Aus dem Lübecker Institut für experimentelle Dermatologie (LIED)

der Universität zu Lübeck

Direktor: Prof. Dr. med. Ralf Ludwig

Klassifikation und Kontrolle von RNA

Inauguraldissertation

zur Erlangung der Doktorwürde der Universität zu Lübeck

- Aus der Sektion Medizin -

vorgelegt von

Mareike Witte

aus Hannover

Lübeck 2016

1. Berichterstatter: Prof. Dr. med. Ralf Ludwig

2. Berichterstatter: Prof. Dr. rer. nat. Frank Kaiser

3. Berichterstatter: Prof. Dr. med. T. Tüting

Tag der mündlichen Prüfung: 11.05.2017

Zum Druck genehmigt. Lübeck, den 11.05.2017

-Promotionskommission der Sektion Medizin-

Klassifikation und Kontrolle von RNA

Einleitung

Ribonukleinsäure (RNA) ist eine Nukleinsäure, die sich als Polynukleotid aus einer Kette von Nukleotiden zusammensetzt. Sie ist an verschiedenen biologischen Prozessen beteiligt und lässt sich anhand ihrer Funktion in verschiedene Subklassen unterteilen. Als messenger-RNA (mRNA) codiert sie genetische Information, daher ist mRNA auch als „codierende“ RNA bekannt. Daneben existiert die sogenannte „nicht-codierende“ RNA (ncRNA), deren Aufgabenspektrum vielfältig und teilweise noch nicht verstanden ist.

Beispiele für ncRNA sind ribosomale RNA (rRNA), transfer-RNA (tRNA), mikro-RNA (miRNA), small nuclear RNA (snRNA) und small nucleolar RNA (snoRNA). Einige dieser Arten von ncRNA sind an zentralen zellulären Prozessen beteiligt, wie zum Beispiel der Translation. So bestehen Ribosomen, in denen die Proteinbiosynthese stattfindet, aus rRNA-Molekülen. In den Ribosomen werden die Codons der mRNA durch die tRNA erkannt und in die entsprechenden Aminosäuren umgesetzt. Das geschieht mithilfe des Anticodons in der tRNA, einer Region aus drei Nukleotiden, die die zu einem Codon komplementären Basen besitzt. Andere Arten von ncRNA sind an post-transkriptionellen Prozessen wie der Modifikation oder Replizierung von DNA beteiligt. Dazu zählen snoRNA, snRNA, Ribonuclease P (RNaseP), Ribonuclease MRP (RNase MRP), Y RNA und telomerase RNA: SnoRNAs sind meist in den Nukleoli anzutreffen. Sie führen Modifikationen in die rRNA ein, die essentiell für die Funktion des Ribosoms sind. SnRNAs sind an dem Prozess des Splicings beteiligt, sie bilden nach ihrer Synthese kleine Ribonukleoproteine (snRNPs). RNase P ist ein Enzym, das eine Vorläufersequenz von tRNA abspaltet. Daher ist es für die Funktion der tRNA essentiell. Daneben ist sie an der Transkription von tRNA, rRNA und snRNA beteiligt. RNase MRP ist in Mitochondrien direkt an der Initiation der Prozessierung von mitochondrialer RNA beteiligt. Im Nukleus ist es in die Prozessierung von Vorläufer-rRNA involviert. Y RNA ist ein Initiationsfaktor für die DNA-Replikation. Telomerase RNA ist ein Bestandteil des gleichnamigen Enzyms, das die Endstücke der Chromosomen, die sogenannten Telomere, nach der Zellteilung wiederherstellt. Eine weitere nicht-codierende Art von RNA stellt die mikro-RNA (miRNA) dar. Neben ihrer Bedeutung im Zellzyklus und dem Zellmetabolismus hat miRNA eine wichtige Funktion in der Regulation der Genexpression. Wie die anderen RNA-Arten wird auch die miRNA selbst wiederum durch Abschnitte auf dem Genom reguliert.

Noch immer werden neue Gensequenzen identifiziert, die für RNA codieren. Durch den Einsatz moderner Sequenzierungsverfahren werden fortlaufend neue Datensätze produziert. Für die Zuordnung von DNA-Abschnitten zu ihrem Transkriptionsprodukt gibt

es mehrere Möglichkeiten: Eine besteht darin, die Sequenzierungsprodukte mit bekannten Sequenzen abzugleichen, die in Datenbanken zu finden sind. Dazu wurden bereits Programme entwickelt, wie beispielsweise „RNAhybrid“ und „snoGPS“. Aufgrund der großen Variabilität der Sequenzen, die für eine RNA-Art codieren, kann bislang jedoch nur ein kleiner Teil von genomischer DNA mit diesem Verfahren ihrem Transkriptionsprodukt zugeordnet werden. Neue, zunächst nicht klassifizierbare Sequenzen können dann durch experimentelle Analyse ihrer Eigenschaften, wie zum Beispiel ihrer Faltung, mittels Kernspinresonanzspektroskopie (NMR-Spektroskopie) einer Art zugeordnet werden. Dieses Verfahren ist jedoch aufwändig und teuer. Eine zuverlässige und schnelle Alternative bieten die *in silico*-Verfahren, die mithilfe von bioinformatischen Methoden unbekannte DNA-Sequenzen ohne Abgleichung mit bekannten Sequenzen zuverlässig einordnen können. In der Literatur lassen sich dazu Klassifikationssysteme finden, wie zum Beispiel „snoReport“, welches snoRNA aus DNA-Sequenzen erkennt.

Diese Dissertation soll neue Einblicke in die komplexen Zusammenhänge und Eigenschaften der unterschiedlichen Arten von RNA geben. Zwar haben sich Forscher seit der Entdeckung der RNA in vielen und weitreichenden Studien mit dem Thema der Klassifikation und Kontrolle von RNA auseinandergesetzt, vom vollständigen Verständnis ihrer Vielfalt und der komplexen Mechanismen ist man bislang noch immer weit entfernt. Insgesamt sollen die folgenden Fragen geklärt werden:

Fragestellung 1: Wie kann ncRNA klassifiziert werden und ist es möglich, mithilfe eines *in silico*-Verfahrens unbekannte ncRNA-Sequenzen zu identifizieren?

Fragestellung 2: Wie wird miRNA reguliert und wie ist die miRNA in die Expression anderer Gene involviert?

Fragestellung 3: Inwieweit kann mRNA für die Genexpressionsanalyse verwendet werden?

Fragestellung (1)

Publikation:

Witte, M., Gupta, Y., Möller, S., Ludwig, R.J., Restle, T., Zillikens, D., and Ibrahim, S.M. (2014). ptRNApred: computational identification and classification of post-transcriptional RNA. Nucleic Acids Res. 42, e167.

Inhalt

RNA-Sequenzen lassen sich, basierend auf ihrer Funktion in der Zelle, einer RNA-Art zuordnen. Aufgrund ihrer phylogenetischen Verwandtschaft weisen RNA-Sequenzen innerhalb einer RNA-Art strukturelle Ähnlichkeiten auf. Die strukturellen Eigenschaften von RNA-Sequenzen einer RNA-Art können zur Entwicklung bioinformatischer Methoden verwendet werden, die neue Sequenzen einer RNA-Art zuordnen können. Ziel von Studie (1) war es, verschiedene RNA-Arten anhand ihrer strukturellen Eigenschaften zu charakterisieren. Mithilfe dieser charakteristischen Eigenschaften sollte dann das Programm „ptRNApred“ entwickelt werden, mit dem unbekannte Sequenzen einer bestimmten RNA-Art zugeordnet werden können. Hierbei wurde besonders auf die Charakterisierung von RNA-Arten geachtet, die an post-transkriptioneller Modifikation oder Replikation von DNA beteiligt sind. Zu dieser Gruppe gehörten zum Zeitpunkt dieser Arbeit snoRNA, snRNA, Ribonuclease P (RNaseP), Ribonuclease MRP (RNase MRP), Y RNA und telomerase RNA. Zum besseren Verständnis werden diese RNA-Arten in Studie (1) als „post-transkriptionelle RNA“ („ptRNA“) zusammengefasst.

Um ptRNA zu charakterisieren, wurden zunächst 2040 ptRNA-Sequenzen aus der „NONCODE“ heruntergeladen, einer Datenbank, in der Sequenzen nicht-codierender RNA-Arten gespeichert sind. Darunter fanden sich 268 RNase P-, 14 RNase MRP-, 1443 snoRNA- (davon 13 scaRNA-), 46 telomerase RNA-, 14 Y RNA- und 255 snRNA-Sequenzen. Zwei Drittel dieser Sequenzen wurden zur Identifikation charakteristischer Eigenschaften und zur Entwicklung des Klassifizierungssystems verwendet. Das übrige Drittel wurde zunächst beiseitegelegt, um später die Sensitivität und Spezifität des Klassifizierungssystems testen zu können. Die genaue Anzahl der Sequenzen aus den unterschiedlichen RNA-Arten und ihre Verwendung sind in Tabelle 1 zusammengefasst.

	Trainingssequenzen	Testsequenzen
RNase P	178	90
RNase MRP	9	5
snoRNA+scaRNA	978+9	452+4
Telomerase RNA	29	17
Y RNA	9	5
snRNA	170	85

Tabelle 1: Anzahl der Trainings- und Testsequenzen jeder ptRNA-Subklasse.

Als bioinformatische Methode für ptRNApred wählten wir eine „support vector machine“ („SVM“), ein mathematisches Verfahren der Mustererkennung. Vereinfacht dargestellt kann eine SVM durch Eingabe von Daten und einer dazugehörigen Gruppe mit verschiedenen Eigenschaften „trainiert“ werden. Anschließend können neue Daten anhand ihrer Eigenschaften einer bestimmten Gruppe zugeordnet werden. Zum Training der SVM wurden die in Tabelle 1 aufgeführten Trainingssequenzen und ihre dazugehörige RNA-Art verwendet. Um die Sequenzen anschließend ihrer RNA-Art zuordnen und strukturelle Ähnlichkeiten der RNA-Sequenzen feststellen zu können, wurden Eigenschaften unterschiedlicher ptRNA-Arten definiert und auf ihre Fähigkeit zur Diskriminierung verschiedener ptRNA-Arten getestet. Es wurden Eigenschaften gewählt, die sowohl auf der Primär- als auch auf der Sekundärstruktur basierten. Die Fähigkeit zur Diskriminierung wurde dann mithilfe von „LibSVM“ getestet, das eine Schnittstelle zum vektorbasierten Training von SVMs darstellt. Als Eigenschaften wurden beispielsweise die Anzahl der Basen Guanin, Adenin, Cytosin und Uracil innerhalb einer RNA-Sequenz oder die Anzahl der Haarnadelstrukturen im Rahmen der Faltung herangezogen. Zusätzlich wählten wir thermodynamische Eigenschaften aus, wie beispielsweise die freie Energie. Insgesamt konnten sieben Eigenschaften aus der Primär- und 52 Eigenschaften von der Sekundärstruktur abgeleitet werden, es ergaben sich somit 59 Eigenschaften. Zusätzlich verwendeten wir die 32 Tripleteigenschaften, die zur Klassifikation von miRNA bereits im Programm „miPred“ verwendet werden. Zusammenfassend ergaben sich dadurch 91 Eigenschaften. Eine Liste der Eigenschaften und ihren Stellenwert in der Diskriminierung zwischen den unterschiedlichen ptRNA-Arten in Form eines F-Maßes „F-Score“ ist in Tabelle 2 dargestellt. Das F-Maß kombiniert Genauigkeit (precision) und Trefferquote (recall). Je höher das F-Maß, desto besser ist die Fähigkeit zur Diskrimination. Ein weiteres Maß zum Stellenwert in der Diskriminierung ist der „Gini-Index“, der im Folgenden erläutert wird.

