

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

19. Juni 2018

ABSCHLUSSBERICHT

Auftragnehmer:
ecogenics GmbH

Projektnummer:
123-02.05-20.0007/16-II-C

Förderkennzeichen:
2809BE009

Thema:
Molekulargenetische Sortenprüfung der Deutschen Genbank Obst: Erdbeere

Laufzeit:
Mai 2016 – Mai 2018

Berichtszeitraum:
Mai 2016 – Mai 2018

1. ZIELE UND AUFGABENSTELLUNG DES PROJEKTS

Um die Sammlung und Erhaltung von genetischen Ressourcen bei Obst in wissenschaftliche langfristig abgesicherte und nachhaltige Weise zu gewährleisten, hat sich die Deutsche Genbank Obst (DGO) zum Ziel gesetzt, das Sortenspektrum der Erdbeere mit Bezug zu Deutschland pomologisch und genetisch zu bestimmen. Ecogenics wurde beauftragt, die molekulare Überprüfung aller Erdbeersorten der DGO Sammlung durchzuführen. Die molekulare Überprüfung erfolgte mit der in Horvath et al. (2011) beschriebenen Methode, welche auf den Ergebnissen des Europäischen Forschungsprojektes GenBerry (GEBRES 036) basiert. Zum Zeitpunkt der Auftragserteilung befanden sich ca. 240 verschiedene Erdbeersorten mit mindestens drei Pflanzen je Sorte im Freiland an den Standorten Julius Kühn-Institut (JKI) in Dresden-Pillnitz und Bundessortenamt in Wurzen.

Seite | 2

2. PLANUNG UND ABLAUF DES PROJEKTS

Der Ablauf des Projekts bestand darin, die anzuwendende Methodik nach Horvath et al (2011) bei Ecogenics zu etablieren und mit dem Auftraggeber abzustimmen. In einem zweiten Schritt erfolgte die Probenübernahme und Lagerung der Proben bei $\leq -20^{\circ}\text{C}$ sowie die Überprüfung des Analyseverfahrens anhand Referenzgenotypen. Nach Festlegung und Etablierung des Analyseverfahrens, waren >794 Pflanzenproben von bis zu 227 verschiedene Sorten (je 3 – 6 Pflanzen pro Sorte) zu testen. Die Analyse schließt die DNA Extraktion und die Genotypisierung mittels vorher etablierten Multiplex-Polymerasen-Ketten-Reaktion-(PCR)-Assays ein. Dabei erfolgten die Analysen ohne pomologische Informationen resp. Informationen zu den Sorten. Insgesamt wurden 951 Proben analysiert. Die Anzahl Proben überschreitet die garantierte Probenanzahl gemäß Vertrag um 157 Proben.

In Absprache mit dem Auftraggeber kam es zu einer Änderung und Erweiterung der Referenzgenotypen. Referenzgenotypen die im Set in zwei bis 5 facher Probenwiederholungen für die Etablierung der Methode zur Verfügung gestellt wurden: ‚Annelise‘, ‚Deutsch Evern‘, ‚Königin Luise‘, ‚Korbinskaya rannyya‘. Referenzgenotypen die hinzukamen: ‚Cambridge Favourite‘, ‚Chandler‘, ‚Paul Wallbaum‘, ‚Amazone‘, ‚George Soltwedel‘, ‚Hansa‘, ‚Honeoye‘, ‚Madame Moutot‘, ‚Mara des Bois‘, ‚Mieze Schindler‘, ‚Pink Panda‘. Referenzgenotypen die im Set nicht enthalten waren: ‚Apollo‘, ‚Aprikose‘.

Um die eindeutige Zuordnung der Proben im Labor zu gewährleisten, wurden für die Probenübernahme im Vorfeld Briefumschläge mit Barcode Labels versehen und an die DGO versandt. 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 *Fragaria* Blättern hat sich die Säulenextraktion mit DNeasy Plant Kit (Qiagen) erwiesen. Aus dem vom Auftraggeber zur Verfügung gestellten Material wurde ein Blattstück mit einem Skalpell abgeschnitten, getrocknet, zerkleinert und nach dem Hersteller-Protokoll extrahiert. Im Anschluss an die Isolation wurde ein Aliquot des Isolates für die nachfolgende PCR verdünnt, die unverdünnten Isolate wurden bei -20°C eingelagert.

Die festgelegten Mikrosatellitenmarker wurden hinsichtlich Robustheit auf den Referenzproben evaluiert. Dabei war abzuklären, welche 20 Marker aus den 23 Markern nach Horvath et al. (2011) für die Sortenüberprüfung geeignet sind. Im Gegensatz zu der Arbeit von Horvath et al. (2011), bei der auch mit Gelelektrophorese gearbeitet wurde, werden bei dieser Untersuchung ausschließlich Fluoreszenzfarbstoff modifizierte PCR Primer eingesetzt, damit die Fragmentlängenanalyse mittels

Kapillarelektrophorese erfolgen kann. Die 23 Marker (BFACT037, SF1-E10, UDF018, CFVCT028, UFFa13C07, UFFa11G07, UDF-001, CFVCT016, SF1-A01, UFFa09E12, UAFv7344, UDF004, UDF009, EMFv030, ChFaM036, EMFv017, MCAD FAC_001, UFFa01D03, UAFv8316, EMFv001, UFFa03C04, ChFaM078 und ChFaM138) wurden auf den Referenz-Proben einzeln amplifiziert und mit einer Software auf ihre Eignung analysiert.

Die Amplikonlänge und Anzahl der Allele auf den getesteten Proben wurde bestimmt und das Aussehen des Markers und dessen Auswertbarkeit wurde evaluiert. Drei Marker aus dem Set der 23 Marker waren wenig robust (BFACT037, SF1-E10 und UDF018) und wurden aus diesem Grund in Rücksprache mit dem Auftraggeber von der Analyse ausgeschlossen.

Die 20 ausgewählten Mikrosatelliten Marker wurden aufgrund der vorhergehenden Evaluierung so kombiniert, dass sie in fünf 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 1):

Tabelle 1: Mikrosatelliten Marker Kombination im Multiplex.

<i>PCR Ansatz</i>	<i>Marker</i>	<i>Fluorophor</i>
1	UFFa13C07	6-FAM
	CFVCT028	ATTO 532
	UDF001	ATTO 550
	UFFa11G07	ATTO 565
2	CFVCT016	6-FAM
	UAFv7344	ATTO 532
	SF1-A01	ATTO 550
	UFFa09E12	ATTO 565
3	EMFv030	6-FAM
	ChFaM036	ATTO 532
	UDF009	ATTO 550
	UDF004	ATTO 565
4	EMFv017	6-FAM
	UFFa01D03	ATTO 532
	UAFv8316	ATTO 550
	MCAD FAC_001	ATTO 565
5	ChFaM138	6-FAM
	EMFv001	ATTO 532
	UFFa03C04	ATTO 550
	ChFaM078	ATTO 565

Die Primer-Sequenzen sind im Appendix in der Tabelle 5 aufgeführt. Die zu verwendenden Mikrosatelliten-Marker Systeme wurden so etabliert und angepasst, dass mit den verwendeten DNA Isolaten robuste, reproduzierbare Daten erzeugt werden können.

Grundsätzlich erfolgten sämtliche Schritte im Rahmen der DNA Isolation, PCR Amplifikation und Kapillarelektrophorese in 96-Wellplatten. Die fünf Multiplex-PCR-Assays wurden mit den verdünnten

Isolaten durchgeführt und zusammen mit einem internen Größenstandard auf einem Kapillarelektrophoresegerät analysiert. Jede Platte beinhaltete Extraktionskontrolle (kein Blattmaterial), PCR Kontrolle (keine Templat DNA in PCR), fünf multiplex spezifische Allelleitern und eine Positivkontrolle (vorgängig isolierte *Fragaria* DNA als Templat).

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, 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 Allel Daten waren untereinander abzugleichen und einer Gengruppe zuzuordnen.

Seite | 4

Die Gengruppen wurden basierend auf den Allel Daten 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.

Nach Offenlegung der Sortennamen wurden die Proben auf Ihre Sortenzugehörigkeit überprüft und mit den ermittelten Gengruppen abgeglichen. In einer letzten Phase wären Wiederholungen durchgeführt worden, jedoch waren keine Wiederholungen nötig. Die DNA der 951 Proben und die Primer-Stocklösungen sind am 13.03.2018 beim Auftraggeber angekommen.

3. WISSENSCHAFTLICHER UND TECHNISCHER STAND, AN DEN ANGEKNÜPFT 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 (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 Erdbeeren angewandt.

4. MATERIAL UND METHODEN

Für die Probenaufbewahrung wurden barcodierte Briefumschläge von Ecogenics zu Verfügung gestellt. Vom Auftraggeber wurden pro Pflanze jeweils zwei eher kleine, junge und gesunde Blätter am Stiel gepflückt und in das dazugehörige, entsprechend barcodierte Probencouvert gesteckt. Zudem wurde auf dem Probenbegleitformular die Identität der Pflanze zu dem entsprechenden Barcode notiert. Die Probenamen wurden vom Auftraggeber definiert. Das Probenbegleitformular wurde elektronisch und in Papierform zusammen mit den Proben an Ecogenics via DHL Kurier versandt. Die Proben wurden bei $-20^{\circ}\text{C} \pm 5^{\circ}\text{C}$ gelagert. Für die Extraktion wurden die Proben aufgetaut und ein ca. 1-2cm² Blattstück bei $70^{\circ}\text{C} \pm 5^{\circ}\text{C}$ getrocknet, mit rostfreien Stahlkugeln und einem Tissue Lyser (Qiagen) zerkleinert. DNA Extraktion wurde mit dem DNeasy 96 Plant Kit (Qiagen) gemäß Protokoll des Herstellers durchgeführt. Um keine Vertauschungen zu verursachen, wurden die Probenzuordnung in die 96-Well Platte mittels Scannen des Barcodes dokumentiert. Die DNA wurde mit Quant-iT™ PicoGreen® dsDNA Reagenz (Thermo Fisher Scientific) und einem

Spektralphotometer quantifiziert. Für optimale PCR Amplifikation wurde die DNA Konzentration auf 1 ± 0.5 ng/ul eingestellt.

Der Forward PCR Primer vom jedem Mikrosatelliten-Marker ist am 5'-Ende mit einem Fluorophor modifiziert, um die Fragmentdetektion 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 ATTO 532, ATTO 550 und ATTO 565 verwendet. Am 5' Ende der Reverse PCR Primer Sequenzen wurde eine Pigtail Sequenz GTTCTT 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, 15 min Denaturierung gefolgt von 35 Zyklen mit 94 °C, 0.5 min; 56 °C, 1.5 min; 72 °C, 1 min und einer finalen Elongation bei 72 °C für 30 min.

Seite | 5

Die Fragmentlängenanalysen wurden mit dem Längenstandard GeneScan LIZ500 (Applied Biosystems) auf einem Applied Biosystems 3730xl DNA Analyzer durchgeführt. Folgenden 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.

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 – 20). Die Auswertung erfolgte mit der Software GeneMarker V2.6.4.

Während der Analyse wurde festgestellt, dass der Marker CFVCT016_6-FAM nicht robust amplifiziert und darum nicht auswertbar ist. Aus diesem Grund wurde dieser Marker unter Absprache mit dem Auftraggeber aus den weiteren Analysen ausgeschlossen.

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).

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

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 0.1 bzw. 10% 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 90-100% identisch sind, in dieselbe Gengruppe zugeordnet. Die Clusteranalyse und Definition der Gengruppe wurde mittels der Statistiksoftware R berechnet.

5. AUSFÜHRLICHE DARSTELLUNG DER ERGEBNISSE

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 Allelzuordnung nicht eindeutig möglich. Im Folgenden werden sämtliche Marker kurz illustriert. Es wird jeweils exemplarisch das Chromatogramm der Proben JKI_422, JKI_062, JKI_169, JKI_071 und JKI_340 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 farblich hervorgehobenen Bereiche definieren die Bins.

Seite | 6

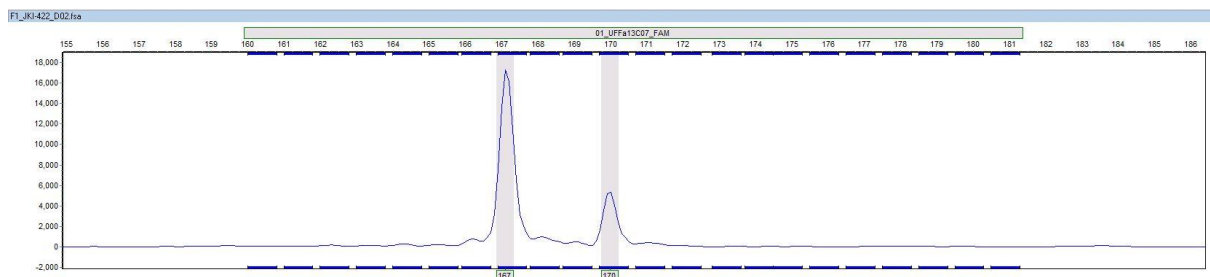


Abbildung 1. Chromatogramm des Mikrosatelliten Markers UFFa13C07 (6-FAM) der Probe JKI_422. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_422 enthält die Allele 167 und 170.

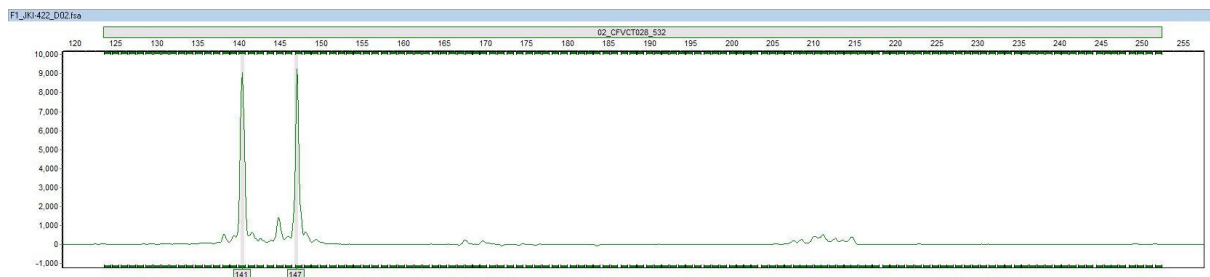


Abbildung 2. Chromatogramm des Mikrosatelliten Markers CFVCT028 (ATTO 532) der Probe JKI_422. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_422 enthält die Allele 141 und 147.

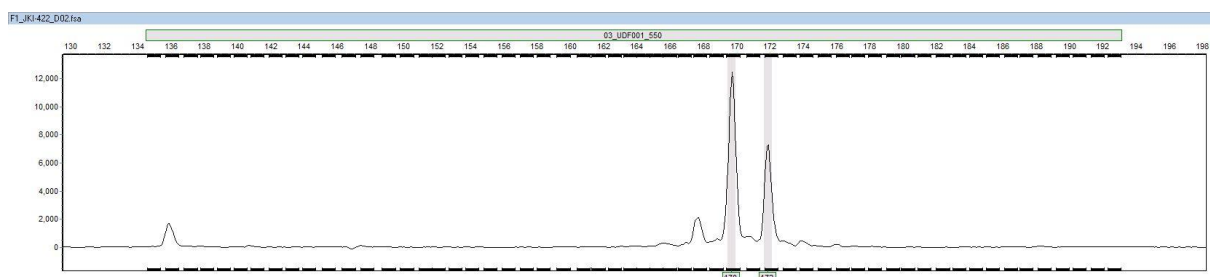
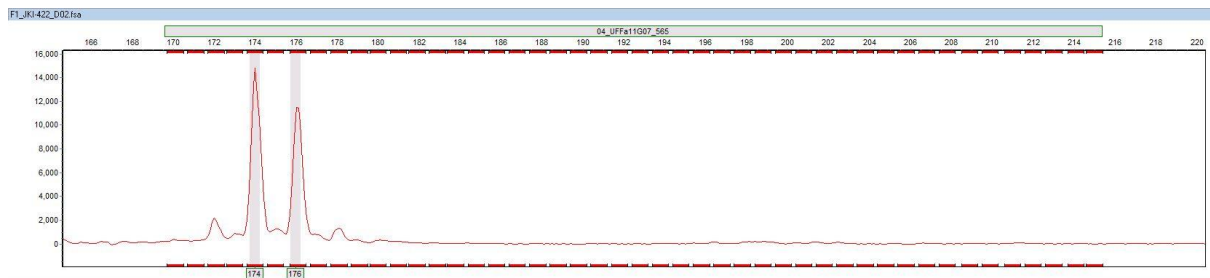


Abbildung 3. Chromatogramm des Mikrosatelliten Markers UDF001 (ATTO 550) der Probe JKI_422. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_422 enthält die Allele 170 und 172.



Seite | 7

Abbildung 4. Chromatogramm des Mikrosatelliten Markers UFFa11G07 (ATTO 565) der Probe JKI_422. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_422 enthält die Allele 174 und 176.

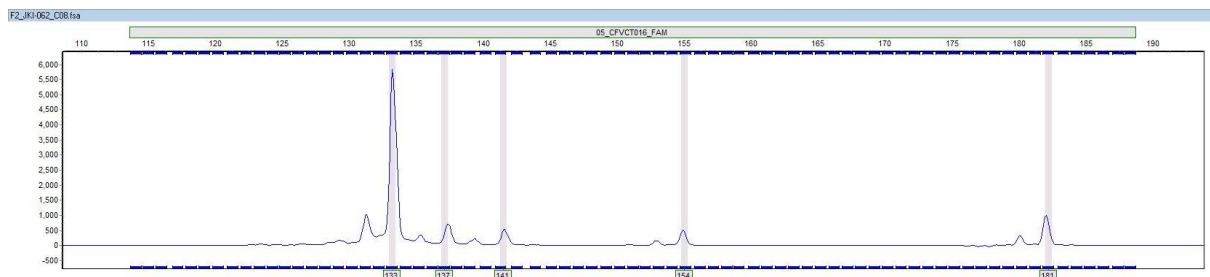


Abbildung 5. Chromatogramm des Mikrosatelliten Markers CFVCT016 (6-FAM) Probe JKI_062. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_062 enthält die Allele 133, 137, 141, 154 und 181. Dieser Marker wurde in der Gengruppenzuordnung nicht berücksichtigt.

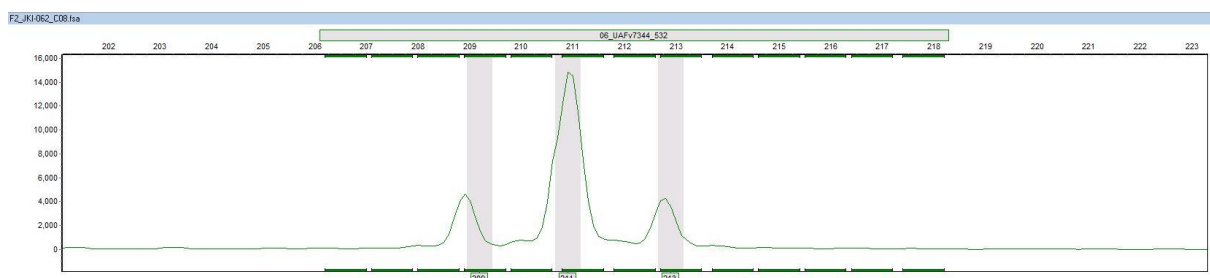


Abbildung 6. Chromatogramm des Mikrosatelliten Marker UAFv7344 (ATTO 532) der Probe JKI_062. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_062 enthält die Allele 209, 211 und 213.

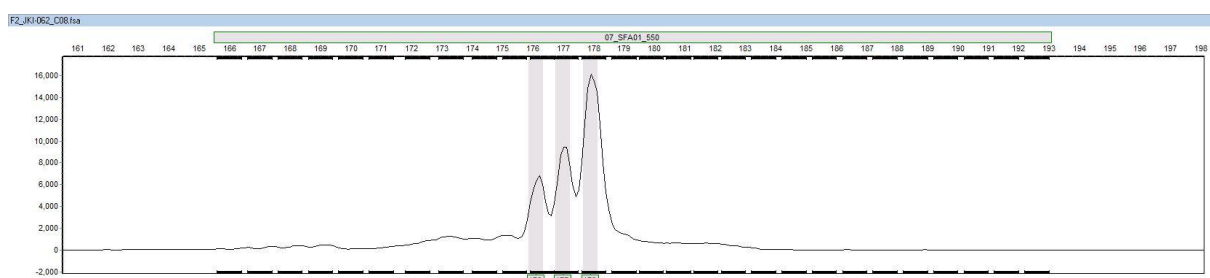
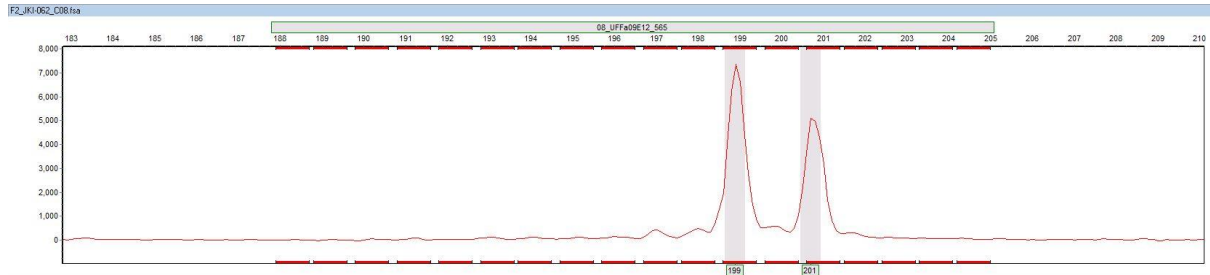


Abbildung 7. Chromatogramm des Mikrosatelliten Marker SF1-A01 (ATTO 550) der Probe JKI_062. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_062 enthält die Allele 176, 177 und 178.



Seite | 8

Abbildung 8. Chromatogramm des Mikrosatelliten Marker UFFa09E12 (ATTO 565) der Probe JKI_062. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_062 enthält die Allele 199 und 201.

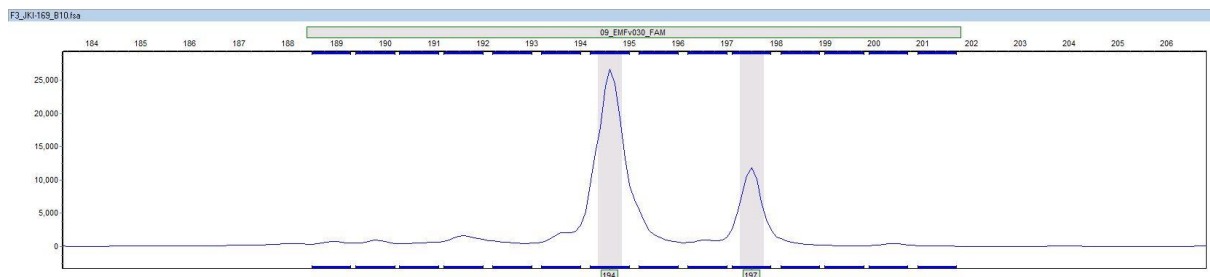


Abbildung 9. Chromatogramm des Mikrosatelliten Marker EMFv030 (6-FAM) der Probe JKI_169. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_169 enthält die Allele 194 und 197.



Abbildung 10. Chromatogramm des Mikrosatelliten Marker ChFaM036 (ATTO 532) der Probe JKI_169. Die Probe JKI_169 enthält das Allele 212.

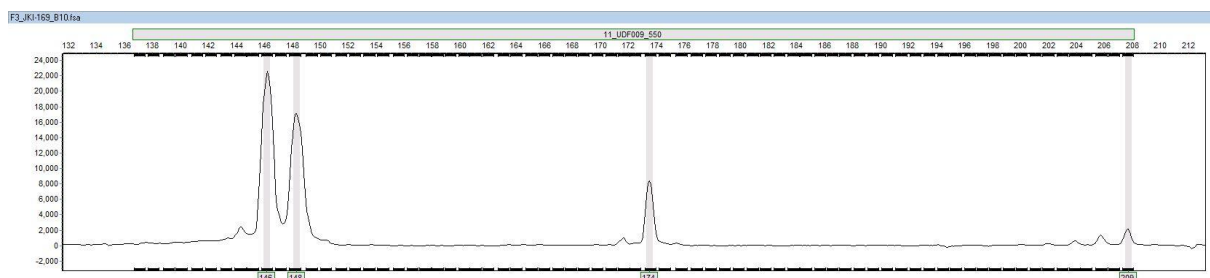
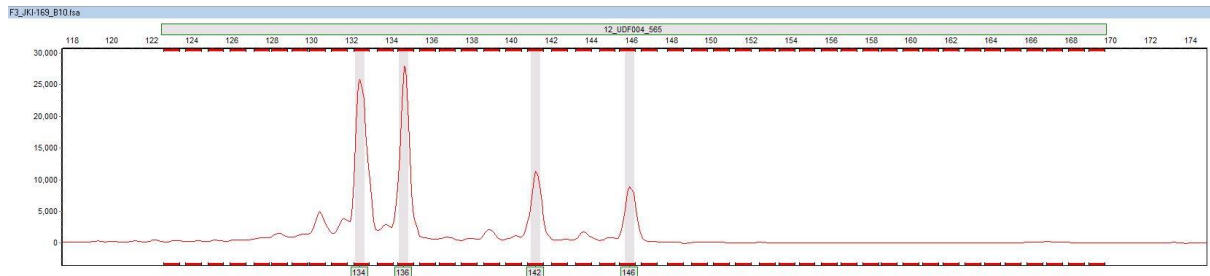


Abbildung 11. Chromatogramm des Mikrosatelliten Marker UDF009 (ATTO 550) der Probe JKI_169. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_169 enthält die Allele 146, 148, 174 und 209.



Seite | 9

Abbildung 12. Chromatogramm des Mikrosatelliten Marker UDF004 (ATTO 565) der Probe JKI_169. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_169 enthält die Allele 134, 136, 142 und 146.

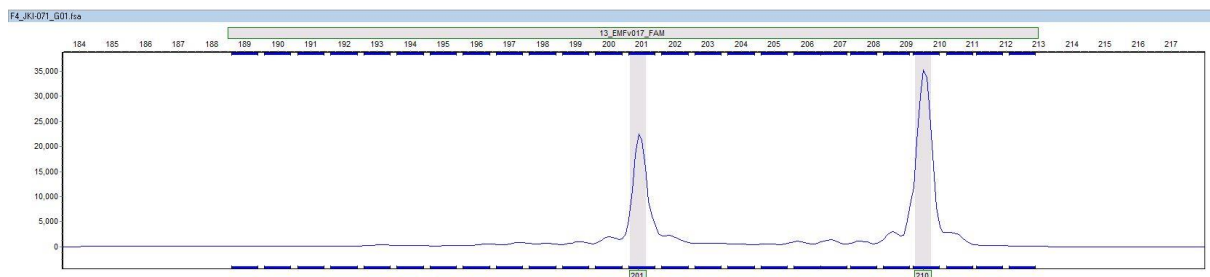


Abbildung 13. Chromatogramm des Mikrosatelliten Marker EMFv017 (6-FAM) der Probe JKI_071. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_071 enthält die Allele 201 und 210.

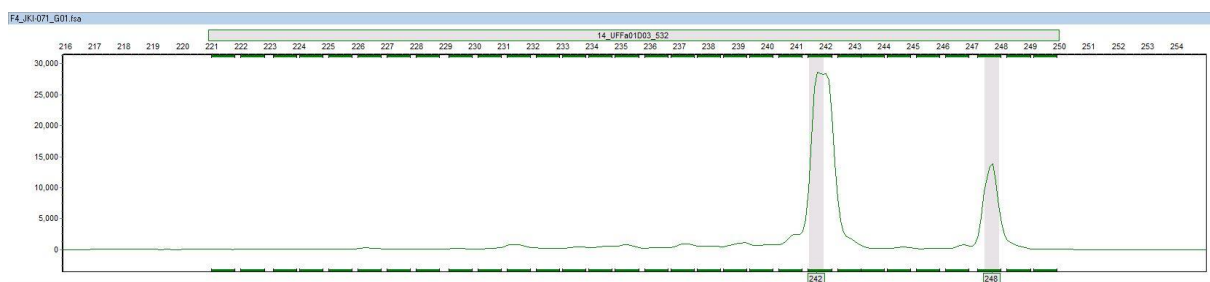


Abbildung 14. Chromatogramm des Mikrosatelliten Marker UFFa01D03 (ATTO 532) der Probe JKI_071. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_071 enthält die Allele 242 und 248.

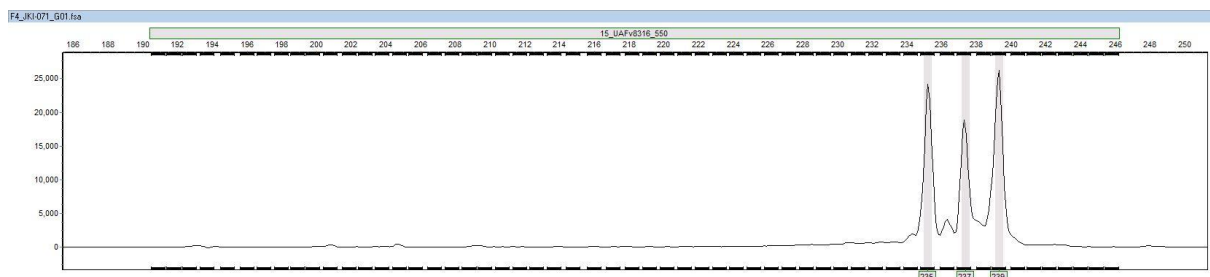
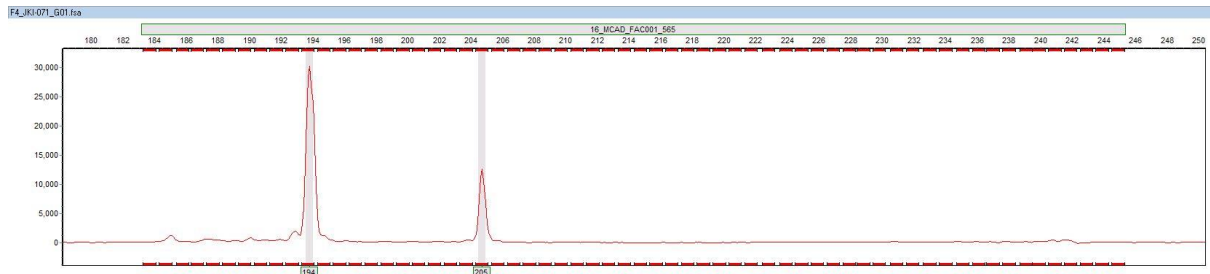


Abbildung 15. Chromatogramm des Mikrosatelliten Marker UAFv8316 (ATTO 550) der Probe JKI_071. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_071 enthält die Allele 235, 237 und 239.



Seite | 10

Abbildung 16. Chromatogramm des Mikrosatelliten Marker MCAD FAC_001 (ATTO 565) der Probe JKI_071. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_071 enthält die Allele 194 und 205.

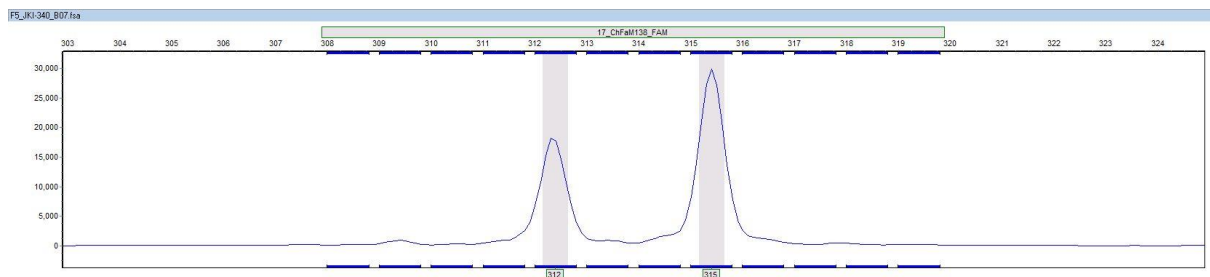


Abbildung 17. Chromatogramm des Mikrosatelliten Marker ChFaM138 (6-FAM) der Probe JKI_340. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_340 enthält die Allele 312 und 315.

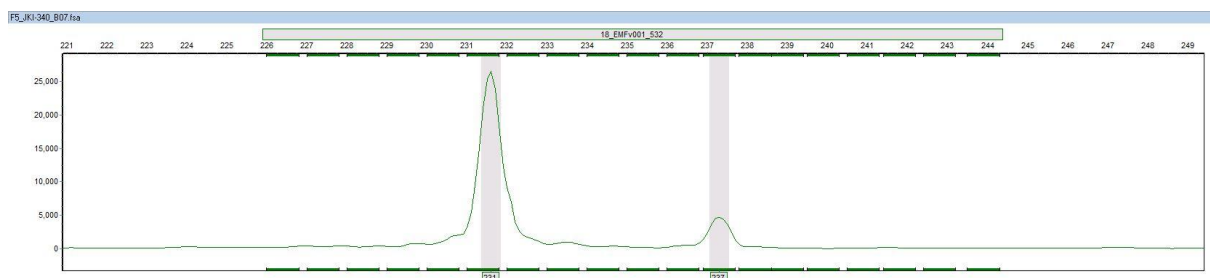


Abbildung 18. Chromatogramm des Mikrosatelliten Marker EMFv001 (ATTO 532) der Probe JKI_340. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Probe JKI_340 enthält die Allele 231 und 237.

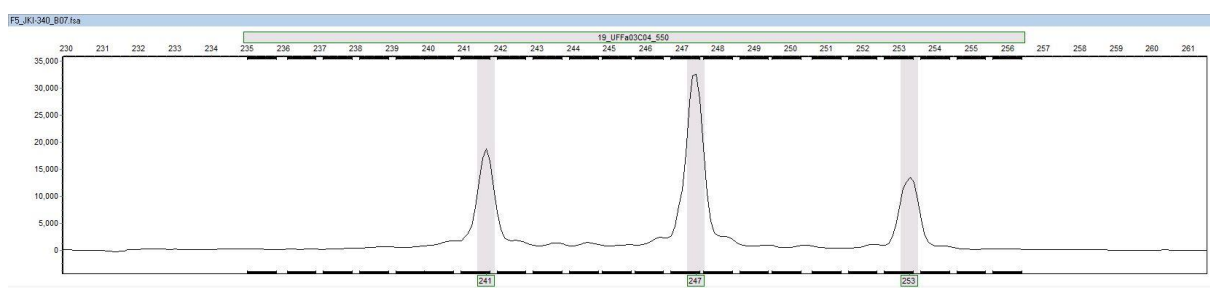
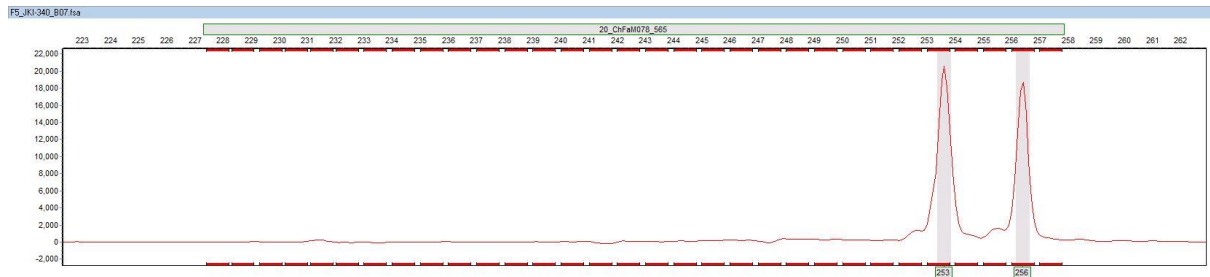


Abbildung 19. Chromatogramm des Mikrosatelliten Marker UFFa03C04 (ATTO 550) der Probe JKI_340. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Proben JKI_340 enthält die Allele 241, 247 und 253.



Seite | 11

Abbildung 20. Chromatogramm des Mikrosatelliten Marker ChFaM078 (ATTO 565) der Probe JKI_340. Im Chromatogramm ist der für den verwendeten Marker relevante Größenbereich dargestellt. Die Proben JKI_340 enthält die Allele 253 und 256.

Die Mikrosatelliten basierte Genotypisierung birgt aus verschiedenen Gründen das Risiko von Null-Allelen und Allelausfällen (allel drop outs). 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. In beiden Fällen wurde sowohl die Isolation, wie auch die betroffene Multiplex PCR wiederholt. Eine höhere Datensicherheit und Qualität könnte durch Replikate erreicht werden.

Nach dieser Analyse wurde festgestellt, dass der Marker CFVCT016_FAM nicht robust amplifiziert und somit die Auswertung dieses Markers nicht robust durchgeführt werden kann. Dies kommt dadurch zustande, dass dieser Marker ein starkes Allel und mehrere schwache Allele aufweist. Durch zu schwache Amplifikation der ohnehin schon schwachen Allele kann keine konsistente Auswertung dieses Markers durchgeführt werden. Aus diesen Gründen und unter Absprache mit dem Auftraggeber, wurde dieser Marker für die Berechnung der paarweisen Distanzen und somit der Zuordnung der Gengruppen ausgeschlossen.

Die Verteilung der paarweisen Distanzen für den gesamten Datensatz ist in Abbildung 21 dargestellt. Gut ersichtlich sind die Unterscheidung von Eigenvergleichen und sehr ähnlichen Genotypen (erster Balken bei einer Distanz von 0-5%) und die Vergleiche zwischen verschiedenen Genotypen (>10% Distanzen). Basierend auf dieser Beobachtung wurde der Distanz Cut-Off für die Gengruppendefinition bei 0.1 bzw. 10% definiert. Damit fallen alle Proben, welche mehr als 90% identische Allele aufweisen, in eine Gengruppe. Die Gengruppenzuordnung mittels Clusteranalyse ist in der Abbildung 22 zu finden.

Der Auftraggeber hat 220 Sortennamen und Sortennummern (20171101_BLE_Erdbeere_Allelreport.xlsx, Nummerierung der Sorten/Spalte E) aufgedeckt.

Da neun Sortennummern eindeutig genetisch verschiedene Proben enthielten, sind diese in 19 Gengruppen unterteilt worden (Siehe Tabelle 2 bzw. 20171101_BLE_Erdbeere_Allelreport.xlsx Blatt "geprüft"/ A9 bzw. Blatt "AlleleReportsortiert", "GenGrou_Eco_Def" /Spalte I).

Die Probe JKI-735_A02 hat ursprünglich eine Ecogenics Gengruppe 187 gebildet (20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "AlleleReportsortiert", "GenGrou_Eco_Def" /J260), jedoch weist diese Probe die identische DGO-Sortennummer (Spalte E) wie die Gengruppe 39 (Spalte I) auf. Nach Überprüfung der Alleldaten wurde eine sehr große Ähnlichkeit dieser Probe zur Gengruppe 39 festgestellt und somit wurde diese Probe schlussendlich in die Gengruppe 39 umgeteilt (Siehe Tabelle 2 bzw. 20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "AlleleReportsortiert", "GenGrou_Eco_Def" / I260). Grund für die Bildung anfänglich eigener Gengruppe der Probe JKI-735_A02 sind zum Beispiel Allelausfälle aufgrund von PCR Inhibition. Eine Erhöhung des Distanz Cut-Off auf beispielsweise 20% würde solche Unsicherheiten abfangen.

Eine Zusammenfassung findet man in der nachfolgenden Tabelle 2 (und in 20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "geprüft"/A9):

Tabelle 2: Folgende Sortennummern umfassen genetisch unterscheidbare Gengruppen.

Nummerierung der Sorten (DGO, Spalte E)	GenGrou_Eco_Def (Ecogenics Spalte I)	Bemerkung
44	23/24	alle JKI-Proben mit der Nr. 44 sind Polka (GenGrou 23) und nicht Elsanta (GenGrou 24)
51	29 / 63	alle BSA-Proben sind Fiesta (GenGrou 29), alle JKI-Proben sind Festiva (GenGrou 63)
108	30/120	alle JKI-Proben mit der Nr. 108 gehören in die Gengruppe 120
116	20/82	alle BSA-Proben gehören zur Gengruppe 20, alle JKI-Proben zur Gengruppe 82
128	31/32/33	BSA-147 ist Sorte Jumbo (GenGrou 32) und nicht Merkur (GenGrou 31), BSA-148 teilt Allele von beiden Sorten (GenGrou 33)
146	28/143	JKI-457 ist Sorte Orion (GenGrou 28) und nicht Pegasus (GenGrou 143)
166	40/189	die BSA-Probe gehört zur Gengruppe 40, die beiden JKI-Proben zur Gengruppe 189
172	154/155	JKI-533 ist Sorte Silvia (GenGrou 154) und nicht Rupine (GenGrou 155)
195	25/167	alle BSA-Proben gehören zur Gengruppe 25, alle JKI-Proben zur Gengruppe 167

69 Sortennummern wurden in 30 Gengruppen zusammengefasst, da diese genetisch identische Proben enthielten (Tabelle 3). Von drei Probensets war jeweils eine Probe einer anderen Gengruppe zuzuordnen (Tabelle 4). Eine Zusammenfassung ist in der nachfolgenden Tabelle 3 und Tabelle 4 (und in 20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "geprüft"/A28) einzusehen:

Tabelle 3: Folgende Sortennummern umfassen genetisch identische Gengruppen.

Nummerierung der Sorten (DGO, Spalte E)	GenGrou_Eco_Def (Ecogenics Spalte I)	Bemerkung
28/137	6	identischer Genotyp
1/56/64/186	8	Sorte 56/64/186 identischer Genotyp, Sorte 1 ist fast identisch, Sorte 1 teilt Allele an jedem der Loci, 6 Allele an 4 Loci sind entweder an- oder abwesend
142/209	12	identischer Genotyp
91/157	16	identischer Genotyp
117/131	18	identischer Genotyp
44/151	23	identischer Genotyp
42/159	26	identischer Genotyp
22/26	27	identischer Genotyp
2/212	39	identischer Genotyp
3/5/87	42	identischer Genotyp
6/139	43	identischer Genotyp
8/112	45	identischer Genotyp
13/38	49	Sorte 13 fast identisch mit Sorte 38, an jedem Loci werden Allele geteilt, 5 Allele an 5 Loci sind an- oder abwesend
15/125	51	identischer Genotyp
19/80	53	identischer Genotyp
23/66/78/109/126/173	54	identischer Genotyp
24/47	55	identischer Genotyp

35/152	58	identischer Genotyp
25/124	83	identischer Genotyp
20/218	94	identischer Genotyp
21/67	95	identischer Genotyp
53/188	105	identischer Genotyp
60/61	110	identischer Genotyp
63/69	111	identischer Genotyp
70/147	113	Sorte 70 fast identisch mit Sorte 147, an jedem Loci werden Allele geteilt, 4 Allele an 4 Loci sind an- oder abwesend
79/81/85/108	120	identischer Genotyp
167/168	150	identischer Genotyp
183/194	162	identischer Genotyp
211/215	179	identischer Genotyp
104/105	184	identischer Genotyp

Tabelle 4: Jeweils eine Probe aus einer Sortennummerngruppe musste in eine andere Gengruppe eingeteilt werden.

Nummerierung der Sorten (DGO, Spalte E)	GenGrou_Eco_Def (Ecogenics Spalte I)	Bemerkung
140/146	28	JKI-457 ist Sorte Orion (140) und nicht Pegasus (146)
100/128	32	BSA-147 ist Sorte Jumbo (100) und nicht Merkur (128)
172/184	154	JKI-533 ist Sorte Silvia (184) und nicht Rupine (172)

Durch minimale aber klare Unterschiede im Genprofil wurden drei Sortennummern 8, 49, 113 in Untergruppen A und B unterteilt (20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "AlleleReportsortiert", Nummerierung der Sorten/1937-1963). Basierend auf diesen Analysen, wurden final 192 Gengruppen festgestellt, wobei die Gengruppe 187 nicht mehr existiert und die Untergruppen von 8A/B, 49A/B, 113A/B auch als eindeutig genetisch unterscheidbar und somit als eigene genetische Gengruppe berücksichtigt wurden (20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "AlleleReportsortiert", "GenGrou_Eco_Def" /J),.

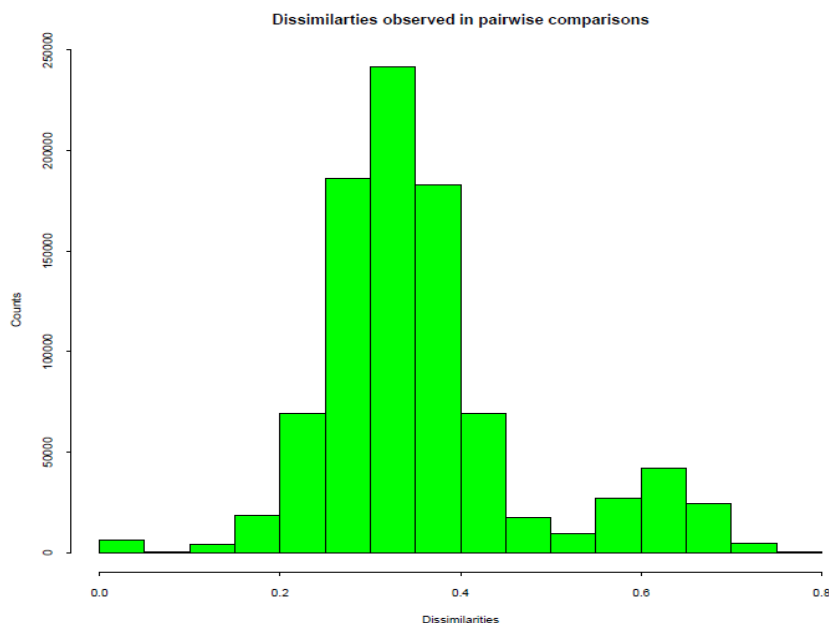


Abbildung 21. Verteilung der paarweisen Distanzen im Erdbeeren Datensatz.

Gene Group Cluster Based on Average Dissimilarities

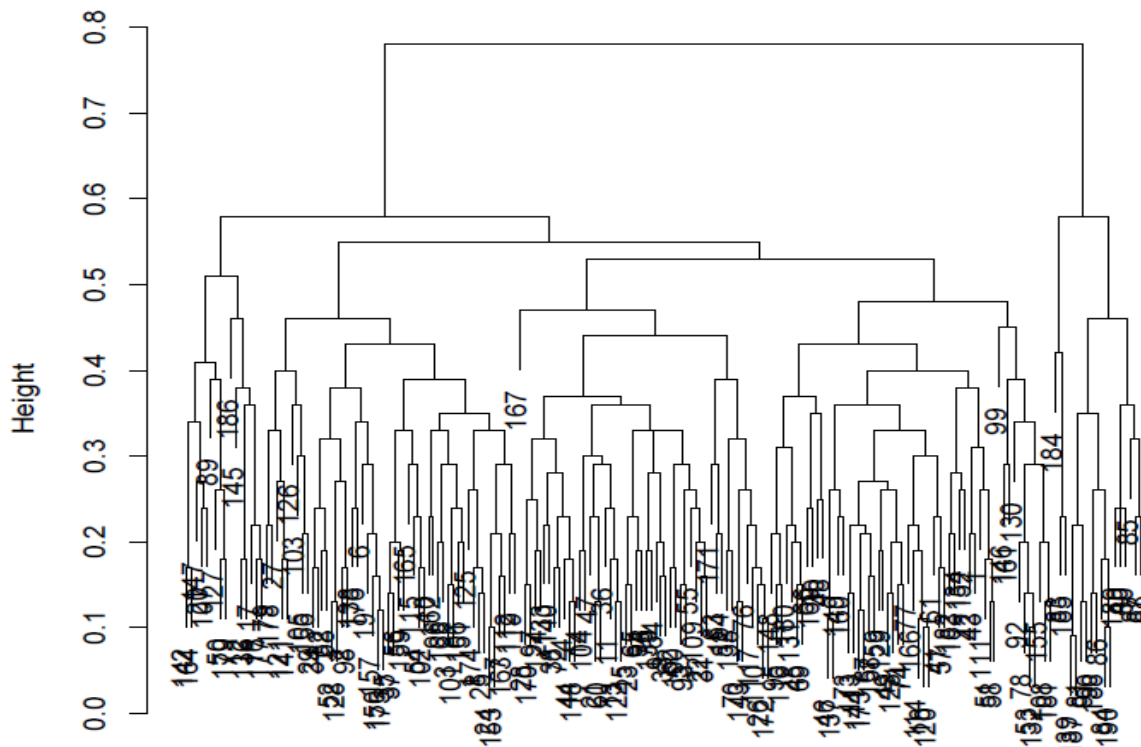


Abbildung 22. Clusteranalyse der definierten Gengruppen basierend auf paarweisen Distanzen. Pro Gengruppe wurde ein Referenz-Genotyp verwendet. Je kürzer der Wert Height (entspricht paarweisen Distanz), desto näher verwandt sind die entsprechenden Gengruppen.

6. NUTZEN UND VERWERTBARKEIT DER ERGEBNISSE

Die im Rahmen dieses Projektes erzeugten Daten von 951 Erdbeerproben und die ermittelten 192 Gengruppen stellen eine sehr gute Grundlage für zukünftige Sortenechtheitsbestimmungen 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 (20171101_BLE_Erdbeere_Allelreport.xlsx, Blatt "Database").

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.

7. ZUSAMMENFASSUNG

951 Blattproben, welche die biologische Vielfalt der Erdbeere in Deutschland repräsentieren, wurden molekulargenetisch analysiert. Die DNA wurde mittels Säulenextraktion aus den Blattproben isoliert. Von jeder Probe wurden 20 bekannte Mikrosatelliten Loci mit fünf 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 unter Ausschluss eines Mikrosatelliten Markers mittels paarweisen Distanzen und Clusteranalyse einer Gengruppe zugeordnet und mit den pomologischen Daten abgeglichen. So konnten 192 Gengruppen definiert werden, welche als Referenzen für zukünftige Sortenechtheitsprüfung verwendet werden können.

Seite | 15

8. GEGENÜBERSTELLUNG DER URSPRÜNGLICH GEPLANTEN ZU DEN TATSÄCHLICH ERREICHTEN ZIELEN; GGF MIT HINWEISEN AUF WEITERFÜHRENDE FRAGESTELLUNGEN

Die ursprünglich geplanten, vertraglich festgehaltenen Ziele wurden weitgehend erreicht. In Abweichung von der ursprünglich gefassten Zielsetzung wurde ein leicht adaptiertes Set von Mikrosatelliten Markern verwendet. Dies geschah im Interesse der Robustheit und in Rücksprache mit dem Auftraggeber.

9. 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.
- 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.
- Horvath, A. , Sánchez-Sevilla, J. , Punelli, F. , Richard, L. , Sesmero-Carrasco, R. , Leone, A. , Hoefler, M. , Chartier, P. , Balsemin, E. , Barreneche, T. and Denoyes, B. (2011), Structured diversity in octoploid strawberry cultivars: importance of the old European germplasm. *Annals of Applied Biology*, 159: 358-371. doi:10.1111/j.1744-7348.2011.00503.x
- 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.

10. APPENDIX

Tabelle 5: Die genauen Mikrosatelliten Primersequenzen der 23 Mikrosatelliten. Am 5' Ende der Reverse Primersequenzen wurde eine Pigtail Sequenz GTTTCTT angebracht. Diese Sequenz sorgt für eine robustere Amplifikation während der PCR und gewährleistet somit eine sauberere Auswertung.

Markername	F Primer Sequenz	Pig Tailed R Primer Sequenz	Ausschluss Multiplex-etablierung	Ausschluss Gengruppen-zuordnung
BFACT037	CGTCCACGATCACATACACAC	<u>GTTTCTT</u> GAAAGGGAGTTGAGCGAAAA	X	
CFVCT016	CACAACGGAGTACACGGCTA	<u>GTTTCTT</u> TTCAGATTCACCTCGGACCTC		X
CFVCT028	GGGAAGAGAGGCCTAAAACC	<u>GTTTCTT</u> CGGCGTCTCAACTTGACC		
ChFaM036	GCAAGCCTCAAGAAGTGAAGG	<u>GTTTCTT</u> CCATCTTGATATCACAGGCATA		
ChFaM078	CAGCCTCATTGCAAATCTGA	<u>GTTTCTT</u> CTTACCGGTTTCGATGTGGT		
ChFaM138	GGCAGTAACCTCCAGAACTCTAA	<u>GTTTCTT</u> TTCGCATTTCCAGCCTTATCT		
EMFv1	GGAGTCCGCCTGCATCTTCTTCA	<u>GTTTCTT</u> CGCGGTCGAGGTGCTAATGCTGTA		
EMFv017	CCGACGAAACCAAGCACCTCCTAC	<u>GTTTCTT</u> TGCATGATCAGCTACGACCTCCTC		
EMFv030	TTCCGGAACAACAGCAACAAAG	<u>GTTTCTT</u> GCGGCGGCACCATCTCG		
MCAD_FAC_001	CCCTCTTCCCCTAAAGTCCA	<u>GTTTCTT</u> CGGACTCACCGGTCACTAGA		
SF1-A01	GGGCAGCAACAAACCAAG	<u>GTTTCTT</u> TAGGATGAACCACACTCTGAA		
SF1-E10	GCGTCTCTCTGTTCTTGCTTTC	<u>GTTTCTT</u> GAGAATTGGTTTCTGGTTCTGG	X	
UAFv7344	TCCTTTGTTTATTTGTATTGTTT	<u>GTTTCTT</u> ATGATTGAAGTGGTGAAGATG		
UAFv8316	CGGTAAACCAGATTACAACCTCTC	<u>GTTTCTT</u> GATCGAGCCCTACCAATTCA		
UDF-001	ATCGAAGCACATGAAGAACG	<u>GTTTCTT</u> TGGGTGGTACTGGTGAGTA		
UDF-004	GCTTGCATTTCAATAGCTGGA	<u>GTTTCTT</u> TTTACTGATGCAGGAGTAGAATGA		
UDF-009	CCTAGAGGAAAACACTGATGACTGA	<u>GTTTCTT</u> AAGGCGAATGCTTTGGTATG		
UDF-018	CGGATCTAAGGCATCTTTTGG	<u>GTTTCTT</u> TTGCACTCCTGCTTTACCTG	X	
UFFa_01D03	TACTGAAATGGGTTTCAGAGC	<u>GTTTCTT</u> GACAGCACAGTCATGGAAGATG		
UFFa_09E12	CGAGGAAGTAACCTCACAGAAA	<u>GTTTCTT</u> GGTGATGGAGAGTGCTGTTAGA		
UFFa_11G07	TCTCTGTGTCTTCCGAAACT	<u>GTTTCTT</u> CTACTGCTCCAACCTCAAATCG		
UFFa_13C07	GGAGTCAACAGTAGTGCAGGTAA	<u>GTTTCTT</u> GGTTTTCTTGCAGTTGGAGTAG		
UFFa_03C04	CGGTTCAAGCAGGAGAATAAAAC	<u>GTTTCTT</u> GCCCCATACTACCATTATGACC		

