

Bundesanstalt für Landwirtschaft und
Ernährung (BLE)
Referat 314
Deichmanns Aue 29
53179 Bonn

22. November 2018

ABSCHLUSSBERICHT

Vertragsnehmer:
ecogenics GmbH
c/o Microsynth AG
Schützenstrasse 15
CH-9436 Balgach

Förderkennzeichen:
FKZ 2816BE004

Projektnummer:
123-02.05-20.0354/16-II-K

Thema:
Zweite molekulargenetische Bestimmung der Kirscharten der Deutschen Genbank Obst

Laufzeit:
14. Februar 2017 – 13. Februar 2019

Berichtszeitraum:
14. Februar 2017 – 03. September 2018

1. Übersicht zu den durchgeführten Tätigkeiten

Um die Sammlung und Erhaltung von genetischen Ressourcen bei Obst in wissenschaftlicher, langfristig abgesicherter und nachhaltiger Weise zu gewährleisten, hat sich die Deutsche Genbank Obst (DGO) zum Ziel gesetzt, das Sortenspektrum der Kirsche pomologisch und genetisch zu bestimmen. Das Sortenspektrum umfasst die deutschen Sorten einschließlich Neuzüchtungen, Sorten mit soziokulturellem, lokalem oder historischen Bezug zu Deutschland und Sorten, die Spender für wichtige obstbauliche Merkmale sind. Der Echtheitsüberprüfung dieser Kirschsornten wird höchste Priorität beigemessen.

Ecogenics wurde beauftragt, die molekulare Überprüfung von 617 Süßkirsch- und 170 Sauerkirschkäumen der DGO durchzuführen. Die molekulare Überprüfung erfolgte mit der in Clarke and Tobutt (2009) beschriebenen Methode, welche nach den Richtlinien des "Fruit Network of the European Cooperative Programme for Plant Genetic Resources (ECPGR)" aufgesetzt wurde. Die Nutzung von Mikrosatelliten als genetische Marker ist eine der ältesten und gleichzeitig immer noch zuverlässigsten Methoden für die Genotypisierung von Individuen (Ellegren 2004). Die Mikrosatelliten-basierte Genotypisierung zur Sortenidentifikation von Kulturpflanzen ist weit verbreitet und hat sich als Methode bewährt (McCouch et al. 1997). Es hat sich gezeigt, dass die PCR-Amplifikation von polymorphen Loci und die anschließende Fragmentlängenanalyse auf einem Kapillarelektrophoresegerät auch zwischen verschiedenen Labors und unter Verwendung von unterschiedlichen Technologien zu reproduzierbaren Resultaten führt (Frey et al. 2007). Bedingung für diese Reproduzierbarkeit ist die gleichzeitige Analyse bekannter Standardsorten, welche erlauben, die Daten aus den verschiedenen Labors zu harmonisieren. Entsprechende Verfahren werden angewendet und wurden beschrieben (Frey et al. 2007) und für die Analyse der Kirschen angewandt.

Der Start des Projekts bestand darin, die anzuwendende Methodik nach Clarke and Tobutt (2009) bei Ecogenics zu etablieren sowie das Analyseverfahren anhand von Referenzgenotypen zu überprüfen. In einem zweiten Schritt erfolgte die Organisation und Durchführung der Sammlung der Blattproben von insgesamt 787 Kirschkäumen direkt auf Trockeneis und die Lagerung dieser Proben bei $\leq -80^{\circ}\text{C}$. Anschließend wurden die gesammelten Proben nach der Extraktion der DNA mit den etablierten Multiplex-Polymerasen-Ketten-Reaktion-(PCR)-Assays genotypisiert. Die Analysen erfolgten ohne Kenntnis der bereits pomologisch bestimmten Sorten.

In Absprache mit dem Auftraggeber kam es zu einer Erweiterung der Referenzgenotypen, da East Malling Research die acht Referenzen nicht fristgerecht liefern konnte. Um die Prüfung des Analyseverfahrens rechtzeitig abschließen zu können, schickte uns Frau Dr. Höfer vom Institut für Züchtungsforschung an Obst am Julius Kühn Institut (JKI) schon vorab sechs Referenzproben. Die Proben, welche vom JKI für die Etablierung der Methode zur Verfügung gestellt wurden, sind: 2 x "Grosse Prinzessin" (Synonym "Napoleon"), *Prunus avium* F12/1, *Prunus incisa* E621, *Prunus mahaleb* SL 64 und *Prunus nipponica* F1292. Die weiteren Referenzen "Noire de Meched", "Goodnestone Black" und 2 x "Noble" wurden zusätzlich zu den 787 angekündigten Proben gesammelt. Die East Malling-Proben trafen am 29.09.2017 bei uns ein. Somit haben wir die Marker mit den JKI-Proben validiert und mit den etwas verspäteten East Malling-Proben zusätzlich verglichen.

Die East Malling Research Referenzen enthielten folgende Sorten: "Napoleon" (Synonym "Grosse Prinzessin"), *Prunus avium* F12/1, *Prunus incisa* E621, "Noire de Meched", "Goodnestone Black", "Noble", *Prunus mahaleb* SL 64 und *Prunus nipponica* F1292. Eine Übersicht der Referenzen ist in der Tabelle 1 ersichtlich. Das verspätete Eintreffen der East Malling Research Proben hatte jedoch keinen Einfluss auf die Resultate. Alle 18 Referenzproben wurden zweifach wiederholt und somit sind 36 Datensätze im Allelreport vom 13.06.2018 vorhanden.

Tabelle 1. Tabellarische Darstellung der 18 Referenzen, deren Herkunft und Probennummern.

<i>Referenz</i>	<i>Vom Standort JKI als Ersatz für East Malling Referenzen geliefert Probennummern</i>	<i>Am Standort JKI gesammelte Referenzen Probennummern</i>	<i>Probennummern der 8 East Malling-Referenzen</i>
"Grosse Prinzessin" (Synonym "Napoleon")	122021_KIA0312_JKI_A_A04 122020_KIA0312_JKI_B_D04	-	A.Nap_B01
<i>Prunus avium</i> F12/1	122019_F12_1_JKI_C04	-	A.F12_1_C01
<i>Prunus incisa</i> E621	122017_E621_JKI_E04	-	I.E621_A01
<i>Prunus mahaleb</i> SL 64	122018_SL64_JKI_E01	-	M.SL64_H01
<i>Prunus nipponica</i> F1292	122022_F1292_JKI_B04	-	N.F1292_D01
"Noire de Meched"	-	110830	A.NdM_E01
"Noble"	-	110802, 110803	A.Nob_F01
"Goodnestone Black"	-	110787	AGB_D01

Am 24.05.2017 erhielten wir den Bescheid von Frau Dr. Höfer, dass jeder der sieben sammlungshaltenden Partner (Bundessortenamt, Prüfstelle Wurzen (BSA); Gemeinde Hagen a.T.W. (Hagen a.T.W.); Julius-Kühn Institut (JKI); Landratsamt Kyffhäuserkreis (Kyffhäuserkreis); Landesbetrieb Landwirtschaft Hessen (LLH); Landesanstalt für Landwirtschaft und Gartenbau Sachsen-Anhalt (LLFG); Stadt Witzenhausen (Witzenhausen)) die Liste mit den zu untersuchenden Bäumen erhalten hat und dass wir mit der Kontaktaufnahme beginnen können. Die Listen der Partner enthielten Informationen zum Standort des Baumes, Sortenname in der Sammlung, Akzessionsnummer und Probennummer der molekularen Sortenprüfung. Die Probenahme vor Ort fand in dem Zeitraum von Juli bis September 2017 und Mai 2018 statt. Um unsererseits eine genaue Rückverfolgbarkeit der Proben zu gewährleisten, arbeiteten wir mit Barcode- und Schlaufenetiketten. Dabei erhielt jedes Probencouvert eine Barcode-Etikette, eine weitere identische Etikette wurde auf die Liste zu der jeweiligen Probennummer geklebt und der jeweilige Baum mit einem Schlaufenetikett und einer identischen Barcodenummer versehen. Die Barcodenummer war mit der Probennummer identisch.

