



BUNDESMINISTERIUM
FÜR GESUNDHEIT

Identifizierung genetischer Risikofaktoren zur Früherkennung und Prognose des Prostatakarzinoms mittels SNP Microarrays

Abschlussbericht



Impressum

Herausgeber, Medieninhaber und Hersteller:

Bundesministerium für Gesundheit, Sektion II
Radetzkystraße 2, A-1030 Wien

Für den Inhalt verantwortlich:

Dr. Ulrich Herzog

Autorin:

Ao. Prof. Dr. Andrea Gsur
Institute of Cancer Research
Department of Medicine I
Medical University of Vienna
Borschkegasse 8a, A-1090 Vienna

+431-40 160-57541

andrea.gsur@meduniwien.ac.at

Erscheinungstermin: September 2012

Inhaltsverzeichnis

Summary.....	3
Zusammenfassung.....	4
1. Einleitung.....	6
2. Ziel der Studie.....	7
3. Material und Methoden	
3.1. Probenselektion und Qualitätskontrolle.....	8
3.2. Statistische Auswertung I.....	9
3.3. KORA Daten.....	9
3.4. Statistische Auswertung II.....	9
4. Ergebnisse	
4.1 Probenselektion und Qualitätskontrolle.....	10
4.2. Statistische Auswertung I.....	10
4.2.1. Qualitätskontrolle.....	11
4.2.2. GWA Analyse.....	18
5. Diskussion.....	22
6. Literatur.....	26

SUMMARY

The high prevalence of prostate cancer and the high incidence/mortality ratio suggesting overdiagnosis and overtreatment explain the interest in identifying genetic markers for not only prostate cancer risk but also prostate cancer aggressiveness. The identification of diagnostic biomarkers indicating prostate cancer aggressiveness at the time of diagnosis is urgently needed to tailor the therapeutic approach and to limit treatment to those patients who require it.

In contrast to colorectal cancer and breast cancer, susceptibility genes with high penetrance account, if at all, only for a small fraction of the observed genetic predisposition to prostate cancer. Rather a number of genetic variants with low penetrance will influence prostate cancer risk. Several genome-wide association studies (GWAS) have identified more than 40 prostate cancer susceptibility loci so far. To identify new and to confirm known prostate cancer susceptibility loci, we have performed a GWAS from our DNA bank using the genome-wide human SNP array 6.0 (Affymetrix). This SNP array contains single nucleotide polymorphisms (SNPs) as well as probes for the detection of copy number variations (CNV). We have screened the DNA of 184 patients with non-aggressive prostate cancer, 184 aggressive prostate cancer cases and 82 benign prostatic hyperplasia (BPH) controls. To enhance the statistical power, we have screened further 200 SNP arrays (100 patients with Gleason Score <7 and 100 patients with Gleason Score \geq 7) that was funded from another project. Furthermore we have used the SNP 6.0 data from 884 men of the „Kooperative Gesundheitsforschung in der Region Augsburg“ (KORA) study as further control group, so the participants were enhanced to 1.334.

Due to the relatively low number in SNP arrays we could not identify risk variants on a genome-wide level of significance. Currently we are under way to analyze the combined data of our prostate patients and controls with the KORA controls. Because of the increase in the number of participants, some SNPs may reach significance. Furthermore several logistic regression models will be calculated for all SNPs where associations with prostate cancer risk have been reported in the literature so far, in regard of prostate cancer risk as well as in progression. Finally

these results will be published in a meta-analysis in cooperation with an international consortium.

ZUSAMMENFASSUNG

Aufgrund der hohen Prävalenz des Prostatakarzinoms und des hohen Inzidenz/Mortalität-Verhältnisses, welches auf eine Überbehandlung des Prostatakarzinoms schließen lässt, besteht der Bedarf genetische Marker nicht nur für das Prostatakarzinom-Risiko sondern auch die aggressive Form des Prostatakarzinoms zu identifizieren. Die Identifizierung von Biomarkern, die zum Zeitpunkt der Diagnose die Differenzierung ermöglichen, welche Männer eine aggressive Form des Prostatakarzinoms entwickeln werden, ist eine dringliche Notwendigkeit, um nur jene den notwendigen Therapien zu unterziehen, wo es auch indiziert ist.

Im Gegensatz zu anderen Krebsarten, wie kolorektalem Karzinom und Mammakarzinom, dürften Suszeptibilitätsgene mit hoher Penetranz für das Prostatakarzinom nur eine untergeordnete Rolle spielen, sondern vielmehr eine Kombination mehrerer genetischer Varianten mit niedriger Penetranz. Durch Genomweite Assoziationsstudien (GWAS) wurden bisher mehr als 40 Suszeptibilitätstoci für das Prostatakarzinom identifiziert.

Wir haben im ersten Schritt unserer Studie die DNA von 184 Patienten mit nicht-aggressivem Prostatakarzinom, 184 aggressive Prostatakarzinomen und 82 benigne Prostatahyperplasien (BPH) aus unserer ca. 2.000 Proben umfassenden DNA Bank mit dem „Human genome-wide SNP 6.0 array“ (Affymetrix) der sowohl „Single nucleotide polymorphisms“ (SNPs) als auch Sonden zur Bestimmung von „Copy number variations“ (CNV) enthält, gescreent. Um die statistische Power zu erhöhen, haben wir weiteren 200 SNP Arrays (100 Patienten mit Gleason Score <7 und 100 Patienten mit Gleason Score ≥7), die aus einem anderem Projekt finanziert wurden, eingesetzt. Unter Verwendung von 884 Kontrollen aus der Kooperative Gesundheitsforschung in der Region Augsburg (KORA) konnte die Zahl der Teilnehmer in dieser Studie auf 1.334 gesteigert werden.

Aufgrund der relativ geringen Anzahl der in diesem Projekt eingesetztem SNP Arrays konnten keine Risikovarianten auf genomweiten Signifikanzniveau identifiziert

werden. Derzeit laufen die Berechnungen der Arraydaten aus unserer Prostatakarzinom DNA-Bank gemeinsam mit den KORA Kontroll-Daten. Durch die weitere Erhöhung der Fallzahl und die Anwendung eines neuen Algorithmus besteht die Möglichkeit, dass dadurch noch SNPs an Signifikanz gewinnen. Von den in der Literatur bereits mit Prostatakarzinomrisiko assoziierten SNPs werden logistische Regressionsmodelle gerechnet. Der Einfluss dieser SNPs soll nicht nur im Hinblick auf das Prostatakarzinomrisiko, sondern auch als Prognosefaktor bewertet werden. Es ist schlussendlich auch noch geplant unsere Ergebnisse in Kooperation mit einem internationalen Prostatakarzinom-Konsortium in einer Metaanalyse zu publizieren.

1. EINLEITUNG

In Österreich ist das Prostatakarzinom die häufigste Krebserkrankung des Mannes und nach dem Lungenkarzinom die zweithäufigste karzinombedingte Todesursache. Die Inzidenz des Prostatakarzinoms ist innerhalb der letzten Jahre in Österreich, wie auch in anderen Industrieländern, deutlich angestiegen, einerseits durch das intensive Prostata-spezifische Antigen (PSA) Screening und durch die höhere Lebenserwartung. Die Inzidenzrate des Prostatakarzinoms ist allerdings in den letzten 3-5 Jahren wieder rückläufig, bedingt durch das routinemäßige PSA Screening und der dadurch vermehrten Diagnosen in Frühstadien innerhalb der letzten Jahren, die Mortalitätsrate ist allerdings weitgehend unverändert geblieben. Die Inzidenz / Mortalitätsrate des Prostatakarzinoms liegt im Bereich von 6-8:1, diese Verhältnis lässt auf eine Überdiagnose schließen. Das PSA Screening hat nicht nur zu einer vermehrten Diagnose des Prostatakarzinoms geführt, sondern auch zu einer Überbehandlung von indolenten Formen. Für diese Patienten wäre ein abwartendes Beobachten sinnvoller anstatt einer Behandlung. Allerdings stehen derzeit keine Biomarker zur Verfügung, die es erlauben diejenigen Männer, die eine aggressive Form des Prostatakarzinoms entwickeln werden zum Zeitpunkt der Diagnose zu identifizieren.

Daher besteht ein dringender Bedarf an verlässlichen genetischen Marker, neben den bekannten klinischen Parametern wie PSA und Histologie, die den weiteren klinischen Verlauf zum Zeitpunkt der Diagnose voraussagen. Obwohl das Prostatakarzinom meist ein relativ langsam wachsender Tumor ist, existieren auch sehr aggressive Formen welche rasch zur Progression und letztendlich zum Tod führen können. Es ist weitgehend ungeklärt welche molekularen Mechanismen für die Progression des Prostatakarzinoms bis hin zur letalen Form verantwortlich sind.

