The Gene Regulatory Landscape of Local Adaptation

Billie A. Gould[a,*], Yani Chen[a], David B. Lowry[a,b,c]

[a]Department of Plant Biology, Michigan State University, East Lansing, MI, USA

[b]Program in Ecology, Evolutionary Biology and Behavior, Michigan State University, East Lansing, MI 48824, USA

[c]Plant Resilience Institute, Michigan State University, East Lansing, MI 48824, USA

Local adaptation is a key driver of ecological specialization. Despite its importance, the evolution of gene regulatory divergence among locally-adapted populations is poorly understood. Here, we evaluate allele-specific gene expression responses for locally-adapted coastal and inland accessions of the plant *Mimulus guttatus* in a field reciprocal transplant experiment. We found 18% of transcripts were characterized by habitat-specific differential expression, a pattern consistent with the minimization of fitness trade-offs through expression plasticity. *Cis*-regulatory variation was pervasive, affecting 92% of differentially expressed genes. Chromosomal inversions were enriched for genes with expression divergence between ecotypes in a habitat-dependent manner, and gene expression heterosis was prevalent across habitats. These results provide multiple new insights into the evolution of transcriptome-wide gene regulatory divergence and plasticity among locally adapted populations.

47 One of the most important forces driving the evolution ecological specialization and the origin of
48 biodiversity on Earth is local adaptation (*1-3*). Local adaptation is characterized by reciprocal
49 home-site advantage, whereby populations perform best in their home habitat while performing
50 poorly in foreign habitats (*4-7*). Some case-studies, such as the evolution of beak size in
51 Darwin's Finches (*8*) and coat color of Beach Mice (*6*), suggest that local adaptation should
52 involve fitness trade-offs at the level of individual genes. However, studies that have combined
53 reciprocal transplant field experiments with quantitative trait locus (QTL) mapping have
54 observed that trade-offs at the individual locus level are rare (*7*). Despite recent progress, we still
55 know little about how transcriptome-wide gene regulatory divergence among locally adapted
56 populations contributes to fitness trade-offs in nature. To understand this important evolutionary
57 process, it is crucial to determine the prevalence of genotype x environment (GxE) interactions
58 affecting the expression of native and non-native alleles across the divergent habitats,
59 characterize the relative contributions of *cis* and/or *trans* regulatory variation to those
60 interactions, and establish the role of genome structure (e.g. inversions) in the evolution of
61 locally adaptive regulatory variation.
62
63 We conducted a field reciprocal transplant experiment in tandem with allele-specific expression
64 analysis to understand how the individual alleles of locally adapted coastal and inland
65 populations of the yellow monkeyflower, *Mimulus guttatus*, respond to divergent habitats (Fig.
66 1). There is a long history of local adaptation research in *M. guttatus.* Across previous reciprocal
67 transplant field experiments, inland plants have survived to flower at 23.9 times the rate of
68 coastal plants in inland habitat, while coastal plants have survived at 2.5 times the rate of inland
69 plants in coastal habitat (*9-11*). Many inland populations have evolved an early flowering annual

70 life-history to escape the harsh summer drought in inland habitats. In contrast, coastal
71 populations experience year-round soil moisture due to summer sea fog, but have had to adapt to
72 toxic oceanic salt spray (*10, 11*). We have previously shown that several regions of the genome,
73 including a chromosomal inversion polymorphism on chromosome 8, contribute to local
74 adaptation of coastal and inland populations (*12*). An additional inversion polymorphism on
75 chromosome 5 also appears to be differentiated between coastal and inland populations (*13*).
76
77 To evaluate allele-specific gene expression of native and foreign transplants across divergent
78 habitats, we planted replicates of an inbred coastal line (SS genotype), an inland line (LL
79 genotype), and their F1 hybrid (SL genotype) into one coastal and one inland field site in the
80 spring of 2016 (Fig. 1). We planted both gardens in early March and collected leaf tissue for
81 RNA sequencing at the end of April, just prior to the summer drought in inland habitat. We
82 aligned reads from field sampled individuals to reference transcriptomes constructed individually
83 for both inbred parental lines. To understand the effect of genotype and environment on gene
84 expression, we partitioned parental expression variation into genotype, site (habitat), and GxE
85 effects using linear models. We exploited allelic polymorphisms between parental lines to
86 quantify allele-specific expression of coastal and inland alleles in the F1 hybrids. This allowed us
87 to partition *cis* and *trans* effects and the interactions of those effects with habitat (site) (*14*).
88
89 Overall, we were able to evaluate gene expression for 10,122 transcripts corresponding to 9,326
90 genes, after transcriptome reconstruction and filtering. Significant expression differences (FDR =
91 0.05) between the inland and coastal field sites (6170 genes; 66%) and between parental
92 genotypes (6460 genes; 69%) were common (Table S1, S2). Most of the expression differences