Gemäß Vertrag sollten wir bei der Probenentnahme eine Liste mit Daten zum Standort des Baumes sowie mit Akzessions- und Probennummern erhalten. Die Probengefäße sollten nur mit der jeweiligen Probennummer gekennzeichnet werden. Da die sammlungshaltenden Partner teilweise nicht genügend zeitliche Kapazität hatten, uns durch ihre Sammlung zu begleiten und uns entsprechende Informationen zu den einzelnen Proben zu geben, haben wir auch diese Aufgabe übernommen.

Es zeichnete sich schon früh ab, dass ein Teil der angemeldeten Proben aufgrund von Rodung oder anderen Gründen nicht gesammelt werden konnte. Aus diesem Grund wurden uns vom JKI bei der Probennahmen im Herbst 2017 sowie im Mai 2018 weitere, bisher nicht gemeldete Akzessionen zur Untersuchung anvertraut. Somit wurden, wie im Vertrag festgelegt, insgesamt 787 unbekannte Proben untersucht. Eine Übersicht der gesammelten Probenanzahl ist in der Tabelle 2 dargestellt.

Tabelle 2. Anzahl der im Herbst 2017 sowie im Mai 2018 gesammelten Proben je Standort sowie Summe aller Proben. Zusätzlich wurden 4 Referenzen vor Ort gesammelt.

<i>Standort</i>	<i>Anzahl angemeldete Süßkirschen</i>	<i>Anzahl angemeldete Sauerkirschen</i>	<i>Total angemeldete Proben</i>	<i>Anzahl Probenaussfall</i>	<i>Zusätzliche Proben während Probensammlung im Herbst 2017</i>	<i>Ersatzproben für Sauerkirschen LLH durch JKI ersetzt</i>	<i>Ersatzproben gesammelt im Mai 2018</i>	<i>Tatsächlich vorhandene Proben</i>
BSA	0	33	33	1	-	-	-	32
Hagen a.T.W.	176	2	178	3	1	-	-	176
JKI	87 und 4 Referenzen	53	144	8	-	20	39	191 und 4 Referenzen
Kyffhäuser	188	26	214	4	-	-	-	210
LLFG	42	31	73	3	3	-	-	73
LLH	104	20	124	40	-	-	-	84
Witzhausen	19	2	21	-	-	-	-	21
Total	620	167	787	59	4	20	39	787 und 4 Referenzen

Um einen großen Teil der Proben zu sammeln, wurden zwei Studenten der Universität Göttingen angestellt. Sie erhielten für diese Aufgabe eine Schulung, um einen konstanten Ablauf beim Sammeln zu gewährleisten. Weitere, am Sammeln beteiligte Personen waren zwei Mitarbeiter der Microsynth AG und Frau Dr. Braun-Lüllemann, letztere als Ansprechpartnerin für den Standort Witzhausen. Frau Dr. Braun-Lüllemann wurde in Absprache mit Frau Dr. Höfer von uns beauftragt, die Proben von 21 Kirschbäumen am Standort Witzhausen zu sammeln, da diese Bäume ohne Kennzeichnung über mehrere Kilometer verteilt standen.

Um eine möglichst optimale Blattqualität zu erhalten, mussten die Blätter bei trockenem Wetter gesammelt werden. Regnerische Augusttage verhinderten das termingerechte Sammeln und verzögerten den Prozess bis Mitte September.

Abgestorbene oder nicht vorhandene Bäume wurden auf der Liste vermerkt und im Nachhinein elektronisch erfasst. Als Probenmaterial wurden jeweils zwei möglichst junge Blätter aus der Baumkrone gesammelt und in ein Probencouvert gegeben. Während des Sammelns wurden die Probencouverts in einem isolierten, mit Trockeneis gekühlten, Rucksack aufbewahrt. Am Ende des Tages wurden alle Proben bei -80°C gelagert.

Bei der Genotypisierung ist die Qualität der aus den Blättern isolierten DNA von großer Bedeutung. Darüber hinaus darf die PCR nicht inhibiert werden und es muss genügend DNA für alle PCRs vorliegen. Zudem muss die DNA-Isolation so erfolgen, dass Probenvertauschungen und Kreuzkontaminationen ausgeschlossen werden können. Um letztere Ziele zu erreichen, wurden

sämtliche Proben mit einer Probennummer und mit einem Barcode versehen und alle Arbeitsschritte in 96-Well-Platten mit entsprechenden Kontrollen durchgeführt.

Bei der ersten Genotypisierung aller Proben und der darauf folgenden Gengruppenzuordnungs-Analyse, wurde festgestellt, dass es zu viele Allel-Ausfälle gegeben hat und somit kein konstanter Datensatz generiert werden konnte. Aus diesem Grund wurde bei der zweiten Genotypisierung nach der Isolation ein zusätzlicher Aufreinigungsschritt gewählt. Als geeignetste Methode zur Isolation genomischer DNA aus Kirschen-Blättern hat sich zum Schluss die "Hot Shot"-Methode und die zusätzliche Aufreinigung über das OneStep PCR Inhibitor Removal Kit (Zymogen) im 96 Well Platten Format erwiesen. Dabei wurde aus dem gesammelten Material eine Blattstanze herausgetrennt, lysiert und nach dem Hersteller-Protokoll auf-gereinigt. In einem ersten Schritt in der "Hot Shot"-DNA-Extraktionsmethode wird eine Blattstanze mit dem Puffer A (0.1 M NaOH, 2% Tween20 in destilliertem Wasser) 15 Minuten lang auf 95°C erhitzt. Der Puffer B (0.1 M HCl, 0.1 M Tris-HCl pH8.0, 0.05 M EDTA in destilliertem Wasser) neutralisiert die so lysierte Blattstanze. Durch die Aufreinigung der neutralisierten Blattstanzenlyse mit dem OneStep PCR Inhibitor Removal Kit war die extrahierte DNA frei von PCR-Inhibitoren. Im Anschluss an die Isolation wurde ein Aliquot des Isolates für die nachfolgende PCR verwendet und die unverdünnten Isolate wurden bei -20°C eingelagert. Jede 96-Well-Platte beinhaltet eine Extraktionskontrolle (Puffer A, Puffer B und kein Blattmaterial) und eine PCR-Kontrolle (nur Multiplex PCR Kit von Qiagen und SSR Marker Primer, keine Templat DNA) sowie drei Multiplex-spezifische Allel-Leitern und eine Positivkontrolle (Blattprobe, mit jeder 96-Wellplatte frisch isoliert).

