Development and utilization of KASP markers for *Prunus persica* (Peach) genetic diversity studies

Caitlin McCann, Master of Science Thesis Proposal

Dr. Dario Chavez

9/18/2023

ABSTRACT

In the U.S., peach production has been declining for almost two decades. During this same period, peach production has been increasing worldwide. In Georgia and most of the Southeastern U.S., this year was one of the worst harvests on record. Georgia peach growers lost up to 90% of their yield due to a late frost. Peach cultivars must be improved to meet changing environmental conditions, as well as threats from disease, pests, and other external factors. However, new sources of genetic diversity must be identified and the genetic tools used to evaluate these sources must be improved to develop better cultivars. To this end, 50k probe targets (Capture-Seq) will be developed and tested to evaluate the genetic variation present in peach populations. SNPs detected through this technology will then be used to study the genetic diversity of peach germplasm collected in Australia, which is being hypothesized to be a novel source of genetic material due to its purported relatedness to ancestral Chinese populations. Peach germplasm available at the UGA Peach Research Program at Dempsey Farm in Griffin, Georgia, plus accessions from the USDA-ARS Germplasm Resources Information Network (GRIN), and previously collected material from Australia will be used. These samples will be genotyped in Rapid Genomics LLC using their proprietary Capture-Seq technology. SNPs from this array will later be developed into KASP markers, which will be used to assess the genetic diversity of newly-acquired Australian samples as compared to known Chinese, American, and European peach genotypes. Population genetics information gained from these analyses will inform future breeding and selection.

INTRODUCTION

Peaches (*Prunus persica*) are a well-known and economically important fruit tree. They were originally domesticated in China thousands of years ago (Su et al. 2015). Today, peaches are grown on several continents, with the top 5 producing countries being China, Italy, Spain, the U.S., and Greece (FAO 2023). As peaches become more widespread and domestic cultivars continue to be refined,

peaches, like many cultivated crops, face threats associated with loss of genetic variation. Lack of variation can render crops vulnerable to disease, drought, pests, and other factors. These could result in reduced yield and economic insecurity. It has been proposed that wild populations of peaches may hold untapped potential for genetic variability, which can be used as a source of new germplasm in breeding programs (Lenne et al. 1991). The Prunus Crop Germplasm Committee issued a statement in 2010 calling for the increase of accessions of new *Prunus* germplasm to strengthen and to improve the U.S. stone fruit industry. Although China, as the original source of domesticated peaches, would be an ideal source of germplasm, it is notoriously difficult to obtain samples from Chinese collaborators. Therefore, peach breeders sought alternative sources of diverse and novel germplasm.

Feral peach populations in Australia are believed to derive from seeds discarded by Chinese immigrants during the gold rush of the 1850s (National Museum Australia 2023). In addition to containing DNA from Chinese landraces, local pressures from fire, drought, and human habitation have potentially produced important stock for scion and rootstock development. To verify the Australian accessions' similarity to those from China, KASP (Kompetitive Allele-Specific PCR) markers will be developed and used to build a pedigree of peach germplasm for use in breeding and selection. Because they use SNPs, which are more abundant in the genome, KASP markers can be used to distinguish between closely related genotypes more efficiently than SSR markers (Steele et al. 2021). Once these markers are developed, they will aid in identification of germplasm, as well as selection for breeding programs. These KASP markers may be also transferable to other *Prunus* species which may provide an added value to these markers. Finally, the current SNP chip arrays for peach are being discontinued and the availability of alternative genotyping tools will be beneficial for researchers working with peach.

BACKGROUND

The Peach

Peach [*Prunus persica* (L.) Batsch] belongs to the Rosaceous family and the genus Prunus. This genus, which encompasses all stone fruits, includes species such as almonds (*P. dulcis* L.), apricots (*P. armeniaca* Scop.), cherries [*P. avium* (L.) Bauhin], and plums (*P. domestica* L.). A peach tree in nature can grow up to 8m in height. Its leaves are lanceolate (narrow and pointed), glabrous (smooth, hairless), and serrate (having serrated edges). Flowers are pink, white, or red, and fruit can be either pubescent (fuzzy) or glabrous (in which case it is called a nectarine). The flesh of the fruit can be white or yellow and comes in melting and non-melting varieties. At the center of the fruit is a stony endocarp which is deeply pitted and contains a cyanogenic glycoside called amygdalin, making the seed bitter and toxic. A peach tree will begin to bear fruit at 2-3 years of age and may begin to decline around 15 years after planting. There are several closely related species of peach that have been used as sources of disease resistance genes. These species include *P. davidiana* Carriere, which is drought tolerant but sensitive to nematodes; *P. ferganensis* Y. Y. Yao, which is resistant to powdery mildew; *P. kansuensis* Rehder, which has frost resistant flowers but bitter fruit; and *P. mira* Koehne, which is considered a possible ancestor of cultivated peaches, and which is itself cultivated in Tibet as the "Tibetan peach" (Layne and Bassi 2008; Rieger et al. 2003; Meader and Blake 1939; Yoshida 1987).

Cultural and Economic Importance

As one of the most widely cultivated fruit trees in the world, peaches are culturally, economically, and scientifically important. First bred in Georgia, the popular Elberta peach was created by Samuel Rumph in 1875, boosting the economy of the American South after the Civil War and helping earn Georgia the name of "Peach State" (Greenlee 2022). Peaches are believed to have originated in China, with evidence of cultivation going back more than 8000 years to the Neolithic period. Recent fossil discoveries in southwestern China have revealed that peaches may have existed in a morphologically modern form since the Pliocene epoch, more than 2.5 million years ago, long before human ancestors

migrated to China (Su et al. 2015). Peaches are also an important cultural symbol in China. The flowers are used in the Spring Festival in south-eastern provinces, and in the past 4000 years, hundreds of unique cultivars have been developed (Layne and Bassi 2008). Peaches were carried from Asia to present day Iran more than 2,000 years ago via the silk road. Iran, formerly known as Persia, was once thought to be the center of origin for peaches, hence the name *persica*. From there, the peach was disseminated to Europe, then to the Americas via Spanish and Portuguese explorers (Byrne et al. 2011). Commercially, peaches can be grown between 25° and 45° N latitude (Layne and Bassi 2008). Most peach production takes place in China, with over 60% of all peaches produced coming from Asia. Italy, Spain, the U.S., and Greece follow China in production volume (FAO 2023). Peaches are the third most produced temperate tree fruit, after apples and pears. As such, they are economically important as well as culturally relevant.

The U.S. is the fourth largest producer of peaches, with a yearly average of 1.13 million tons produced from 1984 to 2021 (FAO 2023). Most peaches in the U.S. come from California, which produced 475 thousand tons in 2022. The next largest producers are South Carolina (67.4 thousand tons) and Georgia (24.8 thousand tons) (USDA 2023). Although peach production is increasing worldwide, production in the U.S. has been declining for decades, with total tons having fallen over 50% since the year 2000 (FAO 2023; USDA 2023). This decline could be caused by several factors including disease, changing climate, and poor fruit quality (Anthony and Minas 2022; Johnson et al. 2022; Parker et al. 2019). These problems could be addressed by the introgression of new traits from other populations, which may improve fruit quality by minimizing the effects of disease and other environmental factors, as well as lead to the production of new cultivars to suit consumer preference. Trait introgression must be preceded by germplasm acquisition and identification, and to that end we must study the genetic makeup of a variety of peach accessions. Luckily, the peach is considered a model genome for many fruit species and is therefore a good subject for future genetic study.

Genetic Importance

Peach is diploid (2n = 2x = 16) and has a relatively small genome (230 Mb). Many morphologically and economically important traits in peaches are highly heritable. This combined with its relatively simple genome, high self-compatibility, and short juvenile period, has made peach a model organism for the *Rosaceae* family (Li, 2013). The peach genome was originally sequenced in 2010 by the Joint Genome Institute (www.peachgenome.org). The cultivar 'Lovell' was sequenced using a doubled haploid, which means its genome was completely homozygous. This homozygosity simplified genomic assembly and allowed for greater coverage during sequencing. As of 2020, dozens of peach genes and QTLs had been identified and connected to agronomically important traits such as fruit size and color, flesh texture, and peach/nectarine character (Li and Wang 2020). Additionally, many major genes in the genus *Prunus* have been mapped, including fruit traits such as glabrous versus pubescent fruit, flat versus round, and melting versus non-melting flesh (Arús et al. 2012; Guo et al. 2020). After thousands of years of selective breeding, limited use of cultivars and the capacity for self-fertilization have resulted in reduced genetic diversity and high homozygosity in peach populations (Mas-Gomez et al. 2021). Although selection has created highly specialized cultivars, it has also limited genetic variation and rendered peach crops vulnerable. Lack of variation means that peach cultivars have less potential to adapt to changing environmental conditions, climate, and consumer preferences. The addition of novel germplasm into the gene pool could increase diversity as well as introduce economically important traits, such as disease resistance (Drogoudi 2023). Recent advances in DNA marker technology have allowed us to genetically categorize genotypes and identify economically important traits, reducing the time, expense, and effort necessary to develop new cultivars and introduce new traits to the gene pool (Arús et al. 2012; Guo et al. 2020).

