



High Levels of Heterozygosity Found for 15 SSR Loci in *Solanum chacoense*

Kathleen G. Haynes¹ · Haitham E. M. Zaki^{1,2} · Christian T. Christensen³ · Elizabeth Ogden¹ · Lisa J. Rowland¹ · Matthew Kramer⁴ · Lincoln Zotarelli³

© The Potato Association of America 2017

Abstract Wild species-related germplasm is widely used to introduce new alleles and/or increase heterozygosity in cultivated species. Twenty-four SSR markers, specifically designed for cultivated potatoes, were evaluated to determine the extent of genetic variation within and among ten accessions of *Solanum chacoense* (*chc*). Fifteen of these markers were informative: there was no polymorphism in one of the markers, four of the markers showed evidence that more than one locus was being amplified, and the other four markers failed to consistently amplify products. Heterozygosity in these 10 accessions ranged from 33% to 87%. Variation among accessions was the largest proportion of variance for three markers, variation among genotypes within accessions was the largest proportion for three markers, and for the other nine markers variation within genotypes (chromosome to chromosome) was the largest proportion. Genetic similarity averaged 29.5% across markers. Where accessions have already been screened and found to possess the trait of interest, multiple genotypes from those accessions should be evaluated to identify genotypes with the greatest expression of the trait.

Resumen El germoplasma relacionado a especies silvestres se usa ampliamente para introducir nuevos alelos y/o aumentar la heterocigocidad en especies cultivadas. Se evaluaron 24 marcadores SSR específicamente diseñados para papas cultivadas, para determinar la amplitud de la variación genética dentro y entre diez introducciones de *Solanum*

chacoense (*chc*). Quince de estos marcadores fueron informativos: no hubo polimorfismo en uno de los marcadores, cuatro de ellos mostraron evidencia de que más de un locus estaba siendo amplificado, y los otros cuatro marcadores fallaron para amplificar consistentemente los productos. La heterocigocidad en estas diez introducciones varió de 33% a 87%. La variación entre las accesiones fue la proporción más grande de varianza para tres marcadores, la variación entre genotipos dentro de las introducciones fue la de mayor proporción para tres marcadores, y para los otros nueve la variación dentro de genotipos (cromosoma a cromosoma) fue la proporción más grande. La similitud genética promedió 29.5% entre los marcadores. En donde las accesiones ya han sido analizadas y que se encontró que poseen el carácter de interés, múltiples genotipos de estas introducciones deberían evaluarse para identificar genotipos con la mayor expansión del carácter.

Keywords Microsatellites · Potato · Heterozygosity · Heterosis · Genetic similarity

Introduction

Cultivated potato, *Solanum tuberosum* L., is an autotetraploid ($2n = 4x = 48$) asexually propagated crop, generally considered to have a narrow genetic base (Mendoza and Haynes

✉ Kathleen G. Haynes
kathleen.haynes@ars.usda.gov

¹ Genetic Improvement of Fruits and Vegetables Laboratory, U.S.D.A.-Agricultural Research Service, Beltsville, MD 20705, USA

² Horticulture Department, Faculty of Agriculture, Minia University, El-Minia 61517, Egypt

³ Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

⁴ Statistics Group, U.S.D.A.-Agricultural Research Service, Beltsville, MD 20705, USA

1974; Hawkes 1978; Love 1999), although the large number of single nucleotide polymorphisms (SNPs) uncovered by Hamilton et al. (2011) could potentially challenge this commonly held assumption. Nevertheless, both mid-parent and high-parent heterosis for total tuber yield has been widely observed in crosses between genetically diverse potato germplasm (Bani-Ameur et al. 1991; Buso et al. 1999a, 1999b, 2000, 2003; De Jong et al. 1981; Ran and Dai 1996; Tai and De Jong 1997; Veilleux and Lauer 1981).

Heterozygosity is thought to be an important factor in the heterotic response observed in crosses between diverse germplasm. Chase (1963) proposed using parthenotes and other diploid stocks to increase heterozygosity at the tetraploid level. Theoretically, a tetraploid could have as many as four different alleles per locus. Sanford and Hanneman (1982) compared complex hybrids containing genomes from *S. tuberosum* Group Tuberosum (one-way), and one (two-way) or two (three-way) other Groups. The two- and three-way hybrid Groups were superior to the one-way Group, but not significantly different from each other, suggesting there might be a heterotic threshold in cultivated potato, beyond which point more heterozygosity would not result in more yield or greater vigor. Bonierbale et al. (1993) tested the maximum heterozygosity hypothesis for tuber yield in crosses among adapted tetraploid germplasm and between adapted and unadapted tetraploid germplasm. Their results supported the maximum heterozygosity hypothesis in crosses involving adapted germplasm, but they found no support for the hypothesis for crosses between adapted and unadapted germplasm.

Genebanks for numerous crops have been established worldwide to preserve genetic diversity. Brown (1989) defined the concept of core collections as a way to capture the genetic diversity of the collection with minimal repetitiveness. Originally core collections were defined based on quantitative and/or qualitative traits and/or geographical site of collection (Huaman et al. 2000). However, del Rio and Bamberg (2002) found that geographic origin was not very useful in gauging interpopulation genetic diversity as measured by RAPDs in a wild species of potato. Jansky et al. (2006) reported that taxonomic relationships and ecogeographic data could not be used to predict where sources of white mold resistance genes would be found. More recently, molecular markers have been employed to construct core collections (Bamberg and del Rio 2014; Bamberg et al. 2016a).

Molecular markers would seem to offer an excellent tool to estimate genetic diversity. However, diversity in such markers depends on whether they are associated with genes under selection or not. Genes undergoing selection have less genetic diversity than neutral genes (McKay and Latta 2002). Using SNPs, Hirsch et al. (2013) found that modern potato breeding efforts have not noticeably changed the percentage of heterozygous loci or the frequency of homozygous, single-dose, and

duplex loci on a genome level in cultivated potato. The only changes they observed were for alleles in biosynthetic pathways important for market class-specific traits, such as pigmentation for specialty market varieties and carbohydrate metabolism for processing varieties. Reed and Frankham (2001) found that the correlation between neutral molecular markers and quantitative variation was weak.