Die festgelegten Mikrosatellitenmarker wurden hinsichtlich ihrer Robustheit anhand von Referenzproben evaluiert und aufgrund der vorhergehenden Evaluierung so kombiniert, dass sie in drei Multiplex-PCR-Ansätzen analysiert werden können. Die Marker und die Fluorophore wurden so verteilt, dass eine möglichst geringe gegenseitige Signalstörung zu erwarten ist. Folgende Marker-Kombinationen wurden deshalb gewählt (Tabelle 3):

Tabelle 3. Drei Multiplex PCR Ansätze für 16 SSR Marker

<i>PCR Ansatz</i>	<i>Marker_Fluorophor</i>	<i>Kommentar</i>
C1	BPPCT037_FAM	
	CPPCT022_FAM	
	EMPA002_At532	
	EMPaS06_At532	
	EMPA003_At550	
	EMPaS14_At556	
C2	CPPCT006_FAM	
	UDP98-412_At532	
	EMPaS01_At532	PCR separat, multiload mit Ansatz C2
	PS05C03_At550	PCR separat, multiload mit Ansatz C2
	EMPaS02_At565	
C3	EMPaS12_FAM	
	EMPA017_FAM	
	EMPA026_At532	
	PceGA34_At550	
	EMPaS10_At565	

Die Marker EMPaS01 und PS05C03 aus dem Multiplex-PCR-Ansatz C2 haben sich bei der weiteren Analyse als zu instabil herausgestellt. Aus diesem Grund wurden diese beiden Marker bei einer tieferen Temperatur amplifiziert und dann für die ABI3730XL-Analyse wieder zu den übrigen C2 Markern hinzugefügt. Die Primer-Sequenzen sind im Appendix in der Tabelle 4 aufgeführt. Die zu verwendenden Mikrosatelliten-Marker-Systeme wurden so etabliert und angepasst, dass mit den verwendeten DNA-Isolaten robuste, reproduzierbare Daten erzeugt werden können.

Der Forward-PCR-Primer von jedem Mikrosatelliten-Marker ist am 5'-Ende mit einem Fluorophor modifiziert, um die Fragment-Detektion auf dem Kapillarelektrophoresegerät zu ermöglichen. Im blauen Detektionskanal wurden Oligonukleotide, welche mit 6-FAM (Fluorescein) modifiziert sind, verwendet, für die Kanäle grün, gelb und rot wurden die Rhodamin-Farbstoffe ATTO532, ATTO550 und ATTO565 verwendet. Am 5' Ende der Reverse-PCR-Primer-Sequenzen wurde eine Pigtail-Sequenz GTTTCTT angebracht. Diese Sequenz sorgt für eine robustere Amplifikation während der PCR (Brownstein et al. 1996). Alle Primer wurden bei der Microsynth AG, Schweiz, synthetisiert. Für die PCR wurde das Multiplex-PCR-Kit (Qiagen, Deutschland) verwendet und folgendes Thermocycler-Programm durchgeführt: 95°C, 10 min Denaturierung, gefolgt von 40 Zyklen mit 94 °C, 0.5 min; 55 °C bzw. 48°C für die beiden instabilen Marker aus der Multiplex-Reaktion C2, 1.5 min; 72 °C, 1 min und einer finalen Elongation bei 72 °C für 30 min.

Die Fragmentlängenanalysen wurden mit dem Längenstandard GeneScan LIZ500 (Applied Biosystems) auf einem Applied Biosystems 3730XL DNA-Analyser durchgeführt. Folgende Fragmentlängenanalysebedingungen wurden eingehalten: Injektionszeit 10 s; Injektionsspannung 1.6 kV; Laufzeit 2100 s; Laufspannung 15 kV; Kapillarlänge 50 cm; Polymer POP7; Filter Dye Set G5. Um die Laufqualität des Applied Biosystems 3730XL DNA-Analyzers für jeden einzelnen Lauf zu

überprüfen, wurde zu Projektbeginn mit den auserwählten Referenzgenotypen für jeden Multiplex-PCR-Assay eine projektspezifische Allel-Leiter erstellt. Mit dieser Allel-Leiter können Laufverschiebungen sofort erkannt und behoben werden.

Grundsätzlich erfolgten sämtliche Schritte im Rahmen der DNA-Isolation, PCR-Amplifikation und Kapillarelektrophorese in 96-Wellplatten. Die drei Multiplex-PCR-Assays wurden mit den unverdünnten Isolaten durchgeführt und zusammen mit einem internen Größenstandard auf einem Kapillarelektrophoresegerät analysiert.

Um die Vergleichbarkeit der Daten aller Proben innerhalb der Studie zu gewährleisten, wurden für jeden Marker Regeln für die Interpretation aufgestellt. Jeder Marker erzeugt ein spezifisches Muster. Zur Illustration wurde für jeden verwendeten Marker ein repräsentatives Chromatogramm abgebildet (Abbildungen 1 – 16).

Die Interpretation der Chromatogramme erfolgte mit kommerziell erhältlicher Software Software GeneMarker V2.6.4 und wurde bei jeder Probe auf Korrektheit überprüft. Um die Zuordnung der Allele robust zu gewährleisten, wurden für jeden Mikrosatelliten-Marker in der Analysesoftware sogenannte Bins definiert. Die Datenauswertung mit dem GeneMarker-Programm erfolgte nur, wenn die Kontrollen die Qualitätskontrolle bestanden haben und wenn die Fragmente des verwendeten Größenstandards einwandfrei zugeordnet werden konnten. Die so erzeugten Alleldaten waren untereinander abzugleichen und einer Gengruppe zuzuordnen.

Die Gengruppen werden mittels Clusteranalysen, basierend auf paarweisen Unterschieden zwischen den Proben, zugeordnet.

Um die paarweisen Unterschiede zu ermitteln, wird in einem ersten Schritt die paarweise Übereinstimmung ermittelt (siehe Formel unten). Aufgrund unserer Erfahrung können sehr kleine Längenunterschiede (± 1 bp) nicht zuverlässig zugeordnet werden. Daher wurde eine Unschärfe von ± 1 bp bei der Berechnung der paarweisen Übereinstimmung zugelassen.

$$\text{paarweise Übereinstimmungen} = \frac{2 * \sum \text{Gemeinsame Allele}}{\sum \text{Allele Probe 1} + \sum \text{Allele Probe 2}}$$

Die paarweisen Distanzen ergeben sich, indem man die paarweisen Übereinstimmungen von 1 subtrahiert (siehe unten).

$$\text{paarweise Distanzen} = 1 - \text{paarweise Übereinstimmungen}$$

Als Cut-off für die Gengruppen wurde 0.2 bzw. 20% Unterschiede in den Alleldaten definiert. Dieser Cut-Off-Wert wurde anhand der Verteilung der paarweisen Distanzen im vorhandenen Datensatz ermittelt. Damit werden alle Proben, welche zu 80-100% identisch sind, derselben Gengruppe zugeordnet. Die Clusteranalyse und Definition der Gengruppe wurde mittels der Statistiksoftware R und eines intern entwickelten Python-Scripts berechnet.

2. Darstellung der Projektergebnisse und darüber hinaus gewonnene Erkenntnisse

Jeder verwendete Mikrosatellitenmarker zeigt individuelle Eigenschaften, beispielsweise in den vorkommenden Allelgrößen, Stotterbanden, in der Signalintensität und im Rauschen. Entsprechend sind in einigen Fällen die Interpretation der Chromatogramme und die Allel-Zuordnung nicht eindeutig möglich. Im Folgenden werden sämtliche Marker kurz illustriert. Es wird jeweils

exemplarisch das Chromatogramm der Probe „Napoleon“ im für den jeweiligen Marker relevanten Größenbereich gezeigt (x-Achse, Fragmentlänge in bp; y-Achse, Fluoreszenzintensität in willkürlichen Einheiten). Die Boxen unterhalb der x-Achse zeigen identifizierte Allele an. Die auf den x-Achsen farbig hervorgehobenen Bereiche definieren die Bins.