Property number	Property description	F-Score	Gini-Index
1	Adenine_content	0,064268	2,41E-44
2	Cytosine_content	0,083006	6,03E-15
3	Entropy_1	0,088785	4,34E-19
4	Entropy_2	0,09256	5,93E-21
5	GC_content	0,106083	5,43E-11
6	Guanine_content	0,12987	2,75E-18
7	Hydrophilicity	0,098519	4,37E-23
8	Keto_content	0,131688	4,12E-16
9	Rise	0,091877	8,37E-20
10	Roll	0,098188	6,50E-26
11	Shift	0,057281	5,62E-49
12	Slide	0,09489	4,93E-19
13	Stacking_energy	0,088663	1,80E-30
14	Tilt	0,060113	6,03E-20
15	Twist	0,090472	3,98E-27
16	value_in_3rd_rnafold	0,039964	2,48E-51
17	count_star_bracket_{	0,032699	0,002720791
18	count_comma_in_rnafold	0,032914	0,000482739
19	value_line_no_3_RNAfold	0,037566	2,54E-21
20	value_line_no_3_RNAfold(second value)	0,034666	7,25E-18
21	value_line_number_4	0,037892	1,68E-18
22	value_line_number_4(second value)	0,034268	6,11E-32
23	frequency_of_mfe_structure_in_ensemble	0,035204	1,12E-16
24	ensemble diversity	0,039227	1,21E-11
25	value_MFE_RNAfold	0,038989	1,18E-34
26	Number_of_loops	0,03857	1,12E-12
27	(((	0,032912	3,18E-29
28	((.	0,037534	7,90E-11
29	(..	0,039972	3,42E-07
30	...	0,029854	7,49E-10
31	.((	0,038714	4,07E-11
32	..(	0,040519	0,000102735
33	.(.	0,023965	0,996225164
34	(.(	0,031401	0,003872273
35	A(((	0,003259	5,34E-15
36	A((.	0,003767	1,49E-06
37	A(.(	0,00501	0,037528552
38	A(..	0,014225	2,74E-05
39	A.((	0,001732	4,20E-06
40	A.(.	0,001482	0,119541752
41	A..(	0,015074	2,38E-12
42	A...	0,029134	5,04E-12
43	U(((	0,003228	1,41E-09
44	U((.	0,00721	2,26E-05
45	U(.(	0,00737	0,000196773
46	U(..	0,010258	4,59E-09

47	U.((	0,002741	0,00024243
48	U.(.	0,000326	0,996205638
49	U..(	0,017021	1,90E-21
50	U...	0,093364	3,83E-43
51	G(((	0,021702	3,72E-27
52	G((.	0,024863	0,000494941
53	G(.(	0,005776	3,58E-08
54	G(..	0,010109	6,84E-14
55	G.((	0,003169	0,000144641
56	G.(.	0,001119	0,403681085
57	G..(	0,010458	2,26E-06
58	G...	0,006942	1,12E-30
59	C(((	0,044132	4,38E-07
60	C((.	0,009389	2,98E-09
61	C(.(	0,001738	0,98836726
62	C(..	0,002454	1,76E-08
63	C.((	0,016575	0,017213739
64	C.(.	0,000654	0,000405832
65	C..(	0,006828	0,003771578
66	C...	0,019666	2,19E-12
67	number_of_AU	0,018431	1,55E-20
68	number_of_CG	0,008113	1,63E-10
69	number_of_GU	0,005051	4,01E-09
70	number_of_mistatches_in_sec_struc	0,036243	1,48E-12
71	number_of_buldges_in_sec_struc	0,06064	4,57E-09
72	A_in_buldges	0,059369	2,94E-13
73	G_in_buldges	0,064056	0,168681863
74	C_in_buldges	0,037917	2,02E-05
75	U_in_buldges	0,035724	1,48E-09
76	length_of_hairpin	0,029	4,29E-12
77	number_of_sub_sec_structure	0,025913	1,59E-11
78	number_of_A_hairpin	0,026994	8,85E-09
79	number_of_G_hairpin	0,010075	0,021895501
80	number_of_C_hairpin	0,027103	0,006452735
81	number_of_U_hairpin	0,066878	1,63E-12
82	number_of_A_purine	0,021688	6,41E-07
83	number_of_A_pyramidine	0,055579	1,70E-14
84	number_of_A_in_first_complementary_strand	0,002319	1,32E-16
85	number_of_G_in_first_complementary_strand	0,005858	4,21E-14
86	number_of_C_in_first_complementary_strand	0,001151	1,07E-07
87	number_of_U_in_first_complementary_strand	0,004656	1,13E-34
88	number_of_A_in_second_complementary_strand	0,003052	2,39E-14
89	number_of_G_in_second_complementary_strand	0,003099	0,00106611
90	number_of_C_in_second_complementary_strand	0,012551	2,20E-08
91	number_of_U_in_second_complementary_strand	0,035877	9,40E-16

Tabelle 2: Eigenschaften für die SVM und ihre Fähigkeit zur Diskrimination zwischen den unterschiedlichen RNA-Arten.

Die in Tabelle 2 aufgeführten Eigenschaften wurden als Vektoren für die SVM verwendet, um zwischen den unterschiedlichen ptRNA-Arten zu diskriminieren. Um jedoch ptRNA von anderen RNA-Arten, wie beispielsweise mRNA unterscheiden zu können, wurde ein weiteres Klassifizierungssystem vorgeschaltet, mit dem ptRNA von anderen RNA-Arten abgegrenzt wird. Bei der Eingabe einer unbekannten Sequenz kontrolliert ptRNApred zunächst, ob es sich bei der Sequenz um ptRNA handelt und ordnet diese im zweiten Schritt einer ptRNA-Art zu. Die Klassifikationsfähigkeit der SVM wurde mithilfe der anfangs beiseitegelegten Testsequenzen (Tabelle 1) kontrolliert. Diese ergab eine Sensitivität von 91% und eine Spezifität von 94%. Zusammenfassend stellt ptRNApred eine neue Möglichkeit zur Detektion und Klassifikation unbekannter DNA-Sequenzen dar.

Neben einer SVM gibt es weitere Algorithmen, um die Fähigkeit von Eigenschaften zur Diskriminierung zwischen RNA-Arten zu testen. Ein Beispiel ist der sogenannte „Random Forest“ nach Breiman. Auch wenn, bezogen auf ptRNApred, dieses Verfahren gegenüber einer SVM im Hinblick auf die Treffsicherheit unterlegen war (82% vs. 91% beim Einsatz der SVM), gibt der Random Forest einen weiteren Einblick in die strukturellen Eigenschaften der ptRNA-Arten. In Tabelle 3 ist eine Rangfolge der 25 Eigenschaften angegeben, mithilfe derer am besten zwischen den unterschiedlichen ptRNA-Arten diskriminiert werden kann. Die Fähigkeit zur Diskriminierung wurde anhand der beiden Algorithmen errechnet und ist in Form eines F-Maßes (entsprechend der SVM) bzw. Gini-Indexes (entsprechend dem Random Forest nach Breiman) angegeben. Interessanterweise finden sich unter den 10 am meisten diskriminativen Eigenschaften, gemessen am F-Maß, 9 Dinukleotid-Eigenschaften. Zudem finden sich unter den 25 am meisten diskriminativen Eigenschaften alle 15 für ptRNApred verwendeten Dinukleotid-Eigenschaften. Auch der Gini-Index zeigt, dass 12 Dinukleotid-Eigenschaften zu den 25 diskriminativsten Eigenschaften gehören. Somit haben Dinukleotid-Eigenschaften in der Diskriminierung der unterschiedlichen ptRNA-Arten einen sehr hohen Stellenwert.

Rank of importance for discrimination	Property ranked by F-score	Property ranked by Gini-Index
1	Keto_content	value_in_3rd_rnafold
2	Guanine_content	Shift
3	GC_content	Adenine_content
4	Hydrophilicity	U...
5	Roll	number_of_U_in_first_complementary_strand
6	Slide	value_MFE_RNAfold
7	U...	value_line_number_4(second value)
8	Entropy_2	G...
9	Rise	Stacking_energy
10	Twist	(((
11	Entropy_1	G(((
12	Stacking_energy	Twist
13	Cytosine_content	Roll
14	number_of_U_hairpin	Hydrophilicity
15	Adenine_content	U..(
16	G_in_buldges	value_line_no_3_RNAfold
17	number_of_buldges_in_sec_struc	Entropy_2
18	Tilt	number_of_AU
19	A_in_buldges	Tilt
20	Shift	Rise
21	number_of_A_pyramidine	Entropy_1
22	C(((	Slide
23	..(	value_line_number_4
24	(..	Guanine_content
25	value_in_3rd_rnafold	value_line_no_3_RNAfold(second value)

Tabelle 3: Rangfolge der 25 am meisten diskriminativen Eigenschaften, gemessen am F-Maß und am Gini-Index.

Schlussfolgerung

Einzelne Sequenzen einer RNA-Art ähneln sich nicht nur in ihrer Funktion, sondern auch in ihrer Struktur. Mithilfe ihrer strukturellen Eigenschaften lassen sich RNA-Sequenzen einer RNA-Art zuordnen. Die wichtigsten Eigenschaften zur Diskriminierung unterschiedlicher Arten von RNA basieren auf der Konstellation ihrer Dinukleotide. Durch die Identifikation charakteristischer Eigenschaften unterschiedlicher RNA-Arten war es uns möglich, „ptRNApred“ zu entwickeln, eine bioinformatische Methode, mit der man verschiedene Arten von RNA aus unbekannten genomischen Sequenzen erkennen kann.

Fragestellung (2)

Publikation:

Gupta, Y., Möller, S., Witte, M., Belheouane, M., Sezin, T., Hirose, M., Vorobyev, A., Niesar, F., Bischof, J., Ludwig, R.J., et al. (2016). Dissecting genetics of cutaneous miRNA in a mouse model of an autoimmune blistering disease. BMC Genomics 17, 112.

Inhalt

Mikro-RNA (miRNA) ist eine nicht-codierende RNA, die neben dem Zellzyklus und dem Zellmetabolismus eine wichtige Rolle in der Regulation der Genexpression spielt. Während die Prozesse, die durch miRNA reguliert werden, inzwischen gut verstanden sind, ist die Regulation der Expression der miRNA selbst bisher weitestgehend unbekannt. Ziel dieser Studie war es, Regionen des Genoms zu identifizieren, die die Expression von miRNAs regulieren. Dabei fokussierten wir uns in Studie (2) auf die miRNA-Regulation in der Haut im Rahmen eines Mausmodells zur Epidermolysis bullosa acquisita (EBA), einer blasenbildenden Autoimmunerkrankung mit Autoantikörpern gegen Kollagen VII in der Haut.

Zur Identifikation von Regionen (=Loci) auf dem Genom, die für die Expression von miRNA in der Haut verantwortlich sind, wurde EBA durch Injektion von rekombinantem Kollagen VII in Mäusen der 4. Generation einer Kreuzung aus den Mausstämmen MRL/MpJ, NZM2410/J, BXD2/TyJ und Cast/EiJ induziert. Die Krankheitsaktivität wurde zu verschiedenen Zeitpunkten beurteilt. Anschließend wurden den Mäusen Hautproben entnommen, die Mäuse wurden genotypisiert und die Expression verschiedener miRNAs gemessen. Zur Identifikation genetischer Loci, die an der Regulation von miRNA-Expression in der Haut beteiligt sind, wurden der Genotyp und die quantitative Expression der verschiedenen miRNAs korreliert. Dadurch konnten Loci identifiziert werden, die die Expression verschiedener miRNAs regulieren. Diese Loci werden auch „expression quantitative trait loci“ (eQTL) genannt. Insgesamt wurden dadurch 42 eQTLs identifiziert, die die Expression von 38 verschiedenen miRNAs regulieren. Die höchste Signifikanz ergab sich dabei für die Regulation der miRNA „miR-298“ durch ein eQTL auf Chromosom 9. Eine Liste der eQTLs ist in Tabelle 4 dargestellt.

AffyID	PEAK SNP	Chr	Pos(Mb)	CI	-logP score	α	Transcript Location (Chr:Pos)
miR-322	rs6386920	1	54.1	51.5-58.4	4.797697778	0.02	X:50407432-50407526
miR-431	rs6386920	1	54.1	51.5-59.8	5.231417376	0.006	12:110828657-110828747
miR-26a	rs6265423	2	47	28.1-51.3	4.376166299	0.021	10:126432586-126432669,9:118940914-118941003
miR-291a-3p	CEL.2_5060 5053	2	50.5	28.1-59.9	4.546953451	0.018	7:3218920-3219001
miR-423-3p	rs6265423	2	47	33.9-51.3	4.948513523	0.015	11:76891566-76891674
miR-671-5p	rs13476472	2	45.5	33.9-50.5	5.838570402	0.004	5:24097932-24098029
miR-23b	rs6250599	2	48.5	44.3-51.3	4.335454904	0.022	13:63401792-63401865
miR-409-5p	rs13476874	2	159.5	136-163	4.175101842	0.035	12:110981368-110981446
miR-409-5p	rs13477083	3	43.3	27.5-45.4	4.215972522	0.033	12:110981368-110981446
miR-546	rs13477126	3	56.4	52.9-58.3	4.493016911	0.029	10:126435496-126435616
miR-200a-star	rs3671119	3	126.1	117.4-131.3	4.688008826	0.011	4:155429005-155429094
miR-339-5p	rs3660863	4	7.1	3.7-19.5	4.878132165	0.003	5:139845604-139845699
miR-465c-5p	rs13477873	4	101.1	82.8-118	4.472478928	0.01	X:64079130-64079210
miR-295	rs3663950	4	135.3	129.4-141.1	4.209508065	0.04	7:3220774-3220842
miR-878-3p	rs13478002	4	136.2	135.3-141.1	5.110152979	0.01	X:64054683-64054760
miR-742	rs3673049	5	90.1	87.5-96.6	4.765858909	0.035	X:64033548-64033612
miR-379	rs6208251	6	104.8	98.4-116.7	4.024954772	0.034	12:110947270-110947335
miR-154	rs13479063	6	136.3	133.92-142.4	4.309450382	0.032	12:110976643-110976708
miR-425*	rs3663988	7	146.5	140.2-146.5	4.342220512	0.035	9:108471108-108471192
miR-486	rs13479880	8	89.3	72.5-95.1	4.108040959	0.033	8:24253027-24253154
miR-487b	rs6257357	8	88	77.7-90.2	4.81779528	0.019	12:110965543-110965624
miR-501-3p	rs13479880	8	89.3	81.7-95	4.205948891	0.03	X:6818369-6818477
miR-130b	rs6413270	9	37.7	36.8-44.4	4.222559895	0.032	16:17124154-17124235
miR-298	rs3700596	9	86.2	85-90.6	6.573678024	0.001	2:174093005-174093086
miR-466c-3p	rs3712394	10	17.7	14.1-24.2	4.569565941	0.03	2:10403161-10403244
miR-466c-3p	rs13480563	10	27.9	24.2-38.7	4.591526295	0.029	2:10403161-10403244
miR-126-5p	rs6374078	10	60.6	30.7-65.9	4.177297455	0.031	2:26446877-26446949
miR-681	rs13481076	11	66.5	40.1-71.3	4.055742697	0.043	12:70864822-70864931
miR-20a-star	rs3712881	11	120.9	112-121	4.160612966	0.03	14:115443379-115443485
miR-203	mCV223512 41	12	60	55.0-72.6	3.838027866	0.039	12:113369091-113369166
miR-542-3p	CEL.12_847 50094	12	91.4	79.7-103.8	4.735149963	0.023	X:50402580-50402664
miR-341	gnf13.079.67 1	13	80.5	69.5-88.3	5.260091403	0.008	12:110849710-110849805
miR-449b	rs13482231	14	67.6	50.9-72.3	4.485232377	0.019	13:113827627-113827706
miR-7a	CEL.15_422 2769	15	4.4	3.2-9.7	4.941341377	0.044	13:58494140-58494247, 7:86033163-86033259
miR-7a	rs13482455	15	16.7	14.8-24.5	5.73214998	0.022	13:58494140-58494247, 7:86033163-86033259
miR-337-3p	rs13482549	15	45.5	38.4-53.9	4.298224443	0.029	12:110823999-110824095
miR-673-5p	rs13482549	15	45.5	38.9-56.6	4.592339382	0.014	12:110810200-110810290
miR-136	rs13482914	17	21	16.5-27.6	4.679083445	0.026	12:110833537-110833598
miR-466b-5p	rs13483212	18	12.3	0-21	4.48515896	0.016	2:10395846-10395927
miR-493	rs6211533	19	57.1	50.2-60.2	4.98913787	0.004	12:110818443-110818525
miR-26a	gnfX.023.543	X	36.3	0-49.3	4.718906745	0.01	10:126432586-126432669, 9:118940914-118941003
miR-466e-5p	rs13483712	X	9.1	0-33.5	4.013512497	0.05	2:10398257-10398340

Tabelle 4: Auflistung der eQTLs, die die Expression von miRNAs regulieren.