93 between the coastal and inland lines could be attributed to *cis*-regulatory variation: 92% of DE
94 transcripts had *cis* regulatory differences, while only 39% had detectable trans effects (Fig. 2A
95 and 2B). On average, the magnitude of *cis*-regulatory effects (LFC = 0.291) were larger than
96 *trans* effects (LFC = 0.107; $P < 0.0001$).
97
98 Across the transcriptome we found 7700 (76%) transcripts were differentially expressed (DE)
99 between parental ecotypes in either one or both habitats (Fig. 3). A total of 1809 (18%)
100 transcripts had significant GxE interactions affecting expression (Fig. 3B). Only 94 transcripts
101 had crossing reaction norms (green points, Fig. 3). More often, expression differences between
102 genotypes were significant at one site and non-significant at the other (1137 transcripts, 11%,
103 blue and red points, Fig 3). When there is a cost to gene expression, GxE interactions may allow
104 genotypes a fitness advantage in one habitat without incurring a fitness trade-off in alternative
105 habitats. Similarly, a recent study of corals found that plasticity of gene regulation is important
106 for local adaptation in the field (*15*).
107
108 We used network clustering analysis to examine correlated expression of gene clusters (modules)
109 within habitats for the coastal and inland lines (cor > 0.75, Fig. S1). The coastal line
110 preferentially expressed modules related to root and epidermal development at its home site,
111 while oxidative stress response, chlorophyll metabolism, and circadian rhythm modules were
112 highly expressed in inland habitat (Table S3). The inland line preferentially expressed gene
113 clusters related to nucleoside catabolism, sugar and starch metabolism, pigment accumulation,
114 and photoperiodism at its home site, but had no modules strongly associated with the coastal site.
115

116 Chromosomal inversions are thought to act as adaptation supergenes that hold together long
117 haplotypes containing multiple adaptive polymorphisms through suppressed recombination (*16,*
118 *17*). We thus tested the hypothesis that inversions hold together haplotypes of DE genes. There
119 were significantly more DE genes inside the chromosome 5 inversion than expected by chance
120 and there were significantly more GxE interactions for genes in this inversion (Tables S1). The
121 chromosome 8 inversion also had significantly more DE genes but only when expression was
122 measured at the coastal site (Table S4). The enrichment of gene expression differences is
123 consistent with patterns found in inversions between chimpanzees and humans (*18*) and
124 *Drosophila* (*19*), but contrasts with results for the plant *Boechera stricta* (*20*). Our results are
125 unique among these studies, as they demonstrate supergene associated expression divergence
126 that is dependent on habitat context (GxE).
127
128 Heterosis (hybrid vigor) is commonly found in crosses among locally-adapted populations and
129 can potentially facilitate gene flow between those populations (*21, 22*). Heterosis is typically
130 quantified at the level of an individual trait, including expression traits (15, 23), by comparing
131 the trait value of hybrids to the trait values of the parents (*23*). In this and a previous study, we
132 have found robust heterosis for multiple phenotypic traits in F1 hybrids between coastal and
133 inland lines (Fig. 1F; Lowry et al. 2008). In the current study we also found nearly 73% of DE
134 transcripts had heterotic (non-additive) expression in the F1 (Fig. 2D, Table S5). The percentage
135 of transcripts with heterotic expression did not vary strongly by site (Fig 2D), although only 54%
136 of heterotic genes were shared between sites. Heterosis often causes increased growth and vigor
137 (21). Consistent with this, heterotic genes were enriched for functional terms associated with
138 growth and development at both field sites (Table S1).