Studies using molecular markers have suggested that there is greater genetic diversity among wild species accessions than within (Bamberg and del Rio 2014, 2016; van Treuren et al. 2004). However, several researchers have suggested that the amount of heterozygosity in the wild species is low. Using SNPs, Hirsch et al. (2013) estimated heterozygosity in the wild species at 7%. Also using SNPs, Hardigan et al. (2015) estimated heterozygosity in *S. chacoense* at 6%. However, in the later study, single plants from only three accessions of *S. chacoense* were investigated.

Solanum chacoense (*chc*) is a wild diploid species ($2n = 2x = 24$) that has been widely utilized in potato breeding. Resistance to Colorado potato beetle (Bamberg et al. 1996), verticillium wilt (Lynch et al. 1997), silver scurf (Rodriguez et al. 1995), root knot nematodes (Janssen et al. 1996; De Vito et al. 2003), late blight (Micheletto et al. 2000), soft rot (Bains et al. 1999) and potato leafroll virus (Brown and Thomas 1994) has been reported in *S. chacoense*. *S. chacoense* has also been reported to be tolerant to drought (Ekanayake and DeJong 1992), salinity (Bilski et al. 1988), and heat (Reynolds and Ewing 1989). In addition, Errebhi et al. (1998) found the two accessions of *chc* they screened to have good nitrogen uptake efficiency.

Based on Errebhi et al. (1998) and Bilski et al. (1988) results, we evaluated additional accessions of *chc* for root biomass (associated with N uptake efficiency) (Christensen et al. 2017) and salinity tolerance (Zaki et al. 2016). As part of this effort, SSRs were used to determine if there was greater variability among or within accessions of *chc*. Simple sequence repeats (SSRs) are particularly useful molecular markers as they are ubiquitous, co-dominant, multi-allelic, reproducible, have high levels of polymorphism, and require low amounts of DNA (Mc Gregor et al. 2000; Milbourne et al. 1997).

Simple sequence repeats (SSRs) have been used for an array of applications in potato including diversity and classification (Spooner et al. 2007), tracing germplasm migrations (Rios et al. 2007), fingerprinting (Moisan-Thierry et al. 2005), and genetic linkage mapping (Feingold et al. 2005). Recently, a more selective group of SSRs has been used to fingerprint 742 potato landraces (Ghislain et al. 2009). Potatoes included ranged in ploidy level from diploid to pentaploid. That study developed a new potato genetic identity (PGI) kit based on 24 SSR markers with two markers per chromosome that allowed for the discrimination of 93.5% of the 742 potato races compared to previous studies that used kits with 50 SSR markers (Spooner et al. 2007).

The purpose of this study was to use SSRs to determine how much genetic diversity exists in a random sample of plant introductions of *chc*. This information will be used to guide future breeding strategies utilizing *chc* germplasm.

Materials and Methods

Plant materials and DNA Extraction

The diploid *Solanum chacoense* (*chc*) genotypes used in this study came from ten randomly chosen accessions obtained from the NRSP-6 United States Potato Genebank in Sturgeon Bay, WI (Table 1) (www.ars-grin.gov/npgs/acc/acc_queries.html) with nine random genotypes from each accession for a total of 90 *chc* genotypes. Six of the accessions were collected in Argentina, three in Paraguay, and one in Bolivia. Since the genetic diversity in this material was unknown, we tried to balance the number of accessions with the number of genotypes within an accession. Five cultivars of *Solanum tuberosum* (*tbr*) were also included as positive controls to reference back to the potato genetic identity kit of Ghislain et al. (2009).

True seed of the *chc* accessions were planted in flats of ProMix in the greenhouse in Beltsville, Maryland and transplanted into 15 cm pots of ProMix approximately four weeks later.

Total DNA was extracted from young leaves of these accessions and the cultivars grown in a greenhouse using GenElute™ Plant Genomic DNA Miniprep Kit (Sigma, St. Louis, MO) according to the manufacturer's instructions.

Table 1 List of plant materials used

Accession/Clone	Species	Ploidy	Where Collected/Orioriginated
PI 275136	<i>chc</i>	2×	Jujuy, Argentina
PI 275142	<i>chc</i>	2x	Salta, Argentina
PI 320286	<i>chc</i>	2x	Cordoba, Argentina
PI 320288	<i>chc</i>	2x	Cordoba, Argentina
PI 320294	<i>chc</i>	2x	Buenos Aires, Argentina
PI 414153	<i>chc</i>	2x	Presidente Hayes, Paraguay
PI 537025	<i>chc</i>	2x	Chuquisaca, Bolivia
PI 566738	<i>chc</i>	2x	Asuncion, Paraguay
PI 566739	<i>chc</i>	2x	Asuncion, Paraguay
PI 566743	<i>chc</i>	2x	La Rioja, Argentina
Atlantic	<i>tbr</i>	4x	USA
Dark Red Norland	<i>tbr</i>	4x	USA
Lamoka	<i>tbr</i>	4x	USA
Red Pontiac	<i>tbr</i>	4x	USA
Snowden	<i>tbr</i>	4x	USA

Simple Sequence Repeat Primer Pairs and PCR Conditions

All 24 SSR primer pairs in the potato genetic kit of Ghislain et al. (2009) were used. To amplify the microsatellite loci, PCR was performed in a 20 µl volume containing 2 ng of genomic DNA, 1 µl of each primer (10 mM), 2 µl of dNTP (10 mM), 1 µl of 10 × reaction buffer (Takara, Japan), and 1 unit of *Taq* polymerase (Takara, Japan). PCR amplification followed the protocols provided by Ghislain et al. (2009). Two replicate PCRs were run for each SSR primer pair.

