

Carsten Carlberg

Die molekulare Basis von Gesundheit

Wie Epigenetik und Ernährung unser
Leben beeinflussen



SACHBUCH

 Springer

Die molekulare Basis von Gesundheit

Carsten Carlberg

Die molekulare Basis von Gesundheit

Wie Epigenetik und Ernährung unser
Leben beeinflussen

Prof. Carsten Carlberg
Institute of Animal Reproduction and Food
Research, Polish Academy of Sciences
Olsztyn, Poland

Institute of Biomedicine
University of Eastern Finland
Kuopio, Finland

ISBN 978-3-662-67985-2 ISBN 978-3-662-67986-9 (eBook)
<https://doi.org/10.1007/978-3-662-67986-9>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2023

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Sarah Koch

Springer ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Das Papier dieses Produkts ist recyclebar.

Vorwort

Um gleich mit der Tür ins Haus zu fallen: in diesem Buch geht es darum, die wissenschaftlichen Grundlagen dafür zu präsentieren, wie **ein jeder von uns Verantwortung übernehmen kann, gesund zu bleiben**. Und selbst wenn wir krank werden sollten, können wir selbst viel zum Heilungsprozess beitragen. Dafür ist es essentiell, ein klares Verständnis der physiologischen¹ Funktionen unseres Körpers zu haben. Diese Einsichten werden erweitert durch die molekularen Grundlagen, wie jede einzelne Zelle funktioniert und Information über ihre Umgebung verarbeitet. Letzteres hat sehr viel mit **Epigenetik** zu tun, die ein zentrales Thema in diesem Buch ist. Der Begriff „Epigenetik“ beschreibt die Verpackung und Zugänglichkeit unseres Genoms, d.h. der Gesamtheit der DNA, die wir in jeder der Billionen Zellen tragen, die unseren Körper formen. Die Vorsilbe „epi“ bedeutet „auf“, „über“ oder „jenseits“ und weist darauf hin, dass **epigenetische Prozesse über dem Genom liegen. Damit hat die Epigenetik im Gegensatz zur Genetik keinen Einfluss auf die DNA-Sequenz unseres Genoms**.

Unser Genom hat in der Summe aller 23 verschiedenen Chromosomen eine Länge von etwa einem Meter. Um es in einem Zellkern mit einem Durchmesser von weniger als $10\ \mu\text{m}^2$ unterzubringen, muss die DNA um Komplexe von Histonproteinen gewickelt werden. **Dieser Protein-DNA-Komplex wird als Chromatin bezeichnet und ist das konkrete Gegenstück der Epigenetik**. Die wichtigste Funktion des Chromatins besteht darin, etwa 90 % unseres Genoms Zell- und Gewebe-spezifisch zu verpacken und damit unzugänglich für Transkriptionsfaktoren³ und Polymerasen⁴ zu halten. Mit anderen Worten, **Chromatin fungiert als Wächter für unerwünschte Genaktivierung**. Auf diese Weise verwendet jedes der 400 verschiedenen Gewebe und Zelltypen, aus denen unseren Körper aufgebaut ist, eine andere Teilmenge der 20.000 proteinkodierenden Gene unseres

¹ Physiologie beschreibt die biologischen, chemischen und physikalischen Vorgänge, die der Funktionsweise von Zellen, Geweben, Organen und Organsystem zugrunde liegen.

² μm = Millionstel Meter.

³ Ein Transkriptionsfaktor ist ein Protein, das an bestimmte Sequenzmotive im Genom bindet und die Transkription nahegelegener Gene reguliert.

⁴ Polymerasen sind Enzyme, die DNA in RNA übersetzen (bei der Transkription), oder DNA in DNA (bei der Replikation).

Genoms. **Das heißt damit auch, dass jeder von uns nur ein Genom hat, aber mindestens 400 verschiedene Epigenome.** Epigenetik verhindert, dass sich z. B. eine Nierenzelle über Nacht in ein Neuron verwandelt oder umgekehrt. Auf diese Weise verschafft Epigenetik differenzierten Zellen eine dauerhafte Erinnerung an ihre Identität.

Die Differenzierung embryonaler Stammzellen in spezialisierte Zelltypen erfolgt während der Embryogenese⁵, d.h. in den ersten Lebenswochen eines Fötus. Auch in Erwachsenen gibt es Stammzellen, z. B. im Knochenmark, der Haut und im Darm. Diese Stammzellen teilen sich kontinuierlich und ein Teil dieser Tochterzellen reift zu spezialisierten Zellen aus, um den kontinuierlichen Verlust von Zellen des Immunsystems oder der äußeren und inneren Oberfläche unseres Körpers zu kompensieren. **Der zugrunde liegende Mechanismus all dieser Differenzierungsprozesse ist eine epigenetische Programmierung des Chromatins.** Das bedeutet, dass im Chromatin Information festgeschrieben wird, wie die Zellen funktionieren müssen, um ihren Aufgaben nachzukommen.

Neben ihrer statischen Funktion hat Epigenetik auch einen dynamischen Aspekt, bei dem die Aktivierung intrazellulärer Signalübertragungswege⁶ über extrazelluläre Signale, wie Peptidhormone, Zytokine oder Wachstumsfaktoren, zur Aktivierung von Transkriptionsfaktoren und Chromatin-modifizierenden Enzymen⁷ führt. Viele dieser Signale stammen direkt oder indirekt von unserer Ernährung. Letztere ist das wichtigste externe Signal, dem wir täglich ausgesetzt sind. Das führt zum zweiten zentralen Thema dieses Buches, der **Nutrigenomik**. Die Nahrung, die wir täglich zu uns nehmen, ist weit mehr als eine Ansammlung von Kohlenhydraten, Proteinen, Lipiden, Vitaminen und Mineralien, die Energie und Bausteine für unseren Körper liefert. **Das faszinierende Fachgebiet der Nutrigenomik beschreibt die Wechselwirkung zwischen Nahrungsmolekülen, ihren Metaboliten und unserem (Epi)genom.** Viele Nahrungsbestandteile, wie beispielsweise ungesättigte Fettsäuren, aktivieren direkt Transkriptionsfaktoren oder verursachen Veränderungen in den Konzentrationen von Metaboliten, die die Aktivität von Chromatin-modifizierenden Enzymen beeinflussen. Auf diese Weise „sprechen“ Moleküle unseres Frühstücks, Mittagessens oder Abendessens mit unserem Genom bzw. Epigenom.

Die meisten nichtübertragbaren Krankheiten haben eine genetische, vererbte Komponente sowie eine epigenetische Komponente, die auf Umwelteinflüssen und unserem Lebensstil basiert. Viele Volkskrankheiten, wie Typ 2 Diabetes (T2D), lassen sich nur zu etwa 20 % durch eine genetische Veranlagung erklären. **Wir können die Gene, mit denen wir geboren werden, nicht ändern, aber**

⁵ Embryogenese ist der biologische Prozess, der zur Bildung des Embryos führt und umfasst etwa die ersten 8 Wochen nach der Befruchtung.

⁶ Signalübertragungswege, sind Prozesse, bei denen eine Zelle auf äußere oder innere Signale reagiert und diese in biologische Funktionen überträgt, wie z. B. der Aktivierung eines Gens.

⁷ Ein Enzym ist ein Protein, das den Ablauf von chemischen Reaktionen beschleunigt oder gar ermöglicht.

wir können uns um die restlichen 80 % des Erkrankungsrisikos kümmern, die hauptsächlich auf Veränderungen unseres Epigenoms zurückzuführen sind. Dieser Satz ist eine zentrale Aussage in diesem Buch und bedeutet, dass eine genetische Veranlagung für eine Erkrankung durch einen angemessenen, gesunden Lebensstil ausgeglichen werden kann, der das Epigenom der betroffenen Gewebe moduliert. Kurz, **es ist nie zu spät etwas für seinen Körper und seine Gesundheit zu tun.**

Diese ersten Beispiele zeigen, dass Epigenetik und Nutrigenomik vielfältige Aspekte von Gesundheit und Krankheit beeinflussen. Ich werde die zentrale Bedeutung der Epigenetik während der **Embryogenese** und Zelldifferenzierung sowie im **Alterungsprozess** und dem Risiko von Fehlernährung für die **Entstehung von Krebs** beschreiben. Darüber hinaus wird die Rolle des Epigenoms als molekularer Speicher zellulärer Ereignisse nicht nur im Gehirn, sondern auch in Stoffwechselorganen und im Immunsystem diskutiert. In diesem Zusammenhang werden Auswirkungen von Epigenetik und Ernährung auf **neurodegenerative Erkrankungen** und Autismus, **Stoffwechselerkrankungen** wie T2D, und ein gestörtes Immunsystem, z. B. bei **Autoimmunerkrankungen**, erklärt.