DNA Marker Technology

Even among closely related individuals, there are unique differences across their genomes. We can compare individuals and assess differences in their DNA sequence by characterizing and targeting these locations. These cataloged differences, called “DNA markers,” are valuable tools for genetic “fingerprinting”, breeding programs, QTL (quantitative trait loci) discovery, and genetic characterization. Two types of markers that have been used extensively in agricultural and horticultural studies are SSR (single sequence repeats) and SNP (single nucleotide polymorphisms). Both methods have benefits and drawbacks. Although SSR markers have been favored in the past for their accuracy in germplasm characterization, SNPs have become more prevalent with advancements in high-throughput sequencing technology (Semagn et al. 2014).

SSR Markers

SSRs, also known as microsatellites, are sections of DNA of short nucleotide motif (2-6 base pairs) repeats (Tautz et al. 1986). These repetitive regions mutate at a rate up to ten orders of magnitude higher than point mutations (Gemayel et al. 2012). When SSRs mutate, they differ in the number of times the motif repeats, so SSR lengths can be used to differentiate between individuals (Tautz et al. 1986). These repeat motifs are abundant in the genome. SSR ubiquity along with their polymorphism makes them useful as DNA markers. Linkage maps for several species have been constructed based on SSR markers, including for humans (Dib et al. 1996).

SSR markers are a codominant marker, meaning that they can inform about the presence of different alleles and differentiate between hetero- and homozygotes, unlike dominant markers which can only detect the presence or absence of an allele (Collard et al. 2005). To design an SSR marker, one would start with a source of genetic material, usually DNA. This DNA is then enriched for SSRs and sequenced (Maio and Castro 2013). Known sequences are compared to a reference genome or to one another, and differences in the sequences can be used to identify polymorphisms. Based location of the

SSR within the genome, primers can be designed upstream and downstream of the repeating motif (SSR-containing regions amplified using PCR). These amplified regions can be analyzed via agarose gel electrophoresis (AGE), polyacrylamide gel electrophoresis (PAGE), or capillary electrophoresis. In the gel, bands of different lengths represent different alleles of the SSR marker, allowing researchers to identify the alleles present in the sample (Tautz et al. 1986). Fluorescent markers can also be attached to primers to allow genotyping by capillary electrophoresis (Csencsics et al. 2010; Agarwal et al. 2015). In addition, there are multiplex methods which differentiate SSR alleles at multiple loci simultaneously (Guichoux et al. 2011).

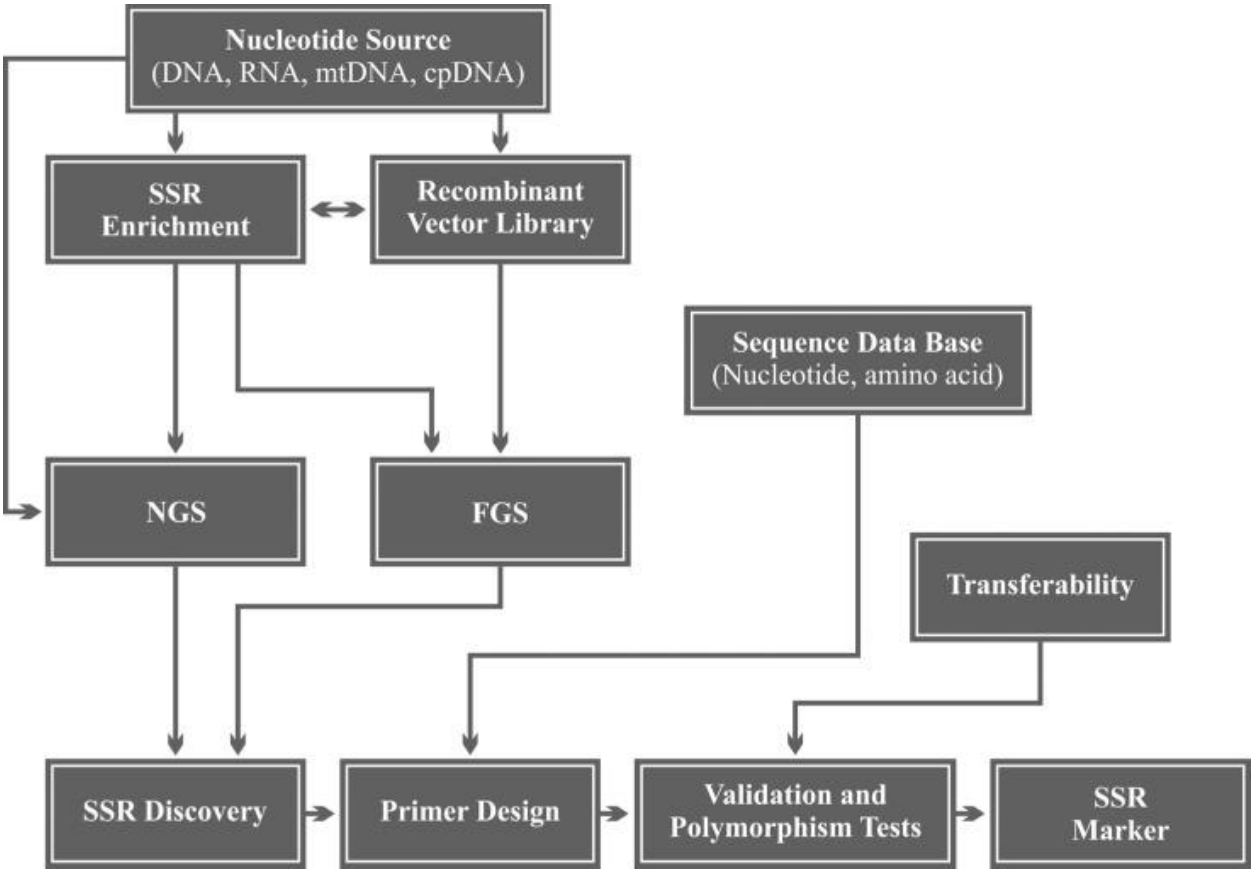


Fig 1. Workflow of how SSR markers are made (Vieira et al. 2016)

SSR Markers have been used in the past to assess genetic diversity and population structure in *Prunus* species. One study used 36 SSR markers to determine the population structure of 195 peach accessions (Chavez et al. 2014). A similar study used SSR data to make pedigree clusters of European plum (*Prunus domestica*) accessions (Antanyniene et al. 2023). Markers can even be utilized across genera, as shown when expressed sequence tags-simple sequence repeat markers (EST-SSR) developed for Himalayan raspberry (*Rubus ellipticus*) were successfully used to analyze genetic diversity of peach cultivars (Sharma 2023). SSR markers have a long history of use in plant breeding, especially in the evaluation of *Prunus* germplasm, but recent advances in Next Generation Sequencing are leading scientists to shift toward SNP markers as a more efficient and cost-effective alternative (Semagn 2014; Zahid et al. 2022).

SNP Markers

SNPs, or single nucleotide polymorphisms, are positions on the genome which vary between individuals by one or multiple base pairs. Unlike SSRs, which can vary in length and therefore have many possible variants, SNPs have only four possible variants, the bases A, C, T, and G. Generally, each individual SNP will have only two variants (A/G or C/T), therefore they are considered “biallelic” (Brookes 1999). Because of their biallelic nature, more SNPs are required to achieve the same level of specificity as SSR markers (Inghelandt et al. 2010). However, SNPs make up for this shortcoming by being common in the genomes of all life, more abundant than SSRs, and capable of high-throughput automation (Mammadov et al. 2012). They are a major source of genetic variation between individuals of the same species, making them useful for population studies and breeding programs (Rafalski 2002).