Neben der Regulation durch einen einzigen Genlocus kann die Expression von miRNA auch durch die Interaktion verschiedener Orte auf dem Genom reguliert werden. Dies wird auch Epistase genannt. Um die Epistase in der Regulation verschiedener miRNAs zu erforschen, wurden jeweils Paare von Variationen im Genom (=single nucleutide Polymorphisms, SNPs) mit der Expression verschiedener miRNAs korreliert. Dabei

ergaben sich insgesamt 200 signifikante SNP-Paare, die 8 verschiedene miRNAs regulieren (Abbildung 1).

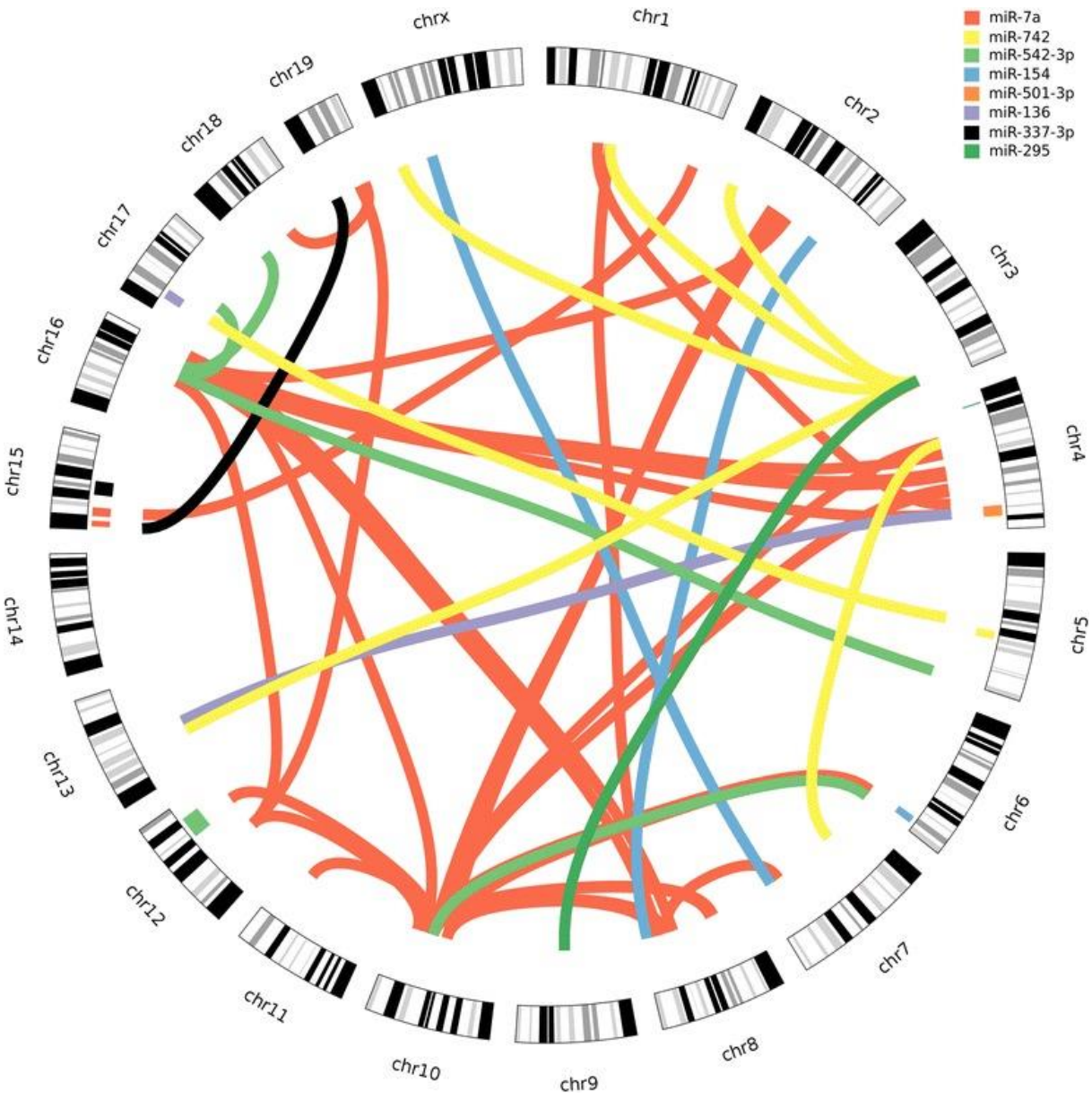


Abbildung 1: Epistatische Regulation der miRNA-Expression.

In vergangenen Studien wurden QTLs identifiziert, die in Verbindung mit der Krankheitsaktivität in dem hier angewendeten Mausmodell zur EBA stehen. Da auch miRNA bei der EBA unterschiedlich exprimiert wird, müssten sich die identifizierten eQTLs für miRNA und die bereits identifizierten QTLs für EBA überlappen. Dies wurde im nächsten Schritt erforscht. Als Ergebnis zeigen sich vier sich überlappende QTLs auf den Chromosomen 9, 12, 14 und 19.

Schlussfolgerung

Die Expression verschiedener miRNAs in der Haut wird sowohl durch einzelne Genloci, als auch durch die Interaktion verschiedener Genloci reguliert. Diese Art von Regulation wird auch Epistase genannt. In Bezug auf EBA, einer Autoimmunerkrankung der Haut, überlappen sich die Genloci, die die Expression verschiedener miRNAs regulieren, mit bereits bekannten Genloci, die mit EBA in Verbindung stehen.

Fragestellung (3)

Publikation:

Iwata, H., Witte, M., Samavedam, U.K.S.R.L., Gupta, Y., Shimizu, A., Ishiko, A., Schröder, T., Seeger, K., Dahlke, M., Rades, D., et al. (2015). Radiosensitive Hematopoietic Cells Determine the Extent of Skin Inflammation in Experimental Epidermolysis Bullosa Acquisita. J. Immunol. Baltim. Md 1950 195, 1945–1954.

Inhalt

Diese Studie verdeutlicht die experimentelle Anwendung von codierender RNA. MRNA kann zur Messung der Expression von Genen herangezogen werden. Gene, die für wichtige Proteine im Rahmen eines bestimmten Prozesses kodieren, werden zur Proteinbiosynthese vermehrt transkribiert. Die dabei gebildete Menge an mRNA in einem bestimmten Gewebe kann anschließend gemessen werden. So kann eine qualitative und quantitative Aussage über die Aktivität verschiedener Gene im Rahmen eines Prozesses, wie beispielsweise der Entzündung, getroffen werden. Zu den geeigneten Verfahren gehören unter anderem die Northern-Blot-Analyse oder die quantitative Reverse Transkriptase-Polymerase-Kettenreaktion (qRT-PCR). Diese Verfahren werden eingesetzt, wenn man eines bzw. nur wenige Transkriptionsprodukte nachweisen möchte. Zum Nachweis einer breiteren Palette an exprimierten Genen in einem Gewebe verwendet man heutzutage sogenannte DNA-Chips (oder DNA-Mikroarrays), mit denen die Expression einer Vielzahl von Genen simultan bestimmt werden kann.

In dieser Studie ging es um Unterschiede in der neutrophilenabhängigen, immunkomplexinduzierten Entzündungsaktivität in zwei unterschiedlichen Mausstämmen. Als Modell für die neutrophilenabhängige, immunkomplexinduzierte Entzündung kam dabei ein Mausmodell zur Epidermolysis bullosa acquisita (EBA, siehe Studie (2)) zum Einsatz. Dabei wurde EBA in zwei verschiedenen Mausstämmen, und zwar C57BL/6J und MRL//MpJ, durch repetitive Injektion von Antikörpern gegen eine immunogene Domäne auf dem Kollagen VII induziert. Die Ausprägung der Entzündung wurde durch ein standardisiertes Scoringsystem zu 6 verschiedenen Zeitpunkten gemessen. Dabei zeigte sich eine hohe Krankheitsaktivität im C57BL/6J-Stamm, während die Mäuse des MRL/MpJ-Stamms kaum Krankheit zeigten. Um zu beweisen, dass dies an Unterschieden in der Neutrophilenaktivität lag, wurden Mäuse beider Stämme zunächst bestrahlt und anschließend Neutrophile des jeweils anderen Stammes transplantiert. Nun zeigte sich ein entgegengesetztes Bild: C57BL/6J-Mäuse mit MRL/MpJ-Neutrophilen blieben nach der Injektion von anti-Kollagen VII-Antikörper gesund, während MRL/MpJ-Mäuse mit C57BL/6J-Neutrophilen eine Krankheitsaktivität zeigen. Um die Entzündungsaktivität der

Neutrophilen zu beurteilen, wurden anschließend Neutrophile isoliert und deren Aktivität mithilfe eines sogenannten „ROS-Release-Assays“ gemessen. Dabei zeigte sich eine höhere Aktivität der Neutrophilen im C57BL/6J-Stamm gegenüber denen des MRL/MpJ-Stamms. Um die beteiligten Gene in den Neutrophilen im Rahmen der immunkomplexinduzierten Entzündung zu identifizieren, wurde die oben beschriebene Genexpressionsanalyse mithilfe des DNA-Chips durchgeführt. Dazu wurde die Expression verschiedener Gene in Neutrophilen vor und nach ihrer Aktivierung mittels Immunkomplexen gemessen. Es ergaben sich 44 differenziell exprimierte Gene in den Neutrophilen des C57BL/6J-Stamms und 90 differenziell exprimierte Gene in den Neutrophilen des MRL/MpJ-Stamms, von denen sich 35 Gene überlappten. Eine genaue Liste der identifizierten Gene findet sich in Tabelle 5.

Common genes	Only MRL/MpJ	Only C57BL/6J
CCL2	ABCG1	CASS4
CXCL2	APLNR	CCNO
DDIT3	ARMC7	FRAT1
DPH1 (OVCA2)	ATF3	P2RY13
DUSP2	CCDC85B	PLA1A
EGR1	CCL3 (CCL3L1, CCL3L3, LOC101060267)	SLCO4C1
EGR2	CCL4	THBD
GADD45G	CCNL1	TNFAIP2
GPR65	CCR1	UBTD1
H6PD	CCR5	
HILPDA	CCRL2	
HMGCS1	CD83	
HYOU1	COQ10B	
IER2	CSTB	
IER5	CXCL3	
IMPACT	DUSP4	
INSIG1	DUSP5	
KCTD7 (RABGEF1)	EGR3	
KLF10	F3	
MAFB	FABP5 (LOC101060453)	
MAFK	FFAR2	
MTUS1	G0S2	
MYC	GJB4	
NAB2	GPRC5A	
OSM	GZMB	
PLK3	HAS2	
PTGER2	HMGCR	
PTGIR	HOMER1	
PTGS2	HSPBAP1	
RGS1	IFRD1	
SPP1	ITPKC	
TNFAIP6	KCNJ2	
XCL1	LFNG	
XCL1 (XCL2)	MKL1	
ZBTB39	NPC1	
	NR4A1	
	NR4A3	
	OLFML3	
	OSGIN2	
	P2RY14	
	PER1	
	PLK2	
	PMAIP1	
	PRICKLE3	
	RAB20	
	RARA	
	RIN2	
	SH3BP2	

SQLE	
TBX6	
TPD52L2	
TRIB1	
VPS37B	
WBP5	
ZFP36	

Tabelle 5: Differenziell exprimierte Gene in den Neutrophilen des C57BL/6J-Stamms und 90 differenziell exprimierte Gene in den Neutrophilen des MRL/MpJ-Stamms nach ihrer Aktivierung mittels Immunkomplexen, von denen sich 35 Gene überlappten.

Um Pathways und die dazugehörigen Schlüsselgene zu identifizieren, die für die Aktivität der Neutrophilen im Rahmen der immunkomplexinduzierten Entzündung eine Rolle spielen, wurden die differenziert exprimierten Gene (Tabelle 5) mithilfe der Programme „Database for Annotation, Visualization and Integrated Discovery“ (DAVID) und „Integrative Pathway Analysis“ (IPA) analysiert. Als Pathways, die in der Aktivität der Neutrophilen im Rahmen der immunkomplexinduzierten Entzündung eine Rolle spielen, wurden dabei für Neutrophile beider Mausstämme die Pathways „cytokine-cytokine receptor interaction“, und „MAPK signalling“ identifiziert. Für Neutrophile des MRL/MpJ-Stammes wurden zusätzlich die Pathways „chemokine signalling“ und „terpenoid backbone signalling“ identifiziert. Die Identifikation dieser Pathways könnte in Zukunft für die Entwicklung spezifischer Inhibitoren verwendet werden. Ein Netzwerk der Gene ist in Abbildung 2 dargestellt.