Die einzigen bisher bekannten Risikofaktoren für die Entwicklung eines Prostatakarzinom sind hohes Alter, ethnische Zugehörigkeit – mit den höchsten Inzidenzraten bei Afroamerikanern – und familiäre Prädisposition. Zwillingsstudien, molekular-epidemiologische Studien und epidemiologische Beobachtungen lassen

einen wesentlichen genetischen Einfluss bei der Entstehung des Prostatakarzinoms vermuten. Im Gegensatz zu anderen Krebsarten, wie kolorektalem Karzinom und Mammakarzinom, konnten noch keine Suszeptibilitätsgene für das Prostatakarzinom mit hoher Penetranz identifiziert werden. Es wird angenommen, dass nur ein kleiner Prozentsatz der genetischen Suszeptibilität des Prostatakarzinoms durch Gene mit hoher Penetranz verursacht wird, sondern vielmehr durch eine Kombination mehrerer genetischer Varianten mit niedriger Penetranz, die zwar nur ein geringes Risiko bewirken, allerdings impliziert deren hohe Frequenz in der Population, dass sie einen beachtenswerten Impact in der Prostatakarzinominzidenz haben könnten.

In den letzten Jahren wurden durch „Genome-wide association studies“ (GWAS) ca. 40 Suszeptibilitätsloci, die mit Prostatakarzinomrisiko assoziiert sind, identifiziert (Abb 1). Die wichtigste Prostatakarzinom Suszeptibilitätsregion ist 8q24, die zuerst durch Linkage Analysen und später durch GWAS identifiziert wurde. Es wurden zumindest fünf benachbarte SNPs in dieser Region identifiziert. Der SNP rs6983267 ist auch mit kolorektalem Karzinom und Ovarialkarzinom, die anderen SNPs dürften nur mit Prostatakarzinomrisiko assoziiert sein. Zwei weitere wichtige Prostatakarzinom Loci befinden sich auch auf 17q und 10q (Abb. 1). Die Risikovorhersage für das Prostatakarzinom kann durch die Identifizierung weiterer genetischer Varianten verbessert werden.

2. ZIEL DER STUDIE

Ziel dieses Forschungsprojektes ist die Identifizierung weiterer und Bestätigung der bisher bekannten Suszeptibilitätsloci für erhöhtes Prostatakarzinomrisiko. Weiters sollen auch genetische Marker identifiziert werden, die die Differenzierung in eine latente oder aggressivere Form der Erkrankung zum Zeitpunkt der Diagnose ermöglichen. Dadurch können Patienten mit einer latenten Form des Prostatakarzinoms belastende Behandlungen erspart werden.

Dazu haben wir die DNA von 184 Patienten mit nicht-aggressivem Prostatakarzinom, 184 aggressive Prostatakarzinomen und 82 benigne Prostatahyperplasien (BPH) aus unserer ca. 2.000 Proben umfassenden DNA Bank mit dem „Genome-wide human SNP array 6.0“ (Affymetrix) der sowohl „Single nucleotide polymorphisms“ (SNPs) als

auch Sonden zur Bestimmung von „Copy number variations“ (CNVs) enthält, gescreent.

Unter Verwendung von SNP 6.0 Array Daten von 884 Kontrollen aus der Kooperative Gesundheitsforschung in der Region Augsburg (KORA) konnte die Zahl der Teilnehmer in dieser Studie auf 1.334 gesteigert werden.

3. MATERIAL UND METHODEN

3.1. Probenselektion und Qualitätskontrolle

Im ersten Schritt wurden insgesamt 250 DNA-Proben, bestehend aus 84 Prostatakarzinompatienten mit einem Gleason Score größer 7, 84 Proben mit Gleason Score kleiner 7 sowie 82 BPH als Kontrollgruppe selektiert.

Für die weiteren 200 DNA-Proben, die aus einem anderem Projekt finanziert wurden, wurden nur Patienten ausgewählt, davon 100 Patienten mit Gleason Score <7 und 100 Patienten mit Gleason Score ≥ 7 . Die Auswahlkriterien waren: eine ähnliche Altersverteilung in allen drei Gruppen, sowie die Vollständigkeit wichtiger Confounding-Daten (Angaben zum Rauchverhalten, Prostata-spezifisches Antigen) in den ausgewählten Datensätzen. Ferner wurde die DNA-Qualität aller Proben in 0,8%igen Agarosegelen überprüft.

Mittels des SNP 6.0 Arrays von Affymetrix können über 1,8 Millionen genetische Varianten einer Person bestimmt werden, davon sind 906.600 SNPs, welche die häufigste Sequenzvariation im menschlichen Genom bilden und sich wegen ihrer Häufigkeit und Stabilität besonders als Marker für genomweite Assoziationsstudien eignen. Zudem trägt dieser Array 946.000 Sonden zur Bestimmung von CNVs, wodurch genetische Veränderungen wie Duplikationen, Deletionen oder Inversionen nachgewiesen werden können.

Die genomische DNA von 450 Patienten und Kontrollen wurde mit „Human genome-wide SNP 6.0“ Arrays hybridisiert. Die Aufzeichnung der Rohdaten der prozessierten Arrays erfolgte durch den GeneChip Scanner 3000 7G (Scanning), als Ergebnis lagen die Daten in *.CEL-Files vor. Die Arraydaten wurden mit der Affymetrix

Genotyping Console Software analysiert (primäre Datenanalyse, QC-Maßnahmen, Normalisierungsverfahren).

3.2 Statistische Auswertung I (Andreas Baierl, Florian Frommlet)

Die sekundäre Datenanalyse wurde von unserem Biostatistiker Dr. Andreas Baierl mit dem R package „GenABEL“ durchgeführt. Dabei handelt es sich um ein „Paket“ zu dem Statistikprogramm „R“, welches speziell zur Analyse von GWAS-Daten entwickelt wurde. Die signifikantesten Ergebnisse (niedrigsten P-Werte) lagen dabei etwas unterhalb des allgemein anerkannten genomweiten Signifikanzniveaus für GWASs. Die Daten wurden von einem weiteren Biostatistiker (DI Dr. Florian Frommlet, Medizinische Universität Wien), einem ausgewiesenen Spezialisten für GWAS-Analyse mittels eines neu entwickelten Verfahrens zur Auswertung von SNP-Arraydaten analysiert. Dies führte zu einem übereinstimmenden Ergebnis.

3.3 KORA Daten

Um die statistische Power der Studie zu erhöhen wurde vom Helmholtzcenter München die Verwendung von Kontrolldaten aus der KORA-Studie beantragt. Zu unseren 450 Patienten konnten so zusätzlich SNP 6.0-Daten von 884 gesunden Männern erhalten und die Gesamtzahl an Individuen in der Studie somit auf 1.334 erhöht werden.

3.4 Statistische Auswertung II (Florian Frommlet)

Daten von 450 Personen unserer Studie „Molekulare Epidemiologie des Prostatakarzinoms“ wurden mit den Referenzdaten von 884 Patienten der KORA-Studie zusammengeführt. Aufgrund von unterschiedlichen Affymetrix-Annotationsfiles kam es bei ca. 13.000 SNPs zu Konflikten bezüglich der Benennung der jeweiligen SNP Allele. Für eine erste Analyse wurden all diese problematischen SNPs nicht berücksichtigt. Derzeit wird ein gemeinsamer Genotyp Call für alle Individuen (unsere Studie + KORA) durchgeführt, der dann die endgültige Analyse aller SNPs

ermöglichen wird. Allerdings wird vor Ablauf der Deadline für diesen Endbericht der gemeinsame Genotyp Call noch nicht abgeschlossen sein.

Mittels PLINK wurden alle verbleibenden SNPs auf Hardy-Weinberg Gleichgewicht untersucht, und jene SNPs, für welche der entsprechende χ^2 -Test signifikant zum Niveau $\alpha = 0.01$ waren wurden ebenfalls für die weitere Analyse nicht berücksichtigt.

Alle verbleibenden SNPs wurden anschließend mittels Cochran Armitage Trend Test auf Assoziation zum Prostatakarzinomrisiko hin getestet. Zusätzlich zu den standardmäßigen Single-Marker Tests wird eine Multilocus-Analyse durchgeführt, wobei Ergebnisse der drei Programme MOSGWA, GWASELECT sowie HYPERLASSO miteinander verglichen werden. Im Wesentlichen basieren alle drei Methoden auf Modellselektion für logistische Regressionsmodelle. MOSGWA verwendet zur Bestimmung des besten Modells spezielle Modifikationen des „Bayesian information criterion“ (BIC), GWASELECT beruht auf der Verwendung von LASSO, während HYPERLASSO einen weitgehend Bayesianischen Ansatz zur Modellselektion verwendet.

4. ERGEBNISSE

4.1 Probenselektion und Qualitätskontrolle

Alle 450 DNAs haben die Qualitätskontrolle bestanden und konnten für die Hybridisierung eingesetzt werden. Die Prozessierung war in 446 von 450 Fällen (99,1%) erfolgreich.

