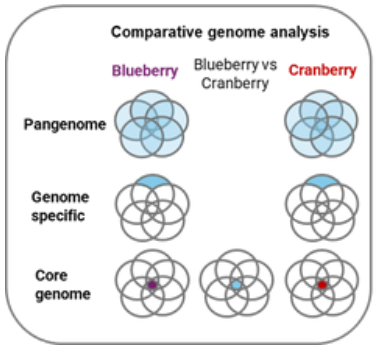
VacCAP
IMPROVING FRUIT QUALITY
ISSUE 10 | JULY 2024

GENOMICS AND GENOTYPING

VacCAP Objective 1 aimed to establish genomic resources to enable effective association mapping studies in blueberry and cranberry



Develop *Vaccinium* pangenome

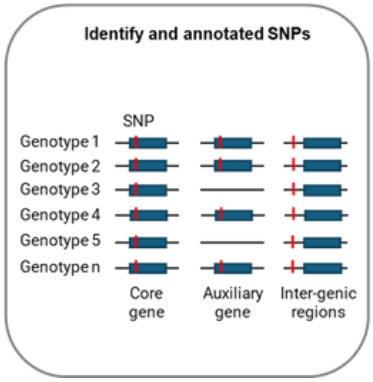
Comparative Genomic Team

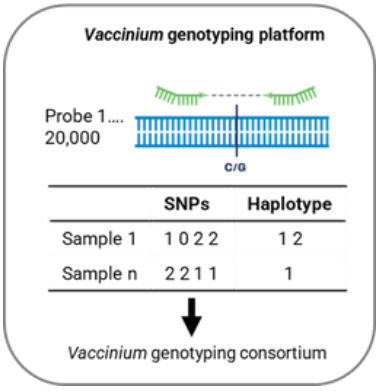
Assemble new blueberry and cranberry genomes for comparative analysis and to construct the Vaccinium pangenome. Identify the core and auxiliary genes. Core gene are conserved across genomes, auxiliary genes are not-conserved and present only in some genomes.

Develop a SNP catalog that combines de-novo with existing SNP sets (within linkage maps or representing QTLs). Annotate the SNPs catalog with SNP location within genes, core or auxiliary genes. This approach will ensure the identification of highly informative SNPs.

Compile a SNP catalog

Genotyping Team





Develop the *Vaccinium* Genotyping Platform

Genotyping Team

Select DNA regions surrounding SNPs to design a genotyping platform. Criteria for SNP selection aim to maximize the representation of core genes, SNPs associated with QTLs and even distribution. Engage *Vaccinium* breeders and geneticist to establish a genotyping consortium that will help to lower the genotyping costs per sample, while ensuring use of the platforms.

ACCOMPLISHMENTS IN GENOMICS AND GENOTYPING

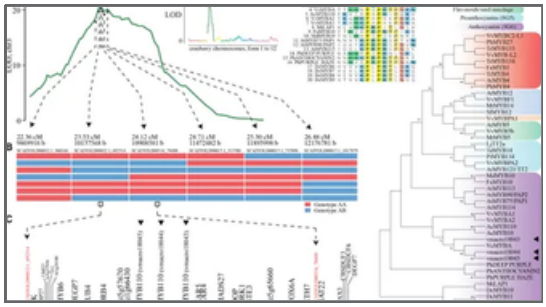
OUTCOMES ARE

- Released **5** linkages maps, **6** genomes and the *Vaccinium* pangenome
- Developed **2** genotyping platforms, established and coordinated the *Vaccinium* Genotyping Consortium
- New genomic resources were made available to the *Vaccinium* and broader research community through The Genome Database for *Vaccinium* (GDV)

EXPANDING GENOMIC RESOURCES

*Reference genomes for two blueberry wild species (*Vaccinium myrtillus *and* V. caesariense*), two cultivated cranberry cultivars ('Stevens' and 'Ben Lear') and two cranberry wild species (*V. microcarpum *and* V. oxycoccos*) were assembled and annotated. In addition, a blueberry and cranberry pangenome was developed.*

These reference genomes were assembled using long read sequencing technology which improved the quality of genomic resources available for the *Vaccinium* community. To build the blueberry and cranberry pangenomes, **20** blueberry and **10** cranberry genomes representing the Northern Highbush (NHB), Southern Highbush (SHB) and cranberry cultivars were resequenced at high coverage and used to build the Vaccinium pangenome. Comparative genome analysis identified the core and auxiliary genes. Core genes are conserved across genomes, auxiliary genes are not-conserved and present only in some genomes. Highlights from these studies are reported below.