Separation of PCR products was achieved using the AdvanCE™ FS96 capillary electrophoresis system (Advanced Analytical Technologies, Inc., Ames, IA, USA). All separations were run using DNF-900 gel for detecting fragments in the 35–500 bp range. Each reaction was run twice and the consistent alleles scored with PROSize™ 2.0 data analysis software (Advanced Analytical Technologies, Inc., Ames, IA, USA). Inconsistent alleles between the two reps were scored as missing.

Data Analysis

Not all markers were deemed reliable. We used a subset of 15 (STG0001, STI0004, STM0037, STM1104, STM5121, STG0010, STI0030, STM1052, STM1106, STM5127, STG0025, STI0033, STM1053, STM5114, STPoAc58), and excluded the known tetraploid cultivars from the genetic analyses. Percent heterozygosity for each accession, genotype within accession and cultivar was calculated as the number of polymorphic loci divided by 15.

Genetic similarity (GS), defined as the probability of randomly choosing two individuals having the same genotype from a population of these ten accessions, was calculated for each marker as the sum of the probabilities of each allelic combination squared. For the simplest example of two alleles (STM1053):

- the probability that one individual is (169,169) = (0.83*0.83) = 0.6889;
- the probability that one individual is (169,174) = (0.83*0.17*2) = 0.2822; and,
- the probability that one individual is (174,174) = (0.17*0.17) = 0.0289.

So, the probability that any two randomly chosen genotypes are the same = the probability that both are (169,169) + the probability that both are (169,174) + the probability that both are (174,174) = $0.6889^2 + 0.2822^2 + 0.0289^2 = 0.5551$. Obviously, this assumes that the population (of ten accessions) is in Hardy-Weinberg equilibrium, which given the population sizes of this study and the method of advancing the population in the genebank (by accession)

would not be true. Nevertheless, it does give some indication of how similar individuals would be if the population was in H-W equilibrium.

Finally, in order to compare *chc*, a diploid, with *tbr*, a tetraploid, each accession of *chc* was theoretically made tetraploid by combining any two of the nine genotypes within that accession in all possible combinations. The expected percent heterozygous loci was then calculated.

For each marker, the amounts of variation among accessions, among genotypes within an accession, and within the genotype (i.e. chromosome to chromosome) were calculated using the R poppr (ver. 2.0.2) package (Kamvar et al. 2014). An analysis of molecular variance (AMOVA) (Excoffier et al. 1992), as implemented in the R ade4 package (Dray and Dufour 2007; called by poppr) was conducted looking at all markers together and then each marker individually.

Results and Discussion

We found one of the markers, STI0014, to be monomorphic in the 90 genotypes of *chc* evaluated. Four of the markers failed to amplify bands in many of the *chc* accessions: STI0012, STI0032, STM0019, and STM0031. As with *chc*, band amplification was inconsistent for three of these same SSR markers in *tbr*: STI0032, STM0019 and STM0031. Published PCR conditions for these four markers may not have been ideal for this germplasm and may require additional refinement. Four of the markers showed evidence that more than one locus in *chc* was being amplified: STG0016, STI0001, STI0003, and STM1064. The inconsistencies between the results of this study and others may be due to cross-species amplification problems. Cross-species amplification problems, whereby markers developed for one species cannot be utilized in another species, have been widely reported for numerous plant species including *Allium* sp. (Lee et al. 2011), *Bambusa* sp. (Nayak and Rout 2005), *Capsicum* sp. (Nagy et al. 2007), *Glycine* sp. (Peakall et al. 1998), *Gossypium* sp. (Guo et al. 2006), *Lolium* sp. (Jensen et al. 2007), *Prunus* sp. (Cipriani et al. 1999), *Rubus* sp. (Lopes et al. 2006), *Solanum* sp. (Salim et al. 2011), *Trifolium* sp. (Kolliker et al. 2001), and *Vicia* sp. (Raveendar et al. 2015).

Ascertainment bias may be another source of the inconsistencies between the results of this study and others. STI0014 is a prime example of this problem. STI0014 was originally chosen because it was informative in the cultivated potato gene pool, but was not informative in this sample of *chc*. Bamberg et al. (2015) found that the relatedness of a species to *S. tuberosum*, for which SNPs for the Infinium 8303 Potato SNP Array were developed, was inversely related to the number of polymorphic SNPs it contained. The same could reasonably be expected to be true for other molecular markers developed for one species but applied to a different one.

Thus, of the 24 SSR markers suggested by Ghislain et al. (2009) in order to standardize testing and reporting across laboratories, only 15 were finally scored and found to be informative. Assuming the SSR markers scored occurred in the same general location in *chc* as reported for the landraces (Ghislain et al. 2009), no markers were scored for chromosomes IV and VII, and only one marker was scored on chromosomes I, V, VI, VIII, and IX.

Solanum chacoense

Of the 15 markers scored, the number of alleles ranged from two (STM1053 and STM5127) to eleven (STM0037) in *chc*, with an average of five (Table 2). All together, 75 alleles were scored. These 75 alleles were able to discriminate among all 90 *chc* genotypes, i.e., each genotype had a unique banding pattern. Nineteen of the alleles occurred with a frequency less than 5%. Two of the alleles were very common, occurring with a frequency > 95%.

In general, the number of alleles we found for the SSR markers in *chc* was lower than reported by Ghislain et al. (2009). The lower number of alleles in this *chc* germplasm may partially be the result of the way seed was increased at the NRSP-6 United States Potato Genebank. Typically, 20 plants per accession were planted periodically for seed increase and inter-mated within the accession. However, in two of the accessions (PI 275142 and PI 566738) a genetic bottleneck occurred during seed increase when only two plants were available for seed increase, and in one accession (PI 566739) only five plants were available for seed increase (John Bamberg, personal communication).