139

140 To identify genes putatively linked to local adaptation, we cross-referenced our results with a
141 recent genomic outlier analysis, which identified 667 genes with a high level of allelic
142 differentiation between coastal and inland populations (*13*). We found that the magnitude of
143 differential expression between lines and between field sites was significantly elevated for
144 candidate genes with highly differentiated SNPs in their promoter regions (Table S6). We were
145 able to examine gene expression in seven top candidate genes highlighted by the outlier analysis
146 and/or previous QTL mapping studies (Fig. 2E-G, Fig. S2). The candidate gene *ABA1* mediates
147 production of abscisic acid, which controls responses to osmotic stress, including drought (*24*).
148 An *ABA1* homolog, Migut.H00431, is located within the chromosome 8 inversion and may
149 contribute to inland drought adaptation. *ABA1* was always more highly expressed in the inland
150 line, and also upregulated in both lines at the inland field site (Fig. 2E). Previously, we showed
151 that coastal *M. guttatus* populations have evolved leaf tissue tolerance to oceanic salt spray (*14*).
152 Leaf salt tolerance is generally mediated by a Na^+/H^+ antiporter (*SOS1*), which is activated by a
153 calcium sensor (*CBL10*; *25*). Both *SOS1* (Migut.E00570) and *CBL10* (Migut.A00138) were more
154 highly expressed in the coastal line than the inland line across habitats (Fig. 2G, F), a pattern
155 consistent with elevated salt tolerance of coastal populations (14). However, because expression
156 of *SOS1* does not appear to be plastic in either ecotype it may be a major contributor to the
157 fitness trade-offs observed between habitats (Fig. 2F).

158

159 In summary, our results demonstrated multiple key patterns regarding the relationship between
160 gene expression and the evolution of local adaptation. Differential expression is substantial
161 between locally-adapted populations in nature and the vast majority of that differential

162 expression is due to *cis*-regulatory divergence. Chromosomal inversions appear to act as
163 supergenes by holding together haplotypes of DE genes, but this pattern depends on habitat
164 context. Heterosis of gene expression is pervasive and heterotic genes are enriched for growth
165 and developmental functions. Expression plasticity of alleles across habitats (GxE) appears to be
166 common and may minimize fitness trade-offs at individual loci that contribute to the overall
167 pattern of local adaptation.
168

169 **REFERENCES AND NOTES**

170 1. J. Clausen (Cornell University Press, Ithaca, NY, 1951).
171 2. P. F. Colosimo *et al.*, *Science* **307**, 1928-1933 (2005).
172 3. J. M. Sobel, G. F. Chen, L. R. Watt, D. W. Schemske, *Evolution* **64**, 295-315 (2010).
173 4. T. J. Kawecki, D. Ebert, *Ecol. Lett.* **7**, 1225-1241 (2004).
174 5. J. Hereford, *Am. Nat.* **173**, 579-588 (2009).
175 6. H. E. Hoekstra, R. J. Hirschmann, R. A. Bundey, P.A. Insel, J. P. Crossland, *Science* **313**, 101-
176 104 (2006).
177 7. S. Wadgymar *et al.*, *Methods Ecol. Evol. In press* (2017).
178 8. A. Abzhanov, M. Protas, B. R. Grant, P. R. Grant, C. J. Tabin, *Science* **305**, 1462-1465
179 (2004).
180 9. M. C. Hall, J. H. Willis, *Evolution* **60**, 2466-2477 (2006).
181 10. D. B. Lowry, R. C. Rockwood, J. H. Willis, *Evolution* **62**, 2196-2214 (2008).
182 11. D. B. Lowry, M. C. Hall, D. E. Salt, J. H. Willis, *New Phytol*. **183**, 776-788 (2009).
183 12. D. B. Lowry, J. H. Willis, *PLOS Biol.* **8**, e1000500 (2010).
184 13. B. A. Gould, Y. Chen, D. B. Lowry, *Mol. Ecol.* **26**, 163-177 (2017).

185 14. J. T. Lovell *et al.*, *Genome Res.* **26**, 510-518 (2016).

186 15. C. D. Kenkel, M. V. Matz, *Nat. Ecol. Evol.* **1**, 0014 (2016).

187 16. M. Kirkpatrick, *PLOS Biol.* **8**, e1000501 (2010).

188 17. M. Joron *et al. Nature* **477**, 203-206 (2011).

189 18. T. Marquès-Bonet *et al.*, *Trends. Genet.* **20**, 524-529 (2004).

190 19. Z. L. Fuller, G. D. Haynes, S. Richards, S. W. Schaeffer, *Genetics* **204**, 287-301 (2016).

191 20. C. R. Lee *et al.*, *Nat. Ecol. Evol.* **1**, 0119 (2017).

192 21. D. Charlesworth, J. H. Willis, *Nat. Rev. Genet.* **10**, 783-796 (2009).