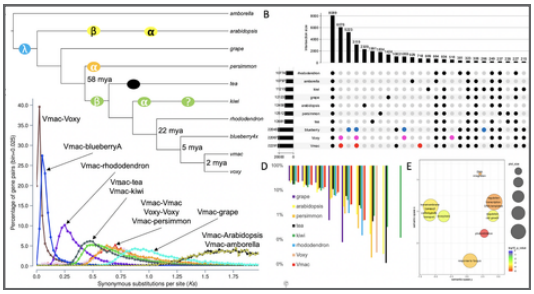
Chromosome-Level Genome Assembly of the American cranberry (*Vaccinium macrocarpon ait.*) and Its Wild Relative *Vaccinium microcarpum*

LINK TO PAPER

In this study Dr. Zalapa's team developed the first chromosome-scale genome assembly of cranberry, cultivar Stevens, and a draft genome of its close wild relative species *Vaccinium microcarpum*. More than 92% of the estimated cranberry genome size (492 Mb) was assembled into 12 chromosomes, which enabled gene model prediction and chromosome-level comparative genomics. Comparative genome analysis revealed two polyploidization events, the ancient γ-triplication, and a more recent whole genome duplication that occurred approximately 61 Mya.

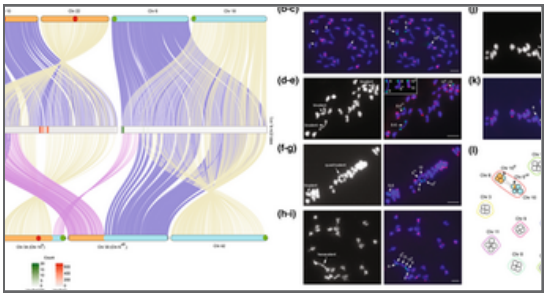
Furthermore, comparative genomics within the *Vaccinium* genus suggested cranberry-*V. microcarpum* divergence occurred 4.5 Mya, following their divergence from blueberry 10.4 Mya. Finally, a cluster of subgroup-6 R2R3 MYB transcription factors were identified within a genomic region spanning a large QTL for anthocyanin variation in cranberry fruit. Phylogenetic analysis suggested these genes likely act as anthocyanin biosynthesis regulators in cranberry. These new cranberry genomic resources facilitate the dissection of the genetic mechanisms governing agronomic traits and further breeding efforts at the molecular level.

Contrasting a reference cranberry genome to a crop wild relative provides insights into adaptation, domestication, and breeding



LINK TO PAPER

In this study Dr. Polashock's and Dr. Michael's teams presented an updated, chromosome-resolved *V. macrocarpon cv* 'Ben Lear' reference genome and the draft genome of the wild cranberry *V. oxycoccos*. The 'Ben Lear' genome assembly improved contiguity of the assembly at contig level. The study confirmed that the Ericaceae has undergone two whole genome duplications that are shared with blueberry and rhododendron. Leveraging resequencing data for 'Ben Lear' inbred lines, as well as several wild and elite selections, common regions that are targets of improvement were identified. These same syntenic regions in *V. oxycoccos*, harbored genes involved in environmental response and plant architecture. The study provided insight into early genomic selection in the domestication of a native North American berry crop.



