Pre-Breeding Climate-Resilient Tree Crops

Andrés J Cortés[1,2]*

[1]Corporación Colombiana de Investigación Agropecuaria - AGROSAVIA, CI La Selva, Km 7 vía Rionegro - Las Palmas, Rionegro, Colombia

[2]Facultad de Ciencias Agrarias - Departamento de Ciencias Forestales, Universidad Nacional de Colombia - Sede Medellín, Medellín, Colombia

*Corresponding author: Andrés J Cortés, Corporación Colombiana de Investigación Agropecuaria - AGROSAVIA, CI La Selva, Km 7 vía Rionegro - Las Palmas, Rionegro, Facultad de Ciencias Agrarias and Departamento de Ciencias Forestales, Universidad Nacional de Colombia - Sede Medellín, Medellín, Colombia

Abstract

The prevalence of tree crops vulnerable to drought and heat stress is on the rise as a consequence of climate change. Therefore, there is an urgent need for the development of innovative analytical approaches to identify the origins of natural tree adaptation and integrate them into the improvement of sustainable populations, with the ultimate goal of ensuring their resilience. In this invited comment to Agricultural Research & Technology, I envision and exemplify how modeling analytical methods via genome-environment associations might be utilized to forecast tree adaptive responses in their natural distribution ranges to better face ongoing climate change. This theoretical exercise will contribute to the advancement of sustainable tree production in the face of extreme weather patterns.

Keywords: Polygenic adaptation; Genome-Environment Associations; Plant genomics; Tree breeding

Editorial

Rethinking Genome-Environment Associations for Tree Species

Tree populations contain natural distinct responses to abiotic stressors, which are seldom found in more advanced breeding generations [1]. Nevertheless, the realization of this potential has proven to be difficult, primarily due to the limitations of phenotyping in natural environments, the highly admixed nature of pedigrees in the wild, and the intricate inheritance patterns associated with key trait variation. For instance, adaptation would often involve multiple loci with infinitesimal effects hard to detect [2]. In order to bridge these limitations, Genome-Environment Association (GEA) inferences inspired in classical Genome-Wide Association Study (GWAS) models [3] could be employed as a way to disclose and leverage naturally existent Linkage Disequilibrium (LD) with bioclimatic variables across tree populations.

Aiming for the integration of GEA models with more empirical estimations of the adaptive potential may also boost predictive improvement for adaptation in tree crops [4]. The utilization of modern genomic evolutionary analysis could further assist interpreting the phenomena of evolutionary conservatism, parallelism, and convergence as part of the genetic foundations governing adaptation to diverse forms of abiotic stresses across a broad spectrum of environments. These factors include drivers summarized by Ellegren and Galtier [5] like ecological disruptive and background selection, gene flow, shared ancestral polymorphism, and variation in mutation and recombination rates. Such processes have been explored in more detailed in other works [6-9], and specifically within the context of GEA predictions [10-16]. After all, uncovering the ultimate molecular and evolutionary mechanisms underlying ecological divergence and its plethora of genomic determinants requires contrasting selection forces with those from demography and randomness.

In this regard, research on tree crops that utilize the GEA paradigm can provide significant confirmations of classical molecular genetics' hypotheses in long-lived species by shedding light on the causals of naturally convergent adaptations. For example, is probable that molecular markers located near significant ecological-associated Quantitative Trait Loci (QTLs) may experience hitchhiking effects [7], as result of a low recombination rate and considerable linkage disequilibrium [17,18], which in turn are expected to be more pronounced

towards centromeric regions. It is important to acknowledge these factors in the study of extensive LD and hitchhiking effects, the ultimate proxies for GEA estimations, even if they are assumed to be minimal in perennial tree crop species with long-distance dispersal and overlapping generation times [19]. Last but not least, advancing GEA studies into the molecular mechanisms of tree evolutionary ecological divergence will require the functional genetic dissection of putative adaptive loci, including the validation of candidate genes, their segmentation from intrinsic genomic features such as adaptive haplotype blocks, and the transgenesis and genetic edition of regulatory mutations that contribute to the genomic basis of adaptation.

