Gene expression evolution between ranges in an invasive weed

Turner K[1], Tian A[2] , Zoubarev A[3], Ch'ng C[2] , Hodgins KA[1]

Department of Botany, Bioinformatics Graduate Program, Center for High-throughput Biology
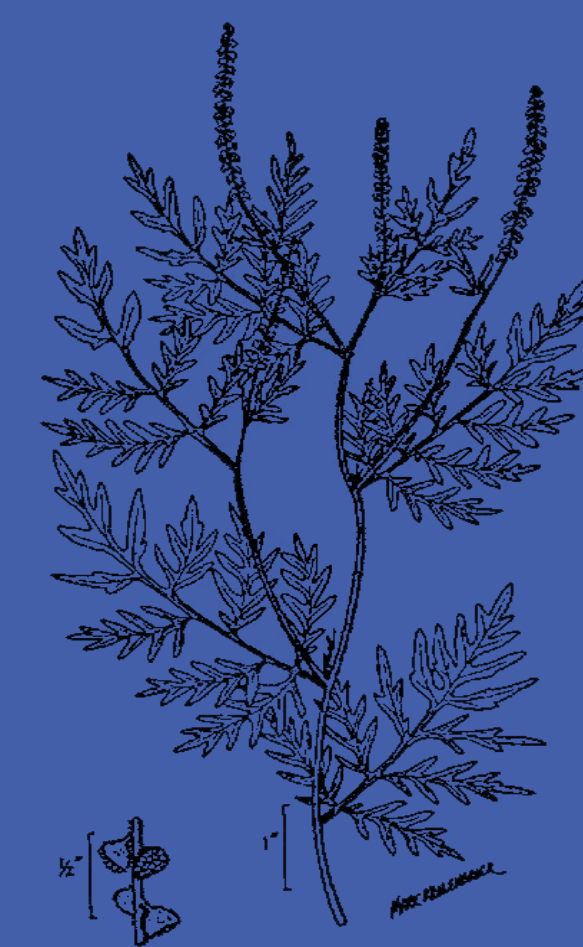
INTRODUCTION

A species is deemed 'invasive' if it is able to come into a new habitat and out-compete species native to that environment, which might generally be thought to be 'well adapted'. The direct costs of weed control and indirect costs of reduced crop production due to weeds are as much as C$40 billion annually in North America (5), and they pose a major threat to biodiversity, rivaling habitat loss. Despite constant human mediated transport, not all species become established, and relatively few become ecologically disruptive in their new habitat.

Invasive individuals are often observed to grow faster and larger, and have different phenotypes than members of the same species from its native range (3). This could be evidence for evolution in the new habitat, for changes that make the species more invasive. However, this could also be explained by differences in latitude and climate between the two ranges. A metanalysis of common garden experiments testing for rapid evolution associated with invasion found that many studies failed to include the effect of latitude, and consequently over-estimated phenotypic differences between ranges (1).

Here, using gene expression data from the invasive Common Ragweed (3; Hodgins, *unpub*), we look changes at the gene expression level which distinguish invasive from native populations of the same species, that are not attributable to latitude or climate alone. We assess if the pattern of differential expression among populations is better explained by latitude or climate variables, such as temperature or precipitation, than it is by range.

DATA

Common Ragweed (*Ambrosia artemisifolia*, Asteraceae) is a highly allergenic annual herb native to North America, but also as an invasive weed in most of Western Europe.

Native Populations   Invasive Populations
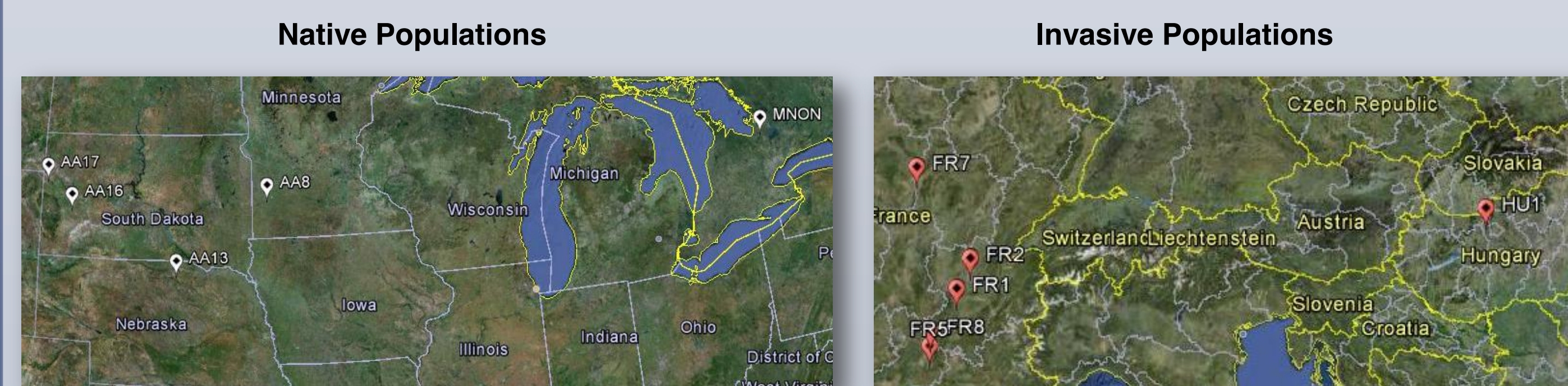
Populations in C1