Autopolyploid inheritance and a heterozygous reciprocal translocation shape chromosome genetic behavior in tetraploid blueberry (*Vaccinium corymbosum*)

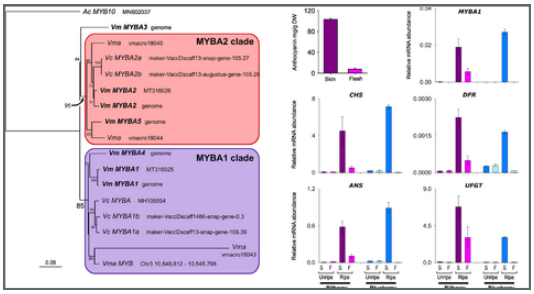
LINK TO PAPER

In this study the Iorizzo, Edger and Iovene teams, described a new high-quality, phased, chromosome-scale genome of a diploid blueberry, clone W85. The genome was integrated with cytogenetics and high-density, genetic maps representing six tetraploid blueberry cultivars, harboring different levels of wild genome admixture, to uncover recombination behavior and structural genome divergence across tetraploid and wild diploid species. Analysis of chromosome inheritance and pairing demonstrated that tetraploid blueberry behaves as an autotetraploid with tetrasomic inheritance. Comparative analysis demonstrated the presence of a reciprocal, heterozygous translocation spanning one homolog of chr-6 and one of chr-10 in the cultivar Draper. The translocation affects pairing and recombination of chromosomes 6 and 10. Besides the translocation detected in Draper, no other structural genomic divergences were detected across tetraploid cultivars and highly inter-crossable wild diploid species.

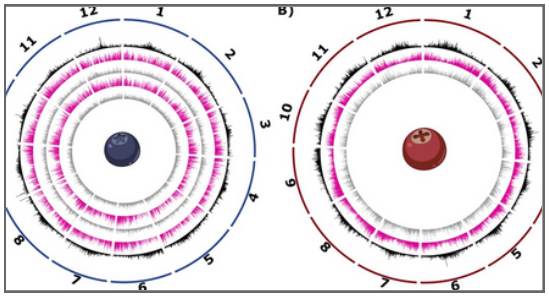
These findings and resources will facilitate new genetic and comparative genomic studies in *Vaccinium* and the development of genomic assisted selection strategy for this crop.

A chromosome-scale assembly of the bilberry genome identifies a complex locus controlling berry anthocyanin composition

LINK TO PAPER



In this study the Chagne and Espley teams presented the first bilberry (*Vaccinium myrtillus* L.) chromosome scale genome assembly. Comparative analysis with other blueberry genomes indicated a high conservation of synteny. A total of 36,404 genes were annotated after nearly 48% of the assembly was masked to remove repeats. The genome unveiled a complex MYBA locus, and identified the key regulating MYB genes that determine anthocyanin production. The new bilberry genome builds on the genomic resources and knowledge of Vaccinium species, to help understand the genetics underpinning some of the quality attributes that breeding programs aspire to improve. The high conservation of synteny between bilberry and blueberry genomes indicates that comparative genome mapping can be applied to transfer knowledge about marker-trait association between these two species, as the loci involved in key characters are orthologous.



Blueberry and cranberry pangenomes as a resource for future genetic studies and breeding efforts

LINK TO PAPER

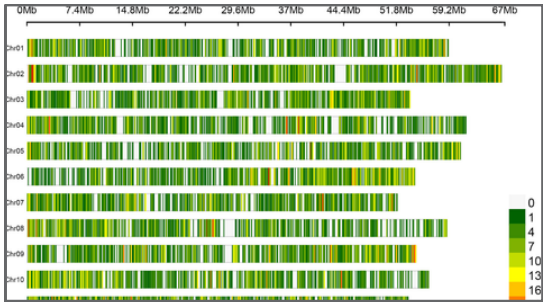
In this study Dr. Edger's team described the construction and analysis of the first pangenome for both blueberry and cranberry. The analysis of these pangenomes revealed both crops exhibit great genetic diversity, including the presence–absence variation of 48.4% genes in highbush blueberry and 47.0% genes in cranberry. Auxiliary genes, those not shared by all cultivars, are significantly enriched with molecular functions associated with disease resistance and the biosynthesis of specialized metabolites, including compounds previously associated with improving fruit quality traits. The discovery of thousands of genes, not present in the previous reference genomes for blueberry and cranberry, will serve as the basis of future research and as potential targets for future breeding efforts.