Dieses Buch wendet sich an den Laien mit einer guten biologischen Grundkenntnis. Ich versuche die zugrunde liegenden Mechanismen der besprochenen Themen vereinfacht darzustellen. Trotzdem möchte ich Verständnis dafür bitten, dass ich trotzdem viele Gene und Proteine benenne. Die ersten 6 Kapitel erläutern die molekularen Grundlagen unseres Genoms, der Epigenetik und der Nutrigenomik, während die folgenden 8 Kapitel Beispiele für die Auswirkungen auf unsere Gesundheit und Erkrankungen liefern. Ein Glossar im Anhang erklärt die wichtigsten Fachbegriffe. Die Abbildungen wurden zu einem großen Teil von meinem Kollegen Prof. Ferdinand Molnár (Astana, Kasachstan) für unsere englischsprachigen Bücher „Human Epigenetics: How Science Works“ (ISBN 978-3-030-22906-1) und „Nutrigenomics: How Science Works“ (ISBN 978-3-030-36948-4) erstellt. Ich danke ihm dafür, die Abbildungen benutzen zu dürfen. Dr. Eunike Velleuer (Düsseldorf, Deutschland) danke ich für die farbliche Überarbeitung der Abbildungen im Stil unseres gemeinsamen Buches „Molecular Medicine: How Science Works“ (ISBN 978-3-031-27132-8).

Ich hoffe, dass die Leser dieses recht visuelle Buch genießen und sich genauso für die molekulare Basis von Gesundheit begeistern lassen wie der Autor.

Inhaltsverzeichnis

1	Das menschliche Genom	1
1.1	Migration und evolutionäre Herausforderungen des <i>Homo sapiens</i>	1
1.2	Vielfalt menschlicher Populationen	3
1.3	Varianten des menschlichen Genoms	8
	Weiterführende Literatur	15
2	Epigenetik, Chromatin und expression	17
2.1	Was ist Epigenetik?	17
2.2	Die epigenetische Landschaft	21
2.3	Nukleosomen: zentrale Einheiten des Chromatins	25
2.4	Chromatinarchitektur: Epigenetik in 3D	26
2.5	Chromatinorganisation	30
2.6	Epigenetik und Genexpression	31
	Weiterführende Literatur	35
3	Epigenetische Regulation	37
3.1	DNA-Methylierung	38
3.2	Genetische Prägung	43
3.3	Histone und ihre Modifikationen: der Histoncode	46
3.4	Genregulation über Chromatin-modifizierende Enzyme	52
3.5	Chromatinremodellierer	54
3.6	Lange ncRNAs als Chromatinorganisatoren	58
	Weiterführende Literatur	61
4	Molekulare Sensoren von Makro- und Mikronährstoffen	63
4.1	Grundlagen des Stoffwechsels	63
4.2	Kohlenhydrate, Aminosäuren und Lipide	66
4.3	Kernrezeptoren als Nährstoffsensoren	73
4.4	Vitamine und ihre Rezeptoren	82
	Weiterführende Literatur	88
5	Nutri genomik und Stoffwechselregulation	89
5.1	Nutri genomik und Nutri genetik	89
5.2	Epigenetische Signalübertragung von Metaboliten des Intermediärstoffwechsel	92

5.3	Zirkadiane Kontrolle von Stoffwechselprozessen	100
	Weiterführende Literatur.	105
6	Ernährung und Volkskrankheiten.	107
6.1	Evolution der menschlichen Ernährung.	107
6.2	Ernährung und Stoffwechselerkrankungen	109
6.3	Ernährung und Krebs	114
6.4	Molekulare Auswirkungen von körperlicher Aktivität	118
6.5	Kalorienrestriktion und der zelluläre Energiestatus.	119
	Weiterführende Literatur.	122
7	Embryogenese und Zelldifferenzierung	125
7.1	Epigenetische Veränderungen während der Embryogenese	125
7.2	Stammzellen und zelluläre Pluripotenz	129
7.3	Epigenetische Dynamik während der Differenzierung	132
	Weiterführende Literatur.	135
8	Das Immunsystem	137
8.1	Epigenetik der Hämatopoese	137
8.2	Akute und chronische Entzündungen	141
8.3	Die Rolle der Epigenetik bei Immunantworten	149
	Weiterführende Literatur.	155
9	Fettleibigkeit und Diabetes.	157
9.1	Fettleibigkeit	158
9.2	Entzündung im Fettgewebe	164
9.3	Reaktion auf metabolischen Stress im ER.	168
9.4	Energiehomöostase und hormonelle Regulation der Nahrungsaufnahme	171
9.5	Glukosehomöostase und Insulinresistenz	176
9.6	β -Zell-Versagen und T2D	183
	Weiterführende Literatur.	195
10	Herz-Kreislauf-Erkrankungen und das metabolische Syndrom	197
10.1	Bluthochdruck und Arteriosklerose	198
10.2	Lipoproteine und Dyslipidämien	204
10.3	Metabolisches Syndrom.	209
	Weiterführende Literatur.	219
11	Epigenetik von Krebs	221
11.1	Epimutationen bei Krebs	221
11.2	Epigenom-weite Störungen als Kennzeichen von Krebs.	224
11.3	Epigenetische Umprogrammierung bei Krebs.	228
11.4	Epigenetische Mechanismen von Krebs	231
	Weiterführende Literatur.	235

12 Neuroepigenetik	237
12.1 Die Rolle der Epigenetik bei der neuronalen Entwicklung	237
12.2 Epigenetische Grundlagen des Gedächtnisses	241
12.3 Epigenetik neurodegenerativer Erkrankungen	242
Weiterführende Literatur	248
13 Ernährung und Altern	249
13.1 Evolutionäre Anpassungen des menschlichen Genoms	249
13.2 Generationsübergreifende epigenetische Vererbung	254
13.3 Konservierte nährstoffassoziierte Signalwege und das Altern	259
13.4 Epigenetik des Alterns	264
Weiterführende Literatur	267
14 Der gesunde Lebensstil	269
14.1 Epigenom-weite Varianten	269
14.2 Personalisierte Medizin und Ernährung	272
14.3 Epigenetische Therapie von Krankheiten	276
Weiterführende Literatur	279
Glossar	281
Stichwortverzeichnis	295

Abkürzungsverzeichnis

1,25(OH) ₂ D ₃	1,25-Dihydroxyvitamin D ₃
25(OH)D ₃	25-Hydroxyvitamin D ₃
3D	dreidimensional
5caC	5-Carboxylcytosin
5fC	5-Formylcytosin
5hmC	5-Hydroxymethylcytosin
5hU	5-Hydroxyuracil
5mC	5-Methylcytosin
5mU	5-Hydroxyuracil
ABC	ATP-Bindekassette
ACACA	„acetyl-CoA carboxylase alpha“
ACAT1	Acetyl-CoA-Acetyltransferase 1
ADAMTS	„ADAM metalloproteinase with thrombospondin“
ADP	Adenosindiphosphat
ADRB3	Adrenozeptor Beta 3
AGRP	Agouti-verwandtes Peptid
AICAR	5-Aminoimidazol-4-Carboxamidribonukleotid
AKT	AKT Serin/Threonin Kinase, wird auch Proteinkinase B (PKB) genannt
AMP	Adenosinmonophosphat
AMPK	Adenosinmonophosphat-aktivierte Proteinkinase
AMY	Amylase
ANGPTL2	Angiopoietin-ähnliches Protein 2
AP1	„activating protein 1“, ein Heterodimer der Onkoproteine Jun und Fos
APEH	„acylaminoacyl-peptide hydrolase“
APO	Apolipoprotein
APPL1	„adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 1“
AR	Androgenrezeptor
ARID1A	„AT-rich interaction domain 1A“
ARL4C	„ADP ribosylation factor like GTPase 4C“
ARNTL	„aryl hydrocarbon receptor nuclear translocator-like“, wird auch BMAL1 genannt