- Ragweed seeds collected from native and invaded ranges
- 5 native, 6 invasive populations (bioclimatically similar)
- Grown in the same greenhouse at UBC (3).
- Control conditions only analyze here
- Leaf tissue sampled 2 weeks apart (C1 and C2)
- Pooled 3 individuals from the same seed-family and population = one biological sample
- 2 to 5samples/population
- Gene expression analyzed using species specific NimbleGen arrays (4)
- Data preprocessed and normalized by Dr. Hodgins.

Sample breakdown for Treatment
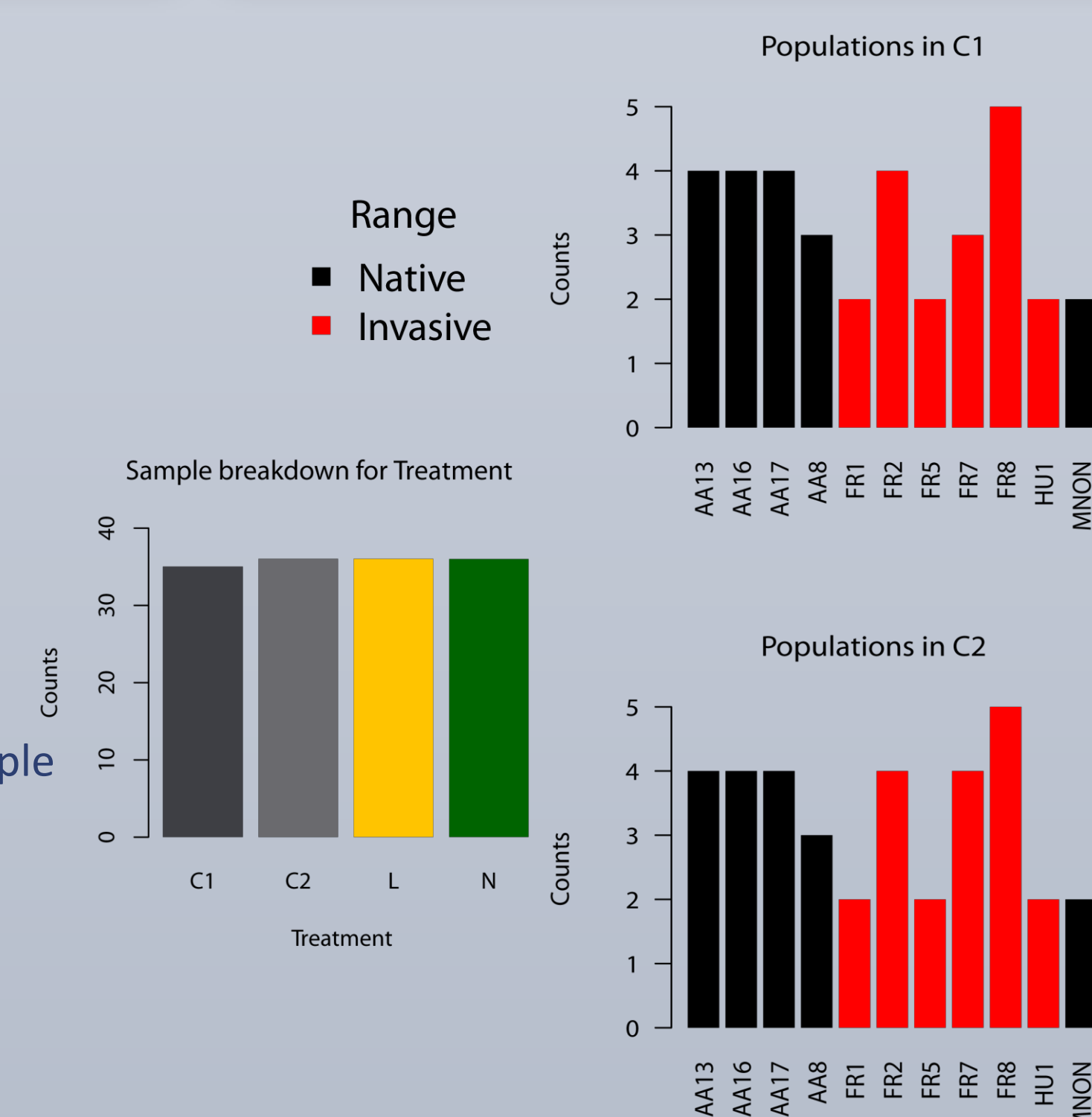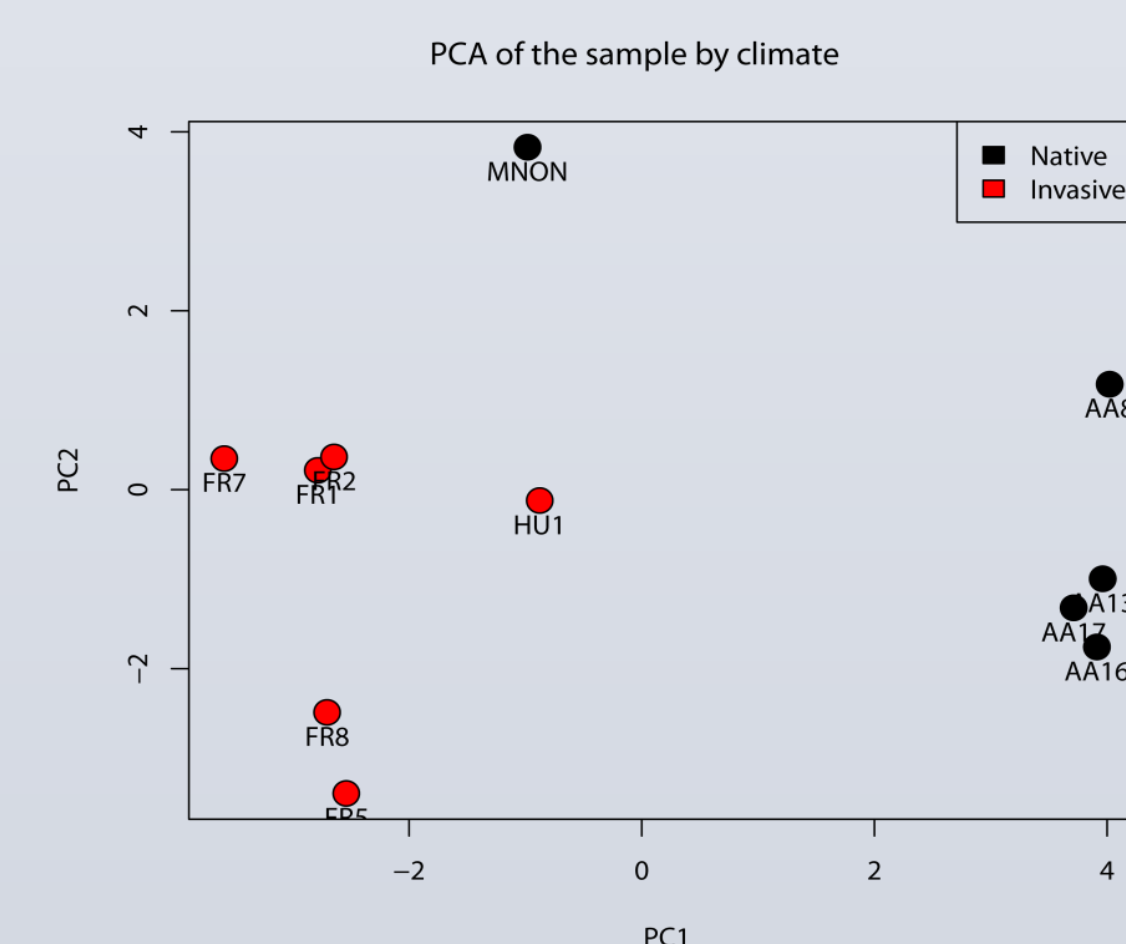
Populations in C2