193 22. E. Baack, M. C. Melo, L. H. Rieseberg, D. Ortiz-Barrientos, *New Phytol.* **207**, 968-984

194 (2016).

195 23. N. M. Springer, R. M. Stupar, *Plant Cell* **19**, 2391-2402 (2007).

196 24. L. Xiong, H. Lee, M. Ishitani, J. K. Zhu, *J. Biol. Chem.* **277**, 8588–8596 (2002).

197 25. H. Ji *et al.*, *Mol. Plant.* **6**, 275-286 (2013).

198

199

200

201 **ACKNOWLEDGMENTS**

202

203 We thank D. Des Marais, J. Tung, J. Friedman, R. Guerrero for comments that helped to improve

204 our manuscript. We thank K. Childs for guidance with expression analyses, B. Blackman and E.

205 Patterson for starting seeds for us at UC Berkeley, and D. Popovic for assistance in field

206 planting. The Bodega Marine Reserve and the Pepperwood Preserve kindly allowed us to

207 conduct this reciprocal transplant experiment. Funding was provided by Michigan State

University. Sequence data are deposited in the Short Read Archive (XX). The authors declare no conflicts of interest.

SUPPLEMENTARY MATERIALS

Materials and Methods

Figs. S1 to S9

Tables S1 to S6

References (26-46)

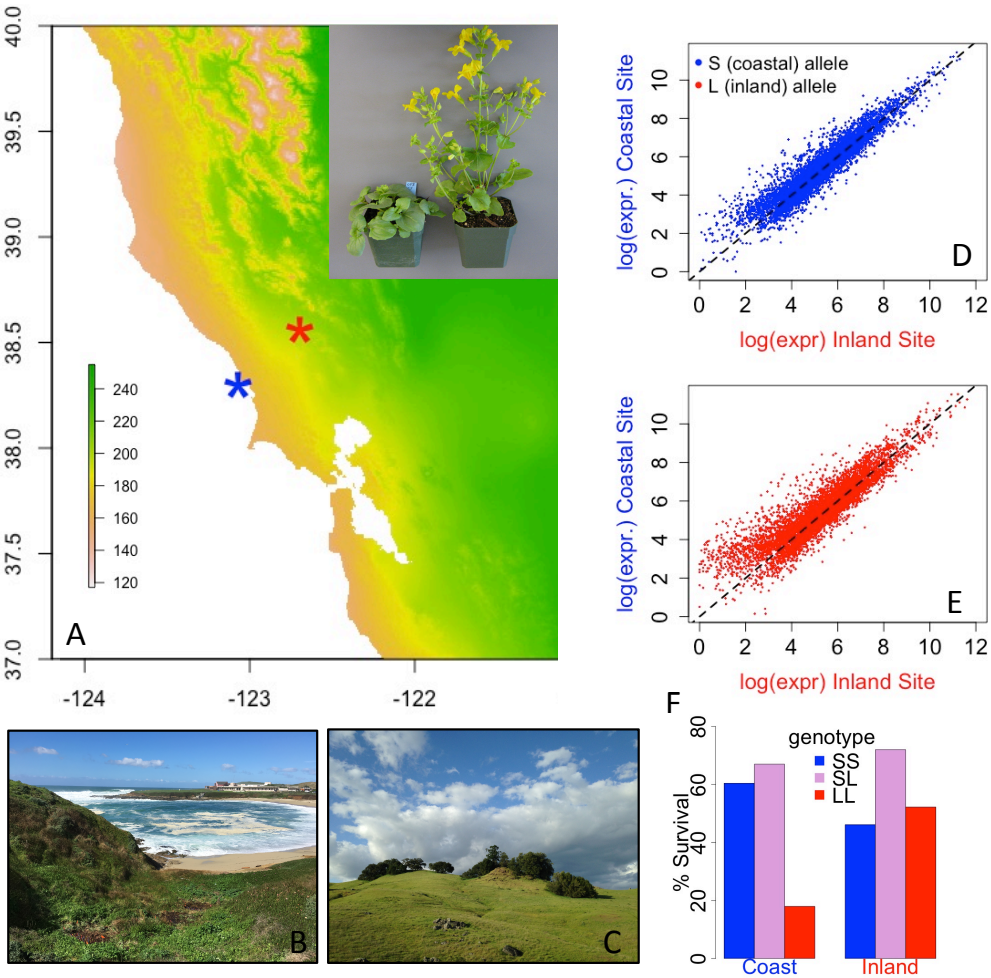


Fig. 1 A) Map of Northern California with the location of the two field transplant sites. Colors represent mean temperature of the warmest quarter in degrees C x 10. Inset: left, coastal ecotype; right, inland ecotype. B) The coastal field site. C) The inland field site. D) SS parental gene expression at two field sites. E) LL parental gene expression at two field sites.F) Survival of parental and F1 (SL) plants at the two field sites.