Overall, *chc* averaged 64% heterozygosity for these 15 loci (Table 3). Among the accessions, heterozygosity ranged from 33 to 80%. The range of heterozygosity within an accession was always lower than for the accession as a whole as a result of the presence of alleles with low frequencies (Table 2). PI 275142, one of the accessions with only two plants initially, was heterozygous at eight of the markers, whereas PI 566738, the other accession with only two plants initially, was heterozygous at 12 of the markers. PI 566739, the accession with only five plants initially, was heterozygous at 13 of the markers (Table 3). Clearly these accessions entered the genebank with a high degree of heterozygosity. Although there was no genetic bottleneck reported during seed increase at the genebank, PI 320294 and PI 566743 were less heterozygous for many of the markers, indicating that some alleles may have been fixed prior to the accessions' introduction into the genebank. Our estimates for *chc* were similar to the estimates obtained by Bamberg and del Rio (2004) for *S. jamesii*, another diploid outcrosser, as is *chc*: They estimated heterozygosity as 70% using RAPDs and 51% using SNPs. In contrast, Hardigan et al. (2015) found that heterozygosity in 12 South American wild diploid species ranged from about 2% to

Table 2 SSR marker, allele sizes and frequencies for ten accessions of *Solanum chacoense* and the estimate of genetic similarity (GS) for any two individuals randomly sampled from the population of these accessions

SSR Marker	Allele Sizes (bp)	Allele Frequencies	GS
STG0001	127, 133, 139, 145	0.09, 0.31, 0.47, 0.13	17.6%
STG0010	149, 156, 162, 166	0.13, 0.47, 0.39, 0.01	23.2%
STG0025	191, 197, 202	0.27, 0.57, 0.15	24.2%
STI0004	80, 83, 86	0.42, 0.51, 0.07	29.1%
STI0030	83, 85, 89, 93, 97, 102, 106, 117, 123	0.08, 0.14, 0.31, 0.13, 0.12 0.11, 0.02, 0.08, 0.01	4.9%
STI0033	100, 107, 113, 117, 121, 127, 133, 141	0.02, 0.03, 0.08, 0.04, 0.24, 0.46, 0.12, 0.01	12.3%
STM0037	71, 73, 80, 84, 87, 91, 95, 100, 114, 119, 125	0.02, 0.05, 0.08, 0.23, 0.12, 0.22, 0.22, 0.02, 0.01, 0.02, 0.01	5.3%
STM1052	197, 207, 212	0.01, 0.92, 0.07	73.3%
STM1053	169, 174	0.83, 0.17	55.5%
STM1104	170, 175, 179, 185, 189	0.24, 0.16, 0.40, 0.12, 0.08	11.0%
STM1106	130, 141, 147, 156, 163, 172	0.15, 0.46, 0.11, 0.07, 0.16, 0.05	11.0%
STM5114	278, 283, 290, 297, 303	0.09, 0.31, 0.46, 0.12, 0.03	16.5%
STM5121	274, 278, 285, 293	0.08, 0.31, 0.54, 0.06	22.2%
STM5127	237, 264	0.99, 0.01	96.1%
STPoAc58	225, 230, 237, 245, 249, 257	0.78, 0.06, 0.03, 0.08, 0.02, 0.02	39.9%
Average			29.5%

13%, depending on species. However, their data was based on a single genotype for each accession of the wild species, with most wild species represented by three accessions. Their values were in the lower ranges for individual genotypes obtained in this study. Hardigan et al. (2015) 'suggest that

Table 3 Percent heterozygous loci for 15 SSR markers in nine accessions of *Solanum chacoense* and five *S. tuberosum* cultivars, and the range in percent heterozygous loci for the same 15 SSR markers for the nine individual genotypes within each accession. Expected percent heterozygous loci for the same 15 SSR markers assuming these nine accessions of *S. chacoense* were tetraploid

Accession	% Heterozygosity	Range	Expected % heterozygosity Assuming Tetraploidy
PI 275136	73	13–43	80
PI 275142	53	20–33	93
PI 320286	53	7–33	67
PI 320288	73	14–29	73
PI 320294	47	7–26	73
PI 414153	73	7–47	93
PI 537025	67	20–47	73
PI 566738	80	21–47	93
PI 566739	87	20–53	87
PI 566743	33	0–27	60
Atlantic	47		
Dark Red Norland	40		
Lamoka	47		
Red Pontiac	60		
Snowden	47		

selection pressures within natural populations and breeding programs support allelic diversity at discrete loci and that the SNPs for which a heterozygous state confers agricultural benefit in cultivars have limited overlap with those for which heterozygosity is selected in wild populations.' Thus, it is possible that unbiased markers may indicate a higher level of heterozygosity in the wild species. Since the majority of the wild potato species are diploid outcrossers, our results for *chc* and Bamberg and del Rio's (2004) results with *S. jamesii* suggest that heterozygosity will likewise be high in other self-incompatible diploid potato species.

The lower number of alleles in this *chc* germplasm may also be the result of the numbers sampled. Based on intra-accession heterogeneity using RAPD markers, Bamberg and del Rio (2004) estimated that 25–30 plants would need to be sampled to adequately determine the genetic diversity within the accession. However, nearly all RAPDs are dominant and unable to distinguish between the homozygous dominant and its heterozygous counterpart. In contrast, SSR markers are co-dominant and thus allow both alleles of the heterozygote to be visualized.

Hosaka and Hanneman (1991) examined phenotypic variation for potato seed protein using electrophoresis in several potato species, including *chc*. Their index of variation, which measured the proportion of segregating protein bands per total protein bands detected among 10 individuals of an S₂ population of *chc* was 49%, and decreased with increasing generations of selfing. This estimate of phenotypic variation is in the range reported for genotypic variation in *chc*.

Genetic similarity for these markers ranged from a low of 4.9% for STI0030 to a high of 96.1% for STM5127 (Table 2), with an average across all markers of 29.5%. These values were calculated assuming that the population was in Hardy-Weinberg equilibrium, which was not true for any of the markers (data not shown). However, with random mating, as population size increases, observed GS values would more closely approximate predicted GS values. Obviously, though, the population is extremely heterogeneous. Bamberg and del Rio (2004) looked at the apparent GS of two random samples from the same population and found that for outcrossing species samples, bulked samples of five or more genotypes would give GS estimates of 90% or more; that is, based on two alleles (RAPD bands present or absent) there was a 90% or greater probability of finding both alleles in the samples. The data in our study clearly demonstrate that estimates of GS depend on the number of alleles and the frequency of the predominant allele(s).