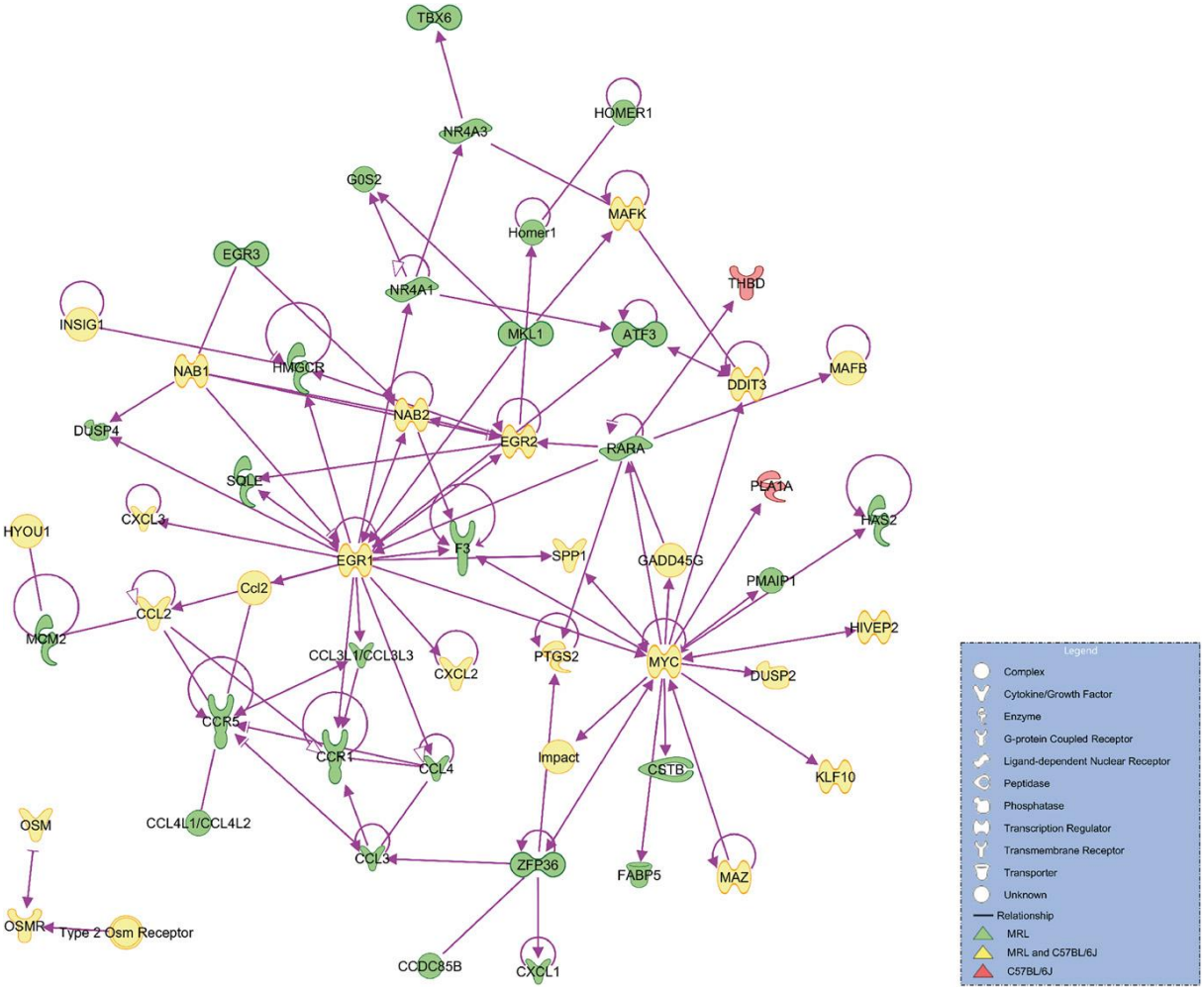


Abbildung 2: Netzwerk der differenziell exprimierten Gene in den Neutrophilen nach Aktivierung mittels Immunkomplexen. Als Schlüsselgene sind vor allem MYC, EGR1 und EGR 2 anzusehen.

Schlussfolgerung

Die Expression von Genen kann durch quantitative Bestimmung von mRNA gemessen werden. In einem Mausmodell zur EBA identifizierten wir durch Genexpressionsanalyse Gene, die an der Pathogenese der Erkrankung und an einer neutrophilenabhängigen, immunkomplexinduzierten Entzündung beteiligt sind. Durch Verknüpfung dieser Gene konnten wir verschiedene Pathways identifizieren, die an der Pathogenese der Erkrankung beteiligt sind. Die Identifikation dieser Pathways könnte in Zukunft für die Entwicklung spezifischer Inhibitoren verwendet werden.

Zusammenfassung

RNA ist eine Nukleinsäure, die sich als Polynukleotid aus einer Kette von vielen Nukleotiden zusammensetzt. Sie lässt sich unterteilen in codierende RNA (mRNA) und nicht-codierende RNA (ncRNA). Darüber hinaus kann ncRNA weiter unterteilt werden. Sequenzen einer RNA-Art ähneln sich nicht nur in ihrer Funktion, sondern auch in ihrer Struktur. Mithilfe ihrer strukturellen Eigenschaften lassen sich RNA-Sequenzen einer RNA-Art zuordnen. Die wichtigsten Eigenschaften zur Diskriminierung unterschiedlicher Arten von RNA basieren auf der Konstellation ihrer Dinukleotide. Mithilfe solcher Eigenschaften erkennt das in Studie (1) präsentierte Programm „ptRNApred“ verschiedene Arten von RNA aus unbekannten genomischen Sequenzen und ordnet sie einer jeweiligen RNA-Art zu. Eine der nicht-codierenden Arten von RNA stellt die mikro-RNA (miRNA) dar. Neben dem Zellzyklus und dem Zellmetabolismus spielt miRNA eine wichtige Rolle in der Regulation der Genexpression. Die miRNA selbst wird wiederum durch Abschnitte auf dem Genom reguliert. Wie Studie (2) zeigt, wird die Expression verschiedener miRNAs in der Haut sowohl durch einzelne Genloci, als auch durch die Interaktion verschiedener Genloci reguliert. Diese Art von Regulation wird auch Epistase genannt. In Bezug auf EBA, einer Autoimmunerkrankung der Haut, überlappen sich die Genloci, die die Expression verschiedener miRNAs regulieren, mit bereits bekannten Genloci, die mit EBA in Verbindung stehen.

Codierende RNA kann verwendet werden, um die quantitative Expression von Genen zu messen. In einem Mausmodell zur EBA identifizierten wir durch Genexpressionsanalyse Gene, die an der Pathogenese der Erkrankung und an einer neutrophilenabhängigen, immunkomplexinduzierten Entzündung beteiligt sind. Durch Verknüpfung dieser Gene konnten wir verschiedene Pathways identifizieren, die an der Pathogenese der Erkrankung beteiligt sind. Die Identifikation dieser Pathways könnte in Zukunft für die Entwicklung spezifischer Inhibitoren verwendet werden. Insgesamt gibt diese Arbeit neue Einblicke in die komplexen Zusammenhänge und Eigenschaften der unterschiedlichen Arten von RNA.

Danksagung

Die Ihnen vorliegende wissenschaftliche Arbeit zum Thema „Klassifikation und Kontrolle von RNA“ konnte nur mit der freundlichen Unterstützung einiger Personen vollendet werden.

Ganz besonderer Dank gilt meinem Doktorvater Prof. Dr. Ralf Ludwig. Durch sein Engagement und seine Ambition für das wissenschaftliche Arbeiten ist es mir möglich gewesen meine Promotion in dieser Art und Weise fertig zu stellen. Er hat mein wissenschaftliches Interesse gefördert und mich mit wertvollen Beiträgen und intensiven thematischen Diskussionen unterstützt. Zudem initiierte er meine Aufnahme ins Graduiertenkolleg 1727 „Modulation von Autoimmunität“, ein Netzwerk aus motivierten und engagierten Wissenschaftlern. Vielen Dank für die vorbildliche Betreuung.

Ein ebenso großer Dank gilt meiner Betreuerin Dr. Katja Bieber. Sie führte mich in das wissenschaftliche Arbeiten ein und brachte mir die nötigen Techniken bei. Vor allem in ihrer effizienten Arbeits- und ihrer logischen Denkweise war sie für mich ein großes Vorbild. Sie stand jederzeit für Fragen und Diskussionen bereit und unterstützte und förderte mich während der gesamten Zeit.

Zudem danke ich Prof. Dr. Saleh Ibrahim, der mein Interesse an der Genetik geweckt hat. In Zusammenarbeit mit seiner Arbeitsgruppe erlernte ich zahlreiche bioinformatische und molekulargenetische Methoden. Professor Ibrahim betreute mich hervorragend bei der Durchführung zahlreicher Projekte.

Außerdem möchte ich Prof. Dr. Zillikens danken, der als Leiter der Klinik für Dermatologie, Allergologie und Venerologie und Leiter des GRK 1727 nicht nur meine Projekte, sondern auch meine wissenschaftliche und medizinische Laufbahn gefördert hat.

Weiterhin danke ich meinen Freunden und Kollegen in den Arbeitsgruppen von Professor Ludwig und Professor Ibrahim und den Mitgliedern des GRK 1727, die mich unterstützt, mit mir diskutiert und die Wissenschaft für mich interessant gemacht haben.

Zum Schluss möchte ich meiner Familie danken, die mich neben diesem äußerst spannenden Abschnitt meiner akademischen Laufbahn auch in allen weiteren Bereichen des Lebens immer unterstützt.

Lebenslauf - MAREIKE WITTE

PERSÖNLICHE DATEN

Name: Mareike Witte
Geburtstag: 02.05.1991
Geburtsort: Hannover

AUSBILDUNG

Seit 08/16	**Assistenzärztin an der Klinik für Dermatologie, Allergologie und Venerologie** UKSH Lübeck
06/16	**Approbation als Ärztin, Universität zu Lübeck**
10/09 – 06/16	**Studium Humanmedizin** Universität zu Lübeck
06/09	**Bilinguales Abitur, Deutsch-Englisch** Gymnasium Schillerschule, Hannover
07/02 – 07/09	**Gymnasium Schillerschule** Hannover
07/00 – 07/02	**Orientierungsstufe** Hannover
08/97 – 07/2000	**Grundschule Mühlenweg** Hannover

PRAKTIKA/TÄTIGKEITEN

12/15 – 04/16	**Praktisches Jahr Dermatologie** UKSH Lübeck
09/15 – 12/15	**Praktisches Jahr Innere Medizin** Segeberger Kliniken
05/15 – 09/15	**Praktisches Jahr Chirurgie** UKSH Lübeck

09/14 – 10/14	**Famulatur Dermatologie** UKSH Lübeck
09/13 – 10/13	**Famulatur Allgemeinmedizin** Paracelsus Gesundheitszentrum, Lübeck
02/13 – 03/13	**Famulatur Allgemeinchirurgie** Hue Central Hospital, Vietnam
07/12 – 08/12	**Famulatur experimentelle Dermatologie** UKSH Lübeck
03/12 – 04/12	**Famulatur Innere Medizin** Vinzenzkrankenhaus, Hannover
2/12 – 3/12	**Famulatur Thoraxchirurgie** UKSH Lübeck
04/12 – 06/12	**Studentische Hilfskraft im Institut für Physiologie,** Universität zu Lübeck
03/08 und 03/11	**Mitarbeit an verschiedenen Forschungsprojekten im Labor der Klinik für Immunologie und Rheumatologie,** MH-Hannover
09/10 – 10/10	**Sozialprojekt im S.O.S.-Kindergarten "Santa Rita"** in Quito, Ecuador
08/09 – 09/09	**Praktikum in der Krankenpflege** Diakoniekrankenhaus Henriettenstiftung, Hannover
07/09 – 08/09 und 02/10 – 03/10	**Praktika in der Krankenpflege** Creighton University Medical Center, Omaha, NE, USA

ERWEITERTE AUSBILDUNG

04/14	**Versuchstierkunde in einem 80 stündigen Kurs der FELASA Kategorie C**
09/13	"**Workshop on AutoDock + MGLtools with lead developers"**
06/12	**Einführung in die Versuchstierkunde und tierexperimentelle Techniken bei der Maus**

VERÖFFENTLICHUNGEN

Dissecting genetics of cutaneous miRNA in a mouse model of an autoimmune blistering disease. Gupta Y, Möller S, **Witte M**, Belheouane M, Sezin T, Hirose M, Vorobyev A, Niesar F, Bischof J, Ludwig RJ, Zillikens D, Sadik CD, Restle T, Häsler R, Baines JF, Ibrahim SM., BMC Genomics. 2016 Feb 16;17:112.

Discovering potential drug-targets for personalized treatment of autoimmune disorders – what we learn from epidermolysis bullosa acquisita. **Witte M**, Koga H, Hashimoto T, Ludwig RJ, Bieber K, Expert Opin Ther Targets. 2016 Aug;20(8):985-98.

Dimethylfumarate Impairs Neutrophil Functions. Müller S, Behnen M, Bieber K, Möller S, Hellberg L, **Witte M**, Hänsel M, Zillikens D, Solbach W, Laskay T, Ludwig RJ, J Invest Dermatol. 2016 Jan;136(1):117-26.

Radiosensitive Hematopoietic Cells Determine the Extent of Skin Inflammation in Experimental Epidermolysis Bullosa Acquisita. Iwata H, **Witte M**, Samavedam UK, Gupta Y, Shimizu A, Ishiko A, Schröder T, Seeger K, Dahlke M, Rades D, Zillikens D, Ludwig RJ, J Immunol. 2015 Sep 1;195(5):1945-54.

ptRNApred: computational identification and classification of post-transcriptional RNA. **Witte M**, Gupta Y, Möller S, Ludwig RJ, Restle T, Zillikens D, Ibrahim SM, Nucleic Acids Res. 2014 Dec 16;42(22):e167.