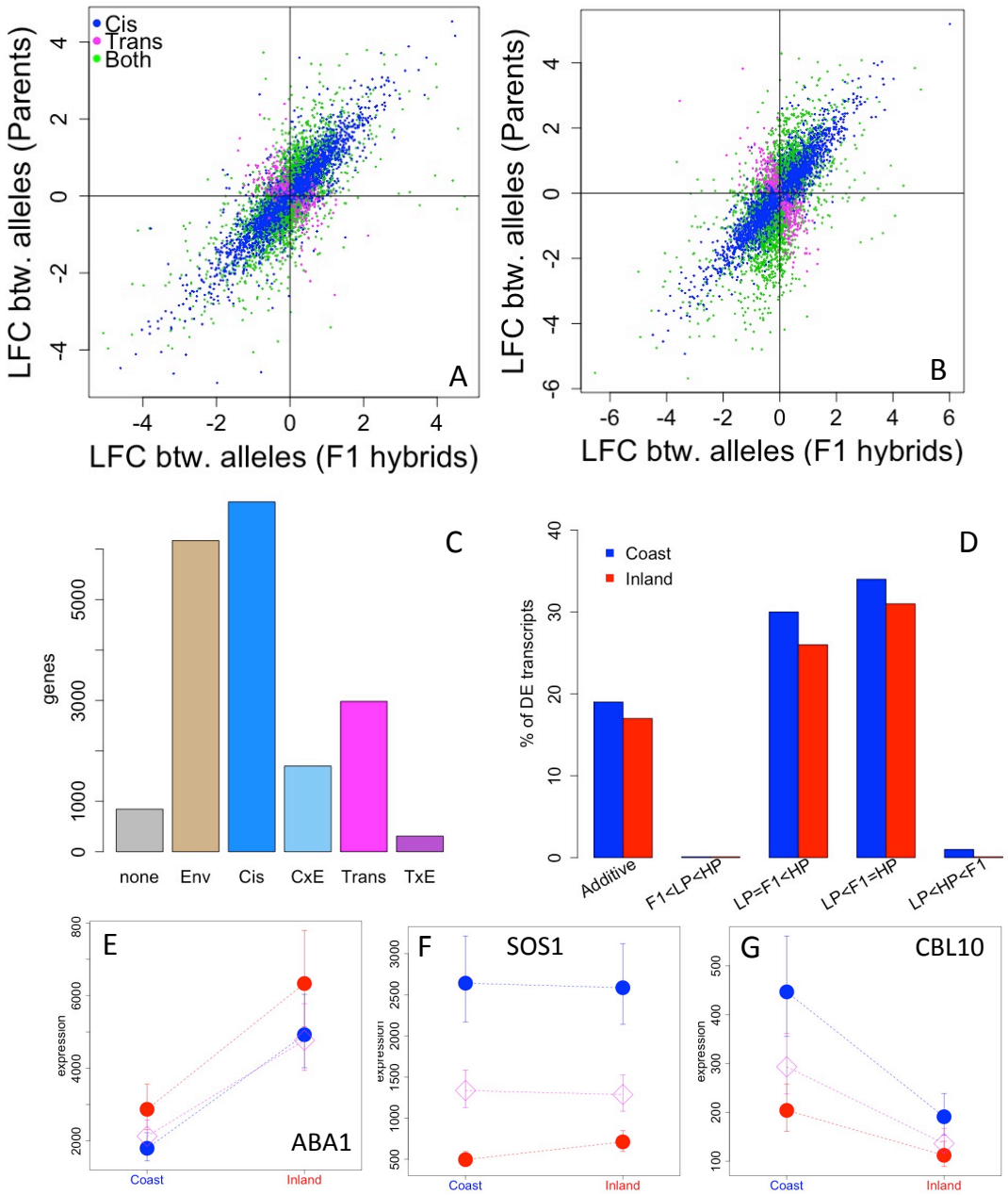
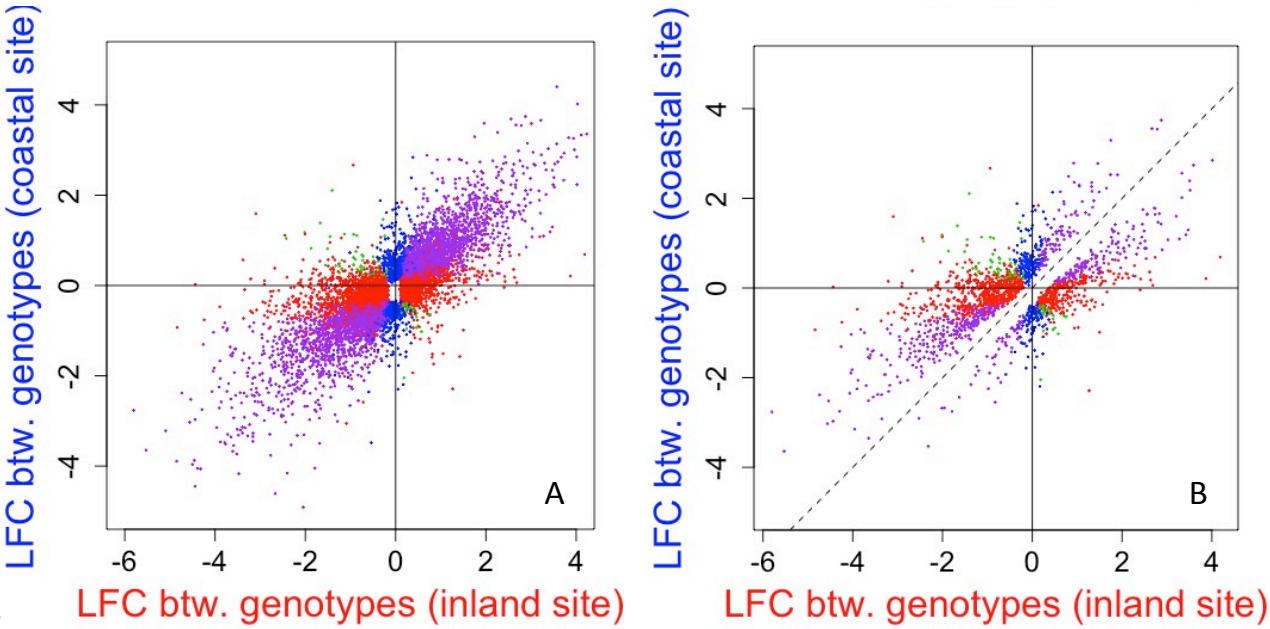
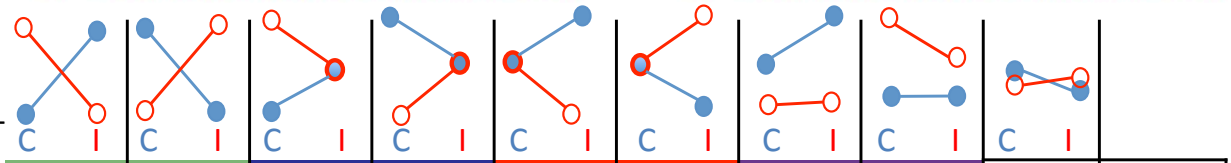


