

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

01. Mai 2019

Abschlussbericht

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

Förderkennzeichen:
FKZ 2816BE008

Projektnummer:
123-02.05-20.0381/16-I-G

Thema:
Zweite molekulargenetische Bestimmung der Apfelsorten der Deutschen Genbank Obst

Laufzeit:
08. Mai 2017 – 07. Mai 2019

Berichtszeitraum:
08. Mai 2017 – 31. Oktober 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 des Apfels pomologisch und genetisch zu bestimmen. Das Sortenspektrum umfasst die deutschen Sorten einschließlich Neuzüchtungen, Sorten mit soziokulturellem, lokalem oder historischem Bezug zu Deutschland und Sorten, die Spender für wichtige obstbauliche Merkmale sind. Der Echtheitsüberprüfung des Apfels wird höchste Priorität beigemessen.

Seite | 2

Ecogenics wurde beauftragt, die molekulargenetische Überprüfung von 5.321 Apfelbäumen der DGO Sammlung durchzuführen. Die molekulare Überprüfung erfolgte nach den Richtlinien des "Fruit Network of the European Cooperative Programme for Plant Genetic Resources (ECPGR)" und mit Markern, welche auf der High-quality Disease Resistant Apples for a Sustainable Agriculture (HiDRAS)-Homepage beschrieben sind.

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 mikrosatellitenbasierte Genotypisierung zur Sortenidentifikation von Kulturpflanzen ist weit verbreitet und hat sich als Methode bewährt (z.B. 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 Äpfel angewandt.

Der Start des Projekts bestand darin, die Methodik, welche schon in der ersten molekulargenetischen Sortenbestimmung von pomologisch bestimmten Apfelsorten (Förderkennzeichen: 2809BE010 (BLE)) in den Jahren 2012 bis 2014 angewandt wurde, auf den Referenzgenotypen zu überprüfen.

In einem zweiten Schritt erfolgte die Organisation und Durchführung der Probensammlung auf Trockeneis von 5.321 Apfelproben und Probenlagerung bei $\leq -80^{\circ}\text{C}$. Anschließend wurden die gesammelten Proben mittels DNA-Isolation und Genotypisierung mit den etablierten Multiplex-Polymerase-Ketten-Reaktion-(PCR)-Assays analysiert. Für jede Probe wurde ein molekulargenetisches Profil erstellt. Die Profile dieser zweiten Sortenprüfung wurden mittels paarweiser Distanzen und Clusteranalyse einer Gengruppe zugeordnet und mit den Daten aus der ersten Sortenprüfung abgeglichen. Dabei erfolgten die Analysen ohne pomologische Informationen resp. Informationen zu den Sorten.

2. Beschreibung der erhobenen Daten

Insgesamt wurden 5.314 Proben zur Überprüfung angemeldet, wobei 5.315 Proben gesammelt wurden. Zudem wurden statt 8, 11 Referenzproben geliefert (Tabelle 1). Die Anzahl der gesammelten Proben unterschreitet die, gemäß Vertrag, garantierte Probenanzahl von 5.321 um 6 Proben, wobei nur 5.314 Proben angemeldet wurden und die 11 Referenzproben in dieser Berechnung nicht eingeschlossen sind. Mit den 11 Referenzproben sind insgesamt 5.325 Proben analysiert worden. Alle

11 Referenzproben wurden 4 Mal analysiert und aus den 5.314 Proben wurden zudem 6 Proben als Blindproben ausgewählt, welche auch 4 Mal analysiert wurden.

Die Proben, welche vom Institute de Recherche en Horticulture et Semences (INRA, Frankreich) für die Etablierung der Methode zur Verfügung gestellt wurden, sind: 'Delicious', 'Fiesta', 'Michelin', 'Prima', 'Worcester Parmain', 'Malus x floribunda 821', 'M. x robusta 5' und 'Malling 9' Die Proben trafen am 29.09.2017 bei uns ein. Die Überprüfung der Referenzen 'Michelin' und 'Malling 9' ergaben ein vertauschtes Genprofil. Zudem entsprach 'Worcester Parmain' nicht dem angegeben Referenzprofil. Diese drei Referenzen wurden von Frau Höfer am 16.04.2018 neu angefordert. Die Vertauschung von 'Michelin' und 'Malling 9' konnte bestätigt werden. Das Genprofil von 'Worcester Parmain' war mit dem ersten Genprofil identisch und stimmte auch bei der zweiten Überprüfung nicht mit dem Referenzprofil überein. Die Vertauschung und die falsche Referenzprobe 'Worcester Parmain' hatten keinen Einfluss auf die Etablierung der Methode.

Tabelle 1. Tabellarische Darstellung der 11 Referenzen, deren Herkunft und Probennummern.

Referenz	Referenzproben von INRA	Neue Proben vom INRA zur Nachprüfung der Referenzproben
'Delicious'	120987_X8508_Delicious_B01	-
'Fiesta'	120988_X4681_Fiesta_C01	-
'Michelin'	120990_X4471_Michelin_E01	125801_X3798_Michelin_R_A01
'Prima'	120992_X2596_Prima_G01	-
'Worcester Parmain'	120994_X2321_Worcester_Pearmin_A01	125803_X2321_Worc.Pearm._R_C01
'Malus x floribunda 821'	120989_X6518_Malus_floribunda_821_D01	-
'M. x robusta 5'	120993_X6022_Robusta_5_H01	-
'Malling 9'	120991_X3798_M9_F01	125802_X4471_M9_R_B01

Die Probennahme vor Ort begann in einem Zeitraum von Juli bis September 2017. Am 24.05.2017 erhielten wir den Bescheid von Frau Dr. Höfer, dass jeder der acht sammlungshaltenden Partner (Bundessortenamt, Prüfstelle Wurzten (BSA); Hermann Cordes Baumschule KG (Cordes); Julius-Kühn Institut (JKI); Kompetenzzentrum Obstbau-Bodensee (KOB); Landratsamt Kyffhäuserkreis (Kyffhäuserkreis); Landwirtschaftliche Lehranstalten Triesdorf Obstlehrgarten (LLA Triesdorf); Landesanstalt für Landwirtschaft und Gartenbau Sachsen-Anhalt (LLFG); Landesamt für Ländliche

Entwicklung, Landwirtschaft und Flurneuordnung Brandenburg, Obstbau Versuchsstation Müncheberg (OGR Müncheberg)) 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, Akzessionsnummer und Probennummer der molekularen Sortenprüfung. Um unsererseits eine genaue Rückverfolgbarkeit der Proben zu gewährleisten, arbeiteten wir mit Barcode- und Schlaufetiketten. Dabei erhielt jedes Probencouvert eine Barcode-Etikette, eine identische Etikette wurde auf die Liste zu der jeweiligen Probennummer geklebt und der jeweilige Baum mit einer Schlaufetikette und einer identischen Barcodenummer versehen. Die Barcodenummer war mit der Probennummer identisch.

Seite | 4

Gemäß Vertrag sollten wir beim Sammeln ausschließlich die Probennummer auf dem Probencouvert notieren und keine weiteren Angaben erhalten. Da die Sammlungshaltenden Partner teilweise nicht genügend zeitliche Kapazität hatten, uns durch ihre Sammlung zu begleiten, konnte dieses Vorgehen nicht umgesetzt werden. Wir erhielten deshalb zur Orientierung Listen mit zusätzlichen Angaben über den Standort der Bäume. Der geplante Zeitraum für die Probensammlung war August. Im Juli wurde mit den Partnern dahingehend Kontakt aufgenommen. Zwei Studenten der Universität Göttingen wurden angestellt, um einen großen Teil der Proben zu sammeln. 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.

Um eine möglichst optimale Blattqualität zu erhalten, mussten die Blätter bei trockenem Wetter gesammelt werden. Regnerische Augusttage verhinderten termingerechtes 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 aufbewahrt.

Im Herbst 2017 konnten, von insgesamt 5.314 angemeldeten Proben, 5.003 Proben gesammelt werden. 139 Proben fehlten auf Grund von abgestorbenen oder nicht vorhandenen Bäumen. Am 05.09.2017 wurde uns von Frau Dr. Höfer mitgeteilt, dass die Beprobung der Apfelbäume am Standort JKI aufgrund eines seit Mai 2017 andauernden Feuerbrand Befalls nicht im Herbst 2017 durchführbar war. Dieser Entscheidung wurde von Frau Prof. Hanke mitgetragen. Darum wurde die Beprobung der gesamten 179 Apfelbäume auf den Frühling 2018 verschoben und konnte im Mai 2018 erfolgreich durchgeführt werden. Die ausgefallenen Proben wurden im Vorfeld vom JKI ersetzt und im Mai 2018 vom Auftragsnehmer gesammelt. Die genauen Zahlen gesammelter Proben pro Standort sind der Tabelle 2 zu entnehmen.

Tabelle 2. Angaben zur Anzahl gesammelter Proben pro Standort.

Standort	Summe angemeldete Proben	Summe fehlender Proben am Sammelstandort	Zusätzliche unangemeldete Proben	Tatsächlich vorhandene Proben
BSA	276	0	0	276
Cordes	76	- 1	0	75
JKI	179	0	133	312
KOB	250	- 7	0	243
Kyffhäuser	436	- 12	0	424
LLA Triesdorf	2.303	- 81	0	2.222
LLFG	67	- 5	0	62
OGR	1.727	- 27	1	1.701
Müncheberg				
Total	5.314	- 133	134	5.315

Bei der Genotypisierung ist die Qualität der DNA von großer Bedeutung, welche aus den Blättern isoliert wird. 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.

Als geeignetste Methode zur Isolation genomischer DNA aus Apfel-Blätter hat sich die "Hot Shot"-Methode im 96-Well-Platten-Format erwiesen.

Aus dem gesammelten Material wurde eine Blattstanze abgeschnitten, lysiert und danach verdünnt. In einem ersten Schritt in der "Hot Shot"-DNA-Isolationsmethode wird eine Blattstanze mit dem Puffer A (0.1 M NaOH, 2% Tween20 in doppeldestilliertem 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 doppeltdestilliertem Wasser) neutralisiert die so lysierte Blattstanze. Im Anschluss an die Isolation wurde ein Aliquot des Isolates für die nachfolgende PCR 1 zu 50 verdünnt und die unverdünnten Isolate wurden bei -20°C eingelagert. Jede 96-Well-Platte beinhaltete eine Isolationskontrolle (Puffer A, Puffer B und kein Blattmaterial), eine PCR Kontrolle (Qiagen Multiplex PCR Kit und SSR Primer, keine Templat DNA), drei multiplex-spezifische Allelleitern und eine Positivkontrolle (Blattprobe, mit jeder 96-Wellplatte frisch isoliert).

Die festgelegten Mikrosatellitenmarker wurden hinsichtlich Robustheit auf den Referenzproben evaluiert und anschließend 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 17 SSR Marker

PCR Ansatz	Marker_Flourophor
M1	CH01h01_FAM
	CH01f07a_FAM
	CH03d07_ATTO532
	CH04c07_ATTO550
	CH05f06_ATTO550
	CH01f03b_ATTO565
M2	GD12_FAM
	CH02d08_ATTO565
	CH05e03_ATTO550
	CH02c09_ATTO550
	CH02g09_ATTO565
	CH01f02_ATTO532
M3	CH04f10_FAM
	COL_ATTO532
	Hi02c07_ATTO550
	CH04e05_ATTO550
	GD147_ATTO565

Die Primer-Sequenzen sind im Appendix in der Tabelle 1 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 vom 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. 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, 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 Analyzer 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 Analyzer für jeden einzelnen Lauf zu überprüfen, wurde zu Projektbeginn mit den auserwählten Referenzgenotypen für jeden Multiplex-PCR-Assay eine projektspezifische Allelleiter erstellt. Mit dieser Allelleiter 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 – 17). Die Auswertung erfolgte mit der Software GeneMarker V2.6.4.

Seite | 7

Die Interpretation der Chromatogramme erfolgte mit kommerziell erhältlicher Software und wurde bei jeder Probe auf Korrektheit überprüft. Um die Zuordnung der Allele robust zu gewährleisten, wurde in der Analysesoftware anhand der Referenzdaten für jeden Mikrosatelliten-Marker und dessen Allele, ein Allel-Raster für jeden Marker festgelegt. 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 wurden basierend auf den Alleldaten ermittelt. Dazu wurden paarweise Distanzen zwischen den Proben berechnet. 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 Distanzen zugelassen.

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

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

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

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

Die paarweisen Distanzen wurden mittels eines intern entwickelten Python-Scripts berechnet. Aufgrund der paarweisen Distanzen können die Gengruppen mittels Clusteranalyse zugeordnet werden. Als Cut-off für die Gengruppen wurde 0.25 bzw. 25% Unterschiede in den Alleldaten definiert. Dieser Cut-Off Wert wurde anhand der Verteilung der paarweisen Distanzen im vorhandenen Datensatz ermittelt (Abbildung 18). Damit werden alle Proben, welche zu 75-100% identisch sind, in dieselbe Gengruppe zugeordnet. Die Clusteranalyse und Definition der Gengruppe wurde mittels der Statistiksoftware R und eines intern entwickelten Python-Scripts berechnet.

3. Darstellung der Projektergebnisse und darüber hinaus gewonnene Erkenntnisse

Jeder verwendete Mikrosatellitenmarker zeigt individuelle Eigenschaften, beispielsweise in den vorkommenden Allelgrößen, Stotterbanden, Signalintensität und 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 X8508 `Delicious` 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.

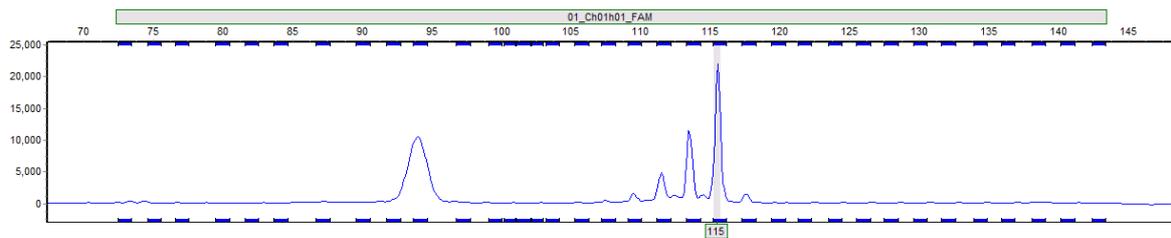


Abbildung 1. Chromatogramm des SSR-Markers Ch01h01 der Probe X8508 Delicious mit den Allelen 115 / 115.

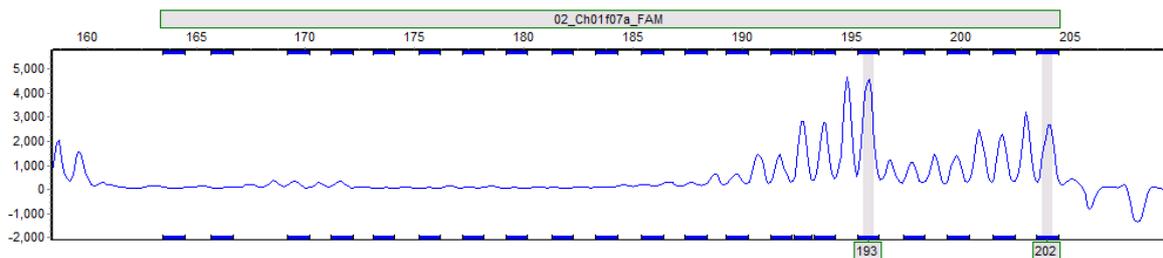


Abbildung 2. Chromatogramm des SSR-Markers Ch01f07a der Probe X8508 Delicious mit den Allelen 193 / 202.

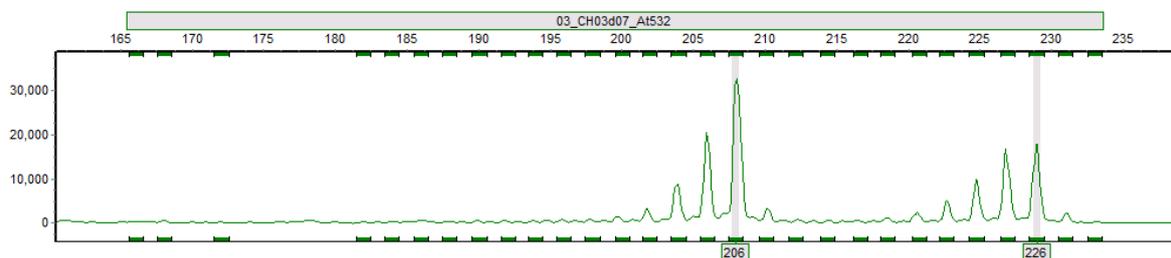


Abbildung 3. Chromatogramm des SSR-Markers CH03d07 der Probe X8508 Delicious mit den Allelen 206 / 226.

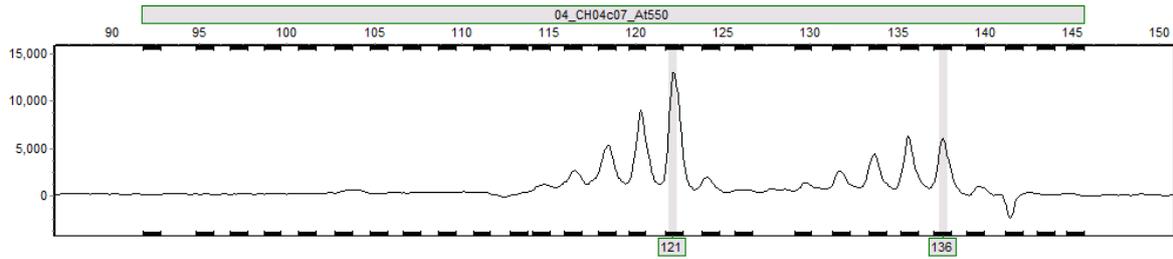


Abbildung 4. Chromatogramm des SSR-Markers CH04c07 der Probe X8508 Delicious mit den Allelen 121 / 136.

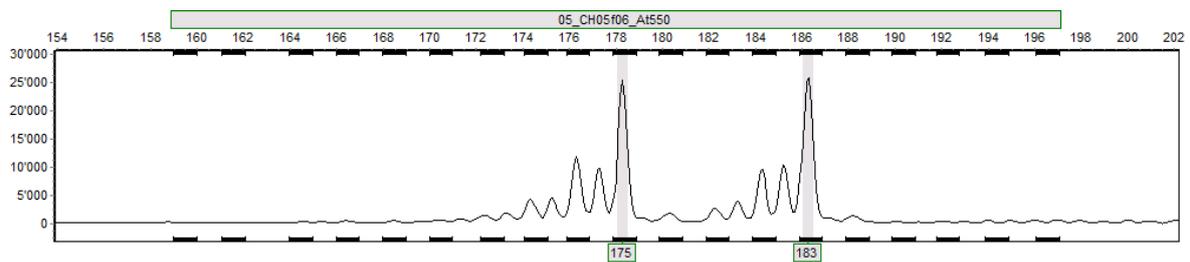


Abbildung 5. Chromatogramm des SSR-Markers CH05f06 der Probe X8508 Delicious mit den Allelen 175 / 183.

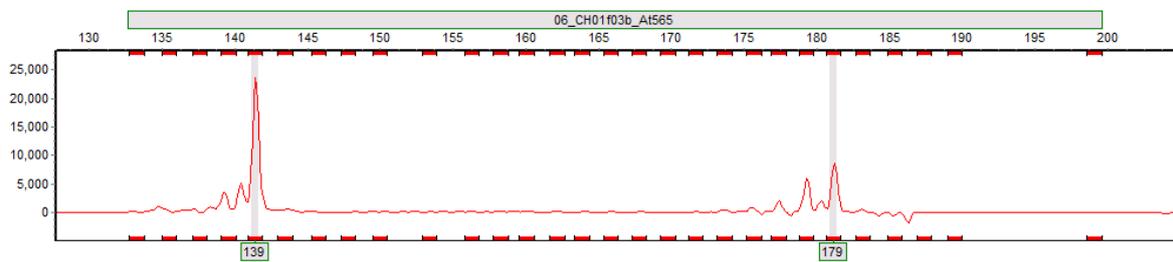


Abbildung 6. Chromatogramm des SSR-Markers CH01f03b der Probe X8508 Delicious mit den Allelen 139 / 179.

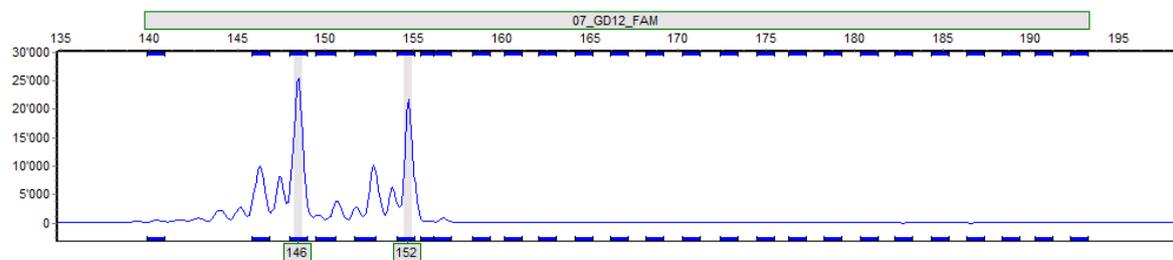


Abbildung 7. Chromatogramm des SSR-Markers GD12 der Probe X8508 Delicious mit den Allelen 146 / 152.

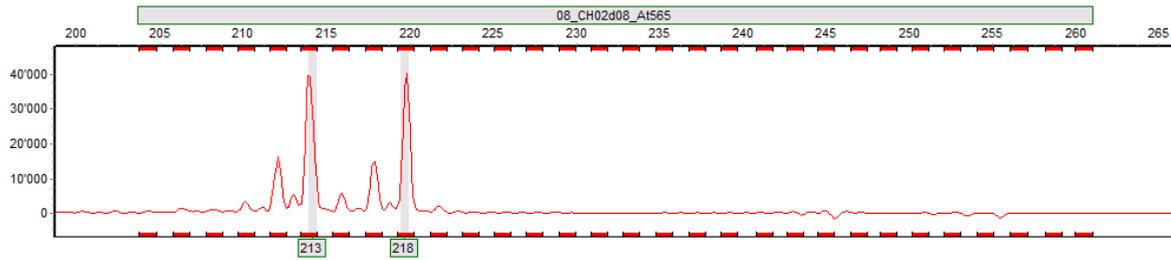


Abbildung 8. Chromatogramm des SSR-Markers CH02d08 der Probe X8508 Delicious mit den Allelen 213 / 218.

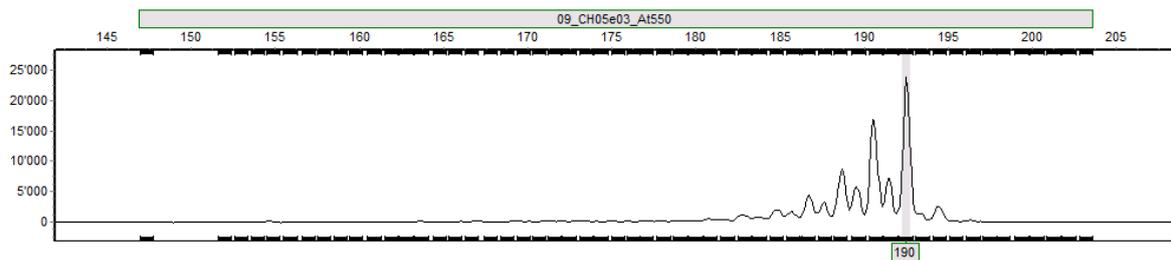


Abbildung 9. Chromatogramm des SSR-Markers CH05e03 der Probe X8508 Delicious mit den Allelen 190 / 190..

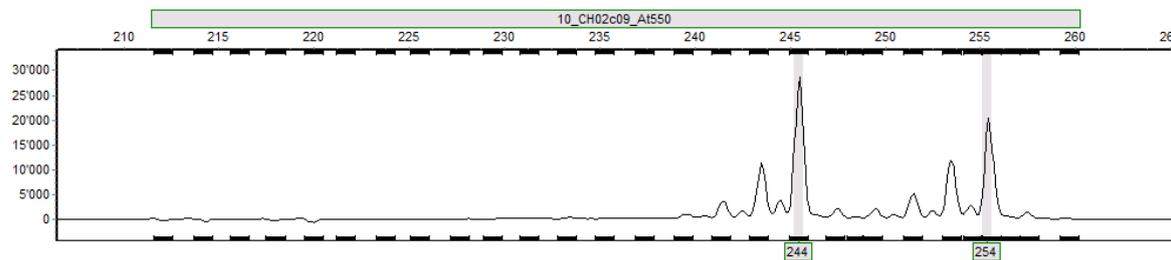


Abbildung 10. Chromatogramm des SSR-Markers CH02c09 der Probe X8508 Delicious mit den Allelen 244 / 254.

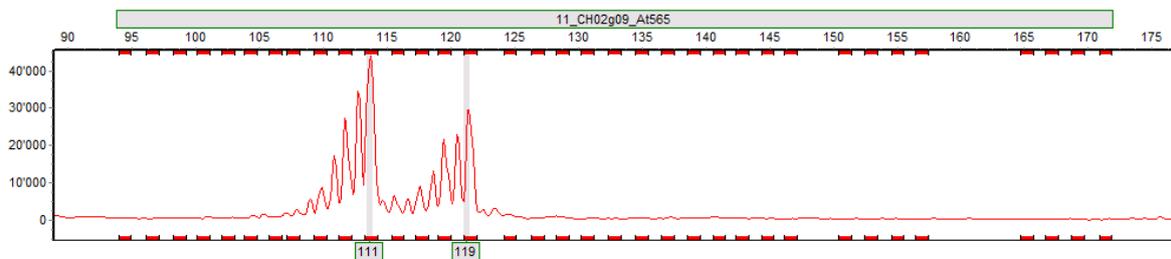


Abbildung 11. Chromatogramm des SSR-Markers CH02g09 der Probe X8508 Delicious mit den Allelen 111 / 119.

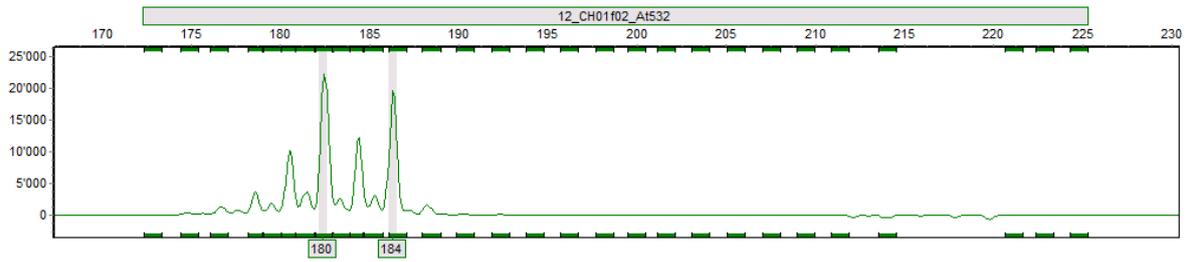


Abbildung 12. Chromatogramm des SSR-Markers CH01f02 der Probe X8508 Delicious mit den Allelen 180 / 184.

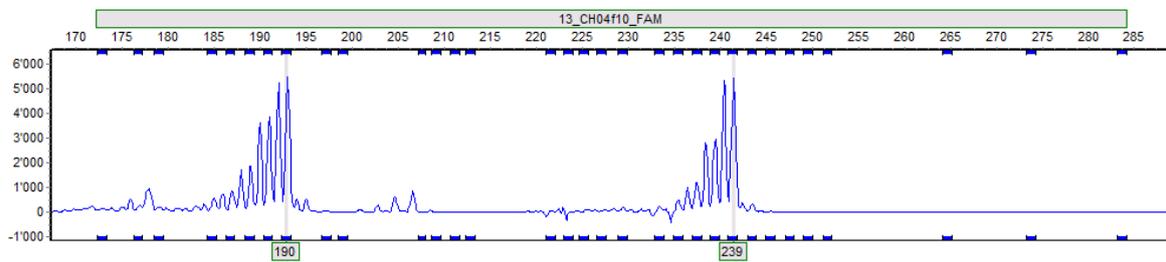


Abbildung 13. Chromatogramm des SSR-Markers CH04f10 der Probe X8508 Delicious mit den Allelen 190 / 239.

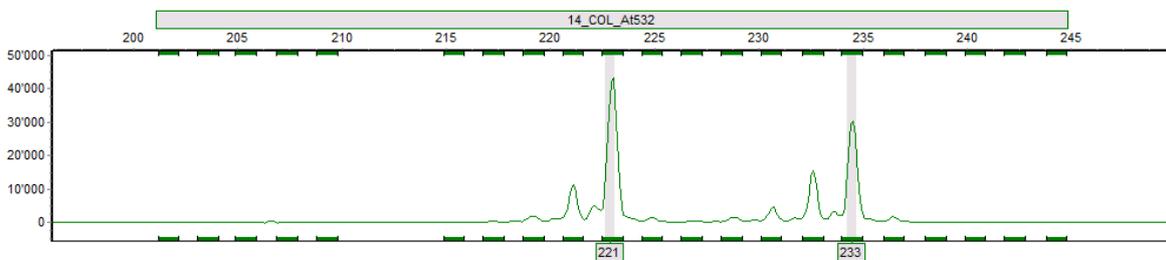


Abbildung 14. Chromatogramm des SSR-Markers COL der Probe X8508 Delicious mit den Allelen 221 / 233.

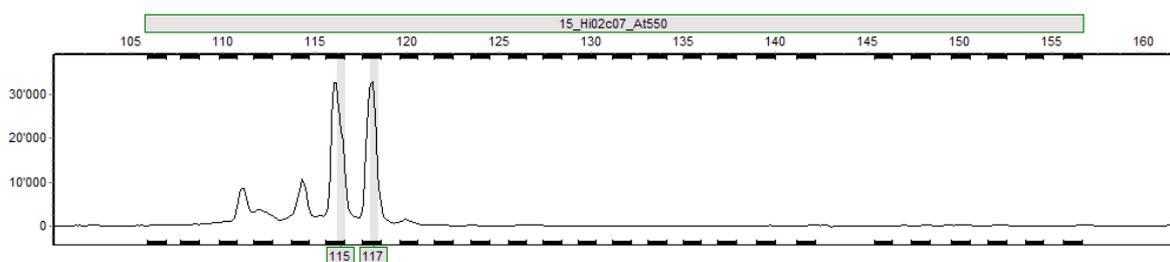


Abbildung 15. Chromatogramm des SSR-Markers Hi02c07 der Probe X8508 Delicious mit den Allelen 115 / 117.

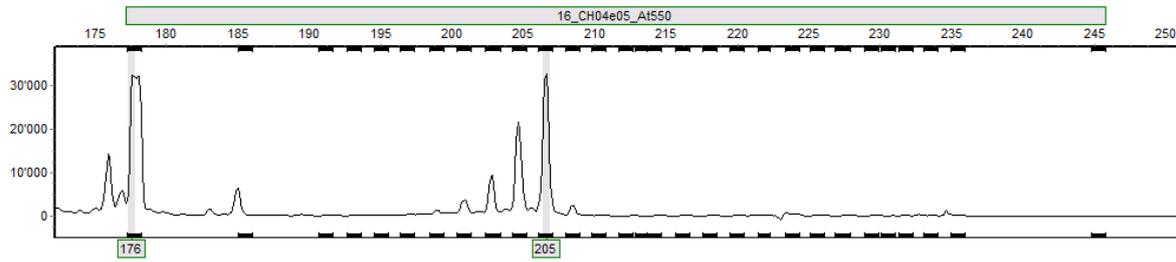


Abbildung 16. Chromatogramm des SSR-Markers CH04e05 der Probe X8508 Delicious mit den Allelen 176 / 205.

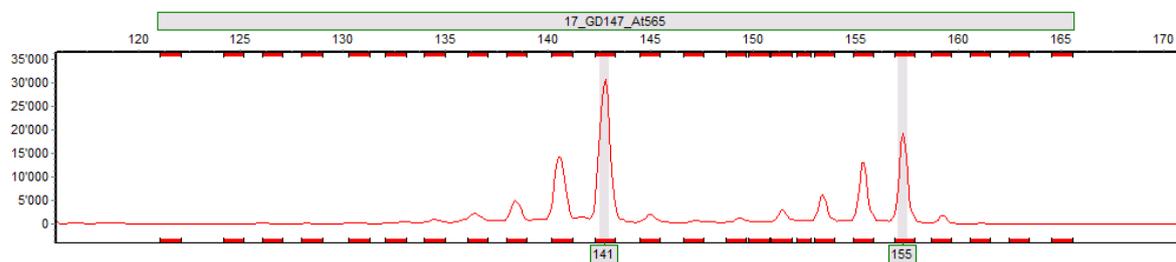


Abbildung 17. Chromatogramm des SSR-Markers GD147 der Probe X8508 Delicious mit den Allelen 141 / 155.

Die Mikrosatelliten-basierte Genotypisierung birgt aus verschiedenen Gründen das Risiko von Null-Allelen und Allel-Ausfällen in sich. 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. Eine noch höhere Datensicherheit und Qualität könnten durch Replikate, d.h. mehrfache Wiederholung der PCR erreicht werden.

Die Gengruppen wurden basierend auf den Alleldaten mittels paarweiser Distanzen zwischen den Proben berechnet. Die Verteilung der paarweisen Distanzen für den gesamten Datensatz ist in Abbildung 18 dargestellt. Gut ersichtlich ist die Unterscheidung von Eigenvergleichen sehr ähnlicher Genprofile (0-0.05 Distanzen bzw. 95-100% identisch) und sehr unterschiedlicher Genprofile (>0.25 Distanzen bzw. 0-75% identisch). Counts beschreibt die Anzahl der paarweisen Vergleiche im Datensatz. Somit fallen die meisten paarweisen Vergleiche der Genprofile in den Bereich ab 0.25 Distanzen bzw. sind weniger als 75% identisch. Basierend auf dieser Beobachtung wurde der Distanz Cut-Off für die Gengruppendefinition bei der Clusteranalyse bei 0.25 Distanzen definiert. Damit fallen alle Probenprofile, welche mehr als 75% identische Allele aufweisen, in eine Gengruppe. Die Clusteranalyse und Zuweisung der Proben erfolgten mit dem Programm R.

Dissimilarities observed in pairwise comparisons

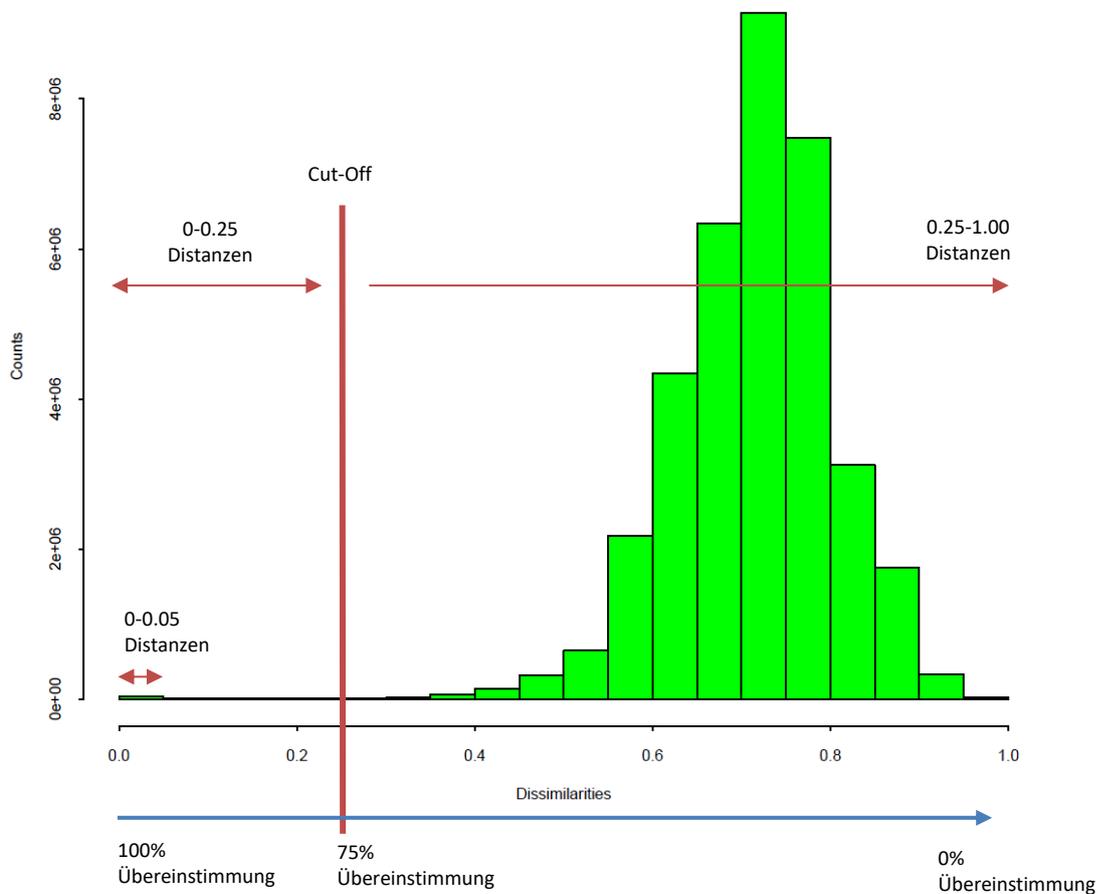


Abbildung 18. Verteilung der paarweisen Distanzen im Apfel-Datensatz. Die X-Achse beschreibt das Resultat von 0.0-1.0 paarweisen Distanzen und die Y-Achse beschreibt die Anzahl der paarweisen Vergleiche: es wurde jedes Probenprofil mit allen Probenprofilen im Datensatz verglichen und in diesem Histogramm dargestellt. Proben welche zu 95-100% identisch sind, sind bei 0-0.05 Distanzen abgebildet. Alle Proben mit einer paarweisen Distanz von 0-0.25 werden zu einer Gengruppe zusammengefasst. Die meisten paarweisen Vergleiche ($>8 \times 10^6$ paarweisen Vergleiche) landen in dem Bereich um 0.75 Distanzen bzw. einer Ähnlichkeit von 25%.

Bereits in den Jahren 2012 bis 2014 wurde eine molekulargenetische Sortenbestimmung von ca. 3.000 pomologisch bestimmten Apfelproben (Förderkennzeichen: 2809BE010 und 2813BE001 (BLE)) bei ecogenics durchgeführt. Die molekularen und pomologischen Daten wurden von Juliane Würdig und Dr. Henryk Flachowsky gesichtet und zusammengeführt. In dieser Metaanalyse konnten 954 Sortengruppen molekular beschrieben werden. Die Liste dieser Gengruppen wurde uns von Frau Höfer am 08. März 2018 zur Verfügung gestellt. Da die zweite Sortenbestimmung identisch wie die erste molekulargenetischen Sortenbestimmung durchgeführt wurde, konnten diese 954 repräsentative Sorten-Datensätze für den Gengruppenabgleich miteingelesen werden. Im Abgleich der 954 Datensätze aus der ersten Prüfung fallen teilweise Sorten-Datensätze in gemeinsame Gengruppen zusammen. Grund für diesen Zusammenschluss ist, dass diejenigen Sorten keine Abweichungen aufweisen oder sich nur in wenigen Allelen (<0.25 Distanzen) im Genprofil unterscheiden und somit mit dem definierten Cut-Off von 0.25 Distanzen als eine Gengruppe zusammengefasst werden

Die im Rahmen der ersten und zweiten molekulargenetischen Bestimmung von Apfelsorten ermittelten 1.533 Gengruppen mittels der Cut-off-0.25-Distanzen 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, gegen welche weitere Proben für die Sortenechtheitsprüfung abgeglichen werden können.

Der Wert der Daten kann noch deutlich erhöht werden, wenn weitere, ähnliche Studien von anderen Auftraggebern mit denselben Markern, oder allenfalls einer Auswahl davon, durchgeführt werden. Durch die Zusammenstellung der Gengruppen in eine Datenbank, können diese Daten als Grundlage für zukünftige Sortenabgleiche benutzt werden.

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

Die ursprünglich geplanten, vertraglich festgehaltenen Ziele wurden weitgehend erreicht. Insgesamt wurden 5.314 Proben zur Überprüfung angemeldet und sogar 5.315 Proben gesammelt. Zudem wurden statt acht 11 Referenzproben geliefert. Alle 11 Referenzproben wurden 4 Mal analysiert und aus den 5.315 Proben wurden zudem 6 Proben als Blindproben ausgewählt, welche auch 4 Mal analysiert wurden.

Seite | 15

Im Rahmen der Auswertung der Ergebnisse wurde eine Validierung der Allelgrößen anhand der Referenzgenotypen durchgeführt. Die Darstellung der Ergebnisse erfolgte in tabellarischer Form und enthält die Standortnummer des Baumes sowie die Probennummer und die Allelgrößen für die ausgewählten Marker. Außerdem wurden die Daten nach identischen Fingerprints sortiert und eine eindeutige Identifikationsnummer vergeben.

Die Ergebnisse wurden am 29.06.2018 erfolgreich der Koordinationsstelle der DGO per E-Mail übermittelt. Die DNA wurde am 12.09.2018 an das JKI zu Händen von Frau Dr. Höfer geliefert.

5. Zusammenfassung

5.315 Blattproben, welche die biologische Vielfalt des Apfels in Deutschland repräsentieren, wurden molekulargenetisch analysiert. Die DNA wurde mittels "Hot Shot" aus den Blattproben isoliert. Von jeder Probe wurden 17 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 dieser zweiten Sortenprüfung wurde mittels paarweiser Distanzen und Clusteranalyse einer Gengruppe zugeordnet und mit den Daten aus der ersten Sortenprüfung abgeglichen. So konnten 1533 Gengruppen definiert werden, welche als Referenzen für zukünftige Sortenechtheitsprüfung verwendet werden können.

6. Literaturverzeichnis

- Ellegren H, 2004. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5(6):435-445.
- 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.
- HiDRAS-Hompage: <http://www.hidras.unimi.it/HiDRAS-SSRdb/pages/index.php>, aufgerufen am 02.07.2018, beschreibt eine Suchmaske für SSR-Markern in Äpfel.

7. Appendix

Tabelle 1. Primer Sequenzen der 17 verwendeten Mikrosatelliten Marker.

Markenname	F Primer Sequenz	R Primer Sequenz
CH01h01	GAAAGACTTGCAAGTGGGAGC	GGAGTGGGTTTGAGAAGGTT
CH01f07a	CCCTACACAGTTTCTCAACCC	CGTTTTGGAGCGTAGGAAC
CH03d07	CAAATCAATGCAAACTGTCA	GGCTTCTGGCCATGATTTTA
CH04c07	GGCCTTCCATGTCTCAGAAG	CCTCATGCCCTCCACTAACA
CH05f06	TTAGATCCGGTCACTCTCCACT	TGGAGGAAGACGAAGAAGAAAAG
CH01f03b	GAGAAGCAAATGCAAAACCC	CTCCCCGGCTCCTATTCTAC
GD12	TTGAGGTGTTTCTCCATTGGA	CTAACGAAGCCGCCATTTCTTT
CH02d08	TCCAAAATGGCGTACCTCTC	GCAGACACTCACTCACTATCTCTC
CH05e03	CGAATATTTTCACTCTGACTGGG	CAAGTTGTTGACTGCTCCGAC
CH02c09	TTATGTACCAACTTTGCTAACCTC	AGAAGCAGCAGAGGAGGATG
CH02g09	TCAGACAGAAGAGGAACTGTATTTG	CAAACAAACCAGTACCGCAA
CH01f02	ACCACATTAGAGCAGTTGAGG	CTGGTTTGTTCCTCCAGC
CH04f10	GTAATGGAAATACAGTTTCACAA	TTAAATGCTTGGTGTGTTTTGC
COL	AGGAGAAAGGCGTTTACCTG	GACTCATTCTTCGTCGCTACTG
Hi02c07	AGAGCTACGGGGATCCAAAT	GTTTAAGCATCCCGATTGAAAGG
CH04e05	AGGCTAACAGAAATGTGGTTTG	ATGGCTCCTATTGCCATCAT
GD147	TCCCGCCATTTCTCTGC	GTTTAAACCGCTGCTGCTGAAC

