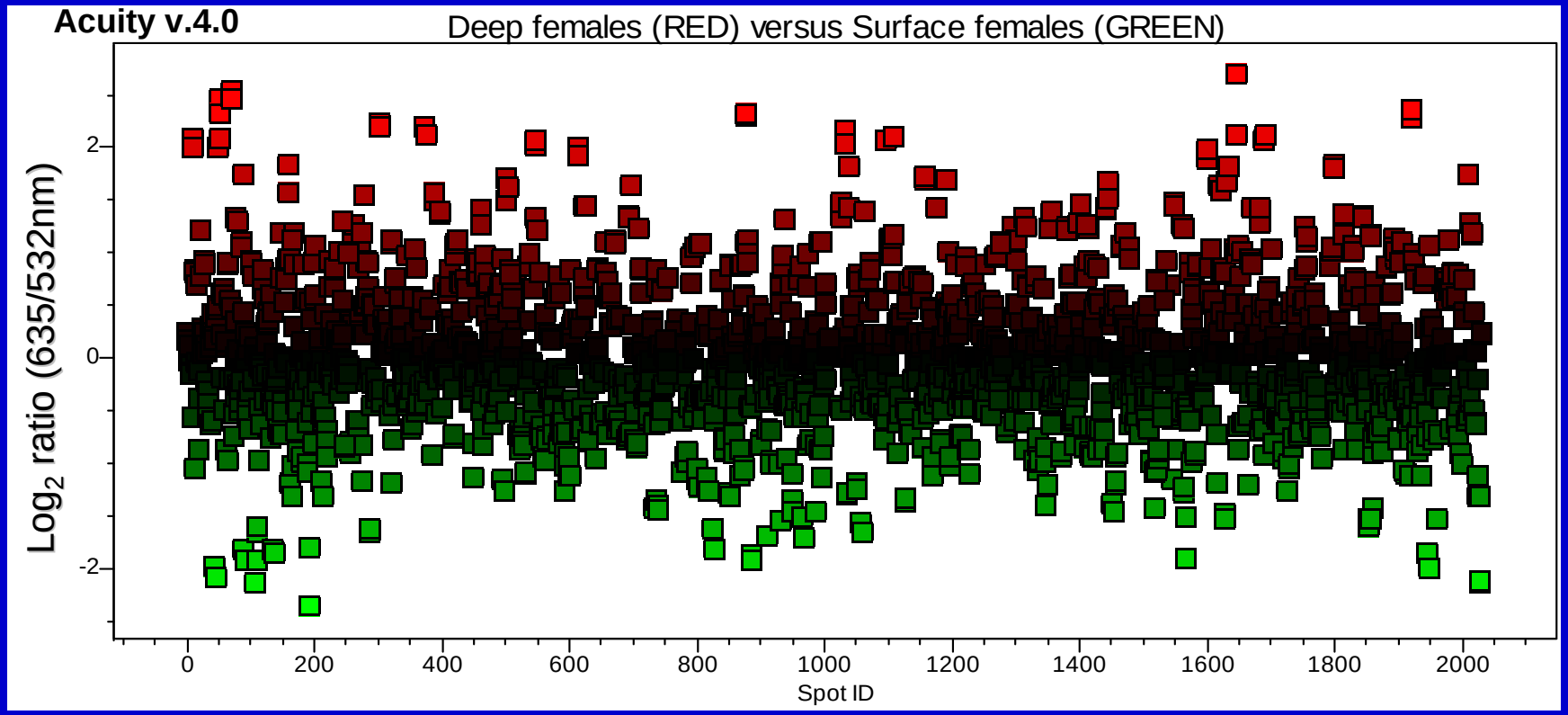
Comparison between female *C. finmarchicus* collected in surface (0-30m) and deep (140-170m) MOC-1 samples from Wilkinson Basin in April 2008.

Significant differences in numerous genes associated with physiological adaptation to small-scale and short-term environmental variation.