The pangenome, as a multiple-sequence alignment, as well as individual annotated genomes, are publicly available for analysis on the Genome Database for *Vaccinium*—a curated and integrated web-based relational database. Lastly, the core-gene predictions from the pangenomes will be useful to develop a community genotyping platform to guide future molecular breeding efforts across the family.

EXPANDING GENOTYPING CAPACITY

The new genomes and the pangenome were used to develop a genotyping platform for blueberry and cranberry that is optimized for performing genome wide association studies (GWAS) and work across the blueberry and cranberry germplasm.

The platforms were made available to the Vaccinium community by forming a consortium that enabled us to lower the costs for genotyping. The *Vaccinium* Genotyping Consortium includes **15** members representing public and private breeding programs, from U.S., France, New Zealand, Canada, and Italy and has genotyped **14,137** samples. The platform has been already used for **23** projects by **24** studies/experiments.



Development of a targeted genotyping platform for reproducible results within tetraploid and hexaploid blueberry

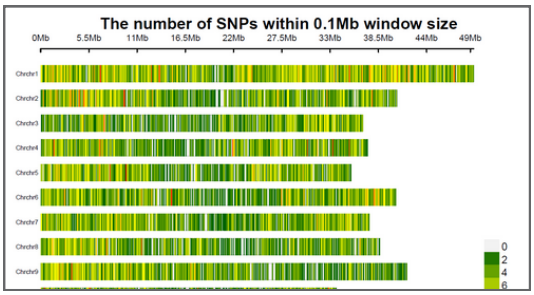
LINK TO PAPER

The development of a standardized genotyping platform that targets a specific set of polymorphic loci can unify the scientific and breeding community toward blueberry improvement. In this study ths Bassil team developed and evaluated a targeted genotyping platform for cultivated blueberries that is affordable, reproducible, and sufficiently high density to warrant large-scale adoption for genomic studies. The Flex-Seq platform was developed in a two-step procedure that resulted in 22,000 loci that yielded 194,365 single nucleotide polymorphisms when assessed in a diversity set of 192 samples including cultivated and other related wild *Vaccinium* species. Locus recovery averaged 89.4% in the cultivated polyploid blueberry (NHB, SHB, and rabbiteye [RE]) and on average 88.8% were polymorphic. While recovery of these loci was lower in the other *Vaccinium* species assayed, recovery remained high and ranged between 60.8% and 70.4% depending on the taxonomic distance to the cultivated blueberry targeted. NHB had the highest mean number of variants per locus at 9.7, followed by RE with 9.1, SHB with 8.5, and a range between 7.7 and 8.5 in other species. As expected, the total number of unique-in-state haplotypes exceeded the total number of variants in the domesticated blueberries. Phylogenetic analysis using a subset of the SNPs and haplotypes mostly conformed to known relationships.

The platform also offers flexibility about the number of loci, depth of sequencing for accurate dosage calling, loci and haplotype reconstruction from increased fragment length. This genotyping platform will accelerate the development and improvement of blueberry cultivars through genomic-assisted breeding tools.

Development of a targeted genotyping platform for cranberry

LINK TO PAPER



The Flex-Seq platform has also been used to develop a standardized genotyping platform for cranberry. The study led by Dr. Nahla Bassil is currently ongoing and the manuscript is in preparation. The cranberry platform aims to replicate the success of the blueberry platform with 17,500 loci distributed evenly across the cranberry genome, and it was designed to be affordable, reproducible, and sufficiently high density for widespread adoption. The platform is being evaluated on 192 cranberry samples that are predominately American cranberry and yielded approximately 60,000 single nucleotide polymorphisms. The mean number of variants per locus is approximately five with unique-in-state haplotypes currently being assessed. The platform will offer similar benefits to the blueberry Flex-Seq platform to accelerate the development and improvement of cranberry cultivars.

ONGOING RESEARCH

- Annotating and mining genes associated with quality characteristics
- Completing annotation of a new tetraploid genome
- Continue coordinating the *Vaccinium* genotyping consortium

Testimonials From the Lab

Dr. Sujeet Verma
Genomics and Molecular Diagnostics Manager
Fall Creek Farm and Nursery, Inc.