The analysis for all markers estimated the among accession variance proportion at 31.4%, genotypes within accessions at 22.1%, and within genotypes (chromosome to chromosome variance) at 46.5% (Table 4). For individual marker analysis, variation among accessions ranged from a low of 6.3% to a high of 51.2%; variation among genotypes within accessions ranged from -6.0% to 64%; and variation within a genotype ranged from 4% to 100%. Clearly genotypes within an

accession are not identical for these markers, though they are more similar to each other than they are to genotypes of other accessions. Van Treuren et al. (2004) found that 91% of the observed variation in AFLPs from the series *Acaulia* could be found among accessions. Individual marker analysis is far less reliable than overall marker analysis, in part due to problems estimating parameters with small data sets and in part to markers with very little variation. Indicative of the problems with individual marker analysis, two markers, STM0037 and STM5127, had a negative variance estimate and should be considered to be zero.

One assumption in studies such as this one is that general neutral marker diversity is expected to be an indicator of general useful phenotypic diversity. Thus, one might expect that higher levels of heterozygosity within an accession would translate into greater phenotypic variation, resulting in both superior and inferior phenotypes. Conversely, lower levels of heterozygosity within an accession would translate into less phenotypic variation, resulting in a smaller range of phenotypic values. Recent research by Christensen et al. (2017) examined rooting characteristics in tissue culture in some of the same accessions used in this study: genotypes were clustered into clusters representing good, moderate, and poor rooting characteristics. PI 566743 and PI 320294 had the lowest levels of heterozygosity in this study at 33% and 47%, respectively. Christensen et al. (2017) found that one genotype of PI 566743 clustered in the group with good rooting characteristics, while the other genotype clustered in the moderate group. One genotype of PI 320294 clustered in the good group, while the other clustered in the poor group. There were three accessions with high levels of heterozygosity in this study: PI 566738 at 80%, PI 275136 at 73%, and PI 320288 at 73%. In the study by Christensen et al. (2017) PI 566738 genotypes clustered in the moderate and poor clusters; PI 275136 genotypes clustered in the good and moderate clusters; and, PI 320288 genotypes clustered tightly in the moderate cluster. Thus, there was no apparent correlation between marker diversity and phenotypic variation for rooting characteristics. This agrees with studies by Spooner et al. (2009) and Jansky et al. (2015) where there appears to be a disconnect between genetic diversity and phenotypic diversity. However, Bamberg and del Rio (2014) established a core collection of *S. microdontum* using AFLPs and found that the core collection contained 25 of 26 of the most desirable phenotypic traits, suggesting that there is a relationship between genetic and phenotypic diversity. In a later study, Bamberg et al. (2016b) found visual impression cogs to be a rapid, low-tech and low-cost first step to detecting pools of genetic diversity within species. Obviously, the relationship between genetic and phenotypic diversity will depend on a number of factors, among them the kinds and number of markers (Ghislain et al. 2006), whether the marker is neutral or associated with the trait in question, the number of

Table 4 Percentage variation among accessions, among genotypes within accessions, and within a genotype (i.e. chromosome to chromosome) for 15 SSR markers individually and over all 15 markers as revealed by AMOVA

SSR Marker	Among accessions	Among genotypes (accessions)	Within genotypes
STG0001	29.0	36.9	34.1
STG0010	29.6	15.4	55.0
STG0025	13.4	25.4	61.1
STI0004	47.1	7.7	45.2
STI0030	51.2	20.9	27.9
STI0033	36.9	22.0	41.1
STM0037	33.8	-6.0	72.2
STM1052	11.8	64.0	24.2
STM1053	37.9	58.1	4.0
STM1104	26.2	19.6	54.2
STM1106	40.1	37.6	22.3
STM5114	26.3	35.5	38.2
STM5121	21.1	16.3	62.6
STM5127	6.3	-6.3	100.0
STPoAc58	19.0	0.6	80.4
Over all	31.4	22.1	46.5

genes governing the trait, the mode of gene action, and whether the trait is quantitative or qualitative.

Another general assumption is that genetic diversity is important for generating phenotypic diversity. Tam et al. (1992) discuss this as one of the two main reasons for utilizing diploid species in breeding. On the other hand, little segregation is expected in crosses between highly homozygous parents. However, Endelman and Jansky (2016) crossed a doubled monoplod with a (theoretically) highly homozygous clone of *S. chacoense*. The progeny were expected to be heterozygous but homogeneous; however, the progeny tubers showed tremendous phenotypic variation. The powerful genetic tools developed and continuing to be developed reveal that potato breeding is full of surprises.

Solanum tuberosum L.

Of the 15 markers scored in common with *chc*, the number of alleles ranged from one to six in *S. tuberosum*. Two of the markers had only one allele across all five cultivars: STM1053 and STM5121. Heterozygosity ranged from 40% to 60% for each cultivar. Over all markers and all cultivars, heterozygosity averaged 48%. This is in close agreement with the 59% reported by Hirsch et al. (2013) for chip processing cultivars. Also, Bamberg and del Rio (2004) reported that genetic heterogeneity in *S. sucrense*, another tetrasomic outcrosser, was 44%.

In order to compare *chc* to *tbr* at the same ploidy level, within each accession *chc* was theoretically made tetraploid by combining any two of the nine genotypes within that accession in all possible combinations. The expected percent heterozygous loci was then calculated (Table 3). The expected percent heterozygous loci at the tetraploid level was higher than the observed value at the diploid level when the homozygous diploid was heterogeneous at that locus, for example when both A_1A_1 and A_2A_2 genotypes were found in that accession. The expected heterozygosity in tetraploid *chc* was usually much higher than the observed heterozygosity in the tetraploid *tbr* varieties included in this study and reported by others (Hirsch et al. 2013). If heterozygosity *per se* is important in the observed heterotic responses frequently observed in potato when unrelated germplasm is crossed, obviously there is considerable room for increasing heterozygosity. However, Bonierbale et al. (1993) found that heterozygosity was important for yield in adapted germplasm but not for unadapted germplasm. Tetraploid *chc* have been synthesized (Sanford et al. 1997) and our results suggest that intercrossing these tetraploids could result in a tetraploid population with greater heterozygosity than is currently present in commercial varieties. Nevertheless, the fact that the heterozygosity of commercial varieties hovers around 50% suggests that after over a century of breeding efforts, not all traits are impacted equally by heterozygosity.