ASH1L	ASH1-ähnliche Histonylinmethyltransferase
ASIP	Agouti-Signalprotein
ATF6	Aktivierender Transkriptionsfaktor 6
ATP	Adenosintriphosphat
BDNF	vom Gehirn stammender neurotropher Faktor
BLK	„BLK proto-oncogene, src family tyrosine kinase“
BMI	Body-Mass-Index
BMI1	BMI1-Protoonkogen, Polycomb-Ringfinger
bp	Basenpaar
BRD	Bromodomäne enthaltend
CAMP	antimikrobielles Cathelicidin-Peptid
CAR	konstitutiver Androstanrezeptor
CBFB	„core-binding factor subunit β “
CBX	Chromobox
CCK	Cholecystokinin
CCL	CC-Chemokin-Ligand
CCR	C-C-Motiv-Chemokinrezeptor
CDKAL1	„CDK5 regulatory subunit associated protein 1 like 1“
CDKN	Cyclin-abhängiger Kinase-Inhibitor
CDX2	„caudal type homeobox 2“
CEBP	CCAAT-bindendes Protein
CEL	Carboxylesterlipase
CETP	Cholesterinester-Transferprotein
CETPD	CETP-Mangel
CHREBP	Kohlenhydrat-responsives Element-Bindeprotein, wird vom <i>MLXIPL</i> -Gen kodiert
CIMP	CpG-Insel-Methylator-Phänotyp
CLOCK	„clock circadian regulator“
CLP	gemeinsame lymphoide Vorläuferzellen
CMF	gemeinsame myeloische Vorläuferzellen
CNV	Kopienzahlvariante
CPT1A	Carnitin-Palmitoyltransferase 1A
CREB3L3	„cAMP responsive element binding protein 3 like 3“
CREBBP	CREB-Bindungsprotein, wird auch KAT3A genannt
CRP	C-reaktives Protein
CRY1	Cryptochrom zirkadiane Uhr 1
CSF2	koloniestimulierender Faktor 2
CTCF	CCCTC-Bindefaktor
CTLA4	zytotoxisches T-Lymphozyten-assoziiertes Protein 4
CTNS	„cystinosis, lysosomal cystine transporter“
CXCL	CXC-Motiv-Chemokin
CXCR	C-X-C-Motif-Chemokinrezeptor
CXXC1	CXXC-Fingerprotein 1
CYP	Cytochrom P450
DAAT1	Diacylglycerol-O-Acyltransferase 1

DACH1	Transkriptionsfaktor 1 der Dackelfamilie
DAG	Diacylglycerin
DALY	behinderungsbereinigtes Lebensjahr
DAMP	Schaden-assoziiertes molekulares Muster
DCT	Dopachromtautomerase
DEFB4	Defensin Beta 4A
DHA	Docosahexaensäure
DNMT	DNA-Methyltransferase
DOHaD	Entwicklungsursprünge von Gesundheit und Krankheit
DOT1L	„DOT1 like histone lysine methyltransferase“
EDAR	„ectodysplasin A receptor“
EED	embryonale Ektodermentwicklung
EGFR	„epidermal growth factor receptor“
EGIR	European Group for the Study of Insulin Resistance
EHMT2	„euchromatic histone lysine methyltransferase 2“
EIF2AK3	Eukaryotischer Translationsinitiationfaktor 2 Alpha Kinase 3
EMT	Epithelial-mesenchymale Transition
ENCODE	Enzyklopädie der DNA-Elemente
EP300	E1A-Bindungsprotein p300, wird auch KAT3B genannt
EPA	Eicosapentaensäure
ER	endoplasmatisches Retikulum
ERN1	„endoplasmic reticulum to nucleus signaling 1“
eRNA	Enhancer-RNA
ES	embryonale Stamm
ESR	Östrogenrezeptor
EZH	„enhancer of zeste homolog“, wird auch genannt KMT6A
FAD	Flavinadenindinukleotid
FANTOM	funktionelle Annotation des Säugetiergenoms
FASN	Fettsäuresynthase
FCH	familiäre kombinierte Hyperlipidämie
FGF	Fibroblastenwachstumsfaktor
FGFR4	FGF-Rezeptor 4
FH	familiäre Hypercholesterinämie
FHC	familiäre Hyperchylomikronämie
FHTG	familiäre Hypertriglyzeridämie
FMR1	„fragile X mental retardation 1“
FOX	Forkhead-Box
FTO	Fettmasse- und Fettleibigkeit-assoziiert
FXN	Frataxin
FXR	Farnesoid-X-Rezeptor
G6P	Glukose-6-Phosphat
G6PC	Glukose-6-Phosphatase
GATA	GATA-Bindeprotein
GCK	Glukokinase
GH1	Wachstumshormon 1

GHR	GH1-Rezeptor
GLP1	Glukagon-ähnliches Peptid 1
GLUT	Glukosetransporter
GMP	Granulozyten-Monozyten-Vorläufer
GPAT	Glyzerinphosphat-Acyl-Transferase
GPR	G-Protein-gekoppelter Rezeptor
GR	Glukokortikoidrezeptor
GSK3	Glykogensynthesekinasen 3
GSV	GLUT4-haltige Speichervesikel
GWAS	Genom-weite Assoziationsstudie
GYS	Glykogensynthase
HAT	Histonacetyltransferase
HBL	Hypobetalipoproteinämie
HDAC	Histondeacetylase
HDL	Lipoprotein hoher Dichte
HGPS	Hutchinson-Gilford-Progerie-Syndrom
HHEX	„hematopoietically expressed homeobox“
HIV	humanes Immunschwächevirus
HLD	Leberlipasemangel
HLP	Hyperlipoproteinämie
HMGCR	3-Hydroxy-3-methylglutaryl-CoA-Reduktase
HNF	Leberzellenkernfaktor
HNRNPU	heterogenes nukleares Ribonukleoprotein U
HOTAIR	„HOX transcript antisense RNA“
HP1	Heterochromatinprotein 1, wird vom <i>CBX5</i> -Gen kodiert
HSC	Hämatopoetische Stammzelle
HSF1	Hitzeschocktranskriptionsfaktor 1
HSP	Hitzeschockprotein
HTG	Hypertriglyceridämie
HTT	Huntingtin
IAP	intrazisternales A-Partikel
ICR	Kontrollregion der genetischen Prägung
IDF	Internationale Diabetes-Föderation
IDH	Isocitratdehydrogenase
IGF	Insulin-ähnlicher Wachstumsfaktor
IGF1R	IGF1-Rezeptor
IKBK	„inhibitor of nuclear factor kappa B kinase subunit beta“
IL	Interleukin
IL1RN	IL1-Rezeptor-Antagonist
ILP	insulinartiges Peptid
IMCL	intramyozelluläres Lipid
Indel	Insertion/Deletion
INF γ	Interferon γ
INO80	INO80-Komplexuntereinheit
INSR	Insulinrezeptor

iPOP	integrative persönliche Omik-Profilierstellung
iPS	induzierte pluripotente Stamm
IRF4	„interferon regulatory factor 4“
IRS	Insulinrezeptorsubstrat
IRX	Iroquois Homöobox 3
ISWI	„imitation switch“
K ^{ATP}	ATP-sensitiver K ⁺ -Kanal
kb	Kilobasenpaare (1000 bp)
KCNQ1	Kaliumspannungs-gesteuerter Kanal, Unterfamilie Q, Mitglied 1
KDM	Lysindemethylase
KLF4	Krüppel-ähnlicher Faktor 4
KMT	Lysinmethyltransferase
LAD	Lamin-assoziierten Domänen
LCAT	Lecithin-Cholesterin-Acyltransferase
LCATD	LCAT-Mangel
LCK	LCK-Proto-Onkogen, Tyrosinkinase der Src-Familie
LCT	Laktase
LDL	Lipoprotein niedriger Dichte
LDLR	LDL-Rezeptor
LEP	Leptin
LEPR	Leptinrezeptor
LINE	langes eingestreutes Element
LIPC	Lipase C
LPCAT3	Lysophosphatidylcholin-Acyltransferase 3
LPL	Lipoproteinlipase
LRH-1	Leberrezeptor-Homolog 1
LRP1	„LDL receptor related protein 1“
LSD1	Lysin-spezifische Demethylase 1, wird auch KDM1A genannt
LTR	„long terminal repeat“
LXR	Leber-X-Rezeptor
M-CFU	koloniebildende Einheiten
MAF	Häufigkeit des selteneren Allels
MAN2A1	„mannosidase alpha class 2A member 1“
MAPK	Mitogen-aktivierte Proteinkinase
Mb	Megabasenpaare (1.000.000 bp)
MBD	Methyl-DNA-Bindedomäne
MC1R	Melanocortin 1 Rezeptor
MC4R	Melanocortin-4-Rezeptor
MCM6	„minichromosome maintenance complex component 6“
mCpH	nicht-CpG-Methylierung
MDP	Makrophagen und Vorläufer von dendritischen Zellen
MECP2	Methyl-CpG-Bindeprotein 2
MEIS1	Meis-Homöobox 1
MEN1	Menin 1
MEP	Megakaryozyten-Erythrozyten Vorläufer