4.2 Statistische Auswertung I (Andreas Baierl)

Die Analyse wurde mit dem Programmpaket „GenABEL“ durchgeführt, das in Form eines Erweiterungspakets der frei verfügbaren statistischen Programmiersprache R unter www.genabel.org bereit gestellt wird. GenABEL umfasst Funktionen zur Analyse genomweiter Assoziationsstudien, sowohl für quantitative als auch qualitative (binäre) Merkmale. Die Analyse folgt weitgehend der empfohlenen

Vorgangsweise im Programmhandbuch (<http://www.genabel.org/sites/default/files/pdfs/ABEL-tutorial.pdf>).

Im ersten Schritt wurden die SNP-Daten in ein spezifisches Dateiformat umgewandelt und eingelesen. Die Phenotyp-Daten (Fall/Kontrolle, Alter, etc.) wurden getrennt eingelesen.

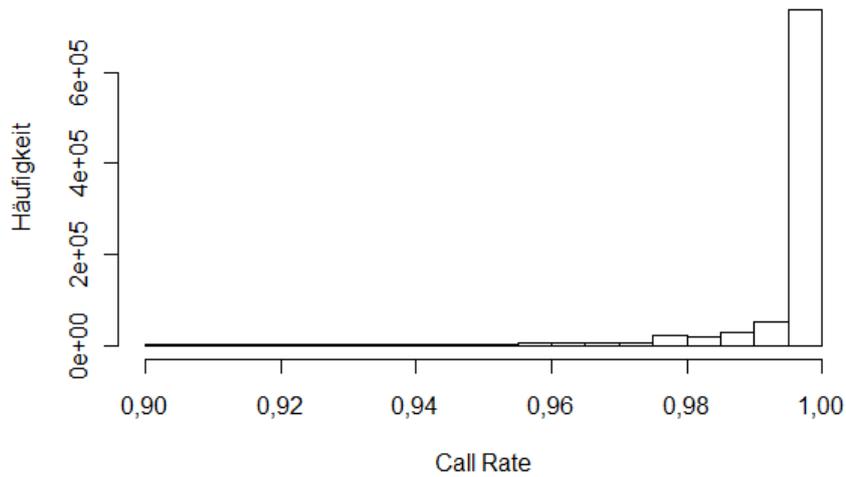
4.2.1 Qualitätskontrolle

Analyse der Call Rate

97% aller SNPs konnten mit einer Call Rate >95% zugeordnet werden.

Call Rate	Anzahl SNPs	in %
kleiner 85%	6969	1%
85-90%	5103	1%
90-95%	14725	2%
95-99%	90878	10%
größer 99%	788007	87%

Histogramm für Call Rate



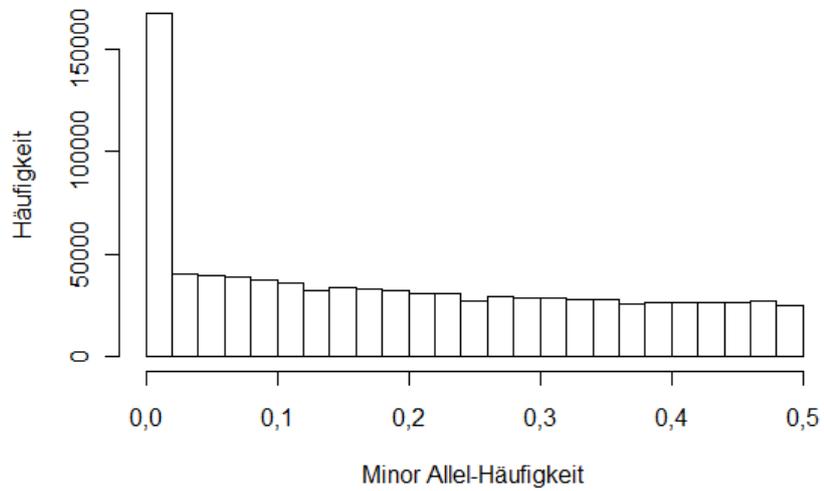
Hardy-Weinberg Gleichgewicht:

p-Wert	gesamtes Sample		nur Kontrollen	
	Anzahl SNPs	in %	Anzahl SNPs	in %
kleiner $5 \cdot 10^{-8}$	27608	3,0%	24423	2,7%
$5 \cdot 10^{-8}$ bis 0,01	10219	1,1%	33248	1,0%
0,01 bis 0,05	23422	2,6%	51796	2,0%
0,05 bis 0,1	30478	3,4%	76419	2,7%
größer 0,1	813955	90%	905682	91,6%

*) $5 \cdot 10^{-8}$ ist der Bonferroni- p-Wert, unter dem von einer Abweichung vom HWE ausgegangen werden kann. Standardmäßig wird für die Analyse des HWE nur die Kontrollgruppe herangezogen. In dieser weichen 2,7% der SNPs signifikant vom HWE ab.

Minor Allele-Häufigkeit

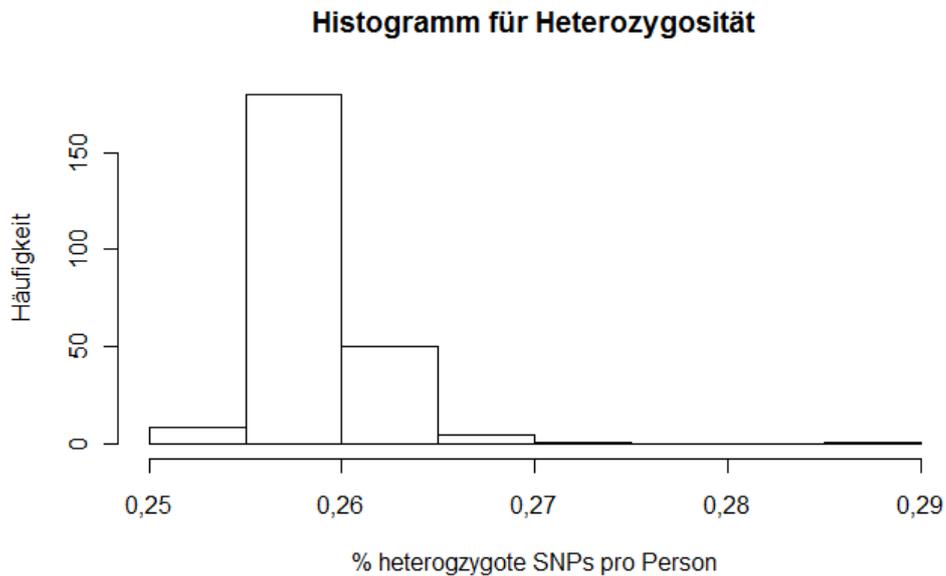
Histogramm der Allel-Häufigkeiten



Allel-Häufigkeit	Anzahl SNPs	in %
0%	81468	9%
0-1%	63491	7%
1-5%	82057	9%
5-10%	96027	11%
10-20%	167166	19%
20-30%	146660	16%
30-40%	136667	15%
40-50%	132146	15%

Heterozygotität

Würde sich der Prozentsatz der heterozygoten SNPs stark von Person zu Person unterscheiden, deutet dies auf Datenfehler/kontaminierte Proben hin.



Bereinigung der Daten

Die Berücksichtigung der Qualitätskriterien findet nicht einzeln statt sondern in Form eines iterativen Verfahrens. SNPs bzw. Beobachtungen werden auf Basis der folgenden Kriterien und dazugehöriger Schwellenwerte ausgeschieden:

- Minor Allel Häufigkeit (Schwellenwert: <1%)
- Call Rate (Schwellenwert: <95%)
- HWE (Schwellenwert: < $5 \cdot 10^{-8}$ = Bonferroni korrigiertes Signifikanzniveau)
- Heterozygotität (Schwellenwert: FDR < 1%)
- Genetische Ähnlichkeit (Schwellenwert: IBS: >0,95)

Ablauf der iterativen Analyse:

RUN 1

905404 markers and 246 people in total

144742 (15,98645%) markers excluded as having low (<1%) minor allele frequency

26519 (2,928969%) markers excluded because of low (<95%) call rate

27607 (3,049136%) markers excluded because they are out of HWE ($P < 5e-08$)

0 (0%) people excluded because of low (<95%) call rate

Mean autosomal HET is 0,3188353 (s.e. 0,003135598)

1 (0,4065041%) people excluded because too high autosomal heterozygosity (FDR <1%)

Excluded people had HET \geq 0,3457924

Mean IBS is 0,7426607 (s.e. 0,00693044), as based on 2000 autosomal markers

1 (0,4065041%) people excluded because of too high IBS ($\geq 0,95$)