Conclusions

Ninety genotypes of *Solanum chacoense*, represented by nine genotypes within each of ten accessions were subject to genetic analyses using SSR markers from the potato genetic identity kit proposed by Ghislain et al. (2009). These markers revealed a high level of heterozygosity in *chc*. Where accessions have already been screened and found to possess the trait of interest, multiple genotypes from those accessions should be evaluated to identify genotypes with the greatest expression of the trait. However, where accessions have not previously been screened for a given trait, sampling should focus on trying to balance a reasonable number of accessions with a reasonable number of genotypes within those accessions.

Acknowledgements The authors are grateful to Drs. Benildo de los Reyes and James Olmstead for their insights and advice on the project. We also thank Dr. John Bamberg and Max Martin at the Potato Gene Bank for experimental material. Thanks, too, to Dr. Bamberg for his many helpful suggestions to improve the manuscript.

References

- Bains, P.S., V.A. Bisht, D.R. Lynch, L.M. Kawchuk, and J.P. Helgeson. 1999. Identification of stem soft rot (*Erwinia carotovora* subspecies *atroseptica*) resistance in potato. *American Journal of Potato Research* 76: 137–141.
- Bamberg, J.B., and A. del Rio. 2004. Genetic heterogeneity estimated by RAPD polymorphism of four tuber-bearing potato species differing by breeding system. *American Journal of Potato Research* 81: 377–383.
- Bamberg, J., and A. del Rio. 2014. Selection and validation of an AFLP marker core collection for the wild potato *Solanum microdontum*. *American Journal of Potato Research* 91: 368–375.
- Bamberg, J., and A. del Rio. 2016. Accumulation of genetic diversity in the US potato genebank. *American Journal of Potato Research* 93: 430–435.
- Bamberg, J.B., C.A. Longtine, and E.B. Radcliffe. 1996. Fine screening *Solanum* (potato) germplasm accessions for resistance to Colorado potato beetle. *American Potato Journal* 73: 211–223.
- Bamberg, J.B., A. del Rio, J. Coombs, and D. Douches. 2015. Assessing SNPs versus RAPDs for predicting heterogeneity and screening efficiency in wild potato (*Solanum*) species. *American Journal of Potato Research* 92: 276–283.
- Bamberg, J., A. del Rio, D. Kinder, L. Louderback, B. Pavlik, and C. Fernandez. 2016a. Core collections of potato (*Solanum*) species native to the USA. *American Journal of Potato Research* 93: 564–571.
- Bamberg, J.B., A. del Rio, and D.A. Navarre. 2016b. Intuitive visual impressions (cogs) for identifying clusters of diversity within potato species. *American Journal of Potato Research* 93: 350–359.
- Bani-Ameur, F., F.I. Lauer, R.E. Veilleux, and A. Hilali. 1991. Genetic composition of 4x-2x potato hybrids – influence of *Solanum chacoense*. *Genome* 34: 413–420.
- Bilski, J.J., D.C. Nelson, and R.L. Conlon. 1988. Response of six wild species to chloride and sulfate salinity. *American Potato Journal* 65: 605–612.
- Bonierbale, M.W., R.L. Plaisted, and S.D. Tanksley. 1993. A test of the maximum heterozygosity hypothesis using molecular markers in tetraploid potatoes. *Theoretical and Applied Genetics* 86: 481–491.