What impact has the genotyping platform had for your breeding program?

The impact of Flex Seq 22K (FS22) genotyping platform (GP) on the Fall Creek (FC) blueberry breeding program is huge because now we have a high-throughput high-density genotyping tool that will give us clarity on genetic architecture of traits. The most valuable impact is the application of genomic selection (GS) and enhanced trait selection. Apart from that, accurate parentage verification and estimation of genetic diversity are other impactful utilities of FS22.

With FS22 application, breeders and geneticists can pinpoint genomic locations of loci controlling traits and differentiate monogenic versus polygenic traits. Blueberry being autotetraploid, the FS22 can help understand the segregation of alleles and their dosage effects better and facilitate the development of molecular markers.

Overall, the implementation of the FS22 genotyping platform will help enhance breeding efficiency and selection at Fall Creek Farm and Nursery, ultimately leading to the development of superior blueberry varieties that meet both grower and consumer needs.

Have you changed anything in your program as a result of using the platform?

We have started implementing genomics selection and parentage verification using the FS22. Moving forward we will make breeding selections based on genomic predictions that will help speed-up the breeding cycle. This will also help inform crossing decisions.

Would you recommend it to other breeding programs?

Yes, I would recommend the FS22 genotyping platform to other breeding programs.

How do you think having this new tool will help your program in the future?

This new tool will help in the development of the marker-assisted breeding program at Fall Creek. It will enhance precision in breeding decisions, greater flexibility in adapting to climate change, sustained genetic diversity, innovative breeding techniques, better collaboration and data sharing, advanced disease management, and future-proofing the breeding program.

continued on next page

Overall, the FS22 genotyping platform will provide a robust foundation for the continuous improvement and innovation of blueberry breeding program, ensuring its success and relevance in the future.

Do you have any other comments?

The FS22 genotyping platform provides an opportunity to establish a stronger collaboration between industry and academia. Having said that, the price per sample genotyping with FS22 is limiting at the moment, and for high throughout genotyping, the price per sample needs to be more affordable! In addition, mid-density and low-density SNP panels should be rolled out for high throughput applications.

Testimonials From the Lab

Dr. Jeffrey Neyhart, Research Geneticist, USDA-ARS
Genetic Improvement for Fruits and Vegetables Laboratory
Rutgers P.E. Marucci Center for Blueberry and Cranberry Research and Extension

What impact has the genotyping platform had for your breeding program?

The high-quality cranberry reference genomes and annotations directly enabled us to identify genomic regions potentially associated with local climate adaptation and abiotic stress tolerance in diverse germplasm. These resources also greatly expedited the development of small- and medium-sized genotyping platforms for rapidly adopting marker-based selection.

Have you changed anything in your program as a result of using the platform?

Our pre-breeding program is very new, which means we can design the program with these genomics resources in mind. For example, we are considering alternative parent and mate selection strategies that use genomic information.

Would you recommend it to other breeding programs?

Based on our experience, I would encourage other programs to consider using these resources.

How do you think having this new tool will help your program in the future?

The goal of our cranberry pre-breeding program is to identify favorable or novel productivity, quality, and abiotic/biotic stress tolerance traits in wild or under-improved germplasm. By using and expanding upon the genomic resources currently available through VacCAP, along with other resources, we hope to catalog the unique haplotypes or alleles that are present in this germplasm and determine which may contribute favorable traits.

Additionally, we think the genotyping platforms developed from the genomic resources will have a profound effect on the resource efficiency of our program. Dealing with wild and underimproved germplasm is genetically “messy,” with many unfavorable traits present alongside those that are favorable. By using markers to predict and select the best individuals from our crosses, we hope to reduce the impact of linkage drag and allocate more resources to those with the greatest potential to be outstanding parents

Do you have any other comments?

We are excited to continue to utilize the genomic resources developed through the VacCAP.

Partner Project

Dr. Moira Sheehan
Senior Research Associate & Director, Breeding Insight
Member of the VacCAP Advisory Panel

How did the VacCAP team collaborate with Breeding Insight?