MGMT	O-6-Methylguanin-DNA Methyltransferase
MHC	Haupthistokompatibilitätskomplex
MHL	gemischte Hyperlipidämie
MIC	MHC-Klasse-I-Polypeptid-verwandte Sequenz
miRNA	micro RNA
MLH1	MutL-Homolog 1
MODY	Erwachsenendiabetes, der in der Jugend auftritt
MPO	Myeloperoxidase
MPP	multipotente Vorläuferzellen
MSR1	„macrophage scavenger receptor 1“
MTHFR	Methylentetrahydrofolatreduktase
MTNR1B	Melatoninrezeptor 1B
MTTP	mikrosomales Triglycerid-Transferprotein
MYLIP	„myosin regulatory light chain interacting protein“
MYO5A	Myosin VA
NAD	Nikotinamidadenindinukleotid
NADP	Nicotinamidadenindinukleotidphosphat
NAFLD	nichtalkoholische Fettlebererkrankung (Fettleber)
NAMPT	Nicotinamidphosphoribosyltransferase
NANOG	Nanog-Homöobox
NCEH1	„neutral cholesterol ester hydrolase 1“
NCOA1	Nuklear Rezeptor Koaktivator 1
NCOR	nuklearer Rezeptor Korepressor
ncRNA	nichtkodierende RNA
NEUROD1	„neuronal differentiation 1“
NFAT	„nuclear factor of activated T-cells“
NFE2L2	NFE2-ähnlicher BZIP-Transkriptionsfaktor 2
NFκB	nuklearer Factor κB
NGS	„next generation sequencing“
NK-Zelle	natürliche Killerzelle
NLRP	NLR-Protein
NO	Stickstoffmonoxid
NOX	NADPH-Oxidase
NPC1L1	„NPC1 like intracellular cholesterol transporter 1“
NPY	Neuropeptid Y
NSD	„nuclear receptor binding SET domain protein“
NTS	<i>Nucleus tractus solitarii</i>
OCA2	„OCA2 melanosomal transmembrane protein“
OCT4	Oktamer-bindender Transkriptionsfaktor 4, wird auch POU5F1 genannt
OGTT	oraler Glukosetoleranztest
OR	„odds ratio“
ORL1	„opioid related nociceptin receptor 1“
PAMP	Pathogen-assoziiertes molekulares Muster
PAX4	„paired box 4“

PBMC	mononukleäre Zelle des peripheren Bluts
PC	Pyruvatcarboxylase
PCK	Phosphoenolpyruvat-Carboxykinase
PCSK	Protein-Convertase
PDCD1	programmierter Zelltod 1, wird auch PD1 genannt
PDE3B	Phosphodiesterase 3B
PDGFRA	Wachstumsfaktor aus Thrombozyten-Rezeptor α
PDH	Pyruvat-Dehydrogenase
PDX1	„pancreatic and duodenal homeobox 1“
PER1	Periode zirkadiane Uhr 1
PFK2	Phosphofruktokinase 2
PFKFB2	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 2
PGC	Urkeimzelle
PGR	Progesteronrezeptor
PI3K	Phosphoinositid-3'-Kinase
PIP3	Phosphatidylinositol-3,4,5-trisphosphat
PKA	Proteinkinase A
PLAU	„plasminogen activator, urokinase“
Pol II	RNA-Polymerase II
POMC	Proopiomelanocortin
POU	POU Homöobox
PPAR	Peroxisom-Proliferator aktivierter Rezeptor
PPARGC1 α	PPAR γ , Koaktivator 1 α
PRC	Polycomb-Repressionskomplex
PRKC	Proteinkinase C
PROP1	„PROP paired-like homeobox 1“
PTEN	Phosphatase und Tensin Homolog
PU.1	Purin-reiche Box 1
PUFA	mehrfach ungesättigte Fettsäure
PXR	Pregnan X Rezeptor
RAPTOR	regulatorisches assoziiertes Protein von TOR
RAR	Retinsäurerezeptor
RB1	RB transkriptioneller Korepressor 1
RBP4	Retinolbindungsprotein 4
RCOR	REST Korepressor
RE	„response element“
REL	REL Protoonkogen, NF κ B-Untereinheit
REST	RE1 unterdrückender Transkriptionsfaktor, wird auch NRSF genannt
ROS	reaktive Sauerstoffspezies
rRNA	ribosomale RNA
RTK	Rezeptortyrosinkinase
RUNX1	„runt-related transcription factor 1“
RXR	Retinoid X Rezeptor
S6K	S6-Kinase

SAH	S-Adenosylhomocystein
SAM	S-Adenosylmethionin
SCAP	„SREBP cleavage-activating protein“
SCARB1	Scavenger-Rezeptor B1
SCD1	Steroyl-CoA-Desaturase 1
SCN	<i>Nucleus suprachiasmaticus</i>
Serpin	Serin-Peptidase-Inhibitor
SETD2	„SET domain containing 2“
SF-1	steroidogener Faktor 1
SFRP5	„secreted frizzled related protein 5“
SHARP	SMRT/HDAC1-assoziiertes Repressorprotein
SIM1	SIM BHLH Transkriptionsfaktor 1
SINE	kurzes eingestreutes Element
SIRT	Sirtuin
SITO	Sitosterolämie
SLE	systemischer Lupus Erythematodes
SMARC	„SWI/SNF-related matrix-associated actin-dependent regulators of chromatin“
SMYD2	„SET and MYND domain containing 2“
SNP	Einzelnukeotidpolymorphismus
SNV	Einzelnukeotidvariante
SOX2	SRY-Box 2
SREBF1	„sterol regulatory element binding transcription factor 1“
STAT	Signaltransduktor und Aktivator der Transkription
SUV39H	„suppressor of variegation 3–9 homolog“
SWI/SNF	„switching/sucrose non-fermenting“
T1D	Typ-1-Diabetes
T2D	Typ-2-Diabetes
TAD	Topologisch-assoziierte Domäne
TAS1R	„taste 1 receptor“
TBC1D	TBC1-Domäne
TCFL2	„transcription factor 7 like 2“
TCGA	Der Krebsgenomatlas
TCR	T Zell Rezeptor
TD	Tangier-Krankheit
TDG	Thymin-DNA-Glykosylase
TERT	Telomerase
TET	„ten-eleven translocation“
TGF	Tumorstromungsfaktor
T _H	T-Helfer
THF	Tetrahydrofolsäure
THRSP	Schilddrüsenhormon-responsiv
TIFIA	Transkriptionsinitiationsfaktor IA
TLR	Toll-like-Rezeptor
TNF	Tumornekrosefaktor

TNFR	TNF-Rezeptor
TOR	„target of rapamycin“
TP53	Tumorprotein p53
TRAF2	„TNF receptor associated factor 2“
T _{REG}	Regulatorische T-Zelle
TRIM	„tripartite-motif-containing protein“
tRNA	transfer RNA
TSC2	Tuberöse Sklerose 2
TSS	Transkriptionsstartstelle
TYR	Tyrosinase
UBR1	„ubiquitin protein ligase E3 component N-recogin 1“
UCP	Entkopplungsprotein
UDP	Uridindiphosphat
UHRF1	Ubiquitin-ähnliche pflanzliche Homöodomäne und RING-Finger-Domäne 1
UPR	„unfolded protein response“
UTR	nichttranslatierte Region
VDR	Vitamin D Rezeptor
VLDL	Lipoprotein sehr niedriger Dichte
WHO	Weltgesundheitsorganisation
WHR	Quotient aus Taillen- und Hüftumfang
XBP1	X-Box Bindungsprotein 1
<i>Xist</i>	X-inaktives spezifisches Transkript
ZNS	zentrales Nervensystem
α -MSH	α -Melanozyten-stimulierendes Hormon



Zusammenfassung

In diesem Kapitel wird zunächst die genetische Anpassung des anatomisch modernen Menschen (*Homo sapiens*) durch Wanderung von Afrika in neue geografische und klimatische Umgebungen in Asien, Europa und Amerika beschrieben. Dazu gehören auch die Herausforderungen, die der Wandel von Jägern und Sammlern hin zu Bauern mit sich gebracht haben. Genetische Unterschiede zwischen menschlichen Populationen sind am deutlichsten in jenen Organen ausgeprägt, die in direktem Kontakt zur Umwelt stehen, wie der Haut, dem Verdauungstrakt oder dem Immunsystem. **Das hat nicht nur zu Unterschieden in der Hautfarbe, sondern auch zu unterschiedlichen Resistenzen gegenüber Krankheiten und einer Diversität der Nahrungspräferenzen und -verträglichkeiten geführt.** Die genetischen Grundlagen der Variation menschlicher Populationen und Individuen wurde von großen Forschungskonsortien, wie dem 1000-Genom-Projekt, untersucht und katalogisiert. Genom-weite Genotypisierung und Sequenzierung des gesamten Genoms ermöglichen die Untersuchung und Analyse komplexer Krankheiten, wie T2D und Herz-Kreislauf-Erkrankungen, auf der Grundlage von Dutzenden bis Hunderten genetischer Varianten, wie Einzelnukleotidvarianten (SNVs) und Kopienzahlvarianten (CNVs).