Fig.2. *Cis* and *trans* regulated transcripts detected at the coastal (A) and inland (B) field sites. Average log2-fold change (LFC) ratios are between inland and coastal alleles in the parental (y-axis) and F1 (x-axis) genetic backgrounds. C) The frequency of significant model effects among expressed genes. D) The frequency of expression heterosis in F1 plants at the inland and coastal field sites. LP, low parent; HP, high parent. E-G) Reaction norms of expression at the two field sites for three candidate adaptation genes. Red, inland parent; Blue, coastal parent; Pink, F1 .



C



point color									N/A	total
allele effect	+/-	-/+	+/0	-/0	0/-	0/+	-/-	+/+	0/0	
# transcript	63	31	412	268	1776	1405	1690	2055	2422	10122
# signif. GxE	63	31	158	105	584	290	333	245	28	1837

Fig. 3 A) Average log2 fold-change (LFC) differences between parental genotypes for expressed transcripts at two field sites. B) LFC differences only for transcripts with significant GxE interactions. C) Expression reaction norms corresponding to colored points in A and B. Row 3 shows the sign of the allelic effect of the L (inland) allele in the coastal and inland habitats. Row 4 lists the number of transcripts with significant allelic affects within sites in the specified direction. Row 5 lists the number of transcripts with significant GxE.