WISSENSCHAFTLICHE AUSZEICHNUNGEN

10/15	**Ausgewählt für 'Meet the Investigator' im Journal of Investigative Dermatology, November 2015**
07/15	**ESDR 2015 Travel Grant**
11/14	**„1st Place Publication Award“** Winter Retreat des GRK 1727 "Modulation of Autoimmunity", Hotel Bad Segeberg
seit 07/14	**Mitglied des Excellence Clusters „Inflammation at Interfaces “**
09/14	**„Science Excellence Stipendium“ der Manfred-Lautenschläger-Stiftung**
03/14	**„AbbVie-Forschungspreis Dermatologie 2014“** für das Projekt „Therapeutisches Potenzial zur Inhibierung eines Integrin/Integrin-Rezeptor-Paares in der Behandlung der Psoriasis“

02/14 **„Kurt und Eva Herrmann Stipendium“ der Alfred-Marchionini-Stiftung**
für das Projekt „Funktion der T Zellen in der immunkomplex-induzierten Gewebsschädigung“

10/12 „**1st Place Presentation Award“**
für die Präsentation „Cellular requirements for autoantibody-induced tissue injury“ auf dem Winter Retreat des GRK 1727 "Modulation of Autoimmunity" in Wedendorf

seit 08/12 **Stipendium des GRK1727 | GRK "Modulation of Autoimmunity"**

BEITRÄGE ZU WISSENSCHAFTLICHEN VERANSTALTUNGEN

09/15 **Postervorstellung „T cells govern immune complex-induced inflammation“**
ESDR 2015 in Rotterdam, Niederlande

01/15 **Vortrag „T cells govern neutrophil-dependent inflammation “**
On-site visit der Reviewer des GRK1727, Universität zu Lübeck

09/14 **Vortrag und Postervorstellung „B-cell independent functions of T cells during immune-complex induced neutrophil-dependent inflammation“**
ESDR 2014 in Kopenhagen, Dänemark

07/14 **Postervorstellung „Cellular requirement for autoantibody-induced tissue injury“**
Poster Walk mit Prof. von Stebut-Borschitz, Universität zu Lübeck

05/14 **Postervorstellung „Using mouse models to dissect the genetics of cutaneous miRNA in autoimmune disease“**
Complex Trait Consortium 2014, Berlin

03/14 **Postervorstellung „Antigen-independent T cell-mediated regulation of neutrophil – dependent immune responses“**
ADF, Köln

02/14 **Postervorstellung „Antigen-independent T cell-mediated regulation of neutrophil – dependent immune responses“**
Poster Walk mit Prof. Pfeilschifter, Universität zu Lübeck

10/13 **Vortrag „Genetics of Psoriasis“**
SPP Medizinische Genetik, Lübeck

09/13 **Präsentation „Cellular requirements for autoantibody-induced tissue injury“**
"Berufsvorbereitendem Symposium" des GRK1727/1 "Modulation von Autoimmunität", Lübeck

07/13 **Postervorstellung „Antigen-independent T cell-mediated regulation of neutrophil – dependent immune responses“**
Poster Walk mit Prof. Scharfetter-Kochanek, Universität zu Lübeck

06/13 **Postervorstellung „Wenn das Abwehrsystem die Fronten wechselt: Die Entstehung von autoimmunen Hauterkrankungen“**
Doktorandentag der Universität zu Lübeck

05/13 **Postervorstellung „Antigen-independent T cell-mediated regulation of neutrophil – dependent immune responses“**
International Pre IID Satellite Symposium on Autoimmune Bullous Diseases, Lübeck

11/12 **Vortrag „Cellular requirements for autoantibody-induced tissue injury“**
Winter Retreat des GRK 1727 "Modulation of Autoimmunity", Hotel Schloss Wedendorf

Anhang

Im Anhang finden Sie die Publikationen, um die es in dieser Arbeit geht:

(1) *Witte, M., Gupta, Y., Möller, S., Ludwig, R.J., Restle, T., Zillikens, D., and Ibrahim, S.M. (2014). ptRNApred: computational identification and classification of post-transcriptional RNA. Nucleic Acids Res. 42, e167.*

(2) *Gupta, Y., Möller, S., Witte, M., Belheouane, M., Sezin, T., Hirose, M., Vorobyev, A., Niesar, F., Bischof, J., Ludwig, R.J., et al. (2016). Dissecting genetics of cutaneous miRNA in a mouse model of an autoimmune blistering disease. BMC Genomics 17, 112.*

(3) *Iwata, H., Witte, M., Samavedam, U.K.S.R.L., Gupta, Y., Shimizu, A., Ishiko, A., Schröder, T., Seeger, K., Dahlke, M., Rades, D., et al. (2015). Radiosensitive Hematopoietic Cells Determine the Extent of Skin Inflammation in Experimental Epidermolysis Bullosa Acquisita. J. Immunol. Baltim. Md 1950 195, 1945–1954.*

Published online 10 October 2014 *Nucleic Acids Research, 2014, Vol. 42, No. 22* ***e167***
doi: 10.1093/nar/gku918

ptRNApred: computational identification and classification of post-transcriptional RNA

Yask Gupta[1,†], Mareike Witte[1,*,†], Steffen Möller[1], Ralf J. Ludwig[1], Tobias Restle[2], Detlef Zillikens[1] and Saleh M. Ibrahim[1,*]

[1]Department of Dermatology, University of Lübeck, 23538 Lübeck, Germany and [2]Institute for Molecular Medicine, University of Lübeck, 23538 Lübeck, Germany

Received September 04, 2013; Revised September 17, 2014; Accepted September 22, 2014

ABSTRACT

Non-coding RNAs (ncRNAs) are known to play important functional roles in the cell. However, their identification and recognition in genomic sequences remains challenging. *In silico* methods, such as classification tools, offer a fast and reliable way for such screening and multiple classifiers have already been developed to predict well-defined subfamilies of RNA. So far, however, out of all the ncRNAs, only tRNA, miRNA and snoRNA can be predicted with a satisfying sensitivity and specificity. We here present *ptRNApred*, a tool to detect and classify subclasses of non-coding RNA that are involved in the regulation of post-transcriptional modifications or DNA replication, which we here call post-transcriptional RNA (ptRNA). It (i) detects RNA sequences coding for post-transcriptional RNA from the genomic sequence with an overall sensitivity of 91% and a specificity of 94% and (ii) predicts ptRNA-subclasses that exist in eukaryotes: snRNA, snoRNA, RNase P, RNase MRP, Y RNA or telomerase RNA. AVAILABILITY: The ptRNApred software is open for public use on http://www.ptrnapred.org/.

INTRODUCTION

There are many different types of RNA with multiple functions in the cell. Some RNA molecules contribute to the translation of genetic information into protein and the regulation of genes. Others function enzymatically by catalyzing biological reactions. While the non-coding regions in the genome were first believed to be dispensable sequences, they have been shown to code for RNA families that play important roles in the eukaryotic cell. These so-called non-coding RNAs (ncRNAs) do not code for protein but are involved in many regulatory processes and can be divided into a tremendous variety of highly plethoric and versatile families that are essential for the cellular function (1). Hence, they form a vast and to a large extent unexplored reservoir of potentially valuable medical biomarkers (2,3). For their identification, modern techniques like next-generation-sequencing and microarray-technologies are being employed (4,5). These techniques provide an immense amount of data and offer ample opportunities to identify novel classes of non-coding RNA. However, the experimental analysis of new sequences is time-consuming and complex, indicating the need to find alternative approaches for their analysis. Promising and auspicious approaches are given by *in silico* methods. Due to phylogenetic relationships, sequences of non-coding RNA show similarities regarding their properties. They can be divided into subclasses based on their conserved properties, meaning sequence conservation and structural conservation (6). Computational methods, such as classification tools, offer a fast and reliable way to analyze and classify sequences by exploiting conserved properties among the sequences (6,7).

Various classification systems have been developed to predict different subsets of RNA, using machine learning and phylogenetic approaches (8–12). So far, tRNAs can be detected reliably using tRNAScan-SE (13). Furthermore, various approaches have been established to detect miRNA (14) and other small RNA subsets. Recently, snoReport was introduced, which is designed to recognize small nucleolar RNA (snoRNA) from the genome without using any target information (15). Most of these systems achieve a satisfying accuracy, however not every RNA family can be predicted. For example, to this point, there is no tool for the prediction of small nuclear RNA (snRNA), Ribonuclease P (RNase P), Ribonuclease MRP (RNase MRP), Y RNA and telomerase RNA. Facing the continuing increase in the number of human RNAs in databases like Rfam (16,17), it is necessary to extend the current possibilities of RNA prediction. SnRNA, RNase P, RNase MRP, Y RNA and telomerase RNA have in common, that, besides snoRNA, they are involved in post-transcriptional modification or DNA replication in eukaryotes (18–23).

*To whom correspondence should be addressed. Tel: +49 451 500 3076; Fax: +49 451 500 5162; Email: mareike.witte@medizin.uni-luebeck.de
Correspondence may also be addressed to Saleh M. Ibrahim. Tel: +49 451 500 5250; Fax: +49 451 500 5162; Email: saleh.ibrahim@uk-sh.de
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

The aim of this study was to develop a tool that can predict and differentiate among sequences coding for these RNA families, for which, for the sake of convenience, we here use the term 'post-transcriptional RNA (ptRNA)'.

PtRNA-subclasses are not only distinct regarding their function, but also regarding their sequence. Therefore, we hypothesized that implementation of certain algorithms would make it possible to identify ptRNA from a genomic sequence. Machine learning algorithms, such as support vector machines (SVMs), have shown high accuracy in the development of classification systems (24–26). Recent advancement in small non-coding miRNA prediction has achieved high performance using machine learning approaches (27–30). Moreover, SVMs are also employed in the prediction of sequence based secondary structure of RNA (31,32). The available tools for designing the machine learning classifier, like SVM-Light (33) and LibSVM (34) are often used for the development of new algorithms for RNA prediction.

With the introduction of 'ptRNApred', this study offers a novel opportunity to detect and classify ptRNA without using any target information. 'ptRNApred' (i) detects RNA sequences coding for post-transcriptional RNA from the genomic sequence and (ii) predicts ptRNA-subclasses that exist in eukaryotes: snRNA, snoRNA, RNase P, RNase MRP, Y RNA or telomerase RNA.

MATERIALS AND METHODS

Dataset

2040 sequences of ncRNA were obtained from the NONCODE database (35), including 268 sequences of RNase P, 14 sequences of RNase MRP, 1443 sequences of snoRNA (1430 + 13 scaRNA), 46 sequences of telomerase RNA, 14 sequences of Y RNA and 255 sequences of snRNA (Table 1). These sequences were used as a dataset for the multiclass-classifier and as a positive set for our binary classifier. The negative set was made up of sequences of tRNA, 5s ribosomal RNA and miRNA that were derived from Rfam (17). Since our classifier focuses on eukaryotes, our selection of miRNA-sequences was restricted to sequences from the species *Homo sapiens*, *Mus musculus* and *Drosophila melanogaster*. The redundancy of the sequences within a set was removed using CD-Hit (36) at a threshold of 0.9 for the positive set and at 0.8 for the negative set of sequences. After removal of redundancy, the ratio of sequences in the negative to positive set was 3:2. Comparable ratios have been frequently used for the generation of SVMs (37,38).

All sets of sequences were divided into two parts: two-third of each set of sequences were used for training the classifiers and one-third was used for testing the performance of the classifier. Table 1 gives an overview on the number of training and testing sequences of each ptRNA-subclass.

A detailed list of the sequences is provided on the classifier's website http://www.ptrnapred.org/. For testing its performance on RNase P, 329 RNase P sequences were downloaded from 'The Ribonuclease P Database' (48). Performance on coding RNA was tested using 10 000 randomly downloaded mRNA sequences from Ensembl (Release 72, June 2013).

Features for classification

Feature selection and SVM training was performed using two sets of input parameters: The first set was based on the primary sequence and the second set considered the secondary structure which was predicted with RNAfold (Version 2.0.7) (39). Primary sequence properties were mainly derived from dinucleotide properties employing DiProGB (40), using a sliding window approach (window size: 2 nucleotides). Some of the properties in DiProGB are highly correlating to each other. The use of highly correlating features for classification would not only be redundant in information, but would overfit the classifier. In order to determine which of the features that we derived from DiProGB were correlating, we determined the Pearson correlation coefficients among all possible dinucleotide properties. Two features were considered as highly correlating when the Pearson coefficient was >0.9. As an example, the dinucleotide property 'stacking energy' was highly correlated to the property 'melting temperature'. Whenever one of two features were highly correlating, one of them was randomly discarded. In our example, we used 'stacking energy' as an input feature and discarded 'melting temperature' from consideration as a feature for classification. A table of the selected dinucleotide properties as well as their dinucleotide values (40) is provided in the supplement (Supplementary Table S1).

Secondary structures of every sequence were calculated via RNAfold (39), accessing the Vienna RNA Package (41). Fifty-two different properties were derived from the secondary structure, e.g. the number of loops, the number of bulges, the number of hairpins or the frequency of nucleotides involved in substructures.

Additionally, we included 32 triplet element properties employed by miPred, a triplet SVM for the classification of miRNA (42): MiPred considers the middle nucleotide among the triplet elements, resulting in 32 (4×8) possible combinations, which are denoted as 'U(((', 'A((.', etc.

Altogether, ptRNApred uses 91 features for classification. A detailed description of the feature selection is provided in the supplement (Supplementary Section S1).