- Brown, A.H.D. 1989. Core collections: a practical approach to genetic resources management. *Genome* 31: 818–824.
- Brown, C.R., and P.E. Thomas. 1994. Resistance to potato leafroll virus derived from *Solanum chacoense* – characterization and inheritance. *Euphytica* 74: 51–57.
- Buso, J.A., L.S. Boiteux, and S.J. Peloquin. 1999a. Multitrait selection system using populations with a small number of interploidy (4x-2x) hybrid seedlings in potato: Degree of high-parent heterosis for yield and frequency of clones combining quantitative agronomic traits. *Theoretical and Applied Genetics* 99: 81–91.
- Buso, J.A., F.J.B. Reifschneider, L.S. Boiteux, and S.J. Peloquin. 1999b. Effects of 2n-pollen formation by first meiotic division restitution with and without crossover on eight quantitative traits in 4x-2x potato progenies. *Theoretical and Applied Genetics* 98: 1311–1319.
- Buso, J.A., L.S. Boiteux, and S.J. Peloquin. 2000. Heterotic effects for yield and tuber solids and type of gene action for five traits in 4x potato families derived from interploidy (4x-2x) crosses. *Plant Breeding* 119: 111–117.
- Buso, J.A., L.S. Boiteux, and S.J. Peloquin. 2003. Tuber yield and quality of 4x-2x (FDR) potato progenies derived from the wild diploid species *Solanum berthaultii* and *Solanum tarijense*. *Plant Breeding* 122: 229–232.
- Chase, S.S. 1963. Analytic breeding in *Solanum tuberosum* L. – a scheme utilizing parthenotes and other diploid stocks. *Canadian Journal of Genetics and Cytology* 5: 359–363.
- Christensen, C.T., L. Zotarelli, K.G. Haynes, and J. Colee. 2017. Rooting characteristics of *Solanum chacoense* and *Solanum tuberosum* in vitro. *American Journal of Potato Research*. <https://doi.org/10.1007/s12230-017-9597-x>.
- Cipriani, G., G. Lot, W.G. Huang, M.T. Marrazzo, E. Peterlunger, and R. Testolin. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: Isolation, characterization and cross-species amplification in *Prunus*. *Theoretical and Applied Genetics* 99: 65–72.
- De Jong, H., G.C.C. Tai, W.A. Russell, G.R. Johnston, and K.G. Proudfoot. 1981. Yield potential and genotype-environment interactions of tetraploid-diploid (4x-2x) potato hybrids. *American Journal of Potato Research* 58: 191–199.
- De Vito, M., N. Greco, D. Carputo, and L. Frusciante. 2003. Response of wild and cultivated clones to Italian populations of root knot nematodes *Meloidogyne* spp. *Nematropica* 33: 65–72.
- del Rio, A.H., and J.B. Bamberg. 2002. Lack of association between genetic and geographic origin characteristics for the wild potato *Solanum sucrense* Hawkes. *American Journal of Potato Research* 79: 335–338.
- Dray, S., and A.B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v022.i04>.
- Ekanayake, I.J., and J.P. DeJong. 1992. Stomatal response of some cultivated and wild tuber-bearing potatoes in warm tropics as influenced by water deficits. *Annals of Botany* 70: 53–60.
- Endelman, J.B., and S.H. Jansky. 2016. Genetic mapping with an inbred line-derived F₂ population in potato. *Theoretical and Applied Genetics* 129: 935–943.
- Errebhi, M., C.J. Rosen, F.L. Lauer, M.W. Martin, J.B. Bamberg, and D.E. Birong. 1998. Screening of exotic potato germplasm for nitrogen uptake and biomass production. *American Journal of Potato Research* 75: 93–100.
- Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.
- Feingold, S., J. Lloyd, N. Norero, M. Bonierbale, and J. Lorenzen. 2005. Mapping and characterization of new EST-derived microsatellites for potato (*Solanum tuberosum* L.). *Theoretical and Applied Genetics* 111: 456–466.
- Ghislain, M., D. Andrade, F. Rodriguez, R.J. Hijmans, and D.M. Spooner. 2006. Genetic analysis of the cultivated potato *Solanum tuberosum* L. Phureja Group using RAPDs and nuclear SSRs. *Theoretical and Applied Genetics* 113: 1515–1527.
- Ghislain, M., J. Núñez, M.R. Herrera, J. Pignataro, F. Guzman, M. Bonierbale, and D.M. Spooner. 2009. Robust and highly informative microsatellite-based genetic identity kit for potato. *Molecular Breeding* 23: 377–388.
- Guo, W.Z., W. Wang, B.L. Zhou, and T.Z. Zhang. 2006. Cross-species transferability of G-arboreum-derived EST-SSRs in the diploid species of *Gossypium*. *Theoretical and Applied Genetics* 112: 1573–1581.
- Hamilton, J.P., C.N. Hansey, B.R. Whitty, K. Stoffel, A.N. Massa, A. van Deynze, W.S. De Jong, D.S. Douches, and C.R. Buell. 2011. Single nucleotide polymorphism discovery in elite North American potato germplasm. *Bio Med Central Genomics* 12: 302.
- Hardigan, M.A., J. Bamberg, C.R. Buell, and D.S. Douches. 2015. Taxonomy and genetic differentiation among wild and cultivated germplasm of *Solanum* sect. *Petota*. *The Plant Genome* 8: 1–16.
- Hawkes, J.G. 1978. Genetic poverty of the potato in Europe. In *Broadening the genetic base of crops*, ed. A.C. Zeven and A.M. van Harten. Pudoc: Wageningen.
- Hirsch, C.N., C.D. Hirsch, K. Felcher, J. Coombs, D. Zarka, A. van Deynze, W. De Jong, R.W. Veilleux, S. Jansky, P. Bethke, D.S. Douches, and C.R. Buell. 2013. Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *Genes, Genomes, Genetics* 3: 1003–1013.
- Hosaka, K., and R.E. Hanneman Jr. 1991. Seed protein variation with accessions of wild and cultivated potato species and inbred *Solanum chacoense*. *Potato Research* 34: 419–428.
- Huaman, Z., R. Ortiz, and R. Gomez. 2000. Selecting a *Solanum tuberosum* subsp. *andigena* core collection using morphological, geographical, disease and pest descriptors. *American Journal of Potato Research* 77: 183–190.
- Jansky, S.H., R. Simon, and D.M. Spooner. 2006. A test of taxonomic predictivity: Resistance to white mold in wild relatives of cultivated potato. *Crop Science* 46: 2561–2570.
- Jansky, S.H., J. Dawson, and D.M. Spooner. 2015. How do we address the disconnect between genetic and morphological diversity in germplasm collections? *American Journal of Botany* 102: 1213–1215.
- Janssen, G.J.W., A. van Norel, B. VerkerBakker, and R. Janssen. 1996. Resistance to *Meloidogyne chitwoodi*, *M-fallax* and *M-hapla* in wild tuber-bearing *Solanum* spp. *Euphytica* 92: 287–294.
- Jensen, L.B., P.B. Holm, and T. Lubberstedt. 2007. Cross-species amplification of 105 *Lolium* perenne SSR loci in 23 species within the Poaceae. *Molecular Ecology Notes* 7: 1151–1161.
- Kamvar, Z.N., J.F. Tabima, and N.J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 4 (2): e2811. <https://doi.org/10.7717/peerj.2811>.
- Kolliker, R., E.S. Jones, M.C. Drayton, M.P. Dupal, and J.W. Forster. 2001. Development and characterisation of simple sequence repeat (SSR) markers for white clover (*Trifolium repens* L.). *Theoretical and Applied Genetics* 102: 416–424.
- Lee, G. A., S. J. Kwon, Y. J. Park, M. C. Lee, H. H. Kim, J. S. Lee, S. Y. Lee, J. G. Gwag, C. K. Kin, and K. H. Ma. 2011. Cross-amplification of SSR markers developed from *Allium sativum* to other *Allium* species. *Scientia Horticulturae* 128: 401–407.
- Lopes, M.S., G. Maciel, D. Mendonca, F.S. Gil, and A.D.C. Machado. 2006. Isolation and characterization of simple sequence repeat loci in *Rubus hochstetterorum* and their use in other species from the Rosaceae family. *Molecular Ecology Notes* 6: 750–752.
- Love, S.L. 1999. Founding clones, major contributing ancestors, and exotic progenitors of prominent North American potato cultivars. *American Journal of Potato Research* 76: 263–272.