We used the R package 'raster' to assemble detailed climate information for each population's location from the WorldClim database version 1.4 (release 3) (http://www.worldclim.org/, 2). As shown in the dot plot matrix, several climate factors seem to group perfectly with range. Climate factor BIO1 – BIO11 describe temperature variables, while BIO12 – BIO19 describe precipitation.
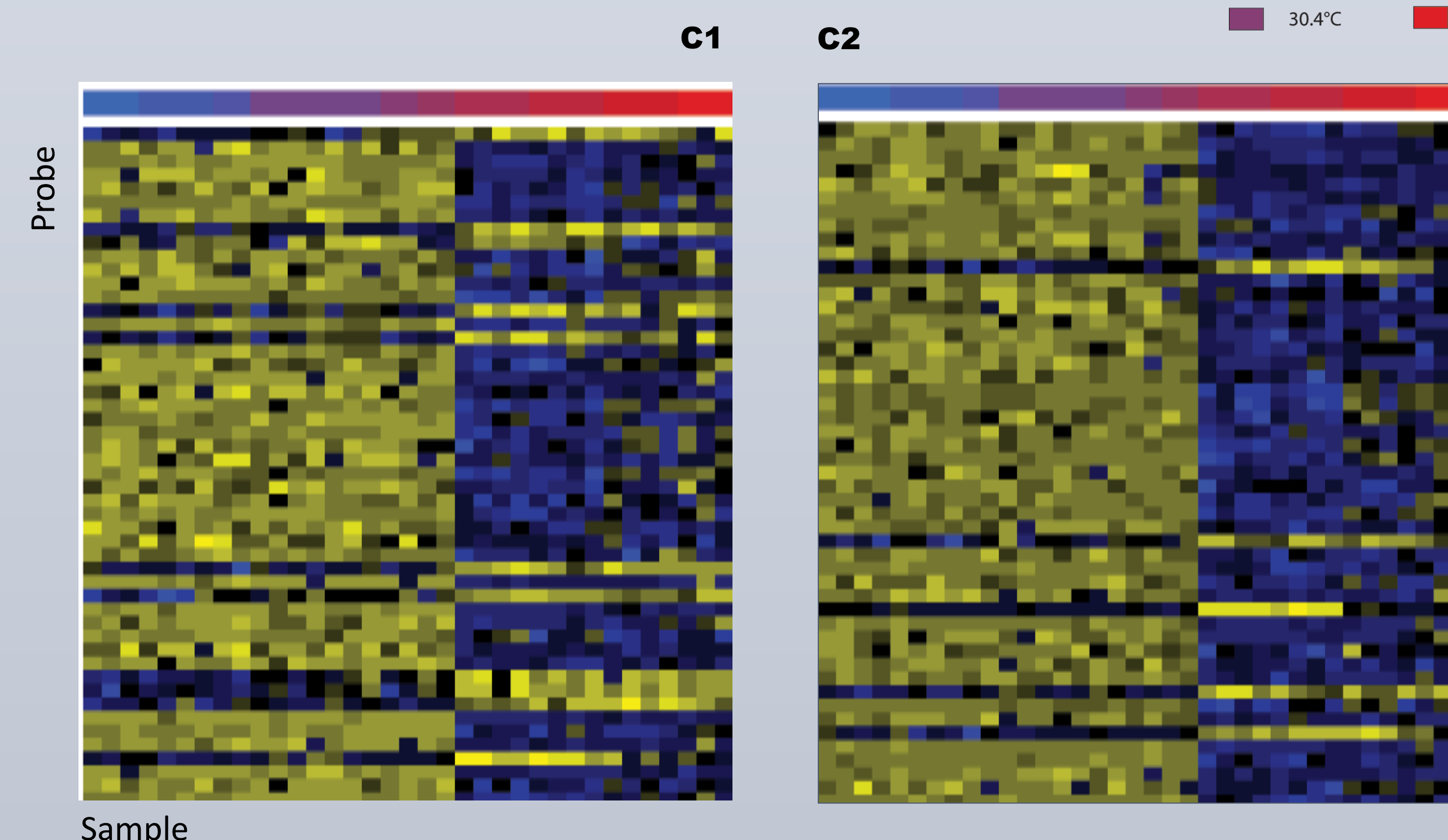
We further investigate the potential impact of these factors on gene expression using various statistical approaches, including two-sample hypothesis testing, principle component analysis (PCA), and linear regression models.