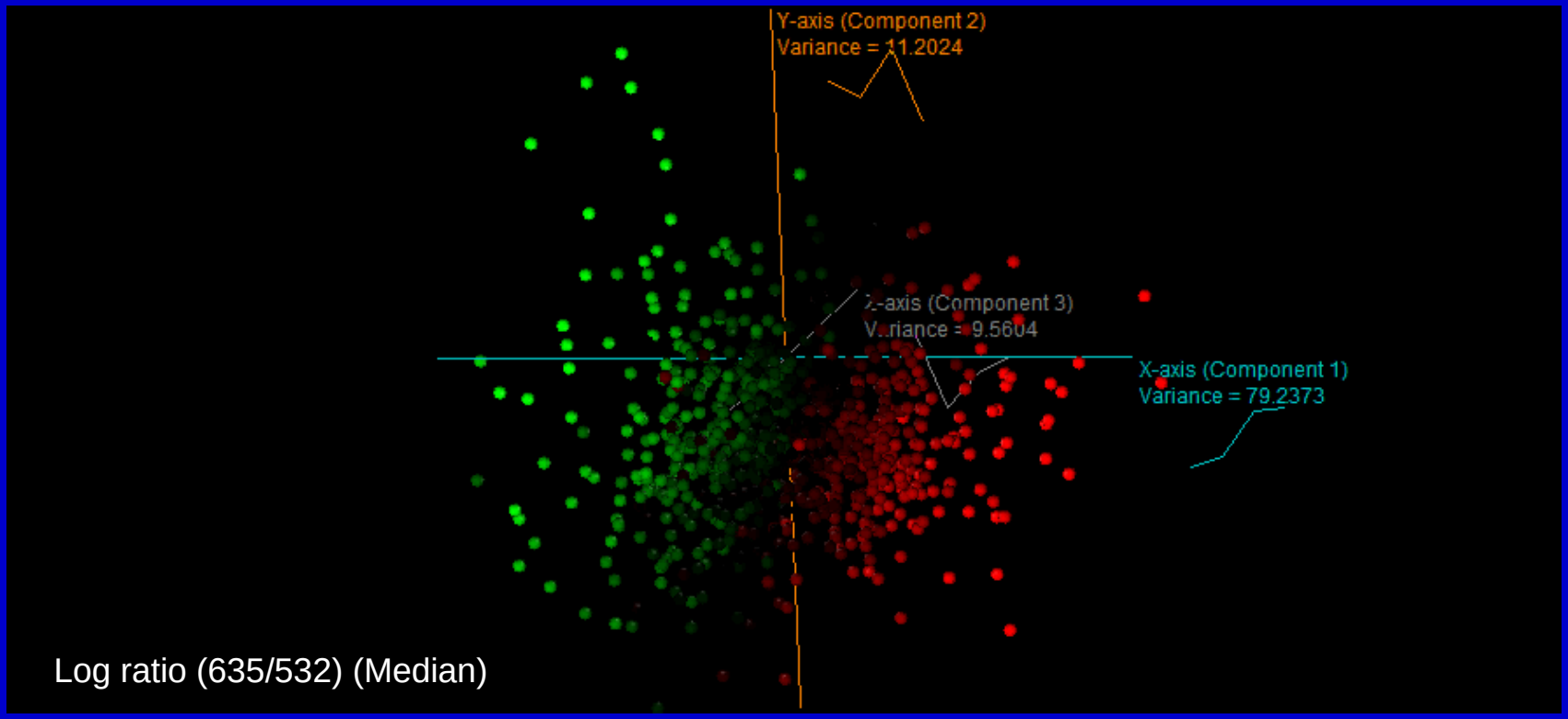
Differential Expression Profile of 1,000 Genes



Log_2 of the ratio of median fluorescence: Index of the fold change in gene expression. Negative values indicate a decrease in expression; positive value indicates an increase. $Log_2 = 2$ is a 4-fold change

Environmental Genomics of *Calanus finmarchicus*

Principal component analysis can be used to identify strongly up- or down-regulated genes. These genes can be targeted for further study of expression levels in samples across the N. Atlantic basin.



Principal component analysis (PCA) shows a low-dimensional analysis of the data set. Genes are plotted in a space defined by the components that are derived from the data.

Environmental Genomics of *Calanus finmarchicus*

Principal component analysis can be used to identify strongly up- or down-regulated genes. These genes can be targeted for further study of expression levels in samples across the N. Atlantic basin.

1. Alkaline phosphatase
2. SPT transcription factor
3. Keto acid dehydrogenase
4. ATP synthase
5. Translation elongation factor g
6. Importin 9
7. Ribonucleotide reductase
8. Septin
9. Vacuolar protein 33B