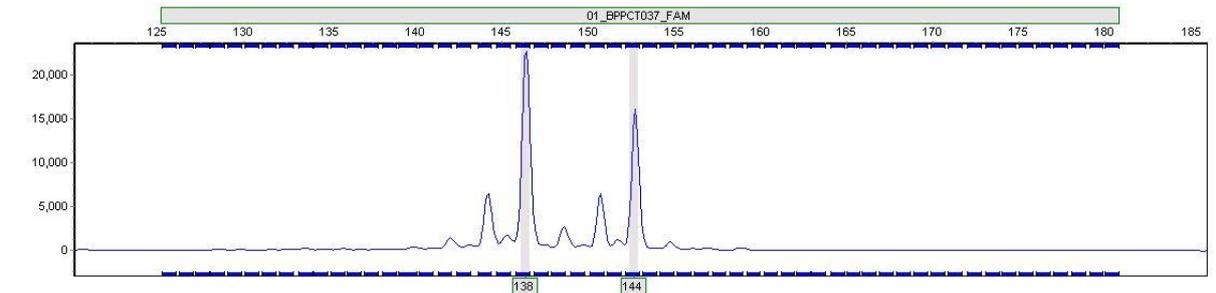


Abbildung 1. Chromatogramm des SSR Markers BPPCT037 der Probe „Napoleon“ mit den Allelen 138/144.

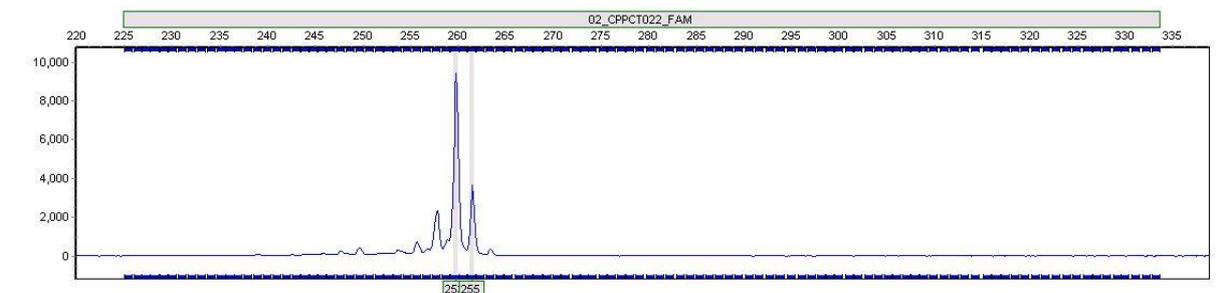


Abbildung 2. Chromatogramm des SSR Markers CPPCT022 der Probe „Napoleon“ mit den Allelen 253/255.

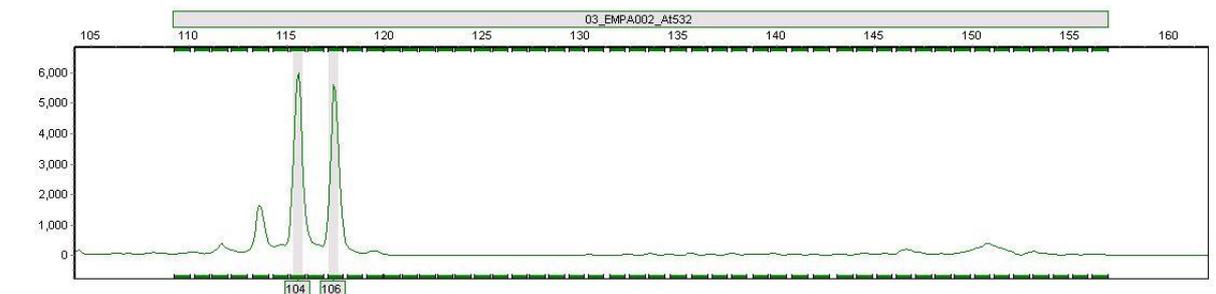


Abbildung 3. Chromatogramm des SSR Markers EMPA002 der Probe „Napoleon“ mit den Allelen 104/106.

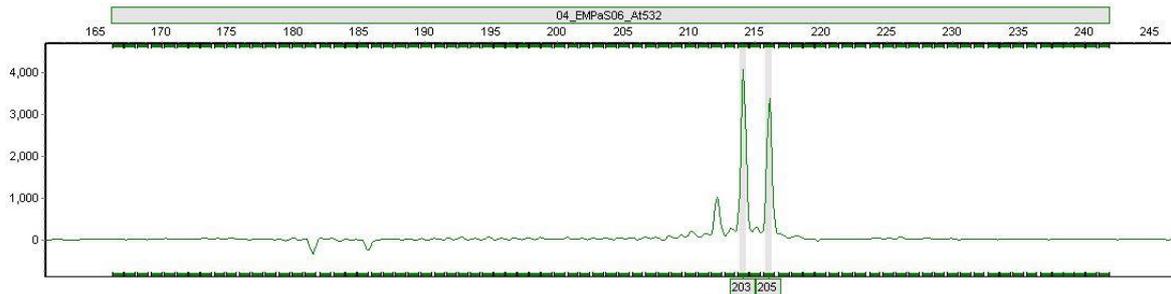


Abbildung 4. Chromatogramm des SSR Markers EMPaS06 der Probe „Napoleon“ mit den Allelen 203/205.

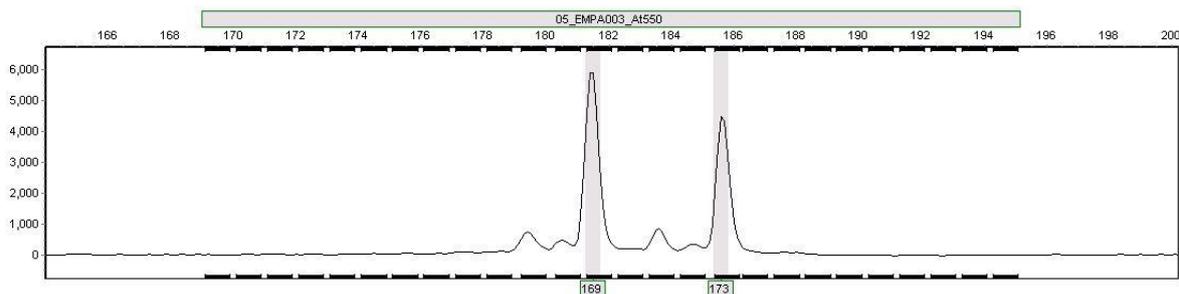


Abbildung 5. Chromatogramm des SSR Markers EMPA003 der Probe „Napoleon“ mit den Allelen 169/173.

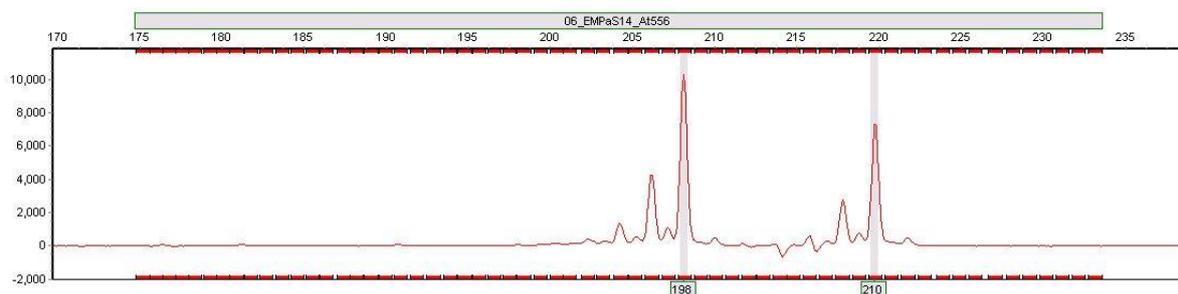


Abbildung 6. Chromatogramm des SSR Markers EMPaS14 der Probe „Napoleon“ mit den Allelen 198/210.