Population by Climate

Range
- ■ Native
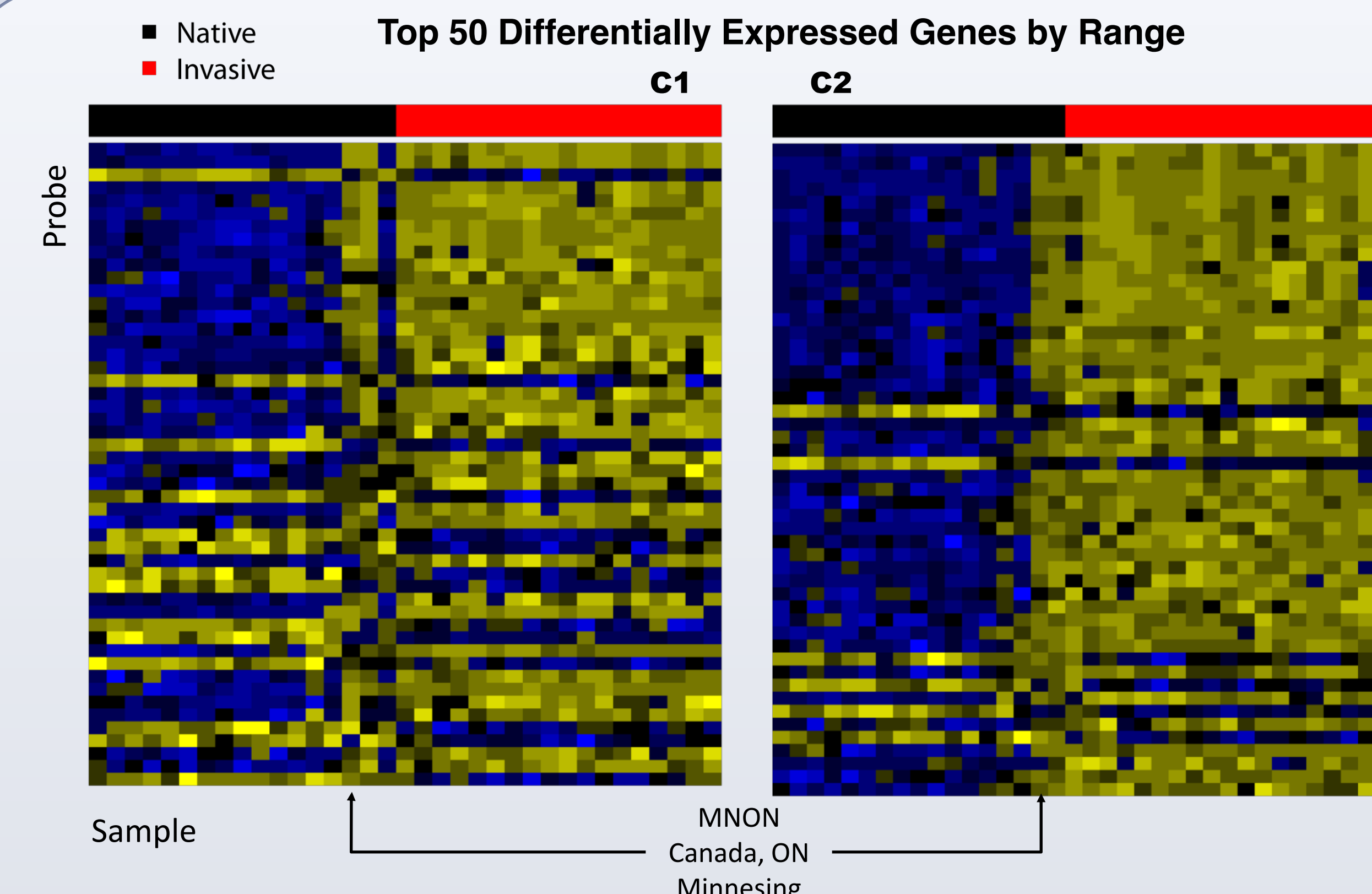