1.1 Migration und evolutionäre Herausforderungen des *Homo sapiens*

Vor etwa 300.000 Jahren entwickelte sich in Ostafrika der anatomisch moderne Mensch (*Homo sapiens*), d. h. Menschen wie heute. Vor etwa 50–75.000 Jahren begannen einige dieser modernen Menschen, nach Asien, Europa und Amerika zu

wandern und ersetzt durch Durchmischung¹ zuvor vorherrschende archaische² Hominiden, wie den Neandertaler (*Homo neanderthalensis*) und den Denisova-Menschen (Abb. 1.1). Beide Arten hatten sich bereits über 400.000 Jahre genetisch an die Umweltbedingungen Eurasiens angepasst. Die Vermischung führte zur Übertragung von Genvarianten und verbesserte die Überlebensfähigkeit außerhalb Afrikas, wie eine Anpassung der Thermoregulation in kälteren Klimazonen, Toleranz gegenüber Sauerstoffmangel (Hypoxie) in großer Höhe und hellerer Haut wegen geringer UV-B-Strahlung in nördlichen Breiten, die zu einem Vitamin D Mangel führen kann (Abschn. 13.1).

Vor etwa 10.000 Jahren begannen unsere Vorfahren, ihre Jäger- und Sammlergewohnheiten aufzugeben und wurden Bauern. Diese signifikante Änderung des Lebensstils basiert zum großen Teil auf neu genutzten Nahrungsmitteln, wie Getreide und Milch (Abschn. 6.1). **Die verbesserte Nahrungsversorgung ermöglichte eine höhere Bevölkerungsdichte, ging jedoch mit einer erhöhten Belastung mit Infektionskrankheiten einher, von denen viele von domestizierten Tieren ausgingen.** Sowohl Ernährungsumstellungen als auch immunologische Herausforderungen verursachten einen dominanten evolutionären Druck und eine verhältnismäßig schnelle genetische Anpassung. Die Folgen dieser genetischen Anpassungen prägten nicht nur die biologische Vielfalt, sondern auch die Gesundheit und das Krankheitsrisiko von uns derzeit 8 Mrd. Menschen.

Weltweit dauerte es mehrere tausend Jahre, also deutlich mehr als 100 Generationen, bis die meisten Menschen von Jägern und Sammlern zu Bauern wurden, aber nur weniger als 50 Jahre, um sie zu Nutzern von Autos und Supermärkten sowie zu Konsumenten von Fastfood zu machen. Der Mensch hatte also ganz einfach keine Zeit, sich genetisch an einen Lebensstil anzupassen, der zu **Fettleibigkeit** (Abschn. 9.1) verleitet und häufig zu weiteren Krankheiten führt, die als **metabolisches Syndrom** (Abschn. 10.3) zusammengefasst werden. Die Kombination aus vorwiegend sitzenden Tätigkeiten und geringer körperlicher Aktivität sowie der leichten Verfügbarkeit energiereicher Nahrung ist eine große Herausforderung für unseren Körper, der von der Evolution auf effiziente Energiespeicherung und große körperliche Mobilität ausgerichtet ist. Deshalb können viele gerade jener genetischen Varianten, die die molekulare Basis eines solcherart optimierten Körpers darstellen, heute zu einem erhöhten Risiko für die Entwicklung chronischer nichtübertragbarer Krankheiten, wie **T2D** (Abschn. 9.6) oder **Herz-Kreislauf-Erkrankungen** (Abschn. 10.1), beitragen.

¹Deshalb besteht das Genom von Europäern bis zu 2,1 % aus DNA-Sequenzen, die vom Neandertaler stammen.

²Heutzutage ausgestorben.

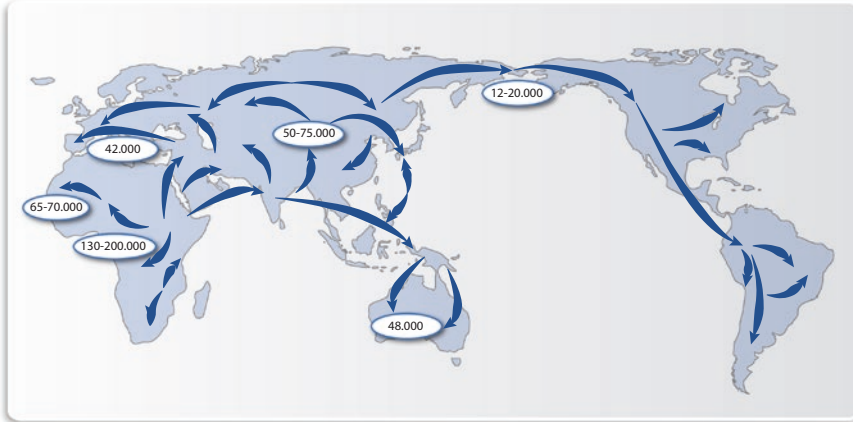


Abb. 1.1 Wanderungen des *Homo sapiens*. Der Ausbreitung des anatomisch modernen Menschen von Ostafrika über den Rest des afrikanischen Kontinents folgte vor etwa 50–75.000 Jahren eine Expansion nach Asien, wahrscheinlich sowohl auf südlichen als auch auf nördlichen Routen. Ozeanien, Europa und Amerika wurden, in dieser Reihenfolge, von Asien aus besiedelt. Die Migrationsmuster basieren hauptsächlich auf Analysen von Veränderungen mitochondrialer DNA

1.2 Vielfalt menschlicher Populationen

Nachdem sich der Mensch auf alle Kontinente verteilt hatte, wurden einige Bevölkerungsgruppen geografisch isoliert, sodass neue Genvarianten nicht mehr auf alle Artgenossen übertragen wurden. Da sich in der Vergangenheit weit voneinander entfernt lebende Bevölkerungsgruppen in der Regel nicht trafen und vermischen konnten, sind Populationen genetisch umso unterschiedlicher, je weiter voneinander entfernt sie sich entwickelt haben. Da der anatomisch moderne Mensch bereits seit rund 300.000 Jahren in Afrika lebt, findet sich auf diesem Kontinent eine größere genetische Vielfalt als im Rest der Welt. **Trotzdem, obwohl es offensichtliche phänotypische Unterschiede von Bevölkerungsgruppen in Bezug auf Hautfarbe, Körpergröße und Gesichtsmerkmale gibt, findet man keine absoluten genetischen Unterschiede zwischen ihnen.** So gibt es, wie es die „Jenaer Erklärung“ von 2019 ausdrückt, „im menschlichen Genom unter den gut 3 Mrd. Basenpaaren keinen einzigen fixierten Unterschied, der zum Beispiel Afrikaner von Nichtafrikanern trennt“. **Bevölkerungsunterschiede basieren auf Varianten an Tausenden von Stellen im Genom.** Das impliziert, dass sich eine Eigenschaft³, wie z. B. die Hautfarbe (Abschn. 13.1), in

³Eine Eigenschaft wird oft als „Merkmal“ bezeichnet.

einer Population ziemlich schnell ändern kann, wenn sich die Häufigkeiten eines Allels⁴, das zum Merkmal beiträgt, verschiebt (Box 1.1).