- Lynch, D.R., M.L. Kawchuk, J. Hachey, P.S. Bains, and R.J. Howard. 1997. Identification of a gene conferring high levels of resistance to *Verticillium* wilt in *Solanum chacoense*. *Plant Disease* 81: 1011–1014.
- Mc Gregor, C.E., C.A. Lambert, M. Greyling, J.H. Louw, and L. Warnich. 2000. A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica* 113: 135–144.
- McKay, J., and R.G. Latta. 2002. Adaptive population divergence: Markers, QTL and traits. *Trends in Ecology and Evolution* 17: 285–291.
- Mendoza, H.A., and F.L. Haynes. 1974. Genetic relationship among potato cultivars grown in the United States. *Hort Science* 9: 328–330.
- Micheletto, S., R. Boland, and M. Huarte. 2000. Argentinian wild diploid *Solanum* species as sources of quantitative late blight resistance. *Theoretical and Applied Genetics* 101: 902–906.
- Milbourne, D., R. Meyer, J.E. Bradshaw, E. Baird, N. Bonar, J. Provan, W. Powell, and R. Waugh. 1997. Comparison of PCR based marker systems for the analysis of genetic relationships in cultivated potato. *Molecular Breeding* 3: 127–136.
- Moisan-Thiery, M., S. Marhadour, M.C. Kerlan, N. Dessenne, M. Perramant, T. Gokelaere, and Y. Le Hingrat. 2005. Potato cultivar identification using simple sequence repeats markers (SSR). *Potato Research* 48: 191–200.
- Nagy, I., A. Stigel, Z. Sasvari, M. Roder, and M. Ganal. 2007. Development, characterization, and transferability to other Solanaceae of microsatellite markers in pepper (*Capsicum annuum* L.). *Genome* 50: 668–688.
- Nayak, S., and G.R. Rout. 2005. Isolation and characterization of micro satellites in *Bambusa arundinacea* and cross species amplification in other bamboos. *African Journal of Biotechnology* 4: 151–156.
- Peakall, R., S. Gilmore, W. Keys, M. Morgante, and A. Rafalski. 1998. Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for the transferability of SSRs in plants. *Molecular and Biological Evolution* 15: 1275–1287.
- Ran, Y., and C. Dai. 1996. Heterosis and genetic analysis of yield and tuber characters of 4x hybrids from 2x-2x, 4x-2x and 2x-4x crosses in potato. *Acta Agronomica Sinica* 22: 745–749.
- Raveendar, S., G.A. Lee, Y.A. Jeon, Y.J. Lee, J.R. Lee, G.T. Cho, J.H. Cho, J.H. Park, K.H. Park, K.H. Ma, and J.W. Chung. 2015. Cross-amplification of *Vicia sativa* subsp *sativa* microsatellites across 22 other *Vicia* species. *Molecules* 20: 1543–1550.
- Reed, D.H., and R. Frankham. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55: 1095–1103.
- Reynolds, M.P., and E.E. Ewing. 1989. Heat tolerance in tuber bearing *Solanum* species – a protocol for screening. *American Potato Journal* 66: 63–74.
- Rios, D., M. Ghislain, F. Rodriguez, and D.M. Spooner. 2007. What is the origin of the European potato? Evidence from Canary Island landraces. *Crop Science* 47: 1271–1280.
- Rodríguez, D.A., G.A. Secor, N.C. Gudmestad, and K. Grafton. 1995. Screening tuber-bearing *Solanum* species for resistance to *Helminthosporium solani*. *American Potato Journal* 72: 669–679.
- Salim, A.P., S. Mohankumar, and Y. Perumal. 2011. Cross species micro-satellite marker amplification in *Solanum lycopersicon* Mill. *International Journal of Tropical Agriculture* 29: 139–143.
- Sanford, J.C., and R.E. Hanneman. 1982. A possible heterotic threshold in the potato and its implications for breeding. *Theoretical and Applied Genetics* 61: 151–159.
- Sanford, L.L., S.P. Kowalski, K.L. Deahl, and S.L. Sinden. 1997. Diploid and tetraploid *Solanum chacoense* genotypes that synthesize leptine glycoalkaloids and deter feeding by Colorado potato beetle. *American Potato Journal* 74: 15–21.
- Spooner, D.M., J. Núñez, G. Trujillo, M. del Rosario Herrera, F. Guzmán, and M. Ghislain. 2007. Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. *Proceedings of the National Academy of Sciences in the United States of America* 104: 19398–19403.
- Spooner, D.M., S.H. Jansky, and R. Simon. 2009. Tests of taxonomic and biogeographic predictivity: resistance to disease and insect pests in wild relatives of cultivated potato. *Crop Science* 49: 1367–1376.
- Tai, G.C.C., and H. De Jong. 1997. A comparison of performance of tetraploid progenies produced by diploid and their vegetatively doubled (tetraploid) counterpart parents. *Theoretical and Applied Genetics* 94: 303–308.
- Tam, T.R., G.C.C. Tai, H. DeJong, A.M. Murphy, and J.E.A. Seabrook. 1992. Breeding potatoes for long-day, temperate climates. *Plant Breeding Reviews* 9: 217–332.
- van Treuren, R., A. Magda, R. Hoekstra, and T.J.L. van Hintum. 2004. Genetic and economic aspects of marker-assisted reduction of redundancy from a wild potato germplasm collection. *Genetic Resources and Crop Evolution* 51: 277–290.
- Veilleux, R.E., and F.I. Lauer. 1981. Breeding behavior of yield components and hollow heart in tetraploid-diploid vs conventionally derived potato hybrids. *Euphytica* 30: 547–561.
- Zaki, H.E.M., K.G. Haynes, and B.T. Vinyard. 2016. In vitro screening of *Solanum chacoense* for salinity tolerance. *American Journal of Potato Research* 93: 149.