Classification system

To create optimal conditions for classification, we compared the outcome of two different algorithms.

On the one hand, we employed a Random Forest according to Breiman (43) as a sophisticated classification method. Random Forests operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes output by individual trees.

On the other hand, we employed LibSVM (34), a library of SVM, which serves as an interface to train and build SVMs based on certain vectors. LibSVM was a superior machine learning algorithm for training our classifiers, since it gave us a better accuracy than Random Forest prediction. Further information is provided in the 'Validation of the algorithm' part of the 'Results' section.

Table 1. Total number of test and training sequences

	Training sequences	Testing sequences
RNase P	178	90
RNase MRP	9	5
snoRNA + scaRNA	978 + 9	452 + 4
telomerase RNA	29	17
Y RNA	9	5
snRNA	170	85

The table displays the number of training and testing sequences of each ptRNA-subclass used for the SVM.

Implementation of LibSVM

The SVM algorithm defines a hyperplane in the feature space with maximum margin distinguishing positive instances from negative (44). LibSVM provides a python script to optimize the grid parameters C and gamma (γ). C stands for cost function, i.e. penalty for the misclassification in the training set and γ is a free parameter deciding upon the impact of each training vector. The parameters are thus used for designing the classifier from a training set. Therefore, we used a (Gaussian) Radial Basis Function (RBF) (45) kernel for classification.

In order to decide the parameters C and γ, LibSVM obtains cross validation (CV) accuracy for each possible parameter setting. Regarding our binary classifier, the highest CV accuracy was achieved when C was set to 32768 and γ was set to 0.008 (Figure 1a). These parameters were used to train the whole training set and to generate the final model.

For multi-class classification, under a given (C, γ), LibSVM uses the one-against-one, one-against-all and sparse method to build hyper planes and to obtain the CV accuracy. Hence, the parameter selection tool suggests the same (C, γ) for all $k(k-1)/2$ decision functions. Yuan *et al.* (46,47) discuss issues of using the same or different parameters for the $k(k-1)/2$ two-class problems. In our case, $C = 4$ and $\gamma = 0.5$ as well as usage of the sparse method gave us the best CV accuracy for the multi-class classification (Figure 1b).

Work flow and output of ptRNApred

The web server implementation accepts sequences in a FASTA-format as an input that can be either uploaded as a file or pasted into the text box (Supplementary Figure S1). By checking 'Post-Transcriptional RNA', an in-built Perl script calculates input vectors for the pre-trained model to predict whether or not the input sequence belongs to the group of post-transcriptional RNA. Additionally selecting 'RNA family', the server also predicts the RNA-subclass.

Altogether, the output includes the prediction for ptRNA as well as the classification of the RNA class within the ptRNA. Additionally, it displays the minimum free energy using RNA-fold (39) as well as the secondary structure, using VARNA (Version 3.1) (48) (Supplementary Figure S2). The output can directly be downloaded.

RESULTS

We created a two-step classifier to distinguish sequences of ptRNA and non-ptRNA in a binary classification, and for the prior separate six classes of post-transcriptional RNA (snRNA, snoRNA, RNase P, RNase MRP, Y RNA or telomerase RNA) in a multi-class classification. Since the binary and multi-class classifiers were trained with separate data and thus function independently, the accuracy was calculated for each individual classifier.

In a 5-fold cross validation, using balanced amounts of randomly selected sequences throughout the positive and negative sets, the binary classifier yields an accuracy of 93% within the training set (Figure 1a) and the multi-class classifier yields a 5-fold cross validation accuracy of 87% (Figure 1b).

When testing the classifiers with the test set of sequences, the binary classifier showed an accuracy of 93%, with a sensitivity of 91%, a specificity of 94% and an overall precision of 90%. The multi-class classifier showed an accuracy of 91%.

The fact that the accuracy in the test set is higher than the 5-fold cross validation accuracy in the training set suggests that an increase in the number of sequences leads to a more accurate prediction. Regarding this matter, we observed an increase of the accuracy of the multi-class classification when adding more sequences to the training set.

Validation of the method

In order to validate our tool by comparing it to existing tools that can predict ptRNA, we found snoReport (15) as an advanced tool for the prediction of snoRNA. snoReport can predict orphan snoRNA without using target information and is therefore similar to our approach. To compare our tool to snoReport, we derived snoRNA sequences from a mouse genome from Ensembl (49) and used them as an independent set. In total, we used an input of 1603 sequences of snoRNA. As a result, snoReport identified 733 sequences correctly. In contrast, our classifier identified 1589 sequences correctly. Furthermore, we abstracted a human dataset with 1641 sequences of snoRNA from Ensembl. While snoReport identified 852 of the snoRNA-sequences correctly, our tool identified 1611 (Table 2).

In order to analyze the low sensitivity of snoReport we inspected the sequences that it fails to classify. We found that snoReport was not able to detect a major snoRNA-subclass snoU13. snoU13 was identified in 1989 (50). It has been well characterized in 35 species by both functional assay and prediction. It is involved in the nucleolytic cleavage at the 3′ end of 18S rRNA where it works as a trans-acting factor (51).

snoReport was unable to assign any of 245 snoU13-sequences in a human cohort to snoRNA. Our tool ptRNApred however identified all of them correctly (Table 2).

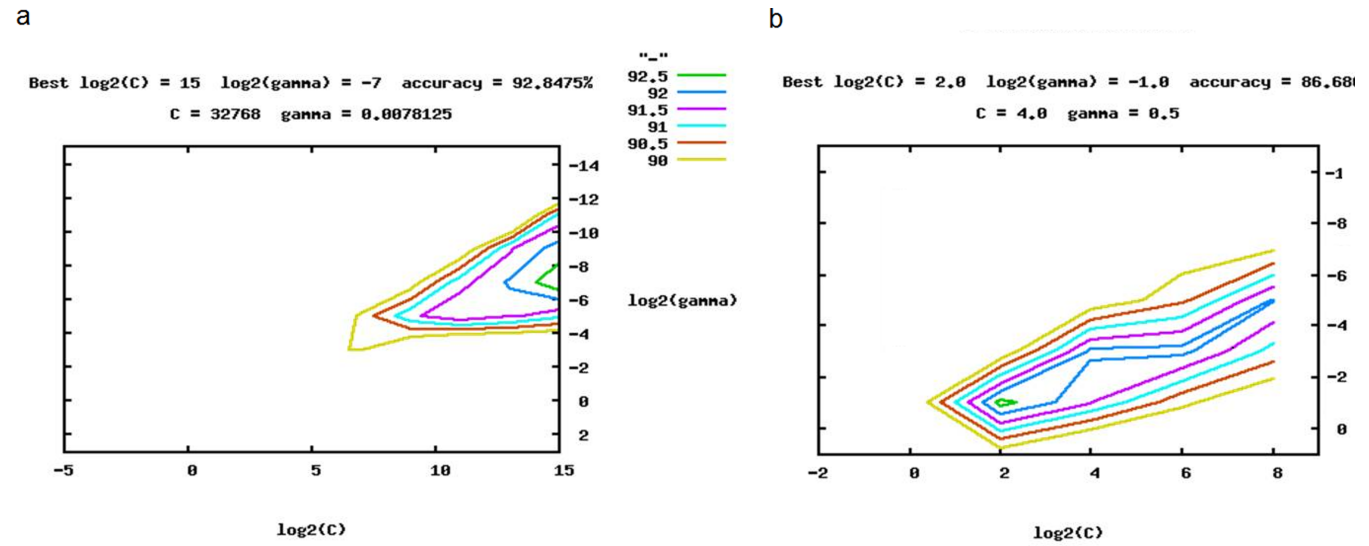


Figure 1. C and γ determination and 5-fold cross validation using LibSVM. The figure shows graphs for different values of the parameters C (a trade-off for misclassification) and γ (inverse width of RBF kernel) on a logarithmic X and Y axis. The ranges of the axes describe the different values that were tested, searching the optimal C and γ values in the grid space. The different colors in the diagram display the different accuracies obtained while optimizing C and γ values. We chose the C and γ values according to the green graphs, respectively, representing the C and γ value with the highest accuracy. (**a**) C and γ determination and 5-fold cross validation of the two-class SVM. The green graph represents the optimal values for C and gamma. In this case, the highest 5-fold cross validation accuracy (92.89%) is achieved when $C = 32768$ and $\gamma = 0.008$. (**b**) C and γ determination and 5-fold cross validation of the multi-class SVM. The green graph represents the optimal values for C and gamma. In this case, the highest 5-fold cross validation accuracy (86.69%) is achieved when $C = 4$ and $\gamma = 0.5$.

Table 2. Comparison between snoReport and ptRNApred

Organism	RNA class	Total number of sequences[a]	Number of sequences identified by snoReport (% of total number of sequences)	Number of sequences identified by ptRNApred (% of total number of sequences)
M. musculus	snoRNA	1603	737(46%)	1589(99%)
H. sapiens	snoRNA	1641	852(52%)	1611(98%)
	snoU13[b]	245[b]	0(0%)[b]	245(100%)[b]

A murine and a human dataset of snoRNA was abstracted from Ensembl (49) and performance of ptRNApred was compared to snoReport as a well-established tool for snoRNA prediction. ptRNApred achieved higher sensitivity than snoReport (99 versus 46% on the murine and 98 versus 52% on the human set of sequences). Regarding snoU13, a member of the snoRNAs, there is an even larger difference in the sensitivity (100 versus 0%).
[a]Total number of snoRNA-sequences downloaded from Ensembl (49).
[b]snoU13 among the human snoRNA sequences.

Current approaches to identify different RNA families heavily rely on their secondary structure conservation. Consequently, these approaches are accurate as long as the RNA families show high secondary structure conservation. However, as soon as a RNA-family lacks a secondary structure because it e.g. has no complementary sequences within itself, it will be misclassified. This explains why snoReport failed to identify snoU13 and possibly also the remaining ∼550 snoRNA sequences: snoU13 does not form any secondary structure conservation as it forms a loop.

Validation of the algorithm

As mentioned in the 'Materials and Methods' section, we compared the algorithm implemented in our tool to Random Forest classification. As a result, implementation of Random Forest yielded an overall accuracy of 82%. In comparison, using LibSVM, our multi-class classifier developed yields an accuracy of 91%. Detailed results of the performance of Random Forest based prediction of the ptRNA-subclasses can be found in the Supplement (Supplementary Table S2).

Variable importance

Classification features have an individual impact on the differentiation of RNA-classes. To determine the importance of each of the 91 features for classification of ptRNA, an *F*-score was calculated for each feature, using LibSVM. *F*-scores can be interpreted as a weighted average of the precision and recall, where an *F*-score reaches its best value at 1 and worst score at 0. Supplementary Table S3 depicts the *F*-score corresponding to every feature.

Additionally, even though Random Forests did not contribute to ptRNApred predictions, those provide useful information on feature importance. One of the measures of variable importance in Random Forest is the mean decrease in accuracy, calculated using the out-of-bag sample. The difference between the prediction accuracy on the untouched out-of-bag sample and that on the out-of-bag sample per-

muted on one predictor variable is averaged over all trees in the forest and normalized by the standard error. This gives the mean decrease in accuracy of that particular predictor variable which has been permuted. Thus, the importance of the predictor variables can be ranked by their mean decrease in accuracy. Supplementary Table S3 depicts the Gini-Index corresponding to every feature.

Interestingly, comparing the 25 most discriminative feature variables according to *F*-score and Gini-Index (Supplementary Table S4), dinucleotide properties achieve high ranks: 9 of the 10 most discriminative features according to the *F*-score are composed of dinucleotide properties. Furthermore, all of the 15 dinucleotide properties can be found among the 25 most discriminative properties. According to the Gini-Index, 12 properties can be found among the 25 most discriminative properties, whereas only three of them can be found among the top 10, indicating the importance of the secondary structure.

Validation of the feature number

As mentioned in the 'Materials and Methods' section, a general concern for all machine learning approaches is that one has too many features, i.e. that one trains on features that are not relevant—referred to as overfitting. This was excluded by the above mentioned cross validation test. On the other hand, too few features would lead to loss of (overall) accuracy. In order to confirm that using less features would lead to loss of accuracy, we selected the 78 most discriminative features based on Random Forest prediction, using the R package 'Boruta'. When using these 78 features instead of 91 features, the 5-fold cross validation accuracy decreased from 92.89 to 74.46% (Supplementary Figure S3).

Performances on a non-eukaryotic system

Even though ptRNApred is designed to primarily predict eukaryotic sequences, ptRNApred was tested for performance on RNase P sequences, using 329 RNase P sequences from 'The Ribonuclease P Database' (52). RNAse P has not only been described in eukaryotic systems (53), but rather distributes among different organisms (20). Interestingly, our tool predicted the RNAse P sequences with an accuracy of 97.3%.

Performances on coding RNA

Over the last few years, several tools have been developed to distinguish coding from non-coding RNA (54–57). Even though our aim was to develop a novel tool that can differentiate between subclasses of non-coding RNAs and not to distinguish between coding and non-coding RNAs, ptRNApred was tested for performance on mRNA. Therefore, ptRNApred was challenged by 10 000 mRNA randomly downloaded sequences from Ensembl. Surprisingly, only 15 of the sequences were misclassified as ptRNA. Therefore, the accuracy of separating out mRNA is 99.85%.

DISCUSSION

RNA classes that are involved in post-transcriptional modification or DNA replication in eukaryotes not only have functional similarities, but rather form a distinct group of ncRNA with sequence and conformational similarities, that make it possible to accurately distinguish them from other RNA classes.

We here present a novel user-friendly tool that employs discriminative properties to (i) distinguish what we here call 'post-transcriptional RNA' from other classes of ncRNA and (ii) discriminate between the different types of ptRNA.