Box 1.1: Natürliche Selektion

Die positive natürliche Selektion, d. h. die Kraft, die die Verbreitung vorteilhafter Eigenschaften antreibt, hat in der Vergangenheit während der Evolution des Menschen eine zentrale Rolle gespielt. Individuen mit Vorteilen (sogenannten **adaptiven Merkmalen**) sind tendenziell erfolgreicher in ihrer Fortpflanzung, d. h. sie tragen mit mehr Nachkommen zur nächsten Generation bei als andere. Aufgrund der Vererbung von einer Generation zur anderen, erhöht der Selektionsprozess die Prävalenz⁵ des jeweiligen adaptiven Merkmals. Unter anhaltendem Selektionsdruck können solche adaptiven Merkmale Schritt für Schritt in der betroffenen Bevölkerungsgruppe universell werden. Zu den Faktoren, die die Selektion begünstigen, also zum sogenannten **evolutionären Druck** beitragen, gehören z. B. Ressourcenbeschränkungen, wie Nahrungsknappheit, oder Bedrohungen, wie das Auftreten von Krankheitserregern. Eine Veränderung in der Häufigkeit einer Genvariante in einer Population kann auch durch zufällige Veränderungen erfolgen, also durch eine genetische „Drift“, die nicht auf Selektion beruht (Abschn. 13.2). Auf diese Weise kann sogar ein schädliches Allel für die Mitglieder einer Population universell werden, z. B. unter dem Einfluss einer schwachen Selektion oder in kleinen Populationen, in denen sich zufällige Änderungen stärker auswirken.

Während der Evolution des Menschen unterlagen bis zu 10 % aller proteinkodierenden Gene, also etwa 2000 Gene, einer positiven natürlichen Selektion. Insbesondere Gene, die für das Immunsystem, den Verdauungstrakt und die Haut (einschließlich Haaren, Schweißdrüsen und Sinnesorganen) wichtig sind, waren von einer positiven Selektion betroffen (Abschn. 13.1). Das liegt daran, dass diese Organsysteme in größerem Maße in Kontakt mit der Umwelt stehen als andere Teile unseres Körpers. Eine solche positive Selektion finden sich beim angeborenen und adaptiven Immunsystems (Box 1.2), etwa in Form von Varianten von Genen, die für Membranimmunrezeptoren kodieren und unter positivem Selektionsdruck durch Wechselwirkung des Körpers mit pathogenen Mikroben und Viren standen. Ein interessantes Beispiel ist das Membranprotein CCR5 (C–C-Motiv-Chemokinrezeptor 5), das als Korezeptor für das humane Immun-

⁴Ein Allel ist die exakte Sequenz an einer definierten Stelle im Genom. In unserem diploiden Genom haben wir in der Regel 2 Kopien eines Allels. Ausnahmen sind das X und Y Chromosom bei Männern.

⁵Prävalenz beschreibt Mengen von Personen in einem definierten Zustand (z. B. des Krankseins).

schwächevirus (HIV) fungiert und für den Eintritt des Virus in T-Zellen essenziell ist. Eine vermutlich seit mehreren Tausend Jahren in menschlichen Populationen auftretende Deletion im *CCR5*-Gen (32 Basenpaare (bp) fehlen) führt dazu, dass CCR5 Protein nicht auf T-Zellen exprimiert⁶ wird und dass homozygote⁷ Träger dieser Mutation vor einer HIV-Infektion geschützt sind. Diese Mutation steht in Populationen, in denen HIV-Infektionen in größerem Umfang auftreten, wie in Südafrika, derzeit unter positivem Selektionsdruck.

Box 1.2: Angeborenes und adaptives Immunsystem

Das Immunsystem besteht aus biologischen Strukturen wie Lymphknoten, Zelltypen wie Monozyten, Makrophagen, T- und B-Zellen und Proteinen, wie Komplementproteinen und Antikörpern, die den Organismus vor Infektionskrankheiten schützen. Das Immunsystem erkennt verschiedenste Moleküle, sogenannte **Antigene**, potenziell pathogenen Ursprungs, beispielsweise auf der Oberfläche von Mikroben, und unterscheidet sie von körpereigenen Strukturen des gesunden Gewebes. Man unterscheidet angeborenes und adaptives Immunsystem (Abschn. 8.1). Das **angeborene Immunsystem** ist evolutionär älter, und neben vor Infektionen schützenden anatomischen Barrieren, wie Haut und Schleimhäuten, basiert es auf Zellen, wie natürlichen Killerzellen (NK-Zellen), Makrophagen und Neutrophilen. Neben dem Mechanismus der Phagozytose⁸ dienen im angeborenen Immunsystem auch antimikrobielle Peptide, beispielsweise aus dem Komplementsystem, der Abwehr von Krankheitserregern. Das **adaptive Immunsystem** wendet ausgefeilte Abwehrmechanismen an, bei denen T- und B-Zellen hochspezifische Oberflächenrezeptoren wie T-Zell-Rezeptoren (TCRs) und B-Zell-Rezeptoren gegen Antigene verwenden. Die B-Zell-Rezeptoren verwandeln sich in sezernierte Antikörper, wenn B-Zellen in Plasmazellen differenzieren. Darüber hinaus erzeugt das adaptive Immunsystem im Anschluss an eine spezifische Reaktion auf ein Antigen ein **immunologisches Gedächtnis**, das zu einer verstärkten Reaktion auf nachfolgende Begegnungen mit demselben Antigen führt.

Weitere Beispiele für Gene, die positivem Selektionsdruck unterliegen, sind einige Allele, die durch Kreuzung mit archaischen menschlichen Spezies in das

⁶Im Zusammenhang mit Genregulation bedeutet „exprimiert“, dass die Transkription eines Gens ermöglicht oder verstärkt wird, sodass mehr Protein produziert wird.

⁷Homozygot bedeutet an einer bestimmten Stelle im Genom die identische Sequenz in beiden Kopien zu haben.

⁸Phagozytose ist ein aktiver Transportprozess, bei dem eine Zelle feste Partikel oder andere Zellen aus dem Extrazellulärraum aufnimmt, indem sich die Zellmembran einstülpt.

Genom des modernen Menschen gelangten⁹. So wurde in Populationen aus Mexiko und Lateinamerika eine mit T2D assoziierte Variante des *SLC16A11*¹⁰-Gens, das für einen Lipidtransporter im endoplasmatischen Retikulum (ER) kodiert, identifiziert. Diese Variante lässt sich auf Beimischung von Neandertaler-Genomabschnitten zurückführen und findet sich häufiger in amerikanischen Ureinwohnern (die ursprünglich aus Sibirien eingewandert sind).

Als vor etwa 500 Jahren das Überqueren der Ozeane möglich wurde, begann eine Migration eines bis dahin nicht gekannten Ausmaßes, die insbesondere in Amerika, aber auch auf anderen Kontinenten, zu erheblichen Bevölkerungsvermischungen führte. Zuvor hatte es in Europa mindestens zwei große Phasen der Vermengung von Populationen gegeben, bei denen sich zuerst vor etwa 9000 Jahren Bauern aus Anatolien mit einheimischen europäischen Jägern und Sammlern vermischten und dann vor etwa 5000 Jahren Yamnaya-Hirten aus der eurasischen Steppe nach Europa einwanderten. Darum ist der Phänotyp heutiger Europäer größtenteils das Produkt der mittel- und jungsteinzeitlichen Kollisionen dieser drei Vorgängerpopulationen.

Hunderte komplexer Merkmale bestimmen den Phänotyp, d. h. wie eine Person aussieht und sich verhält, sowie ihre Risiken, nichtübertragbare Krankheiten zu entwickeln. Darüber hinaus basiert jedes komplexe Merkmal auf Dutzenden bis Hunderten von Genvarianten und Umwelteinflüssen. Die meisten dieser genetischen Varianten sind neutral, d. h. sie tragen nicht zu phänotypischen Unterschieden oder Krankheitsrisiken bei und reichern sich innerhalb der jeweiligen Populationen zufällig an (Box 1.1). Hinsichtlich der Ausprägung eines Phänotyps kann einerseits die Gesamtheit zwar seltener, sich jedoch stark auswirkender Varianten (also Varianten von hoher „Penetranz“) einen maßgeblichen Einfluss haben. Das gilt z. B. bei monogenetischen Erkrankungen (Abb. 1.2, links). Andererseits gibt es darüber hinaus eine große Anzahl häufigerer Varianten mit kleiner bis mäßiger Effektstärke („odds ratio“ (OR)), die bei gängigen komplexen Merkmalen eine dominante Rolle spielen. Das trifft für das genetische Risiko für die meisten Volkserkrankungen zu (Abb. 1.2, rechts).