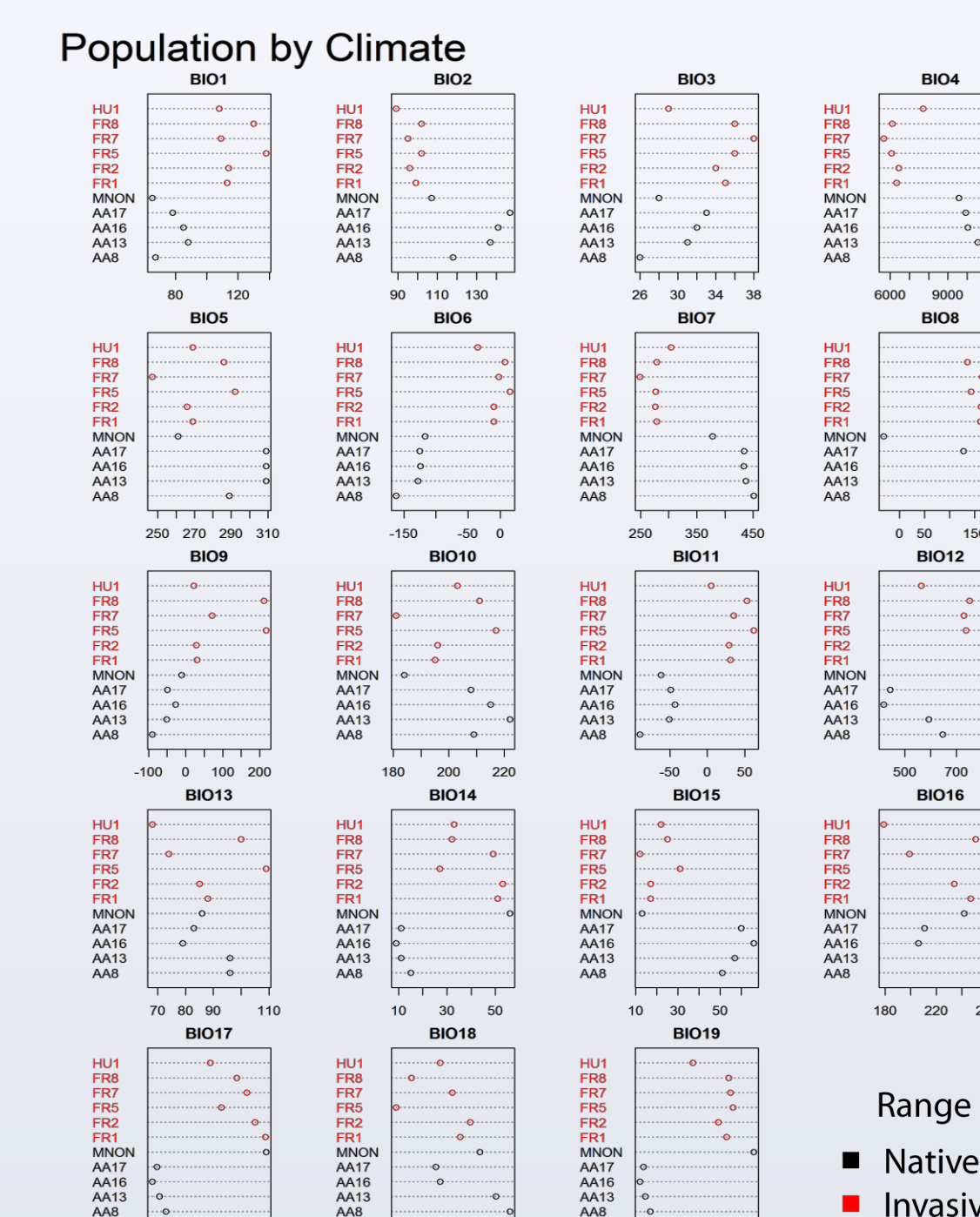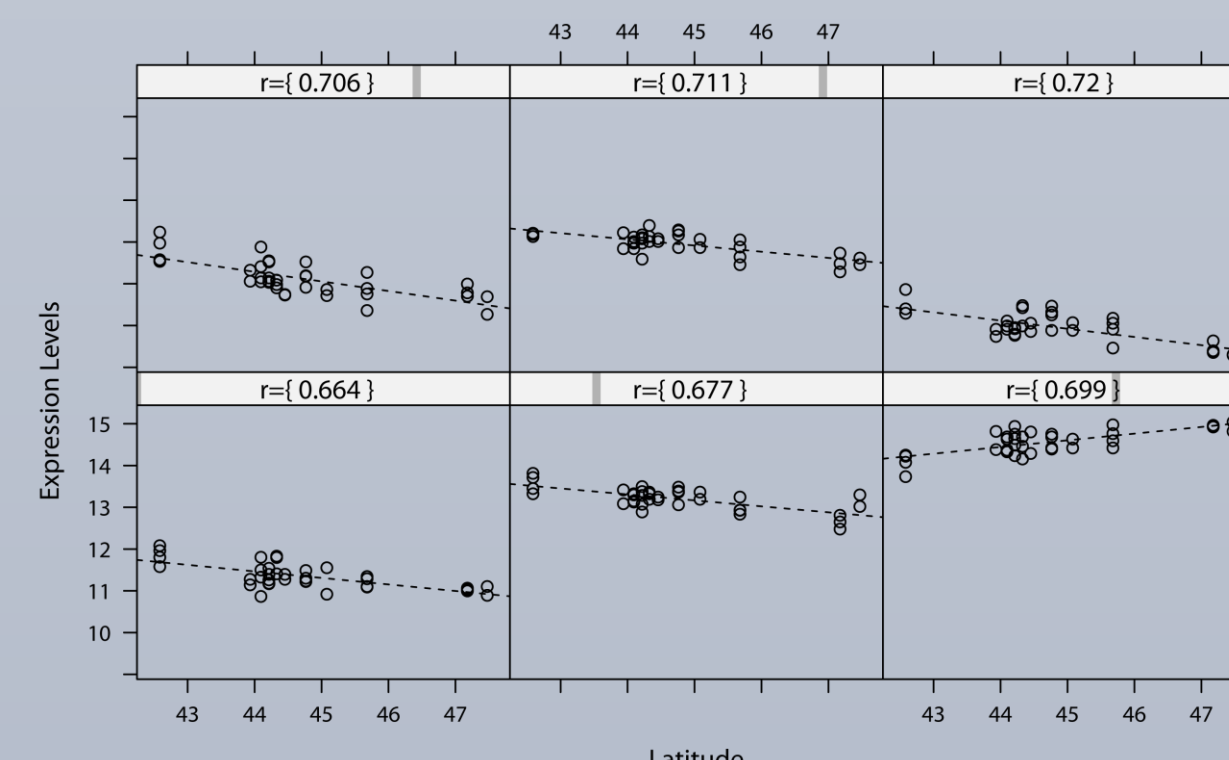- ■ Invasive