Breeding Insight (BI) developed a mid-density genotyping platforms (3,000 SNPs) for blueberry and cranberry that are tailored for routine genotyping in structured breeding populations (including GS populations and QTL populations), MAS introgression, parental or variety verification. These platforms complement the high density platforms that VacCAP team developed. The BI and VacCAP genotyping teams have been collaborating on developing these platforms by sharing sequences and target SNPs.

For blueberry, BI used skim sequencing of 31 parents and available GBS data from USDA-ARS and NCSU to develop a SNP database. This database was ultimately used to identify 10,000 SNP loci from which a final, high-quality set of 3,000 loci were selected for DArTag assay production. These 3,000 loci were also shared back with VacCAP and all were included in the blueberry FlexSeq panel produced by the VacCAP team.

For cranberry, BI used skim sequencing of 53 and available GBS data from USDA-ARS to develop a SNP database. Given that the cranberry FlexSeq panel was already created or in progress, when BI selected a set of 3,000 SNPs to target on DArTag, we maximized the overlap of markers with those in the FlexSeq panel. In other words, BI targeted 3,000 genomic regions where we'd like a marker, and if the VacCAP FlexSeq panel already had a nearby marker, we used that same marker rather than a different one.

Development of these two platforms provides the *Vaccinium* community with options that they can choose depending on their application needs. Shared SNP across the platforms facilitate comparison across studies.

Where do BI and VacCAP collaborate in the future?

BI and VacCAP are in a very unique position in the plant breeding community. Both are examples of how genomic technologies from major crops have been adopted in a clonal propagated perennial fruit crop with success. One place VacCAP and BI could push ahead of the major crops is by using multi-allelic data in genomic analysis. All modern SNP analyses require the data to be bi-allelic, which works well in diploids but fails to produce good results in polyploids. The capture, databasing, and utilization of multi-allelic genomic data in polyploids by BI and VacCAP for blueberry and cranberry could be the first of its kind.

continued on next page

Already, BI has evidence that using microhaplotypes in polyploids allows genomic analyses to achieve higher significance (with lower error) for marker-trait associations, due to the greater information content contained in microhaplotypes. BI is already planning the architecture for a new community resource: one microhaplotype database for the blueberry 3K panel and one for the cranberry 3K DArTag panel. These databases will hold all the observed microhaplotypes with fix-allele identifications and associated sample metadata.

BI and VacCAP can work together on what breeders want to see in these databases as well as how they can use them to find new alleles that might be missing in their germplasm stocks.

New Genomic Resources Shared Through the Genome Database for *Vaccinium* (GDV)

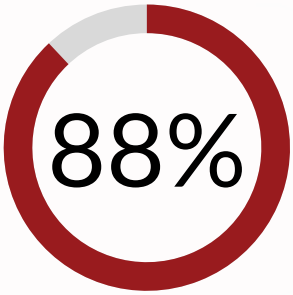
GDV houses and integrates genomic, genetic and breeding data for blueberry, cranberry and other Vaccinium *species.*

According to the GDV team, genome data had a very large increase in 2022 thanks to VacCAP additions. The Pangenome Project contributed **22** blueberry and **10** cranberry genomes. VacCAP also provided **4** other cranberry genomes, and new blueberry and bilberry genomes.

- Chromosome-Level Genome Assembly of the American Cranberry (*Vaccinium macrocarpon* Ait.) and Its Wild Relative *Vaccinium microcarpum*
- Contrasting a Reference Cranberry Genome to a Crop Wild Relative Provides Insights Into Adaptation, Domestication, and Breeding
- A Chromosome-Scale Assembly of the Bilberry Genome Identifies a Complex Locus Controlling Berry Anthocyanin Composition

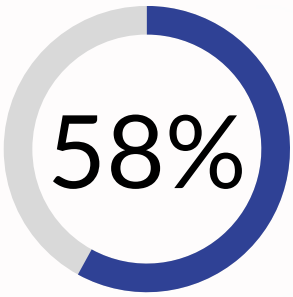
Genomes

VacCAP contributed **37** out of **42** genomes to the GDV



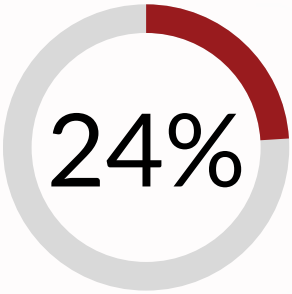
Genetic Markers

VacCAP provided **143,786** out of **247,457** genetic markers available on the GDV