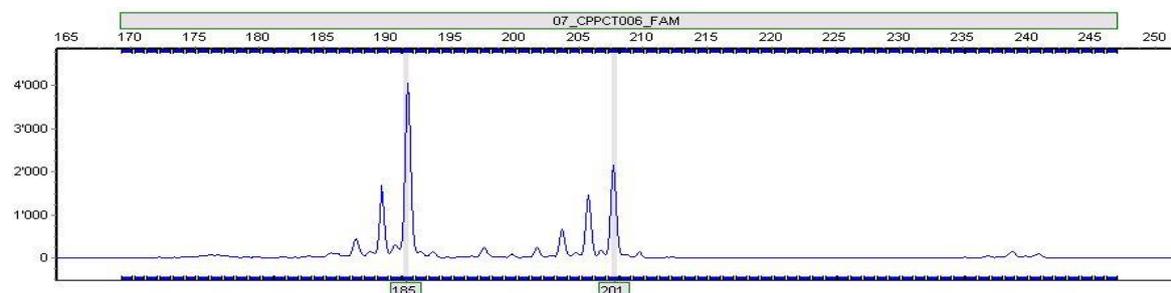


Abbildung 7. Chromatogramm des SSR Markers CPPCT006 der Probe „Napoleon“ mit den Allelen 185/201.

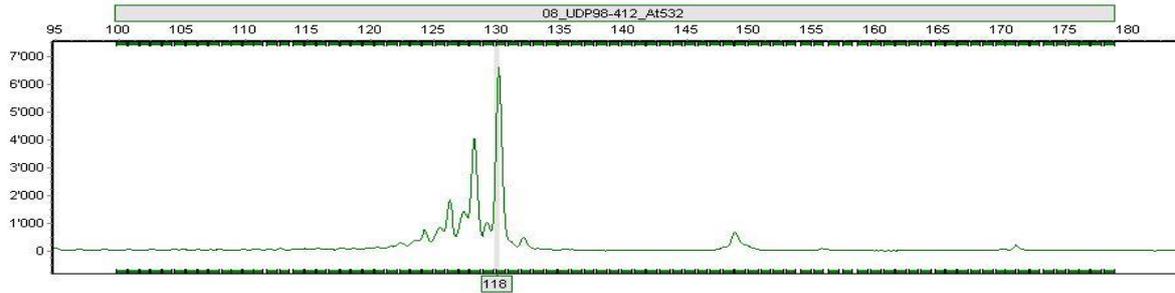


Abbildung 8. Chromatogramm des SSR Markers UDP98-412 der Probe „Napoleon“ mit den Allelen 118/118.

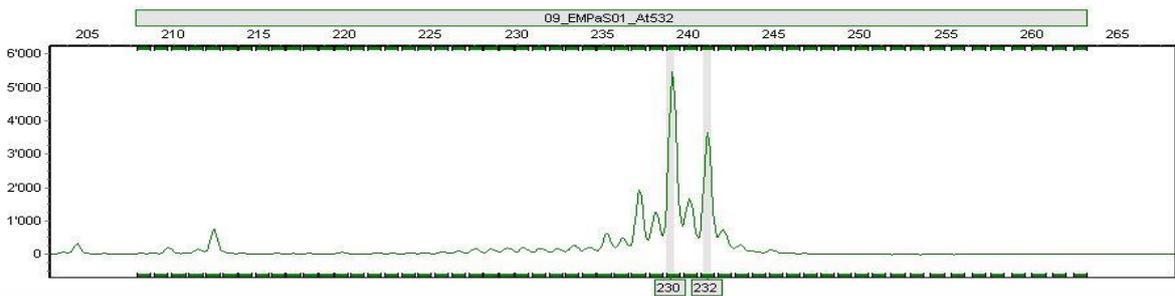


Abbildung 9. Chromatogramm des SSR Markers EMPaS01 der Probe „Napoleon“ mit den Allelen 230/232.

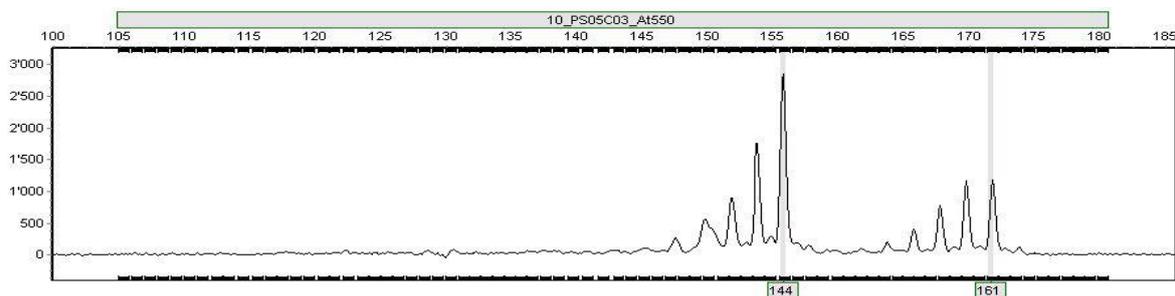


Abbildung 10. Chromatogramm des SSR Markers PS05C03 der Probe „Napoleon“ mit den Allelen 144/161.

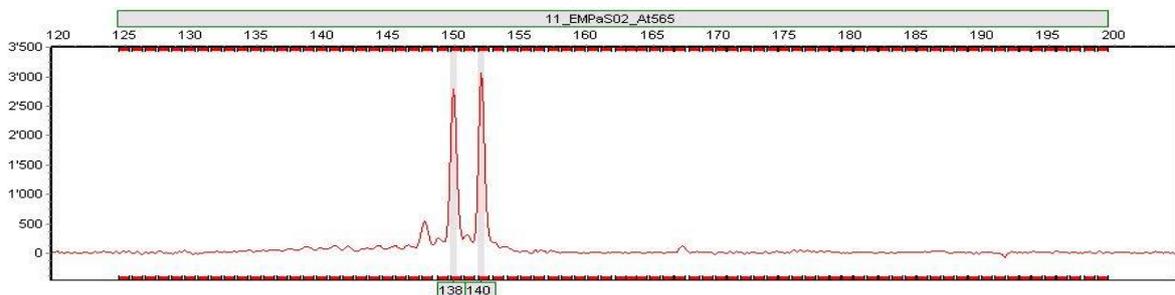


Abbildung 11. Chromatogramm des SSR Markers EMPaS02 der Probe „Napoleon“ mit den Allelen 138/140.

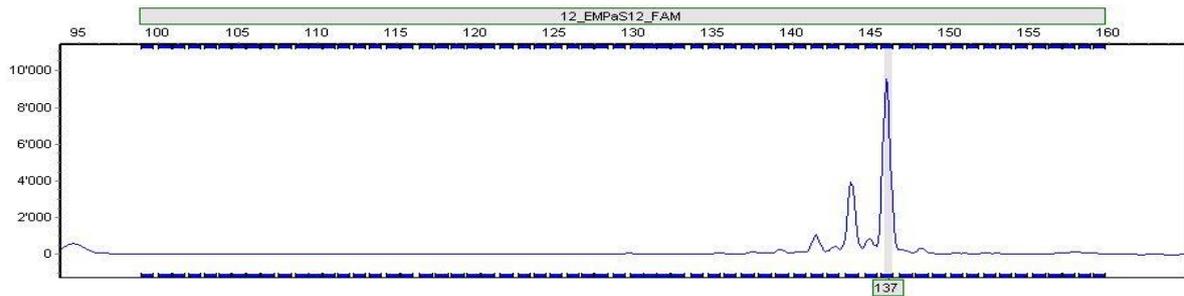


Abbildung 12. Chromatogramm des SSR Markers EMPaS12 der Probe „Napoleon“ mit den Allelen 137/137.

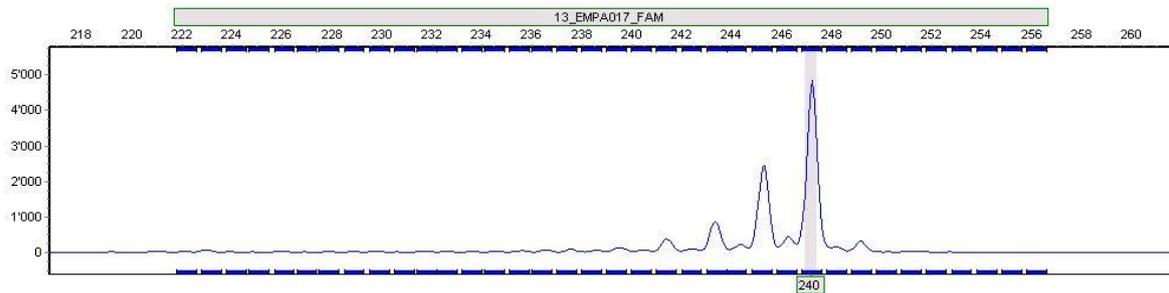


Abbildung 13. Chromatogramm des SSR Markers EMPA017 der Probe „Napoleon“ mit den Allelen 240/240.

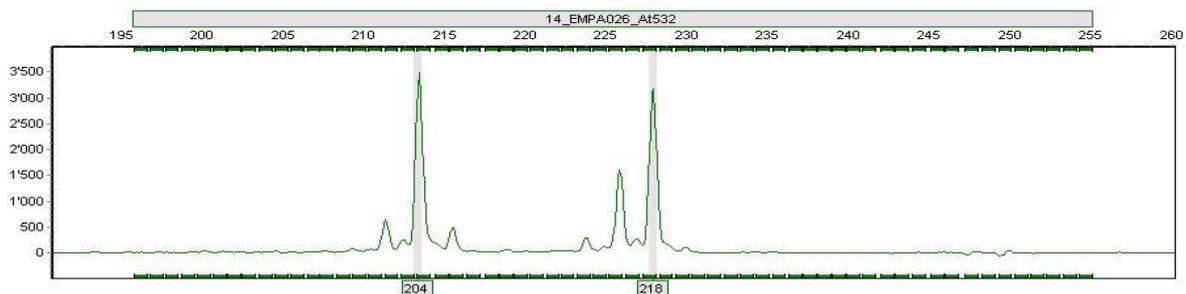


Abbildung 14. Chromatogramm des SSR Markers EMPA026 der Probe „Napoleon“ mit den Allelen 204/218.

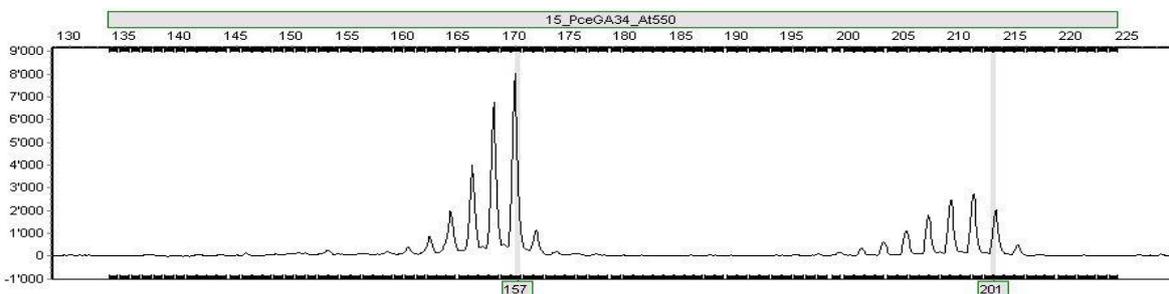


Abbildung 15. Chromatogramm des SSR Markers PceGA34 der Probe „Napoleon“ mit den Allelen 157/201.

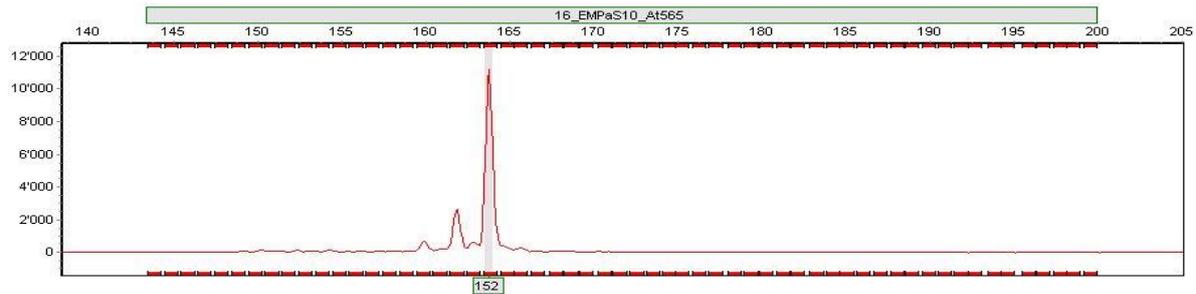


Abbildung 16. Chromatogramm des SSR Markers EMPaS10 der Probe „Napoleon“ mit den Allelen 152/152.

Die Mikrosatellitenbasierte Genotypisierung birgt aus verschiedenen Gründen das Risiko von Null-Allelen und Allelausfällen. Aufgrund der vorliegenden Daten kann nicht zwischen echten Null-Allelen (z.B. Mutation innerhalb der Primerbindungsstelle) und Ausfällen einzelner Marker aufgrund von ungeeigneten experimentellen Bedingungen (z.B. PCR Inhibition) unterschieden werden. Im Falle von gehäuften Ausfällen (d.h. kein Allel konnte zugeordnet werden) eines einzelnen Samples in mehreren Multiplex PCRs kann die Isolation oder die Qualität des Ausgangsmaterials der Grund für die Ausfälle sein. Bei mehreren Ausfällen innerhalb nur eines PCR-Systems kann ein PCR-Artefakt nicht ausgeschlossen werden. Bei der Generierung des ersten Allelreports wurden diese oben beschriebenen Allelausfälle außerordentlich oft beobachtet. Der Grund für die Allelausfälle konnte der offenbar ungeeigneten DNA-Extraktionsmethode und den dadurch im Extrakt verbliebenen PCR-Inhibitoren zugeschrieben werden. Alle Proben, welche mit der oben genannten "Hot Shot"-Extraktionsmethode bearbeitet wurden, wurden über OneStep PCR Inhibitor Removal-Säulen zusätzlich aufgereinigt. Die aufgereinigte DNA wurde nochmals analysiert und die Daten sind deutlich robuster interpretierbar. Eine noch höhere Datensicherheit und Qualität könnte durch Replikate erreicht werden.

Schlussendlich konnten insgesamt 787 Proben einfach wiederholt analysiert werden. Zusätzlich wurden aus den 787 Proben 6 Blindproben (3 Süßkirschen: Proben 110782, 110792, 110800 und 3 Sauerkirschen: Proben 111256, 111258, 111264) gewählt. Diese wurden 4-fach wiederholt analysiert, was wiederum in 24 Datensatzreplikaten resultierte. 36 Datensätze sind Wiederholungsproben der 18 Referenzen (Tabelle 1), wobei jede Referenz 2-fach wiederholt wurde. Die Ecogenics intern verwendete Positivkontrolle ist zehn Mal im Datensatz aufgeführt. Die Probe 110882 wurde wiederholt und ist aus diesem Grund zweimal im Datensatz vorhanden. Eine Verschiebung bei der Anzahl der ursprünglich angemeldeten 617 Süßkirsch- und 170 Sauerkirschproben hat sich aufgrund der Ergebnisse ergeben. Anhand des diploiden Genotyps wurden 607 Proben als Süßkirschen und anhand des tetraploiden Genotyps 179 Proben als Sauerkirschen identifiziert (Allelreport vom 13.06.2018, Spalte L). Die Analyse mit den kirschenspezifischen Mikrosatelliten-Markern ergab bei Probe 110899_D11 kein Resultat. Da die Mikrosatelliten-Marker kirschenspezifisch sind und bei der DNA-Analyse von anderen Obstsorten nicht funktionieren, wurde beim betroffenen Standortbetreuer nachgefragt, von welcher Obstart diese Blattprobe stammte. Als Herkunft wurde ein Apfelbaum identifiziert.

Die Verteilung der paarweisen Distanzen für den gesamten Datensatz ist in Abbildung 17 dargestellt. Gut ersichtlich sind die Unterscheidung von Eigenvergleichen und sehr ähnlichen Genotypen (erster Balken bei einer Distanz von 0.05) und die Vergleiche zwischen verschiedenen Genotypen (Distanzen > 0.2). Basierend auf dieser Beobachtung wurde der Distanz Cut-Off für die Gengruppendefinition bei 0.2 bzw. 20% definiert. Damit fallen alle Proben, welche mehr als 80% identische Allele aufweisen, in eine Gengruppe.

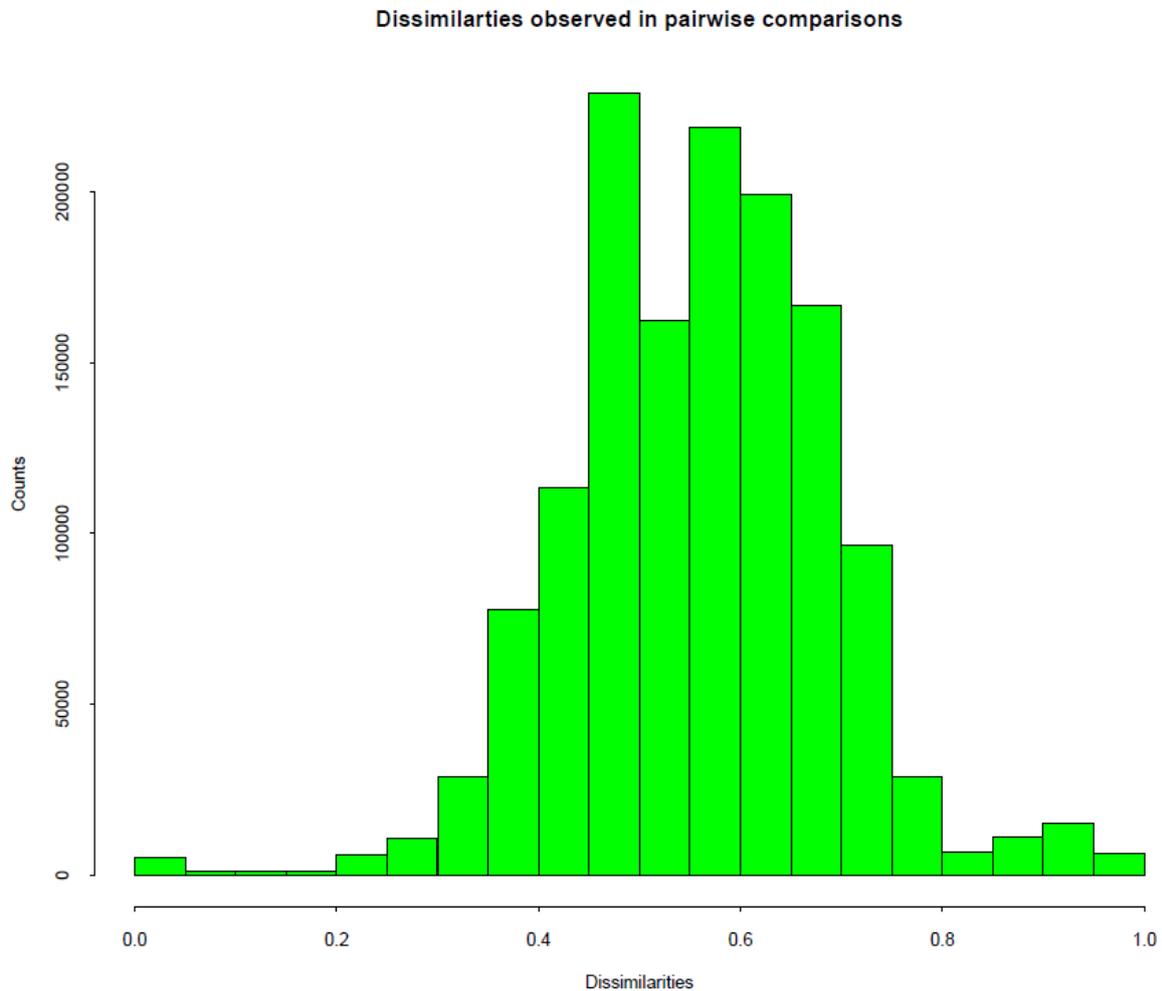


Abbildung 17. Verteilung der paarweisen Distanzen im Kirschen-Datensatz.

Frau Dr. Höfer hat uns gebeten, den Datensatz aus der ersten molekulargenetischen Bestimmung der Kirschsornten der Deutschen Genbank Obst, bestehend aus 367 Profilen (Allelreport vom 13.06.2018, Spalte K) mit bereits eindeutig zugeteilten Gengruppen Identifikationsnummern, in die zweite Gengruppenzuordnungsanalyse einfließen zu lassen. Die Identifikationsnummern Pc_ bzw. Pa_ im Allelreport vom 13.06.2018 (Spalte O) basieren somit auf den Identifikationsnummern des ersten Datensatzes.

Durch die zweite molekulargenetische Bestimmung der Kirschsornten wird die Datenqualität des gesamten Datensatzes verbessert und möglich falsche Unterschiede zwischen den Profilen aufgedeckt. Aus diesem Grund kann es sein, dass einzelne Profile mit unterschiedlichen Identifikationsnummern nun in dieselbe Gengruppe fallen. In diesen Fällen besteht die Identifikationsnummer aus mehreren zusammengeschlossenen Nummern.

Zusammen mit den Daten aus der ersten Sortenprüfung konnten 409 genetisch unterscheidbare Süßkirschen- und 67 Sauerkirschen-Gengruppen ermittelt werden, wobei auch Gengruppen welche in Untergruppen A und B (Allelreport vom 13.06.2018, Spalte N) unterteilt sind, mitgezählt wurden. Unsicherheiten bei der Gengruppenzuordnung wurden in der Spalte N als „zu überprüfen“ deklariert. Untergruppen wurden gebildet, sofern eindeutig 3 oder mehr Allele abweichen. Weiterhin wurden die Nummern der Gengruppen (Allelreport vom 13.06.2018, Spalte M) zufällig vergeben. Bei diesen Nummern handelt es sich um keine Aufzählung der Gengruppen.

Die Resultate der Süß- und Sauerkirschen sind im Allelreport vom 13.06.2018 in einem ersten Reiter gemeinsam und in zwei weiteren Reitern getrennt voneinander dargestellt.

Für den Aufbau einer Datenbank mit nur einem Profil pro Identifikationsnummer wurde in der Spalte P des ersten Reiters jeweils ein repräsentatives Profil ausgewählt (x). Diese ausgewählten Profile wurden wiederum kopiert und in drei weiteren Reitern für Süß- und Sauerkirschen gemeinsam und jeweils getrennt aufgelistet.

Die im Rahmen der ersten und zweiten molekulargenetischen Bestimmung von Kirschsornten ermittelten 476 Gengruppen stellen eine sehr gute Grundlage für zukünftige Sortenabgleiche dar. Basierend auf den vorhandenen Daten kann eine Datenbank mit Referenz-Genotypen für die verschiedenen Gengruppen / Sorten aufgesetzt werden. Außerdem kann der Wert der Daten noch deutlich erhöht werden, wenn weitere, ähnliche Studien von anderen Auftraggebern mit denselben Markern, oder allenfalls einer Auswahl davon, durchgeführt werden.

Nach unserer Analyse zeichnet sich ab, dass die Unterscheidung von Süß- und Sauerkirschen pomologisch nicht immer ganz eindeutig ist. Um die zukünftige Datenbankabgleiche zu vereinfachen, empfehlen wir eine gemeinsame Datenbank für Süß- und Sauerkirschen anzulegen und diese nicht separat zu behandeln.

3. Bestätigung der Einhaltung vorgegebener Referenzen und Datenstrukturen, sowie der erfolgreichen Übernahme der Ergebnisse, Rohdaten und DNA- Stammlösungen durch die Koordinationsstelle der Deutschen Genbank Obst

Die ursprünglich geplanten, vertraglich festgehaltenen Ziele wurden weitgehend erreicht. In Abweichung wurden die Kirschen mit einer besser geeigneten DNA-Extraktionsmethode extrahiert und neu analysiert. Dies geschah im Interesse der Robustheit und in Rücksprache mit dem Auftraggeber. Die im Vertrag festgelegte Probenanzahl von 787 Proben wurde eingehalten. Lediglich 18 statt 8 Referenzproben wurden getestet, wobei jede Referenz-Probe 2-fach wiederholt wurde. Die DNA wurde am 18.07.2018 an das JKI zu Händen von Frau Dr. Höfer geliefert.

4. Zusammenfassung

787 Blattproben, welche die biologische Vielfalt der Kirschen in Deutschland repräsentieren, wurden molekulargenetisch analysiert. Die DNA wurde mittels Säulenextraktion aus den Blattproben isoliert. Von jeder Probe wurden 16 bekannte Mikrosatelliten-Loci mit drei Multiplex-PCR amplifiziert und mittels Kapillarelektrophorese analysiert. Die so erzeugten Chromatogramme wurden ausgewertet und für jede Sorte ein entsprechendes molekulargenetisches Profil erstellt. Das molekulargenetische Profil wurde mittels paarweiser Distanzen und Clusteranalyse einer Gengruppe zugeordnet und mit den Daten der ersten molekulargenetischen Sortenprüfung von 367 Bäumen abgeglichen. 67 Sauerkirschengruppen und 409 Süßkirschengruppen konnten ermittelt werden. Die insgesamt 476 Gengruppen können als Referenzen für zukünftige Sortenprüfungen verwendet werden.

5. Literaturverzeichnis

- Brownstein M. J., Carpten J. D., Smith J. R. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1006, 1008–1010.
- Clarke J.B., Tobutt K.R. 2009. A Standard set of accessions, microsatellites and genotypes for harmonizing the fingerprinting of cherry collections for the ECPGR. *Acta Horticulturae* 814:615-618.
- Ellegren H, 2004. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5(6):435-45.
- Frey JE, Koller B, Frey B & Bünter M, 2007. Identifikation von Obstsorten: Validierung einer Analyseverfahren. *Agrarforschung* 14(11-12):536-541.
- McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T & Blair M, 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol* 35(1-2):89-99.

6. Appendix

Tabelle 4: Die genauen Mikrosatelliten Primer Sequenzen der 16 Mikrosatelliten. Am 5' Ende der Reverse Primer Sequenzen wurde eine Pigtail Sequenz GTTTCTT angebracht. Diese Sequenz sorgt für eine robustere Amplifikation während der PCR und gewährleistet somit eine sauberere Auswertung.

Markenname	F Primer Sequenz	<u>Pig Tailed R Primer Sequenz</u>
EMPA002	TGACAGGTCATCACCATTG	<u>GTTTCTT</u> CAGGATTAAGCATTGCAAATTA
EMPA003	AGCCATTCTGAAAAGGTGGA	<u>GTTTCTT</u> GCATTGAGCCAAACAAAATCA
EMPA017	ATTTCAATGTGGGGATGAGC	<u>GTTTCTT</u> TGAAGTGAGGGAAATGGAGC
PceGA34	GAACATGTGGTGTGCTGGTT	<u>GTTTCTT</u> TCCACTAGGAGGTGCAAATG
EMPaS12	TGTGCTAATGCCAAAATACC	<u>GTTTCTT</u> AACATGCATTTCAACCCACTC
EMPaS02	CTACTTCCATGATTGCCTCAC	<u>GTTTCTT</u> AACATCCAGAACATCAACACAC
EMPaS06	AAGCGGAAAGCACAGGTAG	<u>GTTTCTT</u> TTGCTAGCATAGAAAAGAATTGTAG
EMPaS10	GCTAATATCAAATCCCAGCTCTC	<u>GTTTCTT</u> TTGAAGAAGTATGGCTTCTGTGG
BPPCT037	CATGGAAGAGGATCAAGTGC	<u>GTTTCTT</u> CTTGAAGGTAGTGCCAAAGC
EMPaS14	TCCGCCATATCACAATCAAC	<u>GTTTCTT</u> TTCCACACAAAACCAATCC
EMPaS01	CAAAATCAACAAAATCTAAACC	<u>GTTTCTT</u> CAAGAATCTTCTAGCTCAAACC
UDP98-412	AGGGAAAGTTTCTGCTGCAC	<u>GTTTCTT</u> GCTGAAGACGACGATGATGA
CPPCT022	CAATTAGCTAGAGAGAATTATTG	<u>GTTTCTT</u> GACAAGAAGCAAGTAGTTTG
PS05C03	AGATCTCAAAGAAGCTGA	<u>GTTTCTT</u> AGCTTATGCATATACCTG
EMPA026	ATTGAAAAGCCAAAGAGCG	<u>GTTTCTT</u> TTTACGGTTTGAAGCAAGTG
CPPCT006	AATTAACTCCAACAGCTCCA	<u>GTTTCTT</u> ATGGTTGCTTAATTCAATGG