In total, 717698 (79,26826%) markers passed all criteria

In total, 244 (99,18699%) people passed all criteria

RUN 2

717698 markers and 244 people in total

432 (0,06019245%) markers excluded as having low (<1%) minor allele frequency

0 (0%) markers excluded because of low (<95%) call rate

1 (0,0001393344%) markers excluded because they are out of HWE ($P < 5e-08$)

0 (0%) people excluded because of low (<95%) call rate

Mean autosomal HET is 0,3189208 (s.e. 0,002620902)

0 people excluded because too high autosomal heterozygosity (FDR <1%)

Mean IBS is 0,7462475 (s.e. 0,006525966), as based on 2000 autosomal markers

0 (0%) people excluded because of too high IBS ($\geq 0,95$)

In total, 717265 (99,93967%) markers passed all criteria

In total, 244 (100%) people passed all criteria

RUN 3

717265 markers and 244 people in total

0 (0%) markers excluded as having low (<1%) minor allele frequency

0 (0%) markers excluded because of low (<95%) call rate

0 (0%) markers excluded because they are out of HWE ($P < 5e-08$)

0 (0%) people excluded because of low (<95%) call rate

Mean autosomal HET is 0,3189208 (s.e. 0,002620902)

0 people excluded because too high autosomal heterozygosity (FDR <1%)

Mean IBS is 0,7385854 (s.e. 0,006801173), as based on 2000 autosomal markers

0 (0%) people excluded because of too high IBS ($\geq 0,95$)

In total, 717265 (100%) markers passed all criteria

In total, 244 (100%) people passed all criteria

Ergebnis:

`Per-SNP fails statistics`

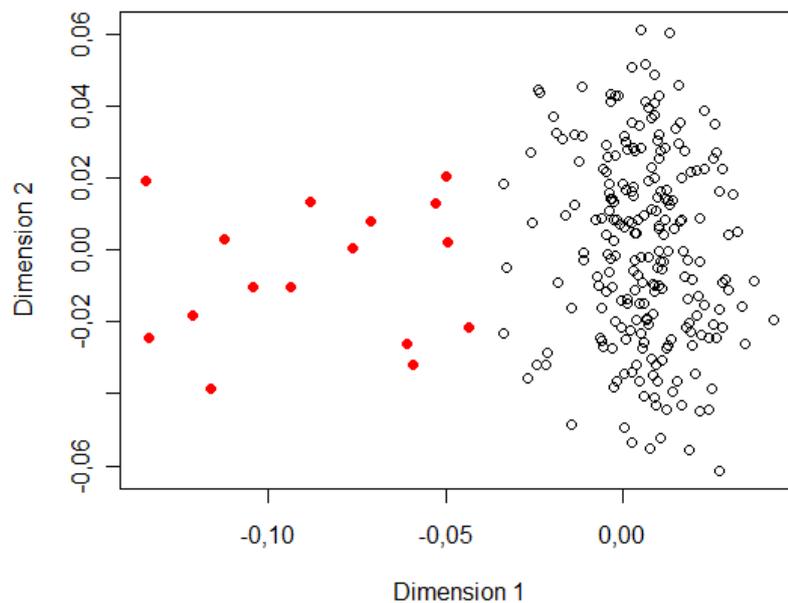
	NoCall	NoMAF	NoHWE	Redundant	Xsnpfail
NoCall	15635	10738	424	0	0
NoMAF	NA	134436	0	0	0
NoHWE	NA	NA	27184	0	0
Redundant	NA	NA	NA	0	0
Xsnpfail	NA	NA	NA	NA	0

`Per-person fails statistics`

	IDnoCall	HetFail	IBSFail	isfemale	ismale	isXXY	otherSexErr
IDnoCall	0	0	0	0	0	0	0
HetFail	NA	1	0	0	0	0	0
IBSFail	NA	NA	1	0	0	0	0
isfemale	NA	NA	NA	0	0	0	0
ismale	NA	NA	NA	NA	0	0	0
isXXY	NA	NA	NA	NA	NA	0	0
otherSexErr	NA	NA	NA	NA	NA	NA	0

Im nächsten Schritt, der 2. Runde der Qualitätskontrolle, wurde untersucht, ob genetische Ausreißer, die nicht zur Population passen, das Ergebnis kontaminieren. Es wurde eine genetische Ähnlichkeitsmatrix berechnet, die mittels Multidimensionaler Skalierung (MDS) in 2 Dimensionen dargestellt werden kann. Jeder Punkt entspricht dabei einem Individuum, die roten Punkte wurden als genetisch unähnlich identifiziert.

Multidimensionale Skalierung



4.2.2 GWA Analyse

In den folgenden Tabellen sind die jeweils signifikantesten SNPs aufgelistet.

Beschreibung der Tabellen:

N: Anzahl der Beobachtungen

1. Test (additiver Effekt)

- effB: OR des additiven (allel-)Effekts,
- se_effB: Standardfehler des additiven Effekts
- Chi2.1df: dazugehöriger Chi2-Wert,
- P1df: dazugehöriger P-Wert.

2. Test (Test für die einzelnen Genotypen)

- effAB: OR des AA vs. AB Effekts
- effBB: OR des AA vs. BB Effekts
- Chi2.2df: Chi2-Wert für den Gesamteffekt des SNPs
- P2df: P-Wert für den Gesamteffekt des SNPs

Analyse der Rohdaten

Ohne Berücksichtigung von Qualitätschecks (Call Rate, HWE, Heterzygotität, Subgruppenstruktur/genetische Ähnlichkeit) und ohne Kovariate.

Ohne korrigierte P-Werte

	Chromosome	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10051190	5	123999278	+	T	C	246	0,310101	0,06520846	22,615	2,0E-06	0,3220	0,0000	23,210	9,12E-06
rs10450286	10	72986492	+	G	A	245	0,4606136	0,10235184	20,253	6,8E-06	0,3862	0,1728	20,265	3,98E-05
rs4645337	5	3279885	+	G	C	246	0,218503	0,04868742	20,141	7,2E-06	0,2203	0,0000	20,144	4,22E-05
rs4866542	5	3282671	+	T	C	246	0,218503	0,04868742	20,141	7,2E-06	0,2203	0,0000	20,144	4,22E-05
rs8026735	15	62125859	+	C	T	246	3,1349329	0,7073134	19,644	9,3E-06	2,3026	12,5000	19,712	5,24E-05
rs11883643	3	185920603	+	G	T	246	0,2112139	0,04868284	18,823	1,4E-05	0,1440	0,4527	22,163	1,54E-05
rs2268844	6	90826709	+	C	T	246	0,4535691	0,10482468	18,722	1,5E-05	0,4655	0,1727	19,273	6,53E-05
rs207268	11	115363695	+	G	T	246	0,1401197	0,03246002	18,634	1,6E-05	0,1495	NA	18,634	1,58E-05
rs314480	11	115382024	+	C	A	246	0,1401197	0,03246002	18,634	1,6E-05	0,1495	NA	18,634	1,58E-05
rs314468	2	239586885	+	A	G	246	0,4656237	0,10822138	18,512	1,7E-05	0,5178	0,1802	19,458	5,95E-05
rs10458597	5	10511590	+	G	A	246	2,6974121	0,63360697	18,124	2,1E-05	2,6005	5,5070	18,636	8,98E-05
rs12565286	11	95977311	+	C	G	244	0,2102787	0,04950512	18,042	2,2E-05	0,2115	0,0000	18,055	1,20E-04
rs12082473	6	90845036	+	C	T	246	0,4633234	0,1092142	17,997	2,2E-05	0,4227	0,1799	18,100	1,17E-04
rs3094315	1	210858403	+	T	C	246	0,4796931	0,11411381	17,671	2,6E-05	0,4261	0,1953	17,702	1,43E-04
rs2286139	8	134527871	+	A	G	246	0,4592999	0,11004529	17,420	3,0E-05	0,3703	0,2066	17,686	1,44E-04
rs11240776	11	115357372	+	C	A	243	0,148647	0,035694	17,343	3,1E-05	0,1566	NA	17,343	3,12E-05
rs2980319	17	49754473	+	G	A	246	0,1329718	0,03202532	17,240	3,3E-05	0,1422	NA	17,240	3,29E-05
rs2980300	2	33703217	+	C	G	246	4,8912323	1,17884441	17,216	3,3E-05	3,8448	Inf	17,351	1,71E-04
rs2905036	16	85992732	+	A	T	245	0,4590878	0,11107088	17,084	3,6E-05	0,3527	0,2405	17,986	1,24E-04
rs11240777	11	95976740	+	C	T	246	0,229422	0,0558557	16,871	4,0E-05	0,2310	0,0000	16,872	2,17E-04
rs4245756	6	166231573	+	G	A	246	2,7296341	0,66478314	16,860	4,0E-05	2,3671	7,0805	16,910	2,13E-04
rs3748597	10	65537753	+	C	G	246	0,4560668	0,11110376	16,850	4,0E-05	0,3208	0,2472	18,541	9,42E-05
rs2341354	8	74039314	+	C	T	246	0,4752045	0,11578683	16,844	4,1E-05	0,4342	0,2018	16,851	2,19E-04
rs4970403	13	108225102	+	T	C	246	0,2207551	0,05380656	16,833	4,1E-05	0,2218	0,0000	16,838	2,21E-04
rs2465126	11	115352894	+	C	T	243	0,1354964	0,03306425	16,793	4,2E-05	0,1451	NA	16,793	4,17E-05