There are different technologies currently used to characterize SNPs across different genotypes. Through those, SNPs are discovered in a genome and deemed reliable for use as markers. There are several reasons that a SNP could be disqualified from being used as a marker. If the SNP is extremely

rare, occurring in less than 1% of a population, it is instead considered a point mutation (Khlestkina et al. 2006). Its rarity makes it less than useful as a method of separating genotypes into groups. A SNP may also occur in non-coding regions of the DNA. SNPs that occur in exonic or regulatory regions of the DNA are often called "functional" because they exist on the part of the DNA which contributes to protein formation and function. These are more useful for characterizing genotypes than "non-functional" SNPs, so functional SNPs are preferred as markers (genomicglossaries.com). Markers can be "trained" to predict the physical characteristics of a plant by comparing the phenotype of an individual to the functional markers present in that genotype (Zhong et al. 2009). This could save time, space, and labor in plant breeding by testing seeds for certain traits, without the need to grow the seeds to discover those traits. SNP arrays have already been developed for apple (Bianco et al., 2014), pear (Xiaolong Li et al. 2019), peach (Verde et al. 2017), grape (Laucou et al. 2018), maize (Xu et al. 2017), and wheat (Sun et al. 2020). There are several different platforms which are used to evaluate SNP markers. Some of those are Integrated DNA Technologies' rhAmp, Thermo Fisher's TaqMan, and KBioscience's KASP (Kompetitive Allele-Specific PCR) (Broccanello et al. 2018). These platforms and their qualifications are listed in Table 1. Note, universal PACE 2.0 Genotyping Master Mix can be substituted in KASP reactions, reducing the Master Mix cost to $762 (3crbio.com/products/).

	rhAmp	TaqMan	KASP
Call-rate	98.10%	97%	97.60%
Cost per assay	$59	$256	$50
Master Mix Cost	$814.0 (25ml)	$586.8 (10ml)	$1083.5 (25ml)
Cost per sample	$0.11	$0.32	$0.12

Table 1. Three prominent SNP genotyping platforms, along with their costs and benefits. (Broccanello et al. 2018)

KASP Markers

KASP is a SNP genotyping platform originally developed by KBioscience, which has since become one of the most well-known SNP platforms (LGC Ltd, Teddington, England). It is uniplex, meaning that it

can analyze one SNP at a time for many samples. Unlike multiplex platforms like Goldengate and Infinium, which are suitable for larger studies, KASP has no minimum sample or SNP requirement. For applications such as quality control, QTL (quantitative trait locus) mapping, and marker assisted selection, scientists are often interested in one or a few SNPs, for which a uniplex approach would be more appropriate. KASP is fluorescence based, meaning that primers used to target allele of a particular SNP will bind to a unique fluorescent dye during PCR. The presence or absence of this dye will be read by a plate reader, which will then inform about which alleles are present in the sample and the zygosity of an individual. A brief overview of the KASP process is outlined in Figure 2. It is also possible to evaluate two SNP loci at a time using additional fluorescent dyes, creating a limited multiplexing capability (Suo et al. 2020). If one is evaluating many SNP loci, this could cut the number of necessary reactions in half. In summary, KASP is cost-effective and well suited to studies with a small number (less than 200) of SNPs.

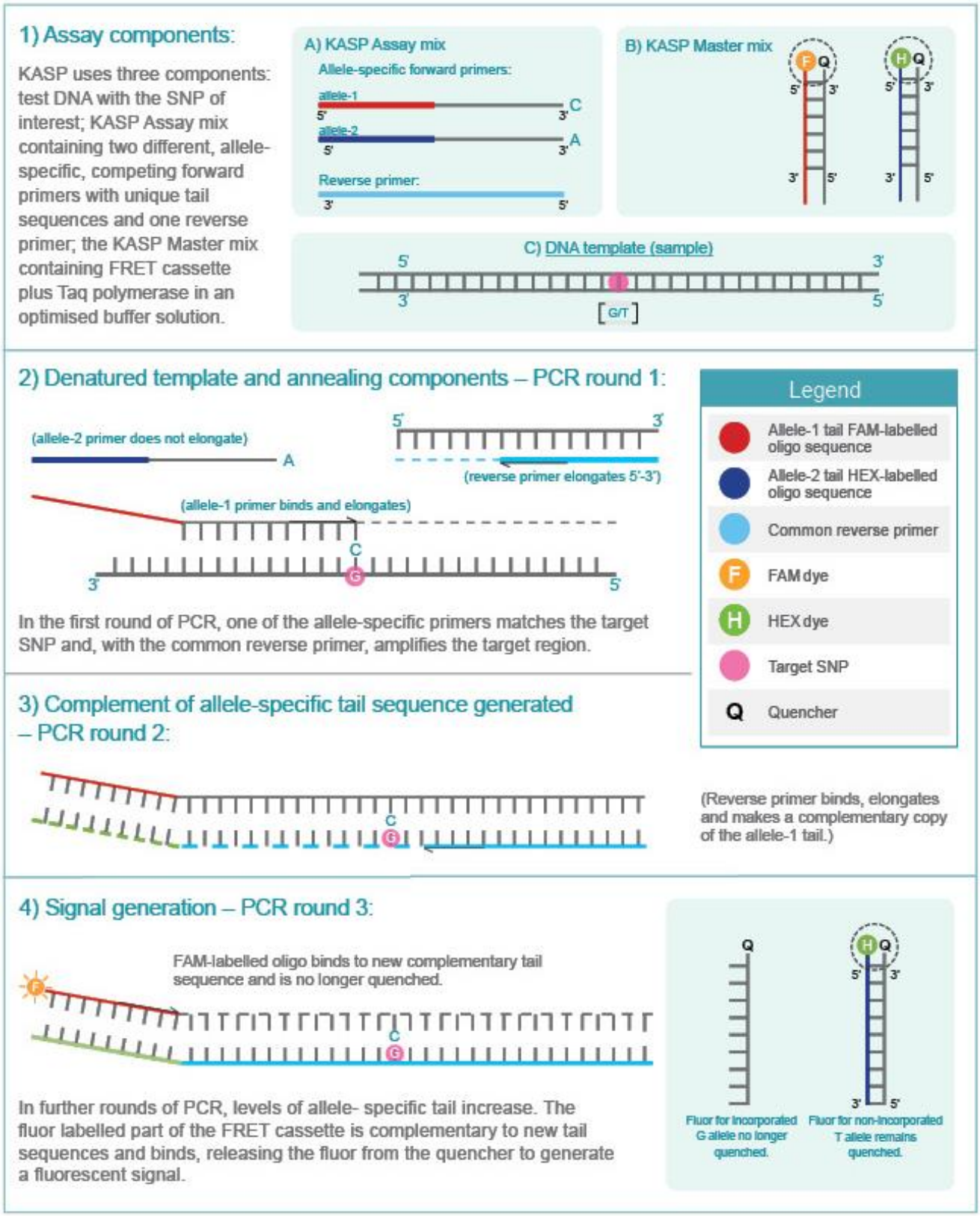


Figure 2. An overview of KASP (LCG Ltd.)

Australian Germplasm

There are over 2000 varieties of peaches worldwide, of which 1569 are registered on fruitandnutlist.org (Medich 2023). Of these registered varieties, only 15, or less than 1%, come from Australia. Peaches in Australia have long been overlooked by research. Peach diversity studies have been conducted on Chinese, European, and American varieties, but not yet on Australian accessions (Li et al. 2013; Verde et al. 2017). The Germplasm Resources Information Network (GRIN), has only 157 accessions of Australian peaches, although it has 348 samples from China and 755 from the U.S. Of these 157 accessions, all of them are categorized as “historic”, meaning they are not available from the National Plant Germplasm System (NPGS) (https://npgsweb.ars-grin.gov/gringlobal/search). Despite the lack of data surrounding Australian peaches, these populations have the potential to be an important source of novel germplasm for U.S. cultivars. This is due not only to their putative relatedness to Chinese populations, but also because local adaptation to the harsh Australian climate may have adapted this germplasm for traits which could aid American peach survival as the climate continues to change. Verifying the phylogenetic relationships of Australian peach populations would constitute the first step toward utilizing them to improve American peach cultivation.

RATIONALE AND SIGNIFICANCE

Peaches are the third most important temperate fruit species in the world, and they are genetically well-studied due to their relatively small genome (Byrne et al. 2012). In especially hot, cold, or humid areas, peaches are susceptible to diseases like the bacterial *Xanthomonas arboricola* or fungal brown rot (*Monilinia* spp.) (Vauterin et al. 1995). Buds and new growth can also die during cold snaps, reducing its yield. Currently, the peach gene pool is relatively small and homogenous compared to other fruit species, making the crop susceptible to factors such as disease and environmental changes. Increased diversity in cultivated peach trees would protect against such factors by providing a source of variation

from which to select more resilient cultivars. We can increase diversity by collecting and integrating peach germplasm from a diverse source, such as Chinese or Australian populations. By verifying the Australian population's diversity and potential for useful traits, we can take the first steps toward adding these useful traits to US populations.