Die verschiedenen Typen genetischer Varianten (auch Polymorphismen genannt) werden nach der Häufigkeit des selteneren Allels (MAF) in der Population eingeteilt. Varianten werden als häufig bezeichnet, wenn die MAF mindestens 1 % beträgt, oder als selten, wenn sie eine MAF von weniger als 1 % haben. SNVs stellen die häufigste Klasse genetischer Varianten dar und etwa 7 Mio. bekannte SNVs weisen in der Gesamtheit untersuchter Populationen eine MAF von mehr als 5 % auf¹¹ und werden als SNPs¹² (Einzelnukleotidpolymorphismen) bezeichnet. Der Einfluss von SNVs auf

⁹ Bis zu 2,1 % unseres Genoms stammt vom Neandertaler.

¹⁰ SLC = „solute carrier family.“

¹¹ www.ncbi.nlm.nih.gov/snp.

¹² SNP = Einzelnukleotidpolymorphism.

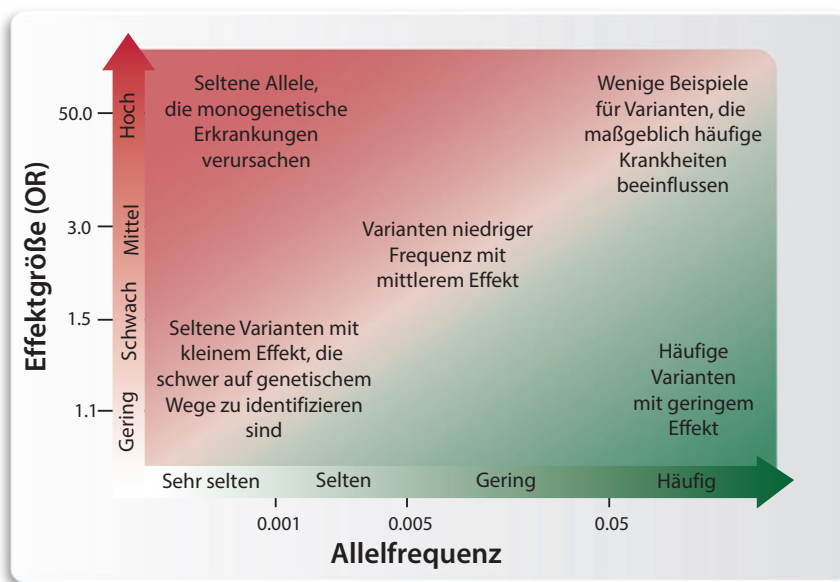


Abb. 1.2 Identifizierung genetischer Varianten nach Häufigkeit des Risikoallels. In der Grafik ist die Effektstärke (OR) über der Häufigkeit einer genetische Variation (Allelfrequenz) aufgetragen

die proteinkodierende Sequenz des menschlichen Genoms ist gut untersucht. **Synonyme Mutationen** verändern das kodierte Protein nicht, während **nicht-synonyme Mutationen** eine Veränderung der Aminosäuresequenz („missense“) bewirken oder ein vorzeitiges Stoppcodon („nonsense“) für die Proteintranslation erzeugen. Letzteres bedeutet, dass ausgehend von der Position der Variante die gesamte Aminosäuresequenz des kodierten Proteins verändert wird. Zusätzlich können Indels als auch CNVs in Exonsequenzen zu einer Verschiebung des Leserasters der Proteintranslation führen. Darüber hinaus können CNVs in Intronsequenzen zu alternativem Spleißen führen.

Das Merkmal „Körpergröße“ ist beispielsweise von mindestens 180 Positionen im Genom abhängig, d. h. es ist ein Paradebeispiel für ein komplexes polygenes Merkmal (Abb. 1.2, rechts) (für weitere Beispiele siehe Abschn. 9.6 und Abschn. 10.3). In Europa hat sich die Körpergröße in den letzten Generationen, nicht zuletzt unter dem Einfluss von Umweltfaktoren wie verbesserter Qualität und Quantität der Ernährung, erheblich verändert (im Durchschnitt eine Zunahme um 10 cm).

1.3 Varianten des menschlichen Genoms

Das menschliche Genom kann in sogenannte Haplotypblöcke unterteilt werden. Das sind Abschnitte unseres Genoms von typischerweise 10–100 kb (Kilobasenpaare) Länge, die von Generation zu Generation in der gleichen Form vererbt werden. Das bedeutet, dass diese Abschnitte des Genoms bisher nicht durch Rekombination während der Meiose¹³, unterbrochen wurden (Abb. 1.3). Im Gegensatz hierzu stellen die Grenzen der Haplotypblöcke Rekombinationsereignisse dar, die vor vielen Generationen stattgefunden haben. **Die Veränderung eines Gens während der Meiose ist etwa 100-mal häufiger als das Entstehen von Punktmutationen und daher ein effizienter Mechanismus für das Fortschreiten der Evolution.** Das zeigt die wichtige Bedeutung sexueller Fortpflanzung für eine effiziente Evolution.

Da afrikanische Populationen schon viel länger existieren als europäische und asiatische (Abb. 1.1), sind ihre Haplotypblöcke kürzer, d. h. in Afrika hatten die Blöcke aufgrund der Häufung von Rekombinationsereignissen durch die höhere Anzahl von Generationen mehr Zeit, zerteilt zu werden. Zudem durchliefen alle nichtafrikanischen Menschen eine Art demografischen Engpass, da sie von einer kleinen Population ostafrikanischer Herkunft abstammen. **Das impliziert, dass die Genome von Nichtafrikanern eine geringere Vielfalt haben.**

Das haploide¹⁴ Referenzgenom des Menschen (Box 1.3) wurde 2001 von den Mitgliedern des Humangenomprojekts (Box 1.4) veröffentlicht und basiert auf wenigen männlicher DNA-Spender. Das Genom jedes Menschen unterscheidet sich an durchschnittlich 4–5 Mio. Positionen von dieser Referenzsequenz. Ein Großteil davon (3,5–4,3 Mio.) geht auf SNVs zurück, also auf Sequenzvarianten, bei denen genau eine Kernbase (A, C, G oder T)¹⁵ verändert ist (Abb. 1.4). Hinzu kommen Strukturvarianten des Genoms (ca. 0,5–0,6 Mio.), die zwar seltener sind, jedoch insgesamt mehr Basenpaare betreffen. Zu den Strukturvarianten gehören vor allem Insertionen, Deletionen, CNVs und Inversionen.

Box 1.3: Das Genom des Menschen

Die vollständige Sequenz des anatomisch modernen Menschen (*Homo sapiens*) und wurde vom „Human Genome Project“ (Box 1.4) entschlüsselt¹⁶. Dieses Referenzgenom ist über verschiedene Genom-Browser

¹³ Meiose ist eine spezielle Form der Zellteilung, die nur in Keimzellen vorkommt und die Genome der Elterngeneration durchmischt.

¹⁴ Haploid bezieht sich auf eine einzelne Kopie des Genoms. Mit Ausnahme von Keimzellen haben unsere Zellen jedoch 2 Kopien des Genoms (von Vater und Mutter) in ihrem Kern, sie sind also diploid.

¹⁵ In DNA gibt es die vier Kernbasen Adenin (A), Guanin (G), Cytosin (C) und Thymin (T) sowie in RNA Uracil (U) statt T.

¹⁶ www.genome.gov/10001772.