An advantage of the tool is its highly accurate and therefore reliable output. This is based on its working principle: More than 90 features that are derived from the primary sequence and secondary structure are used to define properties for characterization and differentiation between the subclasses. Analyzing the most discriminative feature variables according to *F*-score and Gini-Index, the most important features are not only based on the secondary structure, but even more importantly on dinucleotide properties. This might be due to the fact that many nucleic acid properties such as nucleic acid stability, for example, seem to depend primarily on the identity of nearest-neighbor nucleotides (58). Furthermore, the corresponding nearest-neighbor model is also the basis for RNA secondary structure prediction by free-energy minimization (59). It has long been known that also thermodynamic but also conformational nucleotide properties may play a role. It has been shown, for example, that promoter locations can be predicted adopting dinucleotide stiffness parameters derived from molecular dynamic simulations (60). Our tool underscores the value of these properties.

Recently, a focus has been on the characterization of snoRNA. However, there has been no classifier that could predict snRNA, RNase P or RNase MRP, even though these subclasses have conserved secondary structures. Identification of those RNA classes has as yet been dependent on sequence alignment. This technique frequently leads to misidentification, especially if the particular homologous sequence is not present in any database.

Furthermore, ptRNApred can be used to elucidate unknown relations and derivation of RNA classes. Based on the assumption that evolutionarily close RNA families have similar sequence properties, one may speculate that tools like ptRNApred will falsely arrange evolutionarily close RNA families into the same group.

A deficiency of this tool is that its accuracy is dependent on the amount of published ptRNA sequences. Some classes of ptRNA, for example Y RNA, are to this point just rarely available in the NONCODE database, making it hard to define discriminative sequence properties and setting limits to the accuracy. In the current era of high-throughput next generation sequencing, where a large amount of genomic data is generated each day, ptRNA sequences that will be added to the database in the future can be used to increase training and test set, setting a base to improve the classifier. On the other hand, discovery of new candidates for ptRNA requires a method which can classify them rapidly and reliably. Our tool offers a solution to this problem. Also, facing the huge amount of new sequences that are found in Next Generation Sequencing (NGS) or RNA-seq data (61), it is important to include such algorithms into NGS pipelines. For such purposes, we provide a standalone version.

We implemented our method as a web-based server for free public use. Transparent and user-friendly design makes it possible for everyone to understand and employ the tool. Data and scripts for the development of the tool can be downloaded, allowing anyone to acquire the working principle and improve ptRNApred.

Collectively, our tool offers a fast and reliable way to analyze cDNA and RNA sequences and outperforms the existing classifiers. Furthermore, the tool provides comprehensive annotations. Therefore, ptRNApred introduces different opportunities to identify and classify new and unannotated RNA sequences.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Deutsche Forschungsgemeinschaft: the research training programs Modulation of Autoimmunity [GRK1727/1]; Genes, Environment and Inflammation [GRK 1743/1]; Excellence Cluster Inflammation at Interfaces [EXC 306/2]. Funding for open access charge: Deutsche Forschungsgemeinschaft: the research training programs Modulation of Autoimmunity [GRK1727/1]; Genes, Environment and Inflammation [GRK 1743/1]; Excellence Cluster Inflammation at Interfaces [EXC 306/2].
Conflict of interest statement. None declared.

REFERENCES

1. Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
2. Mallardo,M., Poltronieri,P. and D'Urso,O. (2008) Non-protein coding RNA biomarkers and differential expression in cancers: a review. *J. Exp. Clin. Cancer Res.*, **27**, 19.
3. Kim,T. and Reitmair,A. (2013) Non-coding RNAs: functional aspects and diagnostic utility in oncology. *Int. J. Mol. Sci.*, **14**, 4934–4968.
4. Beck,D., Ayers,S., Wen,J., Brandl,M.B., Pham,T.D., Webb,P., Chang,C.-C. and Zhou,X. (2011) Integrative analysis of next generation sequencing for small non-coding RNAs and transcriptional regulation in Myelodysplastic Syndromes. *BMC Med. Genomics*, **4**, 19.
5. Jung,C.-H., Hansen,M.A., Makunin,I.V., Korbie,D.J. and Mattick,J.S. (2010) Identification of novel non-coding RNAs using profiles of short sequence reads from next generation sequencing data. *BMC Genomics*, **11**, 77.
6. Bompfunewerer,A. (2005) Evolutionary patterns of non-coding RNAs. *Theory Biosci.*, **123**, 301–369.
7. Lu,Z.J., Yip,K.Y., Wang,G., Shou,C., Hillier,L.W., Khurana,E., Agarwal,A., Auerbach,R., Rozowsky,J., Cheng,C. *et al.* (2011) Prediction and characterization of noncoding RNAs in C. elegans by integrating conservation, secondary structure, and high-throughput sequencing and array data. *Genome Res.*, **21**, 276–285.
8. Artzi,S., Kiezun,A. and Shomron,N. (2008) miRNAminer: a tool for homologous microRNA gene search. *BMC Bioinformatics*, **9**, 39.
9. Hertel,J. and Stadler,P.F. (2006) Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. *Bioinformatics*, **22**, e197–e202.
10. Laslett,D. and Canback,B. (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.*, **32**, 11–16.
11. Lagesen,K., Hallin,P., Rødland,E.A., Staerfeldt,H.-H., Rognes,T. and Ussery,D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.*, **35**, 3100–3108.
12. Lowe,T.M. and Eddy,S.R. (1999) A computational screen for methylation guide snoRNAs in yeast. *Science*, **283**, 1168–1171.
13. Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **25**, 955–964.
14. Yoon,S. and De Micheli,G. (2005) Prediction and Analysis of Human microRNA Regulatory Modules. *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, **5**, 4799–4802.
15. Hertel,J., Hofacker,I.L. and Stadler,P.F. (2008) SnoReport: computational identification of snoRNAs with unknown targets. *Bioinformatics*, **24**, 158–164.
16. Griffiths-Jones,S., Bateman,A., Marshall,M., Khanna,A. and Eddy,S.R. (2003) Rfam: an RNA family database. *Nucleic Acids Res.*, **31**, 439–441.
17. Burge,S.W., Daub,J., Eberhardt,R., Tate,J., Barquist,L., Nawrocki,E.P., Eddy,S.R., Gardner,P.P. and Bateman,A. (2013) Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.*, **41**, D226–D232.
18. Thore,S., Mayer,C., Sauter,C., Weeks,S. and Suck,D. (2003) Crystal structures of the Pyrococcus abyssi Sm Core and its complex with RNA: common features of binding in archaea and eukarya. *J. Biol. Chem.*, **278**, 1239–1247.
19. Kiss,T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.*, **20**, 3617–3622.
20. Pannucci,J.A., Haas,E.S., Hall,T.A., Harris,J.K. and Brown,J.W. (1999) RNase P RNAs from some archaea are catalytically active. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 7803–7808.
21. Woodhams,M.D., Stadler,P.F., Penny,D. and Collins,L.J. (2007) RNase MRP and the RNA processing cascade in the eukaryotic ancestor. *BMC Evol. Biol.*, **7**(Suppl. 1), S13.
22. Perreault,J., Perreault,J.-P. and Boire,G. (2007) Ro-associated Y RNAs in metazoans: evolution and diversification. *Mol. Biol. Evol.*, **24**, 1678–1689.
23. Lustig,A.J. (1999) Crisis intervention: the role of telomerase. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 3339–3341.
24. Wang,C., Ding,C., Meraz,R.F. and Holbrook,S.R. (2006) PSoL: a positive sample only learning algorithm for finding non-coding RNA genes. *Bioinformatics*, **22**, 2590–2596.
25. Arrial,R.T., Togawa,R.C. and Brigido,M. (2009) Screening non-coding RNAs in transcriptomes from neglected species using PORTRAIT: case study of the pathogenic fungus Paracoccidioides brasiliensis. *BMC Bioinformatics*, **10**, 239.
26. Song,Y., Liu,C., Qu,J. and National Science Foundation (2009) 2009 IEEE INTERNATIONAL CONFERENCE ON BIOINFORMATICS AND BIOMEDICINE WORKSHOPS. *Learning parameters for non-coding RNA sequence-structure alignment*, IEEE , Washington DC, pp. 73–77.
27. Williams,P.H., Eyles,R. and Weiller,G. (2012) Plant MicroRNA prediction by supervised machine learning using C5.0 Decision Trees. *J. Nucleic Acids*, **2012**, 1–10.
28. Sturm,M., Hackenberg,M., Langenberger,D. and Frishman,D. (2010) TargetSpy: a supervised machine learning approach for microRNA target prediction. *BMC Bioinformatics*, **11**, 292.
29. Jha,A. and Shankar,R. (2011) Employing machine learning for reliable miRNA target identification in plants. *BMC Genomics*, **12**, 636.
30. Ding,J., Zhou,S. and Guan,J. (2010) MiRenSVM: towards better prediction of microRNA precursors using an ensemble SVM classifier with multi-loop features. *BMC Bioinformatics*, **11**, S11.
31. Zhao,Y. and Wang,Z. (2008) RNA secondary structure prediction based on support vector machine classification. *Sheng Wu Gong Cheng Xue Bao*, **24**, 1140–1148.
32. Zhao,Y. and Wang,Z. (2008) Consensus RNA secondary structure prediction based on support vector machine classification. *Chin. J. Biotechnol.*, **24**, 1140–1148.
33. Joachims,T. (1999) Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning. MIT-Press., **11**, 41–56.
34. Chang,C.-C. and Lin,C.-J. (2011) LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, **2**, 1–27.
35. Bu,D., Yu,K., Sun,S., Xie,C., Skogerbo,G., Miao,R., Xiao,H., Liao,Q., Luo,H., Zhao,G. *et al.* (2012) NONCODE v3.0: integrative annotation of long noncoding RNAs. *Nucleic Acids Res.*, **40**, D210–D215.

36. Li,W. and Godzik,A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–1659.
37. Han,L.Y. (2004) Prediction of RNA-binding proteins from primary sequence by a support vector machine approach. *RNA*, **10**, 355–368.
38. Bhasin,M. and Raghava,G.P.S. (2004) GPCRpred: an SVM-based method for prediction of families and subfamilies of G-protein coupled receptors. *Nucleic Acids Res.*, **32**, W383–W389.
39. Ding,Y. (2006) Statistical and Bayesian approaches to RNA secondary structure prediction. *RNA*, **12**, 323–331.
40. Friedel,M., Nikolajewa,S., Sühnel,J. and Wilhelm,T. (2009) DiProGB: the dinucleotide properties genome browser. *Bioinformatics*, **25**, 2603–2604.
41. Lorenz,R., Bernhart,S.H., Höner zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
42. Xue,C., Li,F., He,T., Liu,G.-P., Li,Y. and Zhang,X. (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. *BMC Bioinformatics*, **6**, 310.
43. Breiman,L. (1996) Bagging Predictors. Machine Learning., **24**, 123–140.
44. Vapnik,V.N. (1999) An overview of statistical learning theory. *IEEE Trans Neural Netw.*, **10**, 988–999.
45. Kaminski,W. and Strumillo,P. (1997) Kernel orthonormalization in radial basis function neural networks. *IEEE Trans Neural Netw.*, **8**, 1177–1183.
46. Yuan,G.-X., Ho,C.-H. and Lin,C.-J. (2012) Recent advances of large-scale linear classification. Proceedings of the IEEE., **100**.
47. Chen,Y., Lu,B.-L. and Zhao,H. (2012) Parallel learning of large-scale multi-label classification problems with min-max modular LIBLINEAR. *The 2012 International Joint Conference on Neural Networks (IJCNN).*, IEEE, Brisbane, pp. 1–7.
48. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
49. Flicek,P., Ahmed,I., Amode,M.R., Barrell,D., Beal,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fairley,S. *et al.* (2013) Ensembl 2013. *Nucleic Acids Res.*, **41**, D48–D55.
50. Tyc,K. and Steitz,J.A. (1989) U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. *EMBO J.*, **8**, 3113–3119.
51. Cavaillé,J., Hadjiolov,A.A. and Bachellerie,J.P. (1996) Processing of mammalian rRNA precursors at the 3' end of 18S rRNA. Identification of cis-acting signals suggests the involvement of U13 small nucleolar RNA. *Eur. J. Biochem. FEBS*, **242**, 206–213.
52. Brown,J.W. (1999) The Ribonuclease P Database. *Nucleic Acids Res.*, **27**, 314.
53. Jarrous,N. and Reiner,R. (2007) Human RNase P: a tRNA-processing enzyme and transcription factor. *Nucleic Acids Res.*, **35**, 3519–3524.
54. Badger,J.H. and Olsen,G.J. (1999) CRITICA: coding region identification tool invoking comparative analysis. *Mol. Biol. Evol.*, **16**, 512–524.
55. Liu,J., Gough,J. and Rost,B. (2006) Distinguishing protein-coding from non-coding RNAs through support vector machines. *PLoS Genet.*, **2**, e29.
56. Kong,L., Zhang,Y., Ye,Z.-Q., Liu,X.-Q., Zhao,S.-Q., Wei,L. and Gao,G. (2007) CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.*, **35**, W345–W349.
57. Gaspar,P., Moura,G., Santos,M.A.S. and Oliveira,J.L. (2013) mRNA secondary structure optimization using a correlated stem-loop prediction. *Nucleic Acids Res.*, **41**, e73.
58. SantaLucia,J. Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 1460–1465.
59. Mathews,D.H. and Turner,D.H. (2006) Prediction of RNA secondary structure by free energy minimization. *Curr. Opin. Struct. Biol.*, **16**, 270–278.
60. Goñi,J.R., Pérez,A., Torrents,D. and Orozco,M. (2007) Determining promoter location based on DNA structure first-principles calculations. *Genome Biol.*, **8**, R263.
61. Morin,R., Bainbridge,M., Fejes,A., Hirst,M., Krzywinski,M., Pugh,T., McDonald,H., Varhol,R., Jones,S. and Marra,M. (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques*, **45**, 81–94.