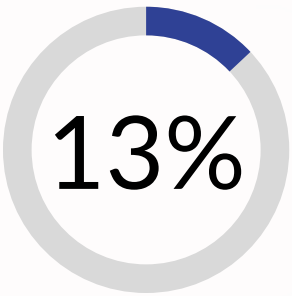
QTL

203 out of **836** quantitative trait loci (QTL) were identified by VacCAP



Genetic Maps

5 of the **39** genetic maps on the GDV were mapped by VacCAP



Genetic Marker and QTL Data Was Entered Into GDV From These Publications:

High-Density Linkage Map Construction and Identification of Loci Regulating Fruit Quality Traits in Blueberry

17,433 markers | 40 QTL | 1 map

18,195 markers | 163 QTL | 1 map

High-Density Linkage Map Construction in an Autotetraploid Blueberry Population and Detection of Quantitative Trait Loci for Anthocyanin Content

Autopolyploid Inheritance and a Heterozygous Reciprocal Translocation Shape Chromosome Genetic Behavior in Tetraploid Blueberry (*Vaccinium Corymbosum*)

108,153 markers | 3 maps

5 markers

A Low Malic Acid Trait in Cranberry Fruit: Genetics, Molecular Mapping, and Interaction With a Citric Acid Locus

Total GDV Data By Year

YEAR	GENOMES	GENES	mRNA	MAPS	MARKERS	QTL	GWAS
2023	42	3193094	3235659	39	247457	836	308
2022	42	3193094	3235659	38	245127	816	-
2021	8	466354	476613	34	136212	465	-
2020	4	245382	253280	31	89951	403	-

Primary citations refer to manuscripts that cite GDV directly. Secondary citations are the number of times the primary citations were cited. Citations increased since the VacCAP project started in 2019.

GDV Peer-Reviewed Citations by Year

YEAR	PRIMARY CITATIONS	SECONDARY CITATIONS
2023	14	38
2022	23	172
2021	22	216
2020	6	73
2011-2019	22	817
TOTAL	**87**	**1,316**

Primary citations refer to manuscripts that cite GDV directly. Secondary citations are the number of times the primary citations were cited. Citations increased since the VacCAP project started in 2019.

GDV Usage by Year

YEAR	SESSIONS	PAGEVIEWS	USERS	COUNTRIES
2023	13,292	334,692	5,660	117
2022	11,235	95,163	6,136	90
2021	10,922	101,998	5,643	94
2020	9,110	66,694	5,465	88
2019	6,832	39,904	1. 4,746	92

Sessions indicate the number of times users visit the site, and pageviews refers to the number of pages they view. Use of GDV as a platform to download and mine genomic and genetic data for Vaccinium species has increased since VacCAP started.

IMPACT

Outcomes of Objective 1 enabled optimization of genetic studies, provided access to genes for genetic studies and functional characterization and have significantly expanded use of DNA tools in the Vaccinium *community.*

To take advantage of the new genotyping platforms, breeding programs are developing genetic stocks (mapping populations and diversity panels) that best fit genetic studies design. As a result, research in *Vaccinium* crops is focusing more on understanding the genetic architecture of important traits and gene identification. This new scenario expands capacity for making new genetic discoveries for application in blueberry and cranberry breeding. Also, the VacCAP project fostered active collaboration among *Vaccinium* research groups. This new dynamic contributed to reducing duplication of work and turning research projects to be complementary to each other. This outcome is a critical step to accelerate genetic advances within the large and expanding *Vaccinium* community.

MORE RESOURCES

Recordings of all Objective 1 webinars are available on the VacCAP Project YouTube Channel. Visit www.vacciniumcap.org or follow us @VacciniumCAP on X (Twitter) to stay updated on the latest VacCAP news.