KASP markers have been compared to other SNP marker platforms and found to be reliable (with a higher call rate than TaqMan), affordable (lower cost per sample than TaqMan), and flexible (Broccanello et al. 2018). The markers developed during this study will be used to assess germplasm for potentially novel and economically important traits, as well as to verify the ancestry of Australian accessions suspected to be Chinese in origin. Once verified and assessed, this new germplasm can be used in breeding programs to strengthen elite cultivars and safeguard to future of peaches in Georgia and across the US.

PROJECT GOALS/OBJECTIVES

Overall Goal: Determine relatedness and genetic diversity of peach accessions from China, Australia, Europe, and US.

Objectives:

1. Create a 50k SNP panel based on 200 peach genotypes available in the UGA peach germplasm collection and USDA
2. Characterize genetic diversity of ~200 Australian peach accessions, as well as current accessions, using 10 KASP markers based on the aforementioned 50k SNP panel.

HYPOTHESIS

The *P. persica* specimens feral in Australia are more diverse than current germplasm available in the US and more closely related to Chinese populations than American or European accessions.

METHODS

Sample Collection, DNA Isolation, and Sequencing

In 2015, 190 peach cultivars and advanced breeding selections were planted as part of the germplasm collection at Peach Research and Extension orchard at Dempsey Farm, University of Georgia, Griffin, GA (33°14’55” N, 84°17’57” W) (Table 2). All trees were grafted onto the peach rootstock “Guardian” and planted in a Cecil sandy loam soil at a planting density of 4.5 m x 6 m (358 trees per ha). A soil amendment with phosphorus, potassium, and lime was applied before the orchard was established according to the guidelines from the 2023 Southeastern Peach, Nectarine, and Plum Pest Management and Culture Guide (Blaauw et al. 2023).

[1] China Pearl	[21] Fireprince	[41] Vulcan	[61] Dixieland	[81] Carolina Red	[101] Late Large 23	[121] Princess Time/ Lovell	[141] Carolina Gold/Guardian
[2] Contender	[22] Flameprince	[42] Winblo	[62] Early Loring Blair	[82] Harrow Beauty	[102] Leafcurl Resistant	[122] Beekman	[142] Challenger/Guardian
[3] Raritan Rose	[23] Garnet Beauty	[43] Amoore	[63] Elegant Lady/ Lovell	[83] 53ZR306/ Lovell	[103] LOV2 - Dhaploid	[123] Flordaking	[143] Contender/Guardian
[4] Reliance	[24] Glohaven	[44] Autumn Red	[64] Fairtime/ Lovell	[84] 7 Ball	[104] LOV2 - Haploid	[124] Green Gage/ Myro29C	[144] NC Yellow/Guardian
[5] Redstar	[25] Jefferson	[45] Bounty	[65] Fantasia	[85] Coronet/Guardian	[105] LOV3 - Dhaploid	[125] Lord Napier	[145] NC97-23/Guardian
[6] Chui Lum Tao (rootstock)	[26] Jerseyqueen	[46] Early August Prince	[66] Flavortop	[86] Snow Gem	[106] LOV5 - Haploid	[126] Burgundy/ Citation	[146] NC97-36/Guardian
[7] Cresthaven	[27] Julyprince	[47] Loring	[67] Gala	[87] Diamond Princess	[107] NJH3-7	[127] Flordadawn	[147] NC97-45/Guardian
[8] Redhaven	[28] Juneprincess	[48] O'Henry/ Guardian	[68] Harvester	[88] Carored - offtype	[108] NJH4-44	[128] Methley/ Myro29C	[148] NC97-48/Guardian
[9] Redhaven/ Lovell	[29] Madison	[49] O'Henry/ Lovell	[69] Hiland	[89] Carored	[109] Zephyr	[129] Panamint	[149] NC98-52/Guardian
[10] Sureprince	[30] Redglobe	[50] Souvenir	[70] Jade	[90] Empress	[110] Desiree	[130] Flordaprince	[150] NC98-67/Guardian
[11] September Snow	[31] Redgold	[51] Summer Beaut	[71] M. A. Blake	[91] Flavorich/Guardian	[111] Desiree/ Lovell	[131] PER2 - Dhaploid	[151] NC98-71/Guardian
[12] 880332	[32] Redrose	[52] White Cloud	[72] Majestic	[92] Le Grand	[112] Lola	[132] Summer Fest	[152] NC-C55-30/Guardian
[13] Augustprince	[33] Roseprincess	[53] White County	[73] Redskin	[93] Messina/ Lovell	[113] Juneprince	[133] Tra-Zee	[153] NC-C55-73/Guardian
[14] Autumnprince	[34] Ruby Pearl/Guardian	[54] White Diamond	[74] Scarletpearl	[94] Rich May	[114] 11 Ball	[134] Var A (Junegold?)	[154] Winblo/Guardian
[15] Belle of Georgia	[35] Rubyprince/Guardian	[55] White River	[75] Snow Queen (aka Karla	[95] Springprince	[115] Caro King	[135] Var B Egla	
[16] Blaze Prince	[36] Scarletprince	[56] Arrington	[76] Summergold	[96] Starlite	[116] Gloria/ Lovell	[136] Var C	
[17] Durbin	[37] Sentry	[57] Bowden	[77] Sunhigh	[97] Suncrest	[117] Gloria/Guardian	[137] Var D	
[18] Early Red free	[38] Springold	[58] Bradley	[78] Sunland	[98] Aneheim	[118] Karla Rose	[138] Var E	
[19] Elberta	[39] Summerprince	[59] Camden	[79] Westbrook	[99] Early Star	[119] Tashkent gold	[139] China Pearl/Guardian	
[20] Encore	[40] Sunprince	[60] Canadian Harmony	[80] White Rock	[100] Lady Nancy	[120] Galaxy	[140] Redglobe/ Guardian	

Table 2. Accessions at Dempsey Farm – showing all unique cultivars and advanced breeding selections.

In addition to this peach germplasm, samples will be obtained from the U.S. National Plant Germplasm Repository in Davis, CA accessed through the USDA-ARS Germplasm Resources Information Network (GRIN). These will include European and Chinese cultivars and advanced breeding selections, as well as Chinese landraces and other non-US accessions. They will represent a diverse genetic pool and enhance the scope of the study.

Previously collected and newly acquired samples from feral Australian populations will also be included in the study. Samples were collected on during expeditions in 2015, 2017, and 2019. These expeditions also served the purpose of identifying locations and available germplasm for future study. Approximately 50 accessions were obtained from the Southeastern coast of Australia between Stanthorpe and Rockhampton in Queensland (Fig 3). Once obtained, feral accessions were kept in a cooler during transportation. Pits were removed, cleaned, and dried, then stored in clear Ziplock bags to be transported from Australia to the U.S. using all necessary labels and phytosanitary permits (USDA, 2021). Budwood was also collected from these trees for propagation at the Maroochy Research Facility (26° 38' 28” S, 152° 56'17” E). In January 2023, additional seed will be obtained from accessions grown at the Maroochy Research Facility. This seed will be shipped to the U.S. according to USDA-APHIS regulations. Pits will be cleaned, dried, and packaged according to the relevant permits (USDA, 2021). Phytosanitary certification will be obtained from a phytosanitary office in Australia. Once shipped or carried to the U.S., seeds determined to be clean will be germinated and tissue samples taken from cotyledons. Samples will also be collected on site from cotyledons or true leaves, depending on what is available from each plant.

Fig 3. A) Queensland, in Southeast Australia. B) Samples were collected from Rockhampton to Stanthorpe.

DNA Isolation

Leaf samples will be stored in a -80°C freezer prior to DNA isolation. Before extraction, 50 mg of leaf tissue per accession will be weighed and placed into 2 mL tubes. These tubes will then be stored at -80°C. The *DNeasy ® Plant Pro Kit* will be used to extract DNA according to the protocols set forth by the manufacturer, Qiagen Inc. (Valencia, California). After 100 µL of DNA is extracted from each sample, the presence of DNA will be confirmed via gel electrophoresis using 1.5% agarose gel. Visual confirmation will be determined according to Lambda DNA standards (Promega Corporation, Madison, WI, USA). DNA quantification will be done on a NanoDrop™ 2000c Spectrophotometer (Thermo Scientific, Waltham, MA, USA). Then, DNA concentration for all the samples will be standardized to 20 ng/µL.

Capture-Seq

Capture-Seq technology uses probe hybridization for targeted DNA sequencing. It is widely utilized in both plant and animal genomics for such tasks as GWAS/QTL mapping, genetic fingerprinting,

SNP discovery, etc. (LCG BioSearch Technologies, Middlesex, UK). Rapid Genomics LLC, owned by LCG BioSearch Technologies, owns the Capture-Seq technology and workflow. Their methods are outlined in Figure 4. In total, 50,000 probes will be created which target evenly spaced exonic regions of the peach genome. These probes will be based on the previously published SNP targets which were used to create the International RosBREED SNP Consortium 16K SNP peach v.2 array, as well as previously published SSR marker flanking regions (data available at Genome Database for Rosaceae; Jung et al. 2019).