Gupta et al. BMC Genomics (2016) 17:112
DOI 10.1186/s12864-016-2455-2

BMC Genomics

RESEARCH ARTICLE Open Access

Dissecting genetics of cutaneous miRNA in a mouse model of an autoimmune blistering disease

Yask Gupta[1], Steffen Möller[1], Mareike Witte[1], Meriem Belheouane[3], Tanya Sezin[2], Misa Hirose[1], Artem Vorobyev[1], Felix Niesar[4], Julia Bischof[1], Ralf J. Ludwig[1], Detlef Zillikens[1,2], Christian D. Sadik[2], Tobias Restle[4], Robert Häsler[5], John F. Baines[3] and Saleh M. Ibrahim[1*]

Abstract

Background: MicroRNAs (*miRNAs*) are small endogenous non-coding RNAs that control genes at post-transcriptional level. They are essential for development and tissue differentiation, and such altered miRNA expression patterns are linked to the pathogenesis of inflammation and cancer. There is evidence that miRNA expression is genetically controlled similar to the transcription of protein-coding genes and previous studies identified quantitative trait loci (QTL) for miRNA expression in the liver. So far, little attention has been paid to miRNA expression in the skin. Moreover, epistatic control of miRNA expression remains unknown. In this study, we characterize genetic regulation of cutaneous miRNA and their correlation with skin inflammation using a previously established murine autoimmune-prone advanced intercross line.

Results: We identified *in silico* 42 eQTL controlling the expression of 38 cutaneous miRNAs and furthermore found two chromosomal hot-spots on chromosomes 2 and 8 that control the expression of multiple miRNAs. Moreover, for 8 miRNAs an interacting effect from pairs of SNPs was observed. Combining the constraints on genes from the statistical interaction of their loci and further using curated protein interaction networks, the number of candidate genes for association of miRNAs was reduced to a set of several genes. A cluster analysis identified miR-379 and miR-223 to be associated with EBA severity/onset, where miR-379 was observed to be associated to loci on chromosome 6.

Conclusion: The murine advanced intercross line allowed us to identify the genetic loci regulating multiple miRNA in skin. The recurrence of trans-eQTL and epistasis suggest that cutaneous miRNAs are regulated by yet an unexplored complex gene networks. Further, using co-expression analysis of miRNA expression levels we showed that multiple miRNA contribute to multiple pathways that might be involved in pathogenesis of autoimmune skin blistering disease. Specifically, we provide evidence that miRNA such as miR-223 and miR-379 may play critical role in disease progression and severity.

Keywords: MicroRNA, Expression QTL, Epistasis, Autoimmune skin blistering disease, Co-expression analysis

Background

The discovery of microRNAs (miRNAs), a class of small endogenous non-coding molecules ranging from 18–24 nt brought a new level of complexity for understanding the mechanisms that constitute various biological processes [1, 2]. The involvement of miRNAs in the control of gene expression has been thoroughly defined for the cell cycle, metabolism, and immune system and in cancer [3–6]. Binding to the 3′ or 5′ UTR region of genes, miRNAs may yield increased or decreased gene expression levels and have been described to affect various molecular pathways [7].

Approximately half the miRNAs are intergenic with few also located in intronic regions [8, 9]. These are understood to have their own enhancers and promoters and are transcribed by RNA polymerase II [10]. However, it remains unclear whether these are produced as

* Correspondence: saleh.ibrahim@uksh.de
[1]Lübeck Institute of Experimental Dermatology, University of Lübeck, Lübeck, Germany
Full list of author information is available at the end of the article

by-products of protein-coding gene transcription or whether their biogenesis has its own machinery [11]. Recent studies in murine and human fibroblasts of liver tissues revealed that miRNA expression can either be regulated by their transcriptional genomic location or by other regions in the genome [12]. Furthermore, mutations in genes involved in miRNA processing, such as *AGO1*, *DGCR8*, and *DICER*, can cause significant changes in the expression of miRNAs, resulting in altered disease susceptibility [13]. Regarding the skin, in vivo studies show that miRNA biogenesis is dependent on both *DICER* and *DGCR8*. A lack of these enzymes causes severe phenotypes [14], underscoring the importance of miRNAs in the regulation of morphogenesis and homeostasis of the skin [15]. Differentially expressed miRNAs are associated with different physiological and pathological processes in the skin such as melanoma, Sézary syndrome, psoriasis and atopic dermatitis [16–19]. Advancements have been achieved in understanding various processes regulated by miRNAs. On the other hand, the mechanisms that regulate their own expression have so far remained unknown. One possible approach towards understanding their regulation is constraining on genetic loci whose variations statistically link the variable phenotypic effect. We thus performed an *in silico* expression QTL-based analysis, which has been widely used for deciphering genetic loci regulating gene expression in various biological processes [20]. Su et al. have further shown that expression QTL-based analysis can provide insights into miRNA regulation [12]. However, no study has yet been performed to determine the impact of genetic variation on miRNAs associated to skin and autoimmune disorders.

This study (i) uses expression QTL-based analysis to define genetic loci that control miRNA expression in the skin and (ii) provides insights into interacting genes to control a single transcript (epistasis). The relevance of these findings is illustrated by employing a mouse model for the autoimmune blistering disease EBA (Epidermolysis Bullosa Acquisita), providing a direct link between miRNA expression and disease phenotype. Taken together, the data provide insights into the complexity of miRNA regulation and possible means to understand various gene interactions altering the expression of miRNAs.

Results

A murine heterogeneous inter-cross line was generated by inter-crossing four parental inbred mouse strains: MRL/MpJ, NZM2410/J and BXD2/TyJ as autoimmunity-prone strains and Cast/EiJ for genetic heterogeneity. As a mouse model for an autoimmune blistering disease, 100 mice from generation 4 of the intercross-line were immunized with recombinant collagen type VII (COL7), an integral component of anchoring fibrils located at the dermal-epidermal junction. This induced a loss of tolerance and production of anti-COL7 autoantibodies in all immunized mice. The incidence of sub-epidermal blisters and the clinical phenotype of epidermolysis bullosa acquisita (EBA) were at 1/3 [21]. All mice were genotyped and miRNA expression profiling was acquired from skin tissue using the Affymetrix's GeneChip miRNA 2.0 array. To reveal genetic loci that regulate miRNA expression, we then performed an association study between the genotype and the respective expression levels of miRNAs. Further we performed co-expression analysis between expression levels of miRNAs and phenotype. The workflow is presented in a flowchart (Additional file 1: Figure S1).

miRNA expression is genetically controlled

We performed a genome-wide scan to detect genetic loci associated with miRNA expression. Expression levels of miRNA were treated as quantitative trait to yield expression quantitative trait loci (eQTL). Genome-wide significance was determined by a permutation test. Using an E-value cutoff of <0.05, 42 eQTL for 38 miRNAs were mapped to the genome, corresponding to 6.83 % of all murine miRNAs present on the above mentioned Affymetrix GeneChip (Table 1 and Fig. 1). Since the wild derived strain CAST/EiJ was incorporated into the advanced intercross line, we investigated the polymorphic sites located in the transcribed miRNAs that may have an effect of probe hybridization which is derived from C57B6/J [22]. We obtained the SNPs and indels in genome of CAST/EiJ from the database [23]. We found that 4/38 (10.53 %) miRNAs (miR-291a-3p, miR-341, miR-449b and miR-681) exhibits indels in CAST/EiJ strain on chromosome 7 (3.2 Mb), 12 (69 and 109 Mb) and 13 (113 Mb), which may suggest false positive associations for those loci (Table 1). The highest -log *P* value of 6.57 was observed for miR-298 on chromosome 9 explaining 20.76 % of the variance. The peak SNP (rs3700596) was found within 1 kb of the *Ube3D* gene, an ubiquitin-conjugating enzyme *E2c* binding protein. We found only trans – eQTL, except for miR-486 (–log $P = 4.10$, rs13479880), which was mapped to the same chromosome of its transcriptional site (Chr 8, ~89 Mb). This indicates that other genes rather than their own transcript may contribute to the regulation of miRNA expression which could be tissue specific.

Prior knowledge suggests classes of gene families that play a major role in the regulation of miRNA expression i.e. Argonaute proteins and helicases which play a crucial role in its regulation [24]. To examine this assumption, we obtained the coordinates of all helicases and other genes involved in pathways of miRNA biogenesis from 'The RNA Helicase Database' and mapped them to the eQTL identified in our study [25]. Helicases like *Ddx39*, *Ddx49*, *Cd97* and *Upf1* were mapped within the

Table 1 eQTL detected for the expression of miRNA

AffyID	PEAK SNP	Chr	Pos(Mb)	CI	-logP score	α	Transcript Location (Chr:Pos)
miR-322	rs6386920	1	54.1	51.5–58.4	4.797697778	0.02	X:50407432-50407526
miR-431	rs6386920	1	54.1	51.5–59.8	5.231417376	0.006	12:110828657-110828747
miR-26a	rs6265423	2	47	28.1–51.3	4.376166299	0.021	10:126432586–126432669,9:118940914-118941003
miR-291a-3p	CEL.2_50605053	2	50.5	28.1–59.9	4.546953451	0.018	7:3218920-3219001
miR-423-3p	rs6265423	2	47	33.9–51.3	4.948513523	0.015	11:76891566-76891674
miR-671-5p	rs13476472	2	45.5	33.9–50.5	5.838570402	0.004	5:24097932-24098029
miR-23b	rs6250599	2	48.5	44.3–51.3	4.335454904	0.022	13:63401792-63401865
miR-409-5p	rs13476874	2	159.5	136–163	4.175101842	0.035	12:110981368-110981446
miR-409-5p	rs13477083	3	43.3	27.5–45.4	4.215972522	0.033	12:110981368-110981446
miR-546	rs13477126	3	56.4	52.9–58.3	4.493016911	0.029	10:126435496-126435616
miR-200a-star	rs3671119	3	126.1	117.4–131.3	4.688008826	0.011	4:155429005-155429094
miR-339-5p	rs3660863	4	7.1	3.7–19.5	4.878132165	0.003	5:139845604-139845699
miR-465c-5p	rs13477873	4	101.1	82.8–118	4.472478928	0.01	X:64079130-64079210
miR-295	rs3663950	4	135.3	129.4–141.1	4.209508065	0.04	7:3220774-3220842
miR-878-3p	rs13478002	4	136.2	135.3–141.1	5.110152979	0.01	X:64054683-64054760
miR-742	rs3673049	5	90.1	87.5–96.6	4.765858909	0.035	X:64033548-64033612
miR-379	rs6208251	6	104.8	98.4–116.7	4.024954772	0.034	12:110947270-110947335
miR-154	rs13479063	6	136.3	133.92–142.4	4.309450382	0.032	12:110976643-110976708
miR-425*	rs3663988	7	146.5	140.2–146.5	4.342220512	0.035	9:108471108-108471192
miR-486	rs13479880	8	89.3	72.5–95.1	4.108040959	0.033	8:24253027-24253154
miR-487b	rs6257357	8	88	77.7–90.2	4.81779528	0.019	12:110965543-110965624
miR-501-3p	rs13479880	8	89.3	81.7–95	4.205948891	0.03	X:6818369-6818477
miR-130b	rs6413270	9	37.7	36.8–44.4	4.222559895	0.032	16:17124154-17124235
miR-298	rs3700596	9	86.2	85–90.6	6.573678024	0.001	2:174093005-174093086
miR-466c-3p	rs3712394	10	17.7	14.1–24.2	4.569565941	0.03	2:10403161-10403244
miR-466c-3p	rs13480563	10	27.9	24.2–38.7	4.591526295	0.029	2:10403161-10403244
miR-126-5p	rs6374078	10	60.6	30.7–65.9	4.177297455	0.031	2:26446877-26446949
miR-681	rs13481076	11	66.5	40.1–71.3	4.055742697	0.043	12:70864822-70864931
miR-20a-star	rs3712881	11	120.9	112–121	4.160612966	0.03	14:115443379-115443485
miR-203	mCV22351241	12	60	55.0–72.6	3.838027866	0.039	12:113369091-113369166
miR-542-3p	CEL.12_84750094	12	91.4	79.7–103.8	4.735149963	0.023	X:50402580-50402664
miR-341	gnf13.079.671	13	80.5	69.5–88.3	5.260091403	0.008	12:110849710-110849805
miR-449b	rs13482231	14	67.6	50.9–72.3	4.485232377	0.019	13:113827627-113827706
miR-7a	CEL.15_4222769	15	4.4	3.2–9.7	4.941341377	0.044	13:58494140–58494247, 7:86033163-86033259
miR-7a	rs13482455	15	16.7	14.8–24.5	5.73214998	0.022	13:58494140–58494247, 7:86033163-86033259
miR-337-3p	rs13482549	15	45.5	38.4–53.9	4.298224443	0.029	12:110823999-110824095
miR-673-5p	rs13482549	15	45.5	38.9–56.6	4.592339382	0.014	12:110810200-110810290
miR-136	rs13482914	17	21	16.5–27.6	4.679083445	0.026	12:110833537-110833598
miR-466b-5p	rs13483212	18	12.3	0–21	4.48515896	0.016	2:10395846-10395927
miR-493	rs6211533	19	57.1	50.2–60.2	4.98913787	0.004	12:110818443-110818525
miR-26a	gnfX.023.543	X	36.3	0–49.3	4.718906745	0.01	10:126432586–126432669, 9:118940914-118941003
miR-466e-5p	rs13483712	X	9.1	0–33.5	4.013512497	0.05	2:10398257-10398340

AffyID Affymentrix ID, *Chr* Chromosome, *Pos (Mb)* Peak Position SNP in Mb, *CI* Confidence Interval (CI) in Mb using 1.5 -log P drop, *α* Genome wide significance