ANALYSIS

We considered three linear models: the first uses two climate factors with no interaction, the second model uses latitude as a covariate, and in the third, range is a factor. Each model was used to identify differentially expressed genes. An FDR cut-off value of 0.05 was used to correct for multiple testing.

PCA of the sample by climate

We used PCA to reduce the number of climate parameters to use in the model. We use the most correlated climate factor for each of the two principal components in the model (BIO7: Temperature Annual Range and BIO18: Precipitation of Warmest Quarter). Using this climate model, we identified 1480 probes from control one and 791 probes from control two that were identified as significantly differently expressed.

Annual Temperature Range

24.9°C	37.8°C
27.6°C	43.3°C
27.7°C	43.4°C
27.9°C	43.7°C
30.4°C	45.1°C

Top 50 Differentially Expressed Genes by Climate

C1   C2

Differentially Expressed Genes in Latitude, C1

r=(0.706)   r=(0.711)   r=(0.72)

r=(0.664)   r=(0.677)   r=(0.699)

A univariate linear model, Y=Xα +ε, was applied to the quantitative covariate, latitude, to identify genes which expression can be predicted with this factor. We obtained 6 statistically significant genes for C1 and none for C2. Preliminary analysis suggests that latitude alone plays a small role in differentiating these populations early in development, and that latitude and range are not confounded in the simple way identified by Colautti et al., 2008. More data and regression diagnostics are needed to justify the assumption of linearity.

Top 50 Differentially Expressed Genes by Range

- ■ Native
- ■ Invasive

C1   C2

MNON
Canada, ON
Minnesing

Our range model identified 1722 significant genes in C1. In C2, 628 genes were differentially expressed. This suggests that these genes have evolved expression level differences between native and invasive ragweed which have allowed the ragweed to be more invasive.
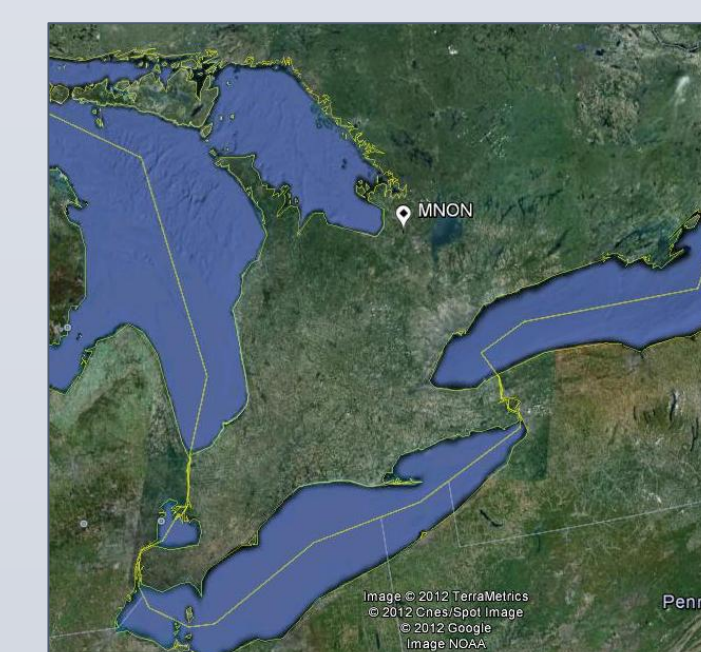
The population MNON, from Ontario, Canada (native range), appears to have an expression profile more similar to the invasive populations. This may be due to similarity in climate, as shown in our PCA analysis. Alternatively, it may represent a source population for the invasion, or a population with some history of return gene flow from the invaded range.