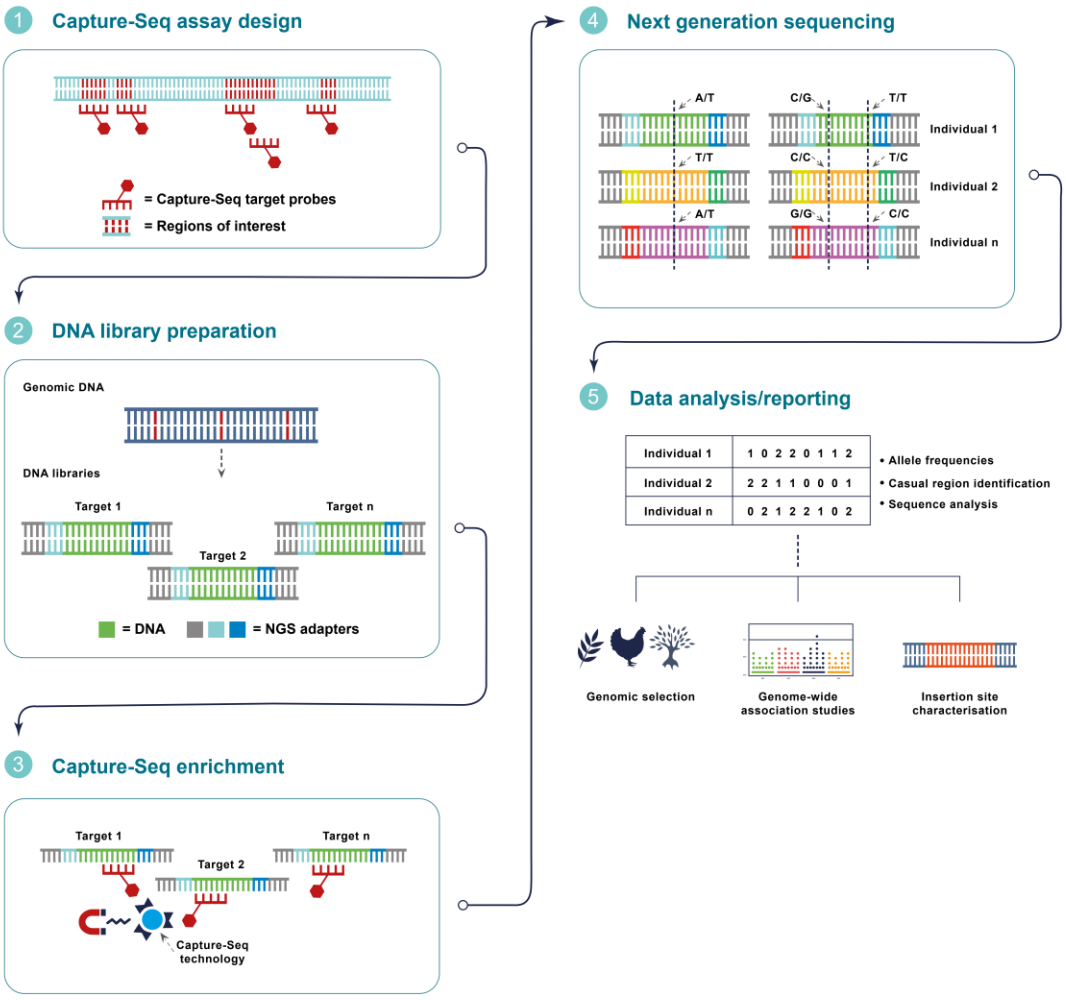


Fig 4. The Capture-Seq workflow (LCG BioSearch Technologies)

Rapid genomics will process the DNA into libraries, which will be sequenced by bonding to Biotinylated 120-mer probes that complement a segment of each sequence. Using these probes, each target locus will be sequenced via Next Generation Sequencing (NGS). Sequence data will be delivered as FASTQ files.

SNP Analysis and Genetic Diversity

The FASTQ files will be used to generate SNP markers through Rapid Genomics' standard bioinformatics pipeline. SNP data will be used for genetic diversity analysis, linkage disequilibrium calculations, genetic structure analysis, and finally for a genome-wide association study (GWAS) as previously described by Mas-Gómez et al. (2022). The GWAS will be conducted using data from the Chavez lab, including 3D scanning data of tree structure and yearly evaluations. Genetic diversity will consist of calculating the fixation index (F_{ST}), G_{ST}, the D_{Jost}, observed heterozygosity (Ho), expected heterozygosity (He), and allelic richness (A_r). Linkage disequilibrium will be evaluated using PLINK (Purcell et al. 2007). The r^2 values will be calculated using SNP data from contiguous SNPs (100) or 5K kbp. These r^2 values will be plotted against genetic distance using the R package ggplot2 (Wickham 2016; v4.1.2 R Core Team 2021). Genetic structure analysis will be performed using fastStructure v.1.0 to ascertain the genetic groupings of all accessions (Raj et al., 2014). K values from 1 to 10 will be used and the k-means algorithm will be used to identify the optimal K cluster using the BIC (Bayesian Information Criterion) (Jombart and Collins 2015).

Primer Design

After obtaining validated SNPs, the area flanking each SNP will be identified using the peach reference genome v2.0 (Verde et al., 2017). The flanking sequences to the desired target SNP(s) will be entered into Primer3Plus software to design forward and reverse primers (Untergasser et al. 2007). The ideal primers will be selected based on product length and annealing temperature (Ta). BLAST will be

used to verify that the primers bind to the expected location on the genome. The tail sequences to bind either FAM or VIC dyes will be attached to the 5' end of their respective primer, while quenchers will be attached to the opposite end. The primers will then be ordered from Sigma-Aldrich Inc. (St. Louis, MO).

PCR and KASP Analysis

PCR will be carried out according to a modified version of the KASP protocol written by Cecilia McGregor in 2015 (Paudel et al. 2019). KASP results will be read on a FRET-capable plate reader and interpreted using KlusterCaller software version 4.1.2.26268 (LCG Biosearch Technologies, Middlesex, UK) to determine the presence of markers in individual genomes. Based on the presence or absence of markers, clusters of peach accessions will be constructed, grouping related genotypes. A subset of KASP markers evenly distributed across the peach genome (approx. 10) will be used to genotype previously obtained samples. Genetic diversity parameters will also be calculated as previously described above.

POTENTIAL PITFALLS

Collecting Samples

Samples will be collected and shipped from different parts of the U.S. and Australia. If samples are collected or labelled incorrectly, DNA could be compromised. Incorrect storage could lead to low quality DNA, which may necessitate resampling. Collection of Australian samples relied partly on locating feral peach trees. External circumstances, like weather, a pandemic, or land development may inhibit researchers' ability to access the trees and collect samples. Researchers will use all applicable permits and follow international regulations on germplasm collection, storage, and transportation. Ample genetic material will be collected to account for possible losses.

DNA Contamination

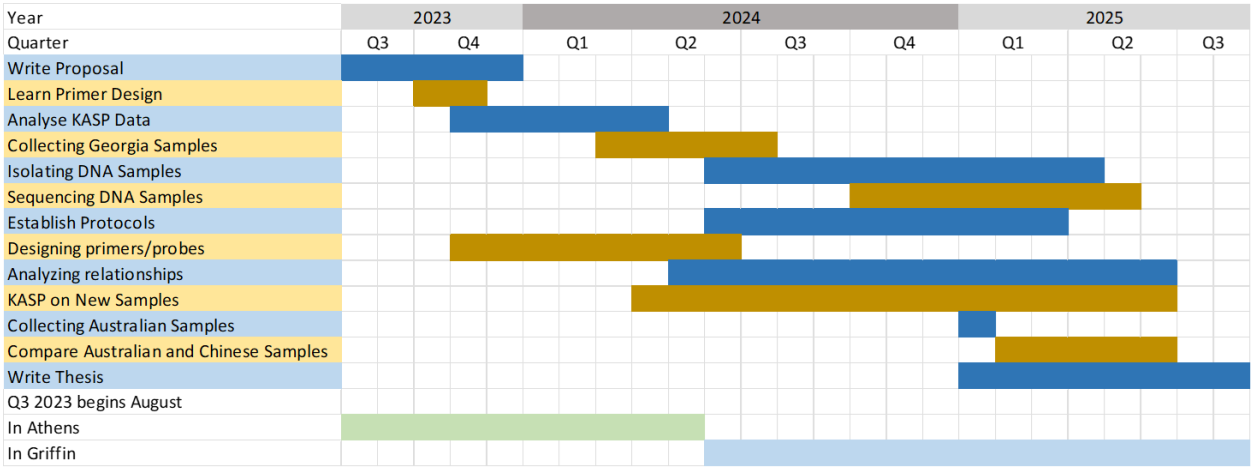
Unclean tools may result on DNA from one sample contaminating another. In closely related accessions, this may lead to confusion about the presence or absence of markers in a sample. Tools will be cleaned regularly and stored appropriately to prevent this.