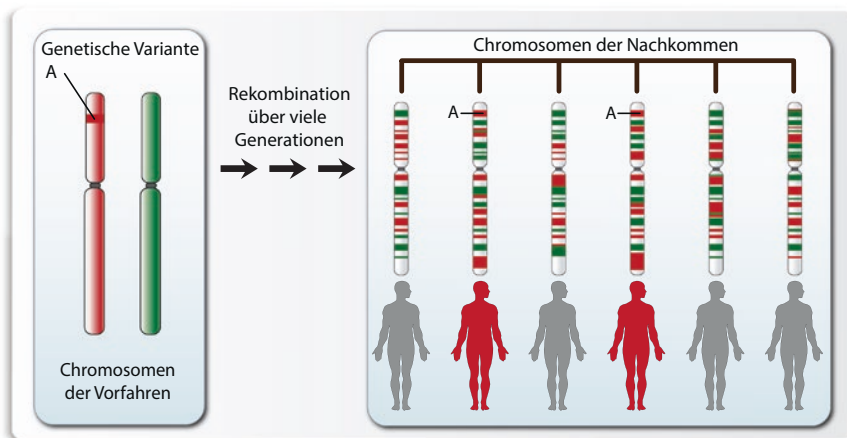


Abb. 1.3 Haplotypen. Zwei vereinfachte Beispielchromosomen von Vorfahren werden durch Rekombination während der Meiose über viele Generationen hinweg vermischt, um die verschiedenen Chromosomen ihrer Nachkommen zu ergeben. Ein typisches Chromosom hat nach 30.000 Jahren mehr als ein Rekombinationsereignis pro 100 kb durchlaufen. Die so entstehenden Blöcke gleichzeitig vererbter genetischer Varianten werden als Haplotypen bezeichnet. Steigt mit einer genetischen Variante (hier gekennzeichnet durch ein A) auf einem angestammten Chromosom das Risiko für eine bestimmte Krankheit, haben die Individuen der aktuellen Generation, die diese Region des Ahnenchromosoms geerbt haben (rot), ein erhöhtes Erkrankungsrisiko. Innerhalb des Haplotypblocks, der die krankheitsverursachende Variante trägt, gibt es viele SNVs, die verwendet werden können, um den genauen Ort der Variante zu identifizieren

im Internet zugänglich und basiert auf 7 jungen männlichen Spendern aus den USA. Mit Ausnahme von Keimzellen, also weiblichen Eizellen und männlichen Spermien, die nur ein haploides Genom haben, und Erythrozyten¹⁷, die keinen Zellkern mehr haben, enthält jede menschliche Zelle ein diploides Genom. Dieses Genom besteht aus $2 \times 3,05$ Mrd. bp, also 3050 Mb¹⁸, und ist auf 2×22 autosomalen Chromosomen sowie zwei X Chromosomen (bei Frauen) oder einem XY-Chromosomensatz (bei Männern) verteilt. Darüber hinaus enthält jedes Mitochondrium 16,6 kb mitochondriale DNA. Das haploide Genom enthält etwa 20.000 proteinkodierende Gene und etwa die doppelte Anzahl von ncRNA (nichtkodierenden RNA)-Genen. Die proteinkodierende Sequenz deckt weniger als 2 % unseres Genoms ab, d. h. **der größte Teil des Genoms ist nichtkodierend und scheint hauptsächlich eine regulatorische Funktion zu haben.**

¹⁷ Erythrozyten sind rote Blutkörperchen, sie repräsentieren 85 % der Zellen in unserem Körper.

¹⁸ 1 Mb = 1.000.000 bp.

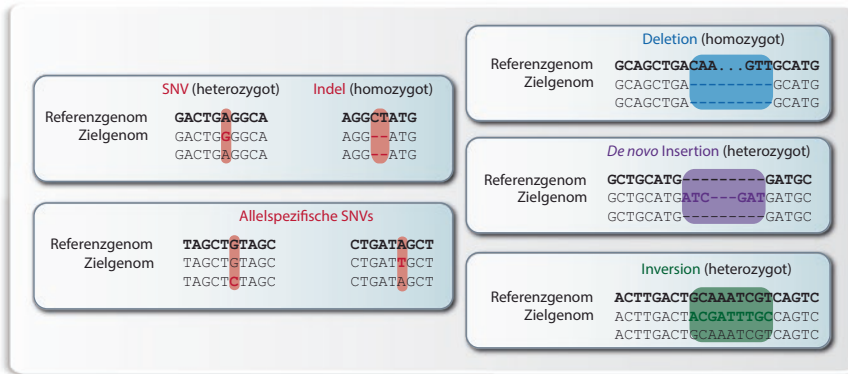


Abb. 1.4 Arten von Varianten in Genomsequenzen. Die jeweils obere Sequenz stellt das haploide Referenzgenom dar, während die beiden darunter stehenden Sequenzen die zwei homologen Sequenzabschnitte (Allele) des diploiden Genoms eines Individuums zeigen. Genetische Varianten können entweder heterozygot oder homozygot sein. Der Begriff der Allelspezifität bezieht sich auf die genaue Zuordnung einer vorliegenden Variante zu einem konkreten Haplotyp, d. h. zu einer der beiden Chromosomen-Kopien im diploiden Genom

Fast 50 % der Sequenz unseres Genoms wird von repetitiver DNA¹⁹ gebildet, die in die folgenden Kategorien eingeteilt wird (in der Reihenfolge ihrer Häufigkeit):

- Lange eingestreute Elemente (LINEs, 500–8000 bp): 20,71 %
- Kurze eingestreute Elemente (SINEs, 100–300 bp): 12,79 %
- Retrotransposonen²⁰, wie „long terminal repeats“ (LTRs, 200–5000 bp): 8,85 %
- Minisatelliten und Microsatelliten (2–100 bp): 4,93 %
- Satelliten²¹ (200–2000 bp): 2,54 %

LINEs und SINEs sind identische oder nahezu identische DNA-Sequenzen, die durch eine große Anzahl von Nukleotiden getrennt sind, d. h. die Wiederholungen sind über das gesamte Genom verteilt. LTRs sind durch

¹⁹Repetitive DNA besteht aus vielen sich wiederholenden Sequenzmotiven.

²⁰Ein Retrotransposon ist ein DNA-Sequenzabschn, der Ähnlichkeit mit dem Genom von Retroviren hat und höchstwahrscheinlich von diesen abstammt.

²¹Der Begriff „Satellit“ stammt vom Laufverhalten bestimmter DNA-Regionen in der Gelelektrophorese. Diese nichtkodierende DNA-Regionen sind über das gesamte Genom verstreut und bestehen aus einer variablen Anzahl von Wiederholungen eines bestimmten Sequenzmotivs. Je nach Größe werden sie als „Satelliten“, „Minisatelliten“ und „Microsatelliten“ bezeichnet.

Sequenzen gekennzeichnet, die an jedem Ende von Retrotransposonen zu finden sind. DNA-Transposonen sind autonome DNA-Sequenzen, die für eine Transposase kodieren, d. h. für ein Enzym, das DNA von einer an eine andere Position im Genom transponiert. Mikrosatelliten sind oft im Zentrum des Chromosoms (dem Zentromer) zu finden und werden durch Tandem-Wiederholungen von 2–10 bp Länge gebildet. Minisatelliten und große Satelliten sind länger (10–60 bp bzw. bis zu 100 bp).

Box 1.4: Internationale Genombiologie-Konsortien („Big Biology Projects“)

Mit rund 20 Jahren Verspätung folgten Molekularbiologen dem Beispiel von Physikern und erkannten, dass einige ihrer Forschungsziele nur durch multi-nationale Kooperationen von Dutzenden bis Hunderten von Forschungsteams und Institutionen in sogenannten „Big Biology“-Projekten erreicht werden konnten. Das erste Beispiel war das „Humangenomprojekt“²², das 1990 gestartet und 2003 abgeschlossen wurde. Zusammen mit Folgestudien hatte das Projekt einen enormen Einfluss auf das Verständnis der Architektur und Funktion menschlicher Gene. Das „HapMap“-Projekt²³ war eines dieser Folgeprojekte, das von den technischen Fortschritten in der Genotypisierung profitierten. Parallel dazu ermöglichten verbesserte Sequenzierungsmethoden²⁴ Bestimmung und Vergleich von Genomen gesunder und (beispielsweise an Krebs) erkrankter Personen. Das mündete in groß angelegte Genomsequenzierungsprogramme wie dem „1000-Genom-Projekt“²⁵ und dem „Der Krebsgenatlas“-Konsortium („*The Cancer Genome Atlas Program*“, TCGA)²⁶. Darüber hinaus konzentrierten sich das „Enzyklopädie der DNA-Elemente“ (ENCODE)-Projekt²⁷ und das Projekt FANTOM5 („*Functional Annotation of the Mammalian Genome*“)²⁸ auf die funktionelle Charakterisierung des menschlichen Genoms. Das ENCODE-Nachfolgeprojekt „*Roadmap Epigenomics*“²⁹ lieferte Referenzsequenzen zum menschlichen Epigenom von mehr als 100 primären menschlichen Geweben und Zelltypen.

²² www.genome.gov/adipokine1772.

²³ <http://hapmap.ncbi.nlm.nih.gov>.

²⁴ NGS = „next-generation sequencing“.

²⁵ www.1000genomes.org.

²⁶ <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>.

²⁷ www.genome.gov/encode.

²⁸ <http://fantom.gsc.riken.jp/5>.

²⁹ www.roadmapepigenomics.org.