Gene lists produced by different models

C1
Range — 1197 — Climate
1722   1480
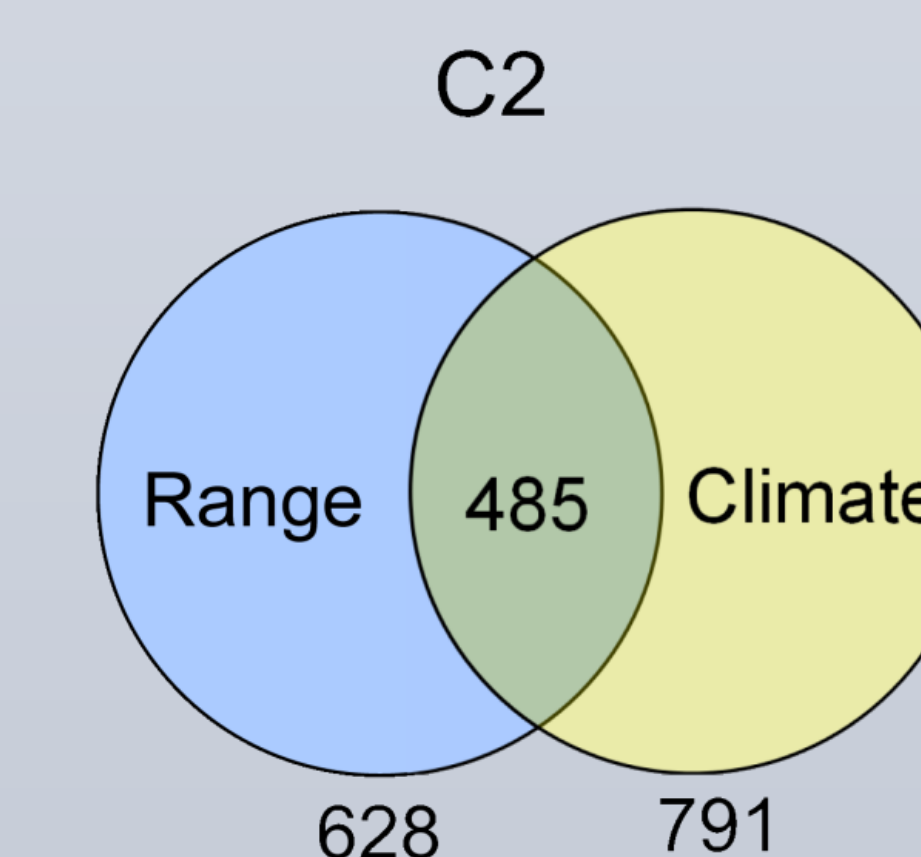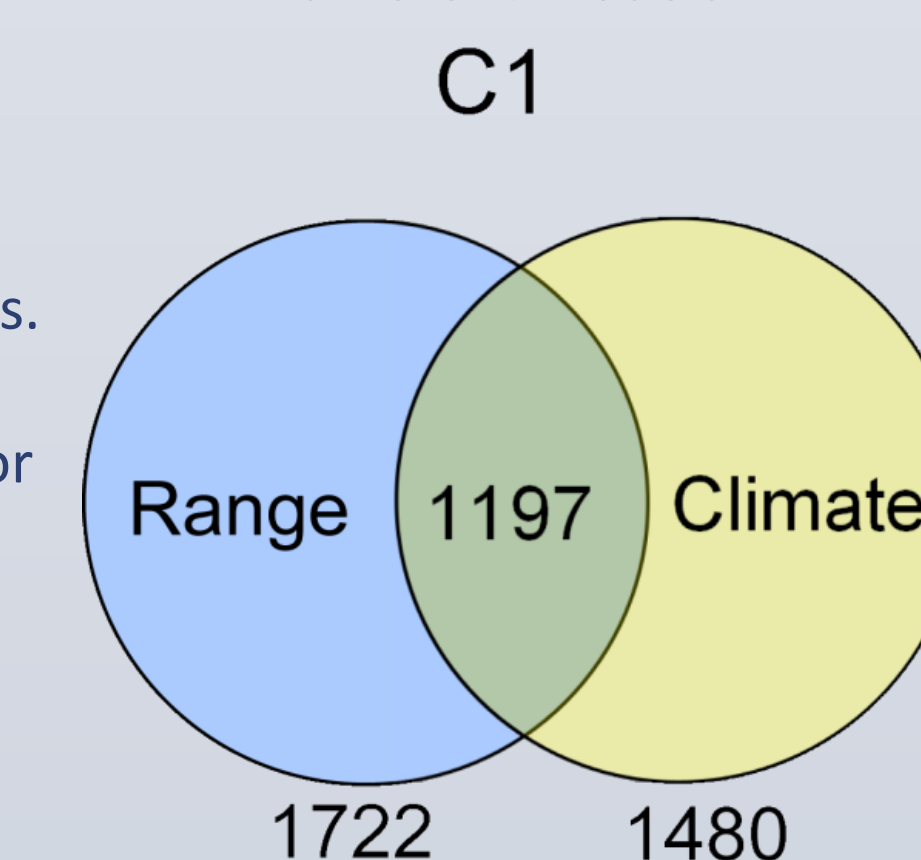
To see if climate and range are substantially confounded, we assessed the overlap between genes identified as significantly differently expressed. Though the two models identify roughly the same number of genes, the two gene lists do not fully overlap. The subset of genes which are shown to be effected by range, but not by climate, may represent evolution in the invasive range. However, the small number of populations in this data set (11) makes it difficult to tease apart the effect of climatic variation and range, per se.

C2
Range — 485 — Climate
628   791

CONCLUSION

Here we identified differentially expressed probes between native and invasive populations of ragweed. In this ragweed data set, we find little evidence of latitude being confounded with range. Unfortunately, there is substantial confounding between range and climate. More work is needed to try to remove climate effects, and focus on "invasiveness genes." This could included GO enrichment to try to interpret the function of these genes. Further, it would be interesting to look further into the unique Ontario population. Overall, differences in gene expression between native and invasive ranges cannot be fully described by the latitude and climate models we have used.

BIBLIOGRAPHY

1. Colautti RI, Maron JL, Barrett SCH. 2009. Common garden comparisons of native and introduced plant populations: latitudinal clines can obscure evolutionary inferences. *Evolutionary Applications* 2: 187–199.
2. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global areas. *International Journal of Climatology* 25.
3. Hodgins KA, Rieseberg L. 2011. Genetic differentiation in life-history traits of introduced and native common ragweed (Ambrosia artemisiifolia) populations. *Journal of Evolutionary Biology* 24: 2731–2749.
4. Lai Z, Kane NC, Kozik A, Hodgins KA, Dlugosch KM, Barker MS, Matvienko M, Yu Q, Turner KG, Pearl SA, et al. 2012. Genomics of Compositae Weeds: EST Libraries, Microarrays, and Evidence of Introgression. *American Journal of Botany* 99: 209–218.
5. Pimentel D, Zuniga R, Morrison D. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics* 52: 273–288.