Primers

DNA primers are known to fail occasionally. It may take several attempts before a primer is designed which reliably bonds to the desired site and can clearly characterize the SNP of interest. Numerous failed primers could delay analysis of the germplasm, as new primers can take several days to design and acquire. Primers will be validated by comparing them against a BLAST search to ensure they do not bind to the wrong location, and multiple primers will be designed to limit delays caused by failed primers.

Analysis

Analysis of marker data and the relatedness of accessions will be a lengthy process. As with any analysis, it will be subject to potential human error, misinterpretation, and faulty reasoning. The researchers will endeavor to limit these risks wherever possible by extensive studying of analytical methods.



Timeline of Project

REFERENCE LIST

"Gold Rushes." National Museum Australia, 2023, 2023, https://www.nma.gov.au/defining-moments/resources/gold-rushes.

Agarwal, Vikram; Bell, George W.; Nam, Jin-Wu; Bartel, David P. "Predicting Effective Microrna Target Sites in Mammalian Mrnas." eLife 4 (2015). https://doi.org/10.7554/eLife.05005.002.

Antanynienė, Raminta, Jūratė Bronė Šikšnianienė, Vidmantas Stanys, and Birutė Frercks. "Fingerprinting of Plum (Prunus Domestica) Genotypes in Lithuania Using Ssr Markers." Plants 12, no. 7 (2023): 1538.

Anthony, Brendon M, and Ioannis S Minas. "Redefining the Impact of Preharvest Factors on Peach Fruit Quality Development and Metabolism: A Review." Scientia Horticulturae 297 (2022): 110919.

Arús, P., Verde, I., Sosinski, B. et al. "The Peach Genome." *Tree Genetics and Genomes* 8 (2012): 531-47. https://doi.org/10.1007/s11295-012-0493-8.

Bianco, Luca, Alessandro Cestaro, Daniel James Sargent, Elisa Banchi, Sophia Derdak, Mario Di Guardo, Silvio Salvi, et al. "Development and Validation of a 20k Single Nucleotide Polymorphism (Snp) Whole Genome Genotyping Array for Apple (Malus× Domestica Borkh)." PloS one 9, no. 10 (2014): e110377.

Blaauw, Brett; Brannen, Phil; Lockwood, David; Schnabel, Guido; Ritchie, David. "2023 Southeastern Peach, Nectarine, and Plum Pest Management and Culture Guide." In UGA Cooperative Extension Bulletin 1171, UGA Extension, 2023.

Broccanello, Chiara, Claudia Chiodi, Andrew Funk, J Mitchell McGrath, Lee Panella, and Piergiorgio Stevanato. "Comparison of Three Pcr-Based Assays for Snp Genotyping in Plants." Plant methods 14 (2018): 1-8.

Brookes, Anthony J. "The Essence of Snps." Gene 234, no. 2 (1999): 177-86.

Byrne, D.H. et al. *Handbook of Plant Breeding.* Vol. 8, Boston, MA: Springer, 2012. doi:10.1007/978-1-4419-0763-9_14.

Chagné, David, Ksenija Gasic, Ross N. Crowhurst, Yuepeng Han, Heather C. Bassett, Deepa R. Bowatte, Timothy J. Lawrence, et al. "Development of a Set of Snp Markers Present in Expressed Genes of the Apple." Genomics 92, no. 5 (2008/11/01/ 2008): 353-58. https://doi.org/https://doi.org/10.1016/j.ygeno.2008.07.008. https://www.sciencedirect.com/science/article/pii/S0888754308001808.

Chavez, Dario J, Thomas G Beckman, Dennis J Werner, and José X Chaparro. "Genetic Diversity in Peach [Prunus Persica (L.) Batsch] at the University of Florida: Past, Present and Future." Tree genetics & genomes 10 (2014): 1399-417.

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. et al. "An Introduction to Markers, Quantitative Trait Loci (Qtl) Mapping and Marker-Assisted Selection for Crop Improvement: The Basic Concepts. ." Euphytica 142 (2005): 169-96. https://doi.org/10.1007/s10681-005-1681-5.

Prunus Crop Germplasm Committee. "Prunus Vulnerability Statement." news release, 2010.

Csencsics, Daniela; Brodbeck, Sabine; Holderegger, Sabine "Cost-Effective, Species-Specific Microsatellite Development for the Endangered Dwarf Bulrush (Typha Minima) Using Next-Generation Sequencing Technology." Journal of Heredity 101, no. 6 (2010): 789-93. https://doi.org/10.1093/jhered/esq069.

Di Maio, A., De Castro, O. "Development and Characterization of 21 Microsatellite Markers for Pancratium Maritimum L. (Amaryllidaceae)." Conservation Genet Resour 5 (2013): 911-14. https://doi.org/10.1007/s12686-013-9931-7.

Dib, Colette; Fauré, Sabine; Fizames, Cécile; Samson, Delphine; Drouot, Nathalie; Vignal, Alain; Millasseau, Philippe; Marc, Sophie; Kazan, Jamile; Seboun, Eric; Lathrop, Mark; Gyapay, Gabor; Morissette, Jean; Weissenbach, Jean "A Comprehensive Genetic Map of the Human Genome Based on 5,264 Microsatellites." Nature 380 (1996): 152-54. https://doi.org/10.1038/380152a0.

Drogoudi, Pavlina et al. "Exploring the Genetic and Morphological Variation and Disease Resistance in Local and Foreign Prunus Persica (L.) Batsch Cultivars." *Agriculture* 13, no. 4 (2023). https://doi.org/10.3390/agriculture13040800.

"Faostat." FAO 2023, 2023, accessed 10/2, 2023, https://www.fao.org/faostat/en/#data/QCL/visualize.

Gemayel, Rita; Cho, Janice; Boeynaems, Steven; Verstrepen, Kevin J. "Beyond Junk-Variable Tandem Repeats as Facilitators of Rapid Evolution of Regulatory and Coding Sequences." Genes 3, no. 3 (2012): 461-80. https://doi.org/10.3390/genes3030461.

Greenlee, Cynthia R. "Reinventing the Peach, the Pimento, and Regional Identity." Issues in Science and Technology 38, no. 4 (2022): 75-83.

GUICHOUX, E.; LAGACHE, L.; WAGNER, S.; CHAUMEIL, P.; LÉGER, P.; LEPAIS, O.; LEPOITTEVIN, C.; MALAUSA, T.; REVARDEL, E.; SALIN, F.; PETIT, R.J. . "Current Trends in Microsatellite Genotyping." Molecular Ecology Resources 11, no. 4 (2011): 591-611. https://doi.org/10.1111/j.1755-0998.2011.03014.x.

Guo, Jian, Ke Cao, Cecilia Deng, Yong Li, Gengrui Zhu, Weichao Fang, Changwen Chen, *et al.* "An Integrated Peach Genome Structural Variation Map Uncovers Genes Associated with Fruit Traits." *Genome biology* 21 (2020): 1-19.

Hao, Xiaopeng; Yang, Tao; Liu, Rong; et al. "An Rna Sequencing Transcriptome Analysis of Grasspea (Lathyrus Sativus L.) and Development of Ssr and Kasp Markers." *Front. Plant Sci.* 8 (2017). https://doi.org/**10.3389/fpls.2017.01873**.

Johnson, Kendall A, Clive H Bock, and Phillip M Brannen. "Phony Peach Disease: Past and Present Impact on the Peach Industry in the Southeastern USA." CABI Agriculture and Bioscience 2, no. 1 (2021): 1-23.

Jombart, Thibaut, and Caitlin Collins. "Analysing Genome-Wide Snp Data Using Adegenet 2.0. 0." 2015.

Jung S, Lee T, Cheng CH, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K, Frank M, Ru S, Hough H, Evans K, Peace C, Olmstead M, DeVetter LW, McFerson J, Coe M, Wegrzyn JL, Staton ME, Abbott AG, Main D. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucleic Acids Res. 2019 Jan 8;47(D1):D1137-D1145. doi: 10.1093/nar/gky1000. PMID: 30357347; PMCID: PMC6324069.

Khlestkina, E. K., and E. A. Salina. "Snp Markers: Methods of Analysis, Ways of Development, and Comparison on an Example of Common Wheat." Russian Journal of Genetics 42, no. 6 (2006/06/01 2006): 585-94. https://doi.org/10.1134/S1022795406060019.