Shifting from Estimation to Prediction of the Adaptive Potential

The implementation of predictive models from the GEA results can be impaired by disregarding the choice of using the complete set of available Single Nucleotide Polymorphism (SNP) genetic markers, in favor of selecting the most predictive SNP dataset according to the GWAS-type analysis [20]. The very same property that makes the polygenic adaptation infinitesimal model unique may be compromised when sub-setting the SNP pool. Ergo, relying solely on SNP markers derived from the GEA inferences will overlook SNPs with minor effects, which are typically undetected by GWAS-type models. Yet, these SNPs may still contribute to the overall unexplained heritability of the adaptive traits.

More effective strategies to utilize alleles responsible for adaptation in tree species include: (1) employing weighted Genomic Prediction (GP) models [21] that incorporate GWAS-based GEA estimates [22], and (2) optimizing the selection of markers by generating saturation curves of the predictive ability using different marker sets ranked according to their beta effects derived from the GP model suggested in (1), rather than relying on estimations obtained only from GWAS-inspired models [23,24]. In the context of tree pre-breeding targets focused on abiotic tolerance, the optimization of marker sets for environmental genotyping might prove advantageous [25]. Such optimization is expected to merge the GEA thinking with speed breeding tactics, as suggested by Migicovsky and Myles [26] and Watson et al. [27]. Marker optimization has historically been used to enhance the efficiency of high-throughput screenings. In addition, there is a current shift in the field of GP from utilizing Best Linear Unbiased Prediction (BLUP)-type models to adopting machine learning (ML) methodologies [28-30].

Finally, climate-based ecophysiological modeling is also promising to categorize tree genotypes based on the stress gradient of the geographic regions from which they originate. For that, climatic data is gathered by utilizing the georeferencing at the time of sample collection, as original highlighted by Hancock et al. [31], but latter on transformed into more biologically meaningful scores [32]. Furthermore, not all tree samples possess the necessary characteristics to effectively calibrate climate-based ecophysiological models from raw bioclimatic variables. Therefore, it is recommended to place emphasis on specific abiotic stress types and resample genotypes accordingly in a bulk-segregant manner.

While it is acknowledged that GEA, GP, ML and speed breeding have proven to be valuable methods for the improvement and deployment of customized adaptive variants [33,34], specially in tree crops [35], it is yet important to note that these techniques are innovations in predictive modeling that may not directly align with methodological advancements in the field [36]. To compensate, *in silico* pre- and breeding models can be advantageous in the optimization of these varietal development tactics. *In silico* breeding simulates and modifies the selection intensity in order to minimize genetic erosion and cycle length, while ensuring that overall genetic gains are not compromised [37].

Concluding Remarks

In essence, GEA enhances our comprehension of the intricate dynamics between tree genomes and their surrounding environment, hence influencing the long-term selection and genetic gain of adaptive phenotypic traits [35]. The mechanistic understanding of how the genomes of perennial tree species function in natural environments holds great potential for utilizing existing standing adaptation to support pre-breeding initiatives targeting drought [38-40] and heat stress [41,42]. This can be achieved through the use of multi-dimensional adaptive scores, which provide a comprehensive assessment of tree adaptability [43], even for biotic antagonistic agents [44,45], while balancing more classical improvement targets in perennials [46,47]. Additionally, GEA helps identifying the factors that either enable or limit future adaptive responses to evolving environmental conditions.

Acknowledgment

References

1. Tester M, Langridge P (2010) Breeding Technologies to Increase Crop Production in a Changing World. Science 327(5967): 818-822.

2. Cortés AJ, López H F, Osorio RD (2020a) Predicting thermal adaptation by looking into populations' genomic past. Front Genet 11: 564515.

3. Morton NE (2005) Linkage disequilibrium maps and association mapping. The Journal of Clinical Investigation 115(6): 1425-1430.

4. Cortés AJ, López HF, Blair MW (2022) Genome-Environment Associations, an Innovative Tool for Studying Heritable Evolutionary Adaptation in Orphan Crops and Wild Relatives. Frontiers in Genetics 13: 910386.

5. Cañas GGP, Sepulveda OS, López H F, Navas AAA, Cortés AJ (2022) Inheritance of Yield Components and Morphological Traits in Avocado cv. Hass From "Criollo" "Elite Trees" via Half-Sib Seedling Rootstocks. Frontiers in Plant Science 13: 843099.

6. Ellegren H, Galtier N (2016) Determinants of genetic diversity. Nat Rev Genet 17: 422-433.

7. Ellegren H, Wolf JBW (2017) Parallelism in genomic landscapes of differentiation, conserved genomic features and the role of linked selection. J Evol Biol 30(8): 1516-1518.

8. Feder JL, Gejji R, Yeaman S, Nosil P (2012) Establishment of new mutations under divergence and genome hitchhiking. Philos Trans R Soc Lond B Biol Sci 367(1587): 461-474.

9. Gompert Z, Comeault AA, Farkas TE, Feder JL, Parchman TL, et al. (2014) Experimental evidence for ecological selection on genome variation in the wild. Ecology Letters 17(3): 369-379.

10. Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, et al. (2017) Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. J Evol Biol 30(8): 1450-1477.

11. Buitrago BMA, Cortés AJ, López HF, Londoño CJM, Muñoz FJE, et al. (2021) Allelic Diversity at Abiotic Stress Responsive Genes in Relationship to Ecological Drought Indices for Cultivated Tepary Bean, Phaseolus acutifolius A. Gray, and Its Wild Relatives. Genes 12(4): 556.

12. Cortés AJ, Blair MW (2018) Genotyping by Sequencing and Genome - Environment Associations in Wild Common Bean Predict Widespread Divergent Adaptation to Drought. Frontiers in Plant Science 9: 128.

13. Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2016) Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. Mol Ecol 25(1): 104-120.

14. López HF, Cortés AJ (2019) Last-Generation Genome-Environment Associations Reveal the Genetic Basis of Heat Tolerance in Common Bean (*Phaseolus vulgaris* L.). Front Genet 10: 22.

15. Pluess AR, Frank A, Heiri C, Lalague H, Vendramin GG, et al. (2016) Genome-environment association study suggests local adaptation to climate at the regional scale in Fagus sylvatica. New Phytol 210(2): 589-601.

16. Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. Mol Ecol 24(17): 4348-4370.

17. Waldvogel AM, Feldmeyer B, Rolshausen G, Exposito AM, Rellstab C, et al. (2020) Evolutionary genomics can improve prediction of species' responses to climate change. Evol Lett 4(1): 4-18.

18. Blair MW, Cortes AJ, Farmer AD, Huang W, Ambachew D, et al. (2018) Uneven recombination rate and linkage disequilibrium across a reference SNP map for common bean (*Phaseolus vulgaris* L.). PLoS One 13(3): e0189597.

19. Kelleher CT, Wilkin J, Zhuang J, Cortés AJ, Quintero ÁLP, et al. (2012) SNP discovery, gene diversity, and linkage disequilibrium in wild populations of Populus tremuloides. Tree Genetics & Genomes 8: 821-829.

20. Berdugo CJA, Cortés AJ, López HF, Delgadillo P, Cerón SI, et al. (2023) Pleistocene dispersion supports a unique native diversity in the Colombian avocado germplasm. bioRxiv 01.27.525883.

21. Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, et al. (2013) Pitfalls of predicting complex traits from SNPs. Nat Rev Genet 14(7): 507-515.

22. Arenas S, Cortés AJ, Mastretta YA, Jaramillo CJP (2021) Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations. Tree Genetics & Genomes 17.

23. Spindel JE, Begum H, Akdemir D, Collard B, Redona E, et al. (2016) Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. Heredity (Edinb) 116(4): 395-408.

24. Grattapaglia D, Silva-Junior OB, Resende RT, Cappa EP, Muller BSF, et al. (2018) Quantitative Genetics and Genomics Converge to Accelerate Forest Tree Breeding. Front Plant Sci 9: 1693.

25. Resende MDV, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, et al. (2012) Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytologist 194(1): 116-128.

26. Cortés AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics 123(5): 827-845.

27. Migicovsky Z, Myles S (2017) Exploiting Wild Relatives for Genomics-assisted Breeding of Perennial Crops. Front Plant Sci 8: 460.

28. Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J, et al. (2018) Speed breeding is a powerful tool to accelerate crop research and breeding. Nature Plants 4(1): 23-29.

29. Cortés AJ, López HF (2021) Harnessing Crop Wild Diversity for Climate Change Adaptation. Genes 12(5): 783.

30. Crossa J, Martini JWR, Gianola D, Perez RP, Jarquin D, et al. (2019) Deep Kernel and Deep Learning for Genome-Based Prediction of Single Traits in Multienvironment Breeding Trials. Front Genet 10: 1168.

31. Montesinos LOA, Montesinos LA, Mosqueda GBA, Montesinos LJC, Crossa J, et al. (2021) A zero altered Poisson random forest model for genomic-enabled prediction. G3 11(2): jkaa057.

32. Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, et al. (2011) Adaptation to Climate Across the Arabidopsis thaliana Genome. Science 334(6052): 83-86.

33. Cortés AJ, Monserrate F, Ramírez VJ, Madriñán S, Blair MW (2013) Drought Tolerance in Wild Plant Populations: the Case of Common Beans (*Phaseolus vulgaris* L.). Plos One 8(5): e62898.

34. Blair MW, Cortés AJ, Penmetsa RV, Farmer A, Carrasquilla GN, et al. (2013) A high-throughput SNP marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics 126: 535-548.

35. Varshney RK, Barmukh R, Roorkiwal M, Qi Y, Kholova J, et al. (2021) Breeding custom-designed crops for improved drought adaptation. Advanced Genetics 2(3): e202100017.

36. Cortés AJ, Restrepo MM, Bedoya CLE (2020b) Modern Strategies to Assess and Breed Forest Tree Adaptation to Changing Climate. Front Plant Sci 11: 583323.

37. Wu X, Islam ASMF, Limpot N, Mackasmiel L, Mierzwa J, Cortés AJ, et al. (2020) Genome-Wide SNP Identification and Association Mapping for Seed Mineral Concentration in Mung Bean (Vigna radiata L.). Frontiers in Genetics 11: 656.

38. Hoyos VV, Arief VN, Yang WH, Sun M, DeLacy IH, et al. (2019) QuLinePlus: extending plant breeding strategy and genetic model simulation to cross-pollinated populations-case studies in forage breeding. Heredity 122(5): 684-695.

39. Blair MW, Cortés AJ, This D (2016) Identification of an ERECTA gene and its drought adaptation associations with wild and cultivated common bean. Plant Science 242: 250-259.

40. Cortés AJ, Chavarro MC, Madriñán S, This D, Blair MW (2012a) Molecular ecology and selection in the drought-related Asr gene polymorphisms in wild and cultivated common bean (*Phaseolus vulgaris* L.). BMC Genetics 13: 58.

41. Cortés AJ, This D, Chavarro C, Madriñán S, Blair MW (2012b) Nucleotide diversity patterns at the drought-related DREB2 encoding genes in wild and cultivated common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics 125(5): 1069-1085.

42. Burbano EE, León PR, Cordero CC, López HF, Cortés A, et al. (2021) Multi-Environment Yield Components in Advanced Common Bean (*Phaseolus vulgaris* L.) × Tepary Bean (*P. acutifolius* A. Gray) Interspecific Lines for Heat and Drought Tolerance. Agronomy 11(10): 1978.

43. López HF, Burbano EE, León PRI, Cordero CCC, Villanueva MDF, et al. (2023) Multi-Environment Genome-Wide Association Studies of Yield Traits in Common Bean (*Phaseolus vulgaris* L.) × Tepary Bean (*P. acutifolius* A. Gray) Interspecific Advanced Lines in Humid and Dry Colombian Caribbean Subregions. Agronomy 13: 1396.

44. Valencia JB, Mesa J, León JG, Madriñán S, Cortés AJ (2020) Climate Vulnerability Assessment of the Espeletia Complex on Páramo Sky Islands in the Northern Andes. Frontiers in Ecology and Evolution 8.

45. Fernández PJ, Cortés AJ, Hernández VCA, Mejía-de-Tafur MS, Rodriguez MC, et al. (2021) Rootstock-Mediated Genetic Variance in Cadmium Uptake by Juvenile Cacao (*Theobroma cacao* L.) Genotypes, and Its Effect on Growth and Physiology. Frontiers in Plant Science 12: 777842.

46. Guevara EM, Osorio AN, Cortés AJ (2021) Integrative Pre-Breeding for Biotic Resistance in Forest Trees. Plants 10(10): 2022.

47. Reyes HPH, Muñoz BL, Velásquez ZV, Patiño L, Delgado POA, et al. (2020) Inheritance of Rootstock Effects in Avocado (Persea americana Mill.) cv. Hass. Front Plant Sci 11: 555071.