Mit korrigierten P-Werten (NA in P2df bedeutet, dass nicht alle 3 Genotypen vorhanden sind)

SNP	Chr	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10051190	5	123999278	+	T	C	246	0,310101	0,06520846	22,615	0,505	0,3220	0,000	23,210	NA
rs10450286	10	72986492	+	G	A	245	0,4606136	0,10235184	20,253	0,865	0,3862	0,173	20,265	NA
rs4645337	5	3279885	+	G	C	246	0,218503	0,04868742	20,141	0,890	0,2203	0,000	20,144	NA
rs4866542	5	3282671	+	T	C	246	0,218503	0,04868742	20,141	0,890	0,2203	0,000	20,144	NA
rs8026735	15	62125859	+	C	T	246	3,1349329	0,7073134	19,644	0,960	2,3026	12,500	19,712	NA
rs11883643	2	239586885	+	A	G	246	0,4656237	0,10822138	18,512	0,985	0,5178	0,180	19,458	NA
rs2268844	3	185920603	+	G	T	246	0,2112139	0,04868284	18,823	0,985	0,1440	0,453	22,163	NA
rs207268	6	90826709	+	C	T	246	0,4535691	0,10482468	18,722	0,985	0,4655	0,173	19,273	NA
rs314480	11	115363695	+	G	T	246	0,1401197	0,03246002	18,634	0,985	0,1495	NA	18,634	NA
rs314468	11	115382024	+	C	A	246	0,1401197	0,03246002	18,634	0,985	0,1495	NA	18,634	NA

Analyse der Daten nach der 1. Runde der Qualitätskontrolle

inklusive Alter als Kovariate

Ohne korrigierte P-Werte

	Chromosome	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10051190	5	123999278	+	T	C	244	0,4326831	0,09366247	21,341	3,8E-06	0,4438	0,0709	22,621	1,2E-05
rs11883643	2	239586885	+	A	G	244	0,538006	0,11715119	21,090	4,4E-06	0,5807	0,2485	22,254	1,5E-05
rs10450286	10	72986492	+	G	A	244	0,5502651	0,12021021	20,954	4,7E-06	0,4808	0,2568	20,978	2,8E-05
rs2268844	3	185920603	+	G	T	244	0,3153717	0,06925358	20,738	5,3E-06	0,2476	0,4064	23,358	8,5E-06
rs8026735	15	62125859	+	C	T	244	2,3159659	0,53085176	19,033	1,3E-05	1,8696	5,8862	19,116	7,1E-05
rs10866475	5	10511590	+	G	A	244	2,1559505	0,49709555	18,810	1,4E-05	2,0472	3,7615	19,208	6,7E-05
rs10167210	2	239601726	+	C	A	244	0,5738864	0,13616412	17,763	2,5E-05	0,5973	0,2938	18,338	1,0E-04
rs10806834	6	166231573	+	G	A	244	2,1881851	0,52115062	17,630	2,7E-05	1,9821	4,4230	17,696	1,4E-04
rs2730577	12	119023969	+	A	G	240	2,0870446	0,49858132	17,522	2,8E-05	1,6590	4,0593	17,716	1,4E-04
rs4568464	6	153995311	+	C	T	244	3,3742804	0,80740627	17,465	2,9E-05	3,0478	10,8339	17,708	1,4E-04
rs2431835	19	58647214	+	G	A	244	2,0142753	0,48204837	17,460	2,9E-05	1,8519	3,5849	17,535	1,6E-04
rs4645337	5	3279885	+	G	C	244	0,3658607	0,08766407	17,418	3,0E-05	0,3466	0,1948	17,507	1,6E-04
rs4866542	5	3282671	+	T	C	244	0,3658607	0,08766407	17,418	3,0E-05	0,3466	0,1948	17,507	1,6E-04
rs7594129	2	239600686	+	C	A	243	0,5769088	0,13861608	17,322	3,2E-05	0,6052	0,2977	17,931	1,3E-04
rs2334025	2	239564114	+	C	T	244	0,57171	0,13781774	17,208	3,3E-05	0,6028	0,2900	17,902	1,3E-04
rs17648308	2	33703217	+	C	G	244	3,1474451	0,75950119	17,174	3,4E-05	2,7784	10,6363	17,312	1,7E-04
rs2345179	2	46113016	+	A	G	243	0,5489011	0,13249756	17,162	3,4E-05	0,3952	0,2766	18,149	1,1E-04
rs1931902	10	7655566	+	G	A	240	2,1335443	0,51659289	17,057	3,6E-05	3,1232	2,2731	26,455	1,8E-06
rs2730570	12	119033790	+	G	A	243	2,0644452	0,49994123	17,052	3,6E-05	1,6552	3,9696	17,218	1,8E-04
rs4628571	1	210858403	+	T	C	244	0,577578	0,13990964	17,042	3,7E-05	0,5181	0,2943	17,052	2,0E-04
rs7752722	6	154006683	+	G	C	244	3,2101649	0,78329525	16,796	4,2E-05	2,8871	10,8304	16,972	2,1E-04
rs6557301	6	154006913	+	G	A	244	3,2101649	0,78329525	16,796	4,2E-05	2,8871	10,8304	16,972	2,1E-04
rs1937603	6	154007005	+	A	G	244	3,2101649	0,78329525	16,796	4,2E-05	2,8871	10,8304	16,972	2,1E-04
rs9397154	6	154007466	+	T	C	244	3,2101649	0,78329525	16,796	4,2E-05	2,8871	10,8304	16,972	2,1E-04
rs9397632	6	154007619	+	C	T	244	3,2101649	0,78329525	16,796	4,2E-05	2,8871	10,8304	16,972	2,1E-04

Mit korrigierten P-Werten ((NA in P2df bedeutet, dass nicht alle 3 Genotypen vorhanden sind))

SNP	Chr	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10051190	5	123999278	+	T	C	244	0,4326831	0,09366247	21,341	0,750	0,444	0,071	22,621	NA
rs11883643	2	239586885	+	A	G	244	0,538006	0,11715119	21,090	0,805	0,581	0,249	22,254	NA
rs10450286	10	72986492	+	G	A	244	0,5502651	0,12021021	20,954	0,810	0,481	0,257	20,978	NA
rs2268844	3	185920603	+	G	T	244	0,3153717	0,06925358	20,738	0,855	0,248	0,406	23,358	NA
rs8026735	15	62125859	+	C	T	244	2,3159659	0,53085176	19,033	0,980	1,870	5,886	19,116	NA
rs10167210	2	239601726	+	C	A	244	0,5738864	0,13616412	17,763	0,995	0,597	0,294	18,338	NA
rs4645337	5	3279885	+	G	C	244	0,3658607	0,08766407	17,418	0,995	0,347	0,195	17,507	NA
rs4866542	5	3282671	+	T	C	244	0,3658607	0,08766407	17,418	0,995	0,347	0,195	17,507	NA
rs10866475	5	10511590	+	G	A	244	2,1559505	0,49709555	18,810	0,995	2,047	3,761	19,208	NA
rs4568464	6	153995311	+	C	T	244	3,3742804	0,80740627	17,465	0,995	3,048	10,834	17,708	NA
rs10806834	6	166231573	+	G	A	244	2,1881851	0,52115062	17,630	0,995	1,982	4,423	17,696	NA
rs2730577	12	119023969	+	A	G	240	2,0870446	0,49858132	17,522	0,995	1,659	4,059	17,716	NA
rs2431835	19	58647214	+	G	A	244	2,0142753	0,48204837	17,460	0,995	1,852	3,585	17,535	NA

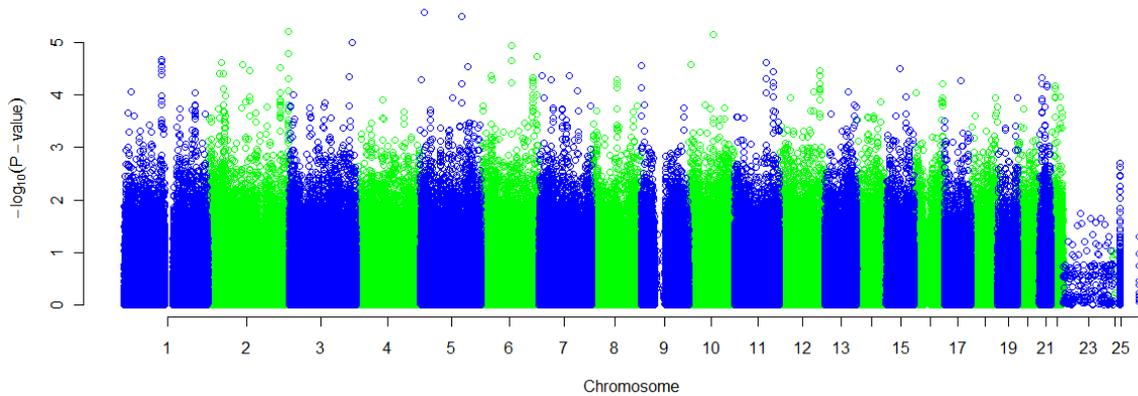
Analyse der Daten nach der 2.Runde der Qualitätskontrolle

inklusive Alter als Kovariate

Ohne korrigierte P-Werte

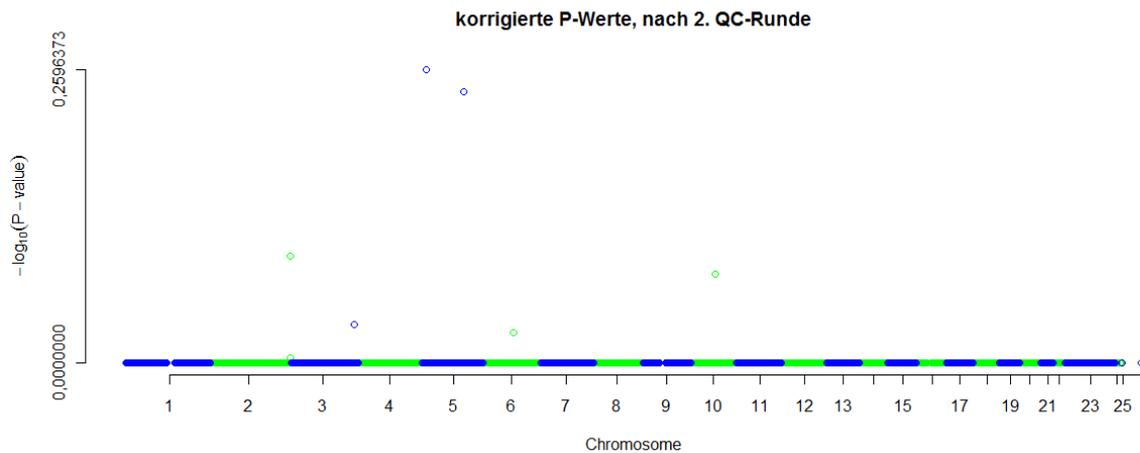
	Chromosome	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10866475	5	10511590	+	G	A	228	2,5891565	0,55127685	22,059	2,6E-06	2,5167	4,9150	22,971	1,0E-05
rs10051190	5	123999278	+	T	C	228	0,3878292	0,08316014	21,750	3,1E-06	0,3959	0,0505	22,600	1,2E-05
rs11883643	2	239586885	+	A	G	228	0,5120464	0,11334342	20,409	6,3E-06	0,5499	0,2243	21,318	2,3E-05
rs10450286	10	72986492	+	G	A	228	0,5196729	0,11573728	20,161	7,1E-06	0,4348	0,2245	20,162	4,2E-05
rs2268844	3	185920603	+	G	T	228	0,2806685	0,06349292	19,541	9,8E-06	0,2073	0,4243	22,611	1,2E-05
rs207268	6	90826709	+	C	T	228	0,5093979	0,11604379	19,270	1,1E-05	0,4637	0,2192	19,388	6,2E-05
rs7594129	2	239600686	+	C	A	228	0,5342037	0,12400772	18,557	1,6E-05	0,5821	0,2500	19,434	6,0E-05
rs10167210	2	239601726	+	C	A	228	0,5342037	0,12400772	18,557	1,6E-05	0,5821	0,2500	19,434	6,0E-05
rs10806834	6	166231573	+	G	A	228	2,4426634	0,57080141	18,313	1,9E-05	2,2409	5,0982	18,523	9,5E-05
rs17688644	1	111395110	+	T	C	227	2,5973952	0,61059608	18,095	2,1E-05	2,1725	8,3647	18,096	1,2E-04
rs2501721	6	90845036	+	C	T	228	0,5216229	0,12281356	18,039	2,2E-05	0,4343	0,2296	18,041	1,2E-04
rs17637944	1	111401704	+	T	C	227	2,5639364	0,60569126	17,919	2,3E-05	2,1312	8,4824	17,927	1,3E-04
rs17026826	2	41123319	+	T	G	228	0,193887	0,04589153	17,850	2,4E-05	0,1852	0,0000	17,907	1,3E-04
rs566598	11	95977311	+	C	G	227	0,2938596	0,06956277	17,845	2,4E-05	0,2788	0,1256	17,954	1,3E-04
rs6759106	2	103468011	+	G	A	228	0,3970064	0,09449699	17,651	2,7E-05	0,3929	0,0676	17,901	1,3E-04
rs1931902	10	7655566	+	G	A	224	2,3270527	0,55423309	17,629	2,7E-05	3,7595	2,5229	28,051	8,1E-07
rs12353411	9	426215	+	C	T	228	0,2521155	0,06012522	17,583	2,8E-05	0,2498	NA	17,583	2,8E-05
rs6677149	1	111393713	+	T	A	228	2,5280275	0,60458291	17,484	2,9E-05	2,1016	8,3647	17,498	1,6E-04
rs17688638	1	111394866	+	T	C	228	2,5280275	0,60458291	17,484	2,9E-05	2,1016	8,3647	17,498	1,6E-04
rs12566518	1	111395388	+	G	T	228	2,5280275	0,60458291	17,484	2,9E-05	2,1016	8,3647	17,498	1,6E-04
rs12567456	1	111401322	+	T	C	228	2,5280275	0,60458291	17,484	2,9E-05	2,1016	8,3647	17,498	1,6E-04
rs6696248	1	111404291	+	T	C	228	2,5280275	0,60458291	17,484	2,9E-05	2,1016	8,3647	17,498	1,6E-04
rs17098673	5	141837250	+	C	T	228	0,2194404	0,05250976	17,464	2,9E-05	0,2194	NA	17,464	2,9E-05
rs4427965	2	214802528	+	C	T	227	0,4884701	0,11707134	17,409	3,0E-05	0,4055	0,2312	17,564	1,5E-04
rs8026735	15	62125859	+	C	T	228	2,4358757	0,58469117	17,356	3,1E-05	1,9406	6,7477	17,446	1,6E-04

unkorrigierte P-Werte, nach 2. QC-Runde



Mit korrigierten P-Werten

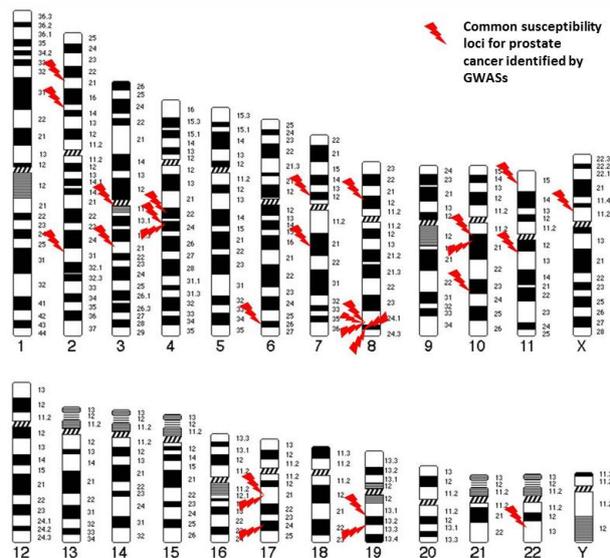
SNP	Chr	Position	Strand	A1	A2	N	effB	se_effB	chi2.1df	P1df	effAB	effBB	chi2.2df	P2df
rs10866475	5	10511590	+	G	A	228	2,5891565	0,55127685	22,059	0,550	2,517	4,915	22,971	NA
rs10051190	5	123999278	+	T	C	228	0,3878292	0,08316014	21,750	0,575	0,396	0,051	22,600	NA
rs11883643	2	239586885	+	A	G	228	0,5120464	0,11334342	20,409	0,805	0,550	0,224	21,318	NA
rs10450286	10	72986492	+	G	A	228	0,5196729	0,11573728	20,161	0,835	0,435	0,224	20,162	NA
rs2268844	3	185920603	+	G	T	228	0,2806685	0,06349292	19,541	0,925	0,207	0,424	22,611	NA
rs207268	6	90826709	+	C	T	228	0,5093979	0,11604379	19,270	0,940	0,464	0,219	19,388	NA
rs7594129	2	239600686	+	C	A	228	0,5342037	0,12400772	18,557	0,990	0,582	0,250	19,434	NA
rs10167210	2	239601726	+	C	A	228	0,5342037	0,12400772	18,557	0,990	0,582	0,250	19,434	NA



5. DISKUSSION

Aufgrund der hohen Prävalenz des Prostatakarzinoms und des hohen Inzidenz/Mortalität-Verhältnisses, welches auf eine Überbehandlung des Prostatakarzinoms schließen lässt, besteht der Bedarf genetische Marker nicht nur für das Prostatakarzinomrisiko sondern auch die aggressive Form des Prostatakarzinoms zu identifizieren.

GWAS stellen für die Identifizierung von Suszeptibilitätsloci ein hervorragendes Tool dar und die Entwicklung in der SNP Array Technologie ermöglicht das Screening von mehr als einer Million genetischer Varianten. In den letzten Jahren wurden durch GWAS ca. 40 Suszeptibilitätsloci, die mit Prostatakarzinomrisiko assoziiert sind, identifiziert (Abb 1). Da beim Prostatakarzinom „high-penetrance“ Mutationen nur eine untergeordnete Rolle, im Gegensatz zum Mammakarzinom und kolorektalem Karzinom, spielen kommen den genetischen Varianten mit niedriger Penetranz eine besondere Bedeutung zu. Um Signaturen solcher „low penetrance“ Varianten zu identifizieren bedarf es weiterer umfangreicher GWAS und deren gemeinsame Auswertung in gepoolten Analysen und Metanalysen.



Locus	Chr.	SNP
ITGA6	2	rs12621278
2p15	2	rs721048
THADA	2	rs1465618
3p12	3	rs2660753
3q21.3	3	rs10934853
TET2	4	rs7679673
PDLIM5	4	rs17021918
PDLIM5	4	rs12500426
6q25	6	rs9364554
7q21	7	rs6465657
JAZF1	7	rs10486567
8p21	8	rs2928679
NKX3.1	8	rs1512268
8q24	8	rs620861
8q24	8	rs10086908
8q24	8	rs445114
8q24	8	rs16902094
8q24	8	rs6983267
8q24	8	rs1447295
8q24	8	rs16901979
CTBP2	10	rs4962416
MSMB	10	rs10993994
11p15	11	rs7127900
11q13	11	rs7931342
HNF1B	17	rs4430796
HNF1B	17	rs11649743
17q24.3	17	rs1859962
KLK2/KLK3	19	rs2735839
19q13.2	19	rs8102476
22q13	22	rs5759167
Xp11	X	rs5945619

modified from Varghese et Easton, 2010
*Genome-wide association studies in
 common cancers-what have we learnt?*

Abb.1 Suszeptibilitätsloci für das Prostatakarzinom

Es wurden wie im Projektantrag geplant, 250 DNA-Proben, bestehend aus 84 Prostatakarzinomen mit einem Gleason Score größer 7, 84 Prostatakarzinome mit Gleason Score kleiner 7 sowie 82 BPH Patienten als Kontrollgruppe, mit dem SNP 6.0 Array prozessiert. Es konnten keine Loci auf genomweitem Signifikanzniveau identifiziert werden. Dies ist darauf zurückzuführen, dass die Anzahl der eingesetzten SNP Arrays zu gering war. Um die statistische Power der Studie zu erhöhen wurden noch weitere 70 SNP 6.0 Arrays – anstatt einer Replikationsrunde wie im Projektantrag ursprünglich geplant – prozessiert. Zusätzlich wurden weitere 130 SNP 6.0 Arrays – die aus anderen Projektmitteln finanziert werden – eingesetzt. Von den insgesamt 200 zusätzlich prozessierten DNA Proben waren 100 von Patienten mit Gleason Score größer als 7, die anderen 100 von Patienten mit Gleason Score kleiner als 7.

Es konnten allerdings auch hier keine statistisch signifikanten Loci identifiziert werden. Deshalb war der nächste Schritt, in Kooperation mit Prof. Dr. Thomas Meitinger vom Helmholtzcenter München zur Analyse 884 KORA Kontrollen der KORA Gen Plattform heranzuziehen. Das entspricht allen verfügbaren männlichen

Teilnehmern der KORA Studie für welche SNP 6.0 Array Daten vorliegen. Diese werden gemeinsam mit den Ergebnissen unserer 450 Patienten derzeit statistisch ausgewertet.

Die Identifizierung von Biomarkern, die zum Zeitpunkt der Diagnose die Differenzierung ermöglichen, welche Männer eine aggressive Form des PC entwickeln werden, ist eine dringliche Notwendigkeit, um nur jene Männern den Therapien zu unterziehen, wo es auch indiziert ist.

Die genetischen Einflüsse, die zu einem aggressivem Verlauf des Prostatakarzinoms führen, sind weitgehend nicht bekannt. Es sind bisher nur wenige potentielle genetische Marker identifiziert, die die Differenzierung zwischen einer aggressiven und indolenten Form des Prostatakarzinoms unterscheiden können. Xu et al. haben einen SNP am Chromosom 17p12 (rs4054823) identifiziert, der ein Marker für eine aggressive Form des Prostatakarzinoms sein könnte. Kürzlich hat diese Gruppe auch eine genetische Variante am Chromosom 20p13, die mit Progression des Prostatakarzinoms assoziiert sein könnte, identifiziert (Jin et al. 2011).

Um genetische Varianten, die mit einer aggressiven Form des Prostatakarzinoms assoziiert sein könnten zu finden haben wir die Arraydaten von Patienten mit einer aggressiven Form des Prostatakarzinoms mit den Daten von Patienten mit nicht-aggressiver Form verglichen. Die Aggressivität des Prostatakarzinoms ist klassifiziert durch klinische Parameter wie PSA Serumspiegel, TNM-Staging und Gleason Score. In der bisherigen Analyse haben sich keine SNPs ergeben, die eine Unterscheidung in eine aggressive und nicht-aggressive Form des Prostatakarzinoms erlauben.

Da derzeit keine verlässlichen Biomarker, außer den bekannten klinischen Parametern wie PSA und Histologie zur Risikoabschätzung, insbesondere für den aggressiven Verlauf des Prostatakarzinoms zur Verfügung stehen, besteht ein weiterer Bedarf an der Durchführung solcher Studien mit größeren Fallzahlen und in unterschiedlichen ethnischen Gruppen.

Ausblick

Aufgrund der relativ geringen Anzahl der in diesem Projekt eingesetzten SNP Arrays konnten keine Risikovarianten auf genomweitem Signifikanzniveau identifiziert werden. Derzeit laufen die Berechnungen der Arraydaten aus unserer Prostatakarzinom DNA-Bank gemeinsam mit den KORA Kontroll-Daten. Durch die weitere Erhöhung der Fallzahl und die Anwendung eines neuen Algorithmus besteht die Möglichkeit, dass dadurch noch SNPs an Signifikanz gewinnen. Weiters werden logistische Regressionsmodelle sämtlicher SNPs, für welche in der Literatur bereits Assoziationen mit dem Prostatakarzinomrisiko gefunden wurden auch in Hinblick auf die Aggressivität des Verlaufs getestet. Es ist schlussendlich auch noch geplant unsere Ergebnisse in Kooperation mit einem internationalen Prostatakarzinom-konsortium in eine Metaanalyse einzubringen.

6. LITERATUR

Duggan D, Zheng SL, Knowlton M, Benitez D, Dimitrov L, Wiklund F, Robbins C, Isaacs SD, Cheng Y, Li G, Sun J, Chang BL, Marovich L, Wiley KE, Bälter K, Stattin P, Adami HO, Gielzak M, Yan G, Sauvageot J, Liu W, Kim JW, Bleecker ER, Meyers DA, Trock BJ, Partin AW, Walsh PC, Isaacs WB, Grönberg H, Xu J, Carpten JD. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. *J Natl Cancer Inst.* 2007 Dec 19;99(24):1836-44.

Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson BE, Haiman CA, Schleutker J, Hamdy FC, Neal DE, Donovan JL, Stanford JL, Ostrander EA, Ingles SA, John EM, Thibodeau SN, Schaid D, Park JY, Spurdle A, Clements J, Dickinson JL, Maier C, Vogel W, Dörk T, Rebbeck TR, Cooney KA, Cannon-Albright L, Chappuis PO, Hutter P, Zeegers M, Kaneva R, Zhang HW, Lu YJ, Foulkes WD, English DR, Leongamornlert DA, Tymrakiewicz M, Morrison J, Arden-Jones AT, Hall AL, O'Brien LT, Wilkinson RA, Saunders EJ, Page EC, Sawyer EJ, Edwards SM, Dearnaley DP, Horwich A, Huddart RA, Khoo VS, Parker CC, Van As N, Woodhouse CJ, Thompson A, Christmas T, Ogden C, Cooper CS, Southey MC, Lophatananon A, Liu JF, Kolonel LN, Le Marchand L, Wahlfors T, Tammela TL, Auvinen A, Lewis SJ, Cox A, FitzGerald LM, Koopmeiners JS, Karyadi DM, Kwon EM, Stern MC, Corral R, Joshi AD, Shahabi A, McDonnell SK, Sellers TA, Pow-Sang J, Chambers S, Aitken J, Gardiner RA, Batra J, Kedda MA, Lose F, Polanowski A, Patterson B, Serth J, Meyer A, Luedeke M, Stefflova K, Ray AM, Lange EM, Farnham J, Khan H, Slavov C, Mitkova A, Cao G; UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology; UK ProtecT Study Collaborators; PRACTICAL Consortium, Easton DF. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet.* 2009 Oct;41(10):1116-21.

Jin G, Sun J, Liu W, Zhang Z, Chu LW, Kim ST, Sun J, Feng J, Duggan D, Carpten JD, Wiklund F, Grönberg H, Isaacs WB, Zheng SL, Xu J. Genome-wide copy-number variation analysis identifies common genetic variants at 20p13 associated with

aggressiveness of prostate cancer. *Carcinogenesis*. 2011 Jul;32(7):1057-62. Epub 2011 May 5.

F. Frommlet, F. Ruhaltinger, P. Twarog, M. Bogdan, Modified versions of Bayesian Information Criterion for genome-wide association studies, 2012, *CSDA*, 56, 1038 - 1051

Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemenev LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet*. 2007 Aug;39(8):977-83.

Gudmundsson J, Sulem P, Rafnar T, Bergthorsson JT, Manolescu A, Gudbjartsson D, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Blondal T, Jakobsdottir M, Stacey SN, Kostic J, Kristinsson KT, Birgisdottir B, Ghosh S, Magnusdottir DN, Thorlacius S, Thorleifsson G, Zheng SL, Sun J, Chang BL, Elmore JB, Breyer JP, McReynolds KM, Bradley KM, Yaspan BL, Wiklund F, Stattin P, Lindström S, Adami HO, McDonnell SK, Schaid DJ, Cunningham JM, Wang L, Cerhan JR, St Sauver JL, Isaacs SD, Wiley KE, Partin AW, Walsh PC, Polo S, Ruiz-Echarri M, Navarrete S, Fuertes F, Saez B, Godino J, Weijerman PC, Swinkels DW, Aben KK, Witjes JA, Suarez BK, Helfand BT, Frigge ML, Kristjansson K, Ober C, Jonsson E, Einarsson GV, Xu J, Gronberg H, Smith JR, Thibodeau SN, Isaacs WB, Catalona WJ, Mayordomo JI, Kiemenev LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong

A, Stefansson K. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. *Nat Genet.* 2008 Mar;40(3):281-3.

Gudmundsson J, Sulem P, Gudbjartsson DF, Blondal T, Gylfason A, Agnarsson BA, Benediktsdottir KR, Magnusdottir DN, Orlygsdottir G, Jakobsdottir M, Stacey SN, Sigurdsson A, Wahlfors T, Tammela T, Breyer JP, McReynolds KM, Bradley KM, Saez B, Godino J, Navarrete S, Fuertes F, Murillo L, Polo E, Aben KK, van Oort IM, Suarez BK, Helfand BT, Kan D, Zanon C, Frigge ML, Kristjansson K, Gulcher JR, Einarsson GV, Jonsson E, Catalona WJ, Mayordomo JI, Kiemeny LA, Smith JR, Schleutker J, Barkardottir RB, Kong A, Thorsteinsdottir U, Rafnar T, Stefansson K. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet.* 2009 Oct;41(10):1122-6.

He, Q., Lin, D., 2011. A variable selection method for genome-wide association studies. *Bioinformatics* 27 (1), 1–8.

Hoggart, C.J., Whittaker, J.C., De Iorio, M., Balding, D.J., 2008. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLOS Genet.* 4 (7), e1000130. doi:10.1371/journal.pgen.1000130

Kote-Jarai Z, Easton DF, Stanford JL, Ostrander EA, Schleutker J, Ingles SA, Schaid D, Thibodeau S, Dörk T, Neal D, Donovan J, Hamdy F, Cox A, Maier C, Vogel W, Guy M, Muir K, Lophatananon A, Kedda MA, Spurdle A, Steginga S, John EM, Giles G, Hopper J, Chappuis PO, Hutter P, Foulkes WD, Hamel N, Salinas CA, Koopmeiners JS, Karyadi DM, Johanneson B, Wahlfors T, Tammela TL, Stern MC, Corral R, McDonnell SK, Schürmann P, Meyer A, Kuefer R, Leongamornlert DA, Tymrakiewicz M, Liu JF, O'Mara T, Gardiner RA, Aitken J, Joshi AD, Severi G, English DR, Southey M, Edwards SM, Al Olama AA; PRACTICAL Consortium, Eeles RA. Multiple novel prostate cancer predisposition loci confirmed by an international study: the PRACTICAL Consortium. *Cancer Epidemiol Biomarkers Prev.* 2008 Aug;17(8):2052-61. Erratum in: *Cancer Epidemiol Biomarkers Prev.* 2008 Oct;17(10):2901.

Lindström S, Schumacher F, Cox DG, Travis RC, Albanes D, Allen NE, Andriole GL, Berndt SI, Boeing H, Bueno-de-Mesquita HB, Crawford D, Diver WR, Gaziano JM, Giles GG, Giovannucci EL, Gonzales C, Henderson B, Hunter DJ, Johansson M,

Kolonel LN, Ma J, Le Marchand L, Pala V, Stampfer MJ, Stram DO, Thun M, Tjonneland A, Trichopoulos D, Virtamo J, Weinstein SJ, Willett WC, Yeager M, Hayes RB, Severi G, Haiman C, Chanock SJ, Kraft P. Common genetic variants in prostate cancer risk prediction - Results from the NCI Breast and Prostate Cancer Cohort Consortium (BPC3). *Cancer Epidemiol Biomarkers Prev.* 2012 Jan 11. [Epub ahead of print]

Lu Y, Zhang Z, Yu H, Zheng SL, Isaacs WB, Xu J, Sun J. Functional annotation of risk loci identified through genome-wide association studies for prostate cancer. *Prostate.* 2011 Jun 15;71(9):955-63. doi: 10.1002/pros.21311.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, 81.

Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, Yu K, Chatterjee N, Welch R, Hutchinson A, Crenshaw A, Cancel-Tassin G, Staats BJ, Wang Z, Gonzalez-Bosquet J, Fang J, Deng X, Berndt SI, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cussenot O, Valeri A, Andriole GL, Crawford ED, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hayes RB, Hunter DJ, Chanock SJ. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet.* 2008 Mar;40(3):310-5.

Varghese JS, Easton DF. Genome-wide association studies in common cancers--what have we learnt? *Curr Opin Genet Dev.* 2010 Jun;20(3):201-9. Review.

Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang Z, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hunter DJ, Chanock SJ, Thomas G. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet.* 2007 May;39(5):645-9.

Yeager M, Chatterjee N, Ciampa J, Jacobs KB, Gonzalez-Bosquet J, Hayes RB, Kraft P, Wacholder S, Orr N, Berndt S, Yu K, Hutchinson A, Wang Z, Amundadottir L, Feigelson HS, Thun MJ, Diver WR, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Crawford ED, Haiman CA, Henderson B, Kolonel L, Le Marchand L, Siddiq A, Riboli E, Key TJ, Kaaks R, Isaacs W, Isaacs S, Wiley KE, Gronberg H, Wiklund F, Stattin P, Xu J, Zheng SL, Sun J, Vatten LJ, Hveem K, Kumle M, Tucker M, Gerhard DS, Hoover RN, Fraumeni JF Jr, Hunter DJ, Thomas G, Chanock SJ. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat Genet.* 2009 Oct;41(10):1055-7.

Xu J, Zheng SL, Isaacs SD, Wiley KE, Wiklund F, Sun J, Kader AK, Li G, Purcell LD, Kim ST, Hsu FC, Stattin P, Hugosson J, Adolfsson J, Walsh PC, Trent JM, Duggan D, Carpten J, Grönberg H, Isaacs WB. Inherited genetic variant predisposes to aggressive but not indolent prostate cancer. *Proc Natl Acad Sci U S A.* 2010 Feb 2;107(5):2136-40.

www.bmg.gv.at

Das Prostatakarzinom ist eine der häufigsten Krebserkrankungen des Mannes. Der derzeit einzige, jedoch sehr unspezifische Biomarker hierfür ist das Prostata-spezifische Antigen (PSA). Im Zuge des vorliegenden Forschungsprojekts wird daher versucht, genetische Marker zu identifizieren, mit denen ein erhöhtes Prostatakarzinomrisiko bzw. eine Unterscheidung zwischen aggressiver und nicht-aggressiver Form des Prostatakarzinoms zum Zeitpunkt der Diagnose erkannt werden kann.