KUMAR, RAJENDER, DC DIMRI, KANCHAN KARKI, KM RAI, NK SINGH, JITENDRA SINGH SHIVRAN, and SWAPNIL BHARTI. "Ssr Marker Based Profiling and Population Structure Analysis in Peach (Prunus Persica) Germplasm." Indian Journal of Agricultural Sciences 93, no. 10 (2023): 1080-85.

Layne, D, and Daniele Bassi. The Peach: Botany, Production and Uses. Cabi, 2008.

Lenne, Jillian M.; Wood, David. "Plant Diseases and the Use of Wild Germplasm." *Annu. Rev. Phytopathol* 29 (1991): 35-63.

Levy, L.; Damsteegt, V.; Welliver, R. "First Report of Plum Pox Virus (Sharka Disease) in Prunus Persica in the United States." *Plant Disease* 85, No. 2 (2007): 202. https://doi.org/10.1094/PDIS.2000.84.2.202B.

Li, Peirong et al. . "Development of a Core Set of Kasp Markers for Assaying Genetic Diversity in Brassica Rapa Subsp. Chinensis Makino." Original. *Plant Breeding* 138, no. 3 (2019): 309-24. https://doi.org/10.1111/pbr.12686. https://onlinelibrary.wiley.com/doi/full/10.1111/pbr.12686.

Li, Xiaolong, Jugpreet Singh, Mengfan Qin, Siwei Li, Xun Zhang, Mingyue Zhang, Awais Khan, Shaoling Zhang, and Jun Wu. "Development of an Integrated 200k Snp Genotyping Array and Application for Genetic Mapping, Genome Assembly Improvement and Genome Wide Association Studies in Pear (Pyrus)." Plant biotechnology journal 17, no. 8 (2019): 1582-94.

Li, Xiong-wei; et al. "Peach Genetic Resources: Diversity, Population Structure and Linkage Disequilibrium." *BMC Genet* 14 (2013): 84. https://doi.org/10.1186/1471-2156-14-84.

Li, Yong; Wang, Lirong. "Genetic Resources, Breeding Programs in China, and Gene Mining of Peach: A Review." *Horticultural Plant Journal* 6, no. 4 (2020): 205-15. https://doi.org/10.1016/j.hpj.2020.06.001.

Mammadov, Jafar, Rajat Aggarwal, Ramesh Buyyarapu, and Siva Kumpatla. "Snp Markers and Their Impact on Plant Breeding." International Journal of Plant Genomics 2012 (2012/12/18 2012): 728398. https://doi.org/10.1155/2012/728398.

Mas-Gómez, Jorge, Celia M. Cantín, María Á. Moreno, Ángela S. Prudencio, Mar Gómez-Abajo, Luca Bianco, Michela Troggio, Pedro Martínez-Gómez, Manuel Rubio, and Pedro J. Martínez-García. "Exploring Genome-Wide Diversity in the National Peach (Prunus Persica) Germplasm Collection at Cita (Zaragoza, Spain)." *Agronomy* 11, 3 (2021): 481. https://doi.org/10.3390/agronomy11030481. https://mdpi-res.com/d_attachment/agronomy/agronomy-11-00481/article_deploy/agronomy-11-00481.pdf?version=1614927671.

Mason, Annaliese S. "Ssr Genotyping." *Methods Mol. Biol.* 1245 (2015): 77-89. https://doi.org/10.1007/978-1-4939-1966-6_6.

Meader, EM, and MA Blake. "Some Plant Characteristics of the Second Generation Progeny of Prunus Persica and Prunus Kansuensis Crosses." Paper presented at the Proc. Amer. Soc. Hort. Sci, 1939.

Medich, Rob, "An All-Encompassing Guide to Peaches." Harry & David (blog), https://www.harryanddavid.com/blog/types-of-peaches#:~:text=More%20than%20300%20types%20of,varieties%20can%20be%20found%20worldwide.

Parker, Lauren E, and John T Abatzoglou. "Warming Winters Reduce Chill Accumulation for Peach Production in the Southeastern United States." Climate 7, no. 8 (2019): 94.

Paudel, Lucky, Josh Clevenger, and Cecilia McGregor. "Refining of the Egusi Locus in Watermelon Using Kasp Assays." Scientia Horticulturae 257 (2019): 108665.

Purcell, Shaun, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, et al. "Plink: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." The American journal of human genetics 81, no. 3 (2007): 559-75.

Rafalski, J Antoni. "Novel Genetic Mapping Tools in Plants: Snps and Ld-Based Approaches." Plant science 162, no. 3 (2002): 329-33.

Raj, Anil, Matthew Stephens, and Jonathan K Pritchard. "Faststructure: Variational Inference of Population Structure in Large Snp Data Sets." *Genetics* 197, no. 2 (2014): 573-89.

Rieger, MARK, R Lo Bianco, and WR Okie. "Responses of Prunus Ferganensis, Prunus Persica and Two Interspecific Hybrids to Moderate Drought Stress." Tree Physiology 23, no. 1 (2003): 51-58.

Scorza, R. and Okie, W. R. . "Peaches (*Prunus*)." *ISHS Acta Horticulture* 290 (1991). https://doi.org/10.17660/ActaHortic.1991.290.5.

Semagn, K., Babu, R., Hearne, S. et al. "Single Nucleotide Polymorphism Genotyping Using Kompetitive Allele Specific Pcr (Kasp): Overview of the Technology and Its Application in Crop Improvement." *Mol. Breeding* 33 (2014): 1-14. https://doi.org/10.1007/s11032-013-9917-x.

Sharma, Samriti, Rajinder Kaur, Krishan Kumar, and Heerendra Sagar. "Cross-Transferability of Rubus Ellipticus Est–Ssr Markers for Genetic Diversity Analysis of Peach (Prunus Persica)." Genetic Resources and Crop Evolution (2023): 1-19.

Sharopova, Natalya; McMullen, Michael D.; Schultz, Linda; et al. "Development and Mapping of Ssr Markers for Maize." *Plant Molecular Biology* 48 (2001): 463-81. https://doi.org/10.1023/A:1014868625533.

SNPs and Other Genomic Variations Glossary: Evolving Terminology for Emerging Technologies, http://www.genomicglossaries.com/content/genetic_variations_gloss.asp.

Suo, Wei, Xiujin Shi, Sha Xu, Xiao Li, and Yang Lin. "Towards Low Cost, Multiplex Clinical Genotyping: 4-Fluorescent Kompetitive Allele-Specific Pcr and Its Application on Pharmacogenetics." PLoS One 15, no. 3 (2020): e0230445.

Steele, K., Tulloch, M.Q., Burns, M. et al. "Developing Kasp Markers for Identification of Basmati Rice Varieties." *Food Anal. Methods* 14 (2021): 663-73. https://doi.org/10.1007/s12161-020-01892-3.

Su, Tao; Wilf, Peter; Huang, Yongjiang; Zhang, Shitao; Zhou, Zhekun "Peaches Preceded Humans: Fossil Evidence from Sw China." *Sci Rep* 5 (2015). https://doi.org/10.1038/srep16794.

Sun, Congwei, Zhongdong Dong, Lei Zhao, Yan Ren, Ning Zhang, and Feng Chen. "The Wheat 660k Snp Array Demonstrates Great Potential for Marker-Assisted Selection in Polyploid Wheat." Plant Biotechnology Journal 18, no. 6 (2020): 1354-60.

Tautz, Diethard; Trick, Martin; Dover, Gabriel A. "Cryptic Simplicity in DNA Is a Major Source of Genetic Variation." Nature 322 (1986): 652-56. https://doi.org/10.1038/322652a0.

Untergasser, Andreas; Harm Nijveen, Xiangyu Rao, Ton Bisseling, René Geurts, and Jack A.M. Leunissen: Primer3Plus, an enhanced web interface to Primer3 Nucleic Acids Research 2007 35: W71-W74; doi:10.1093/nar/gkm3

USDA, APHIS. *Required Documents*. Feb 17, 2021. https://www.aphis.usda.gov/aphis/ourfocus/planthealth/import-information/permits/plants-and-plant-products-permits/plants-for-planting/buying-selling-plants-seeds-online/seeds/required-documents#:~:text=All%20seeds%20imported%20into%20the,See%20Seeds%20with%20Additional%20Requirements).

USDA. "Noncitrus Fruits and Nuts." USDA Economics, Statistics and Market Information System, 2023. https://usda.library.cornell.edu/concern/publications/zs25x846c?locale=en.

Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Paris R, Aramini V, Gazza L, Rossini L, Bassi D, Troggio M, Shu S, Grimwood J, Tartarini S, Dettori MT, Schmutz J (2017) The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. BMC Genomics 18:225 DOI: 10.1186/s12864-017-3606-9

Vieira, Maria Lucia Carneiro; Santini, Luciane; Diniz, Augusto Lima; Munhoz, Carla de Freitas. "Microsatellite Markers: What They Mean and Why They Are So Useful." Genet Mol Biol 39, no. 3 (2016): 312-28. https://doi.org/10.1590/1678-4685-GMB-2016-0027.

Vignal, Alain, Denis Milan, Magali SanCristobal, and André Eggen. "A Review on Snp and Other Types of Molecular Markers and Their Use in Animal Genetics." Genetics Selection Evolution 34, no. 3 (2002/05/15 2002): 275. https://doi.org/10.1186/1297-9686-34-3-275.

"Why Georgia Peaches May Be Hard to Come by This Summer." Food & Wine, 2023, 2023, https://www.foodandwine.com/georgia-peaches-shortage-crop-2023-7561913.

Wickham, Hadley, Winston Chang, and Maintainer Hadley Wickham. "Package 'Ggplot2'." *Create elegant data visualisations using the grammar of graphics. Version* 2, no. 1 (2016): 1-189.

Xu, Cheng, Yonghong Ren, Yinqiao Jian, Zifeng Guo, Yan Zhang, Chuanxiao Xie, Junjie Fu, et al. "Development of a Maize 55 K Snp Array with Improved Genome Coverage for Molecular Breeding." Molecular Breeding 37 (2017): 1-12.

Yoshida, M. "Peach Germplasm." Kajitsu Nippon 42 (1987): 70-74.

Zahid, Ghassan, Yıldız Aka Kaçar, Dicle Dönmez, Ayzin Küden, and Tommaso Giordani. "Perspectives and Recent Progress of Genome-Wide Association Studies (Gwas) in Fruits." Molecular Biology Reports 49, no. 6 (2022): 5341-52.

Zhong, Shengqiang, Jack CM Dekkers, Rohan L Fernando, and Jean-Luc Jannink. "Factors Affecting Accuracy from Genomic Selection in Populations Derived from Multiple Inbred Lines: A Barley Case Study." Genetics 182, no. 1 (2009): 355-64.

Facilities

The University of Georgia Griffin campus is located approximately 40 miles south of Atlanta, GA and 64 miles north-west of Byron, GA (33.26°N, -84.28°W). The facilities to be used for this research project are located within the Griffin Campus. The laboratory to be used is the Peach Research and Extension laboratory that contains standard basic equipment for sample processing with the capabilities of short- and long-term storage. The laboratory has the capabilities for molecular work, including DNA extraction and PCR. Samples processed for molecular work can be shipped directly to the Georgia Genomics Facility at University of Georgia, Athens. Genomic analysis services and advanced computing capabilities are available through the Georgia Advanced Computing Resource Center, which can be accessed in any part of the world. Software required for data analysis is readily available for free use within the server. Training and help sessions are easily accessible if required.

Dempsey Farm, a 12-acre peach satellite orchard for all major varieties, is currently available (33°14'55" N, 84°17'57" W). In addition there are also two greenhouses for general purposes. There are

two large auditoria for presenting workshops or for extension retreats. This location is about 40 min from the major peach growing region in Georgia.

Permits

USDA - Animal and Plant Health Inspection Service - Plant Protection and Quarantine – Application for Permit to Import Plants and Plant Products

- To be applied for

USDA – Permit to Import Plants and Plant Products

- To be applied for

Queensland Government – Maroochy Research Facility Agreement/Consent to Collect

- To be applied for

Budget

	Item	Amount (total)	Year 1	Year 2
Personnel $63,600	Master's Assistant (Caitlin McCann)	$55,600	$27,800	$27,800
	Master's Fellowship	$8,000	$4,000	$4,000
Equipment $29,500	Spark Multimode Microplate Reader, Tecan	$29,500	$29,500	$0
Supplies/Expenses $46,480	Rapid Genomics Capture Seq (250 samples)	$30,250	$30,250	$0
	PACE2.0Genotyping Master Mix ($762/25mL, 5µL/rxn)	$7,620	$7,620	$0
	KASP Primers (FAM, 1 OD)	$300	$150	$150
	KASP Primers (VIC, 1 OD)	$300	$150	$150
	KASP Primers (Rev, 3 OD)	$10	$5	$5
	DNA Extraction Materials, consumables	$6,000	$3,000	$3,000
	Publication Costs	$2,000	$0	$2,000
Travel $12,600	International Travel (Round Trip Plane Ticket x2)	$8,000	$0	$8,000
	Rental Car (one week)	$500	$0	$500
	Gas (one week, 2000 miles, $4.30 USD/gallon)	$500	$0	$500
	Per Diem (one week, two people)	$1,500	$0	$1,500
	Lodging (two rooms, 7 days, $150)	$2,100	$0	$2,100
Indirect Costs	42%	$63,915.60	$43,039.50	$20,876.10
Total Costs		**$216,095.60**	$145,514.50	$70,581.10

Budget Justification

Personnel - $63,600

Master's Assistant – This person will be responsible for data collection and analysis, DNA extraction, KASP reading, and phylogeny assembly.

Equipment - $29,500

Spark Multimode Microplate Reader, Tecan – to read the results of KASP

Supplies/Expenses - $46,480

Capture Seq – Sequence and identify SNPs in 250 accessions. The genomic base against which to compare Australian samples via KASP

KASP Primers– Compare alleles of 200 Australian samples to those previously sequenced. Generate data for use in genomic diversity study.
DNA Extraction materials – to extract DNA from Australian samples
Publishing costs – some journals, especially open access, require a fee to publish or submit a paper.

Travel - $12,600
Travelling to Australia to collect peach germplasm for study.
Visiting Maroochy Research Facility, 100km North of Brisbane in Queensland, and staying for one week to collect, clean, categorize, and ship germplasm.

Indirect Costs - $63,915.60
Indirect costs are calculated from MTDC using the AFRI negotiated rate of 42.0 %.
Year 1: $43,039.50
Year 2: $20,876.10

Caitlin McCann

442 Austin St., Downers Grove, IL 60515
https://redheadcaitlin.quarto.pub/caitlinmccann/
caitlinemccann01@gmail.com
630-432-3077

Professional Summary

Hardworking and committed graduate student eager to learn more about the fields of agriculture and horticulture. Accustomed to working alone and in groups, in busy and calm environments, and reliable under all circumstances. Motivated to accumulate new experiences and grow professionally.

Skills

- R Coding Language
- Laboratory Procedure
- 3D Printing and Modeling
- Hand Pollination
- Microsoft Office
- Smartsheet
- Lucidchart
- Tissue Sampling

Experience

NA R&D Seed Production Intern — *Jun 2022 – Aug 2022*
Syngenta — *Downers Grove, Illinois*

- Updated and improved training processes for North American locations.
- Trained site stewardship coordinators on use of Smartsheet training processes.
- Consolidated data and streamlined data management procedures.
- Visualized data using Smartsheet and Lucidchart.

Undergraduate Researcher — *Aug 2020 - Dec 2021*
University of Alabama - McKain Lab — *Tuscaloosa, Alabama*

- Collected samples with a team of graduate and undergraduate students in the field.
- Isolated DNA using CTAB protocol and constructed DNA libraries.
- Presented posters at URCA and Botany Conference 2021.

Undergraduate Researcher — *May 2021 - May 2023*
University of Alabama - Benstead Lab — *Tuscaloosa, Alabama*

- Organized and isolated samples of macroinvertebrates using a microscope.
- Helped gather data in a study on the environmental impact of stream warming.
- Trained a fellow undergraduate on use of tools and methods of data collection.

Seeds Operations Intern — *May 2023 – Jul 2023*
Syngenta — *Slater, Iowa*

- Led teams of seasonal workers in tissue sample collection and hand pollination.
- Prepared yield experiments on soybean crops.
- Organized and participated in trial field planting.
- Took performance notes on corn and soybean trial fields.

Education

Bachelor of Science: Biology — *Graduated May 2023*
The University of Alabama — *Tuscaloosa, Alabama*

- GPA 3.98, Graduated *Summa Cum Laude* and with Honors
- Minor in Liberal Arts (Blount Undergraduate Initiative)
- Awarded Presidential Scholarship
- Vice President of the Outdoor Adventures Club (2021)
- Studied abroad at Maynooth University, Ireland in Spring 2022

Master's of Science: Plant Breeding, Genetics, and Genomics — *Aug 2023 - current*
The University of Georgia — *Athens, Georgia*

- Assistantship in the lab of Dr. Dario Chavez
- Developing KASP markers for *Prunus persica* (peach) genome
- Awarded UGA Grad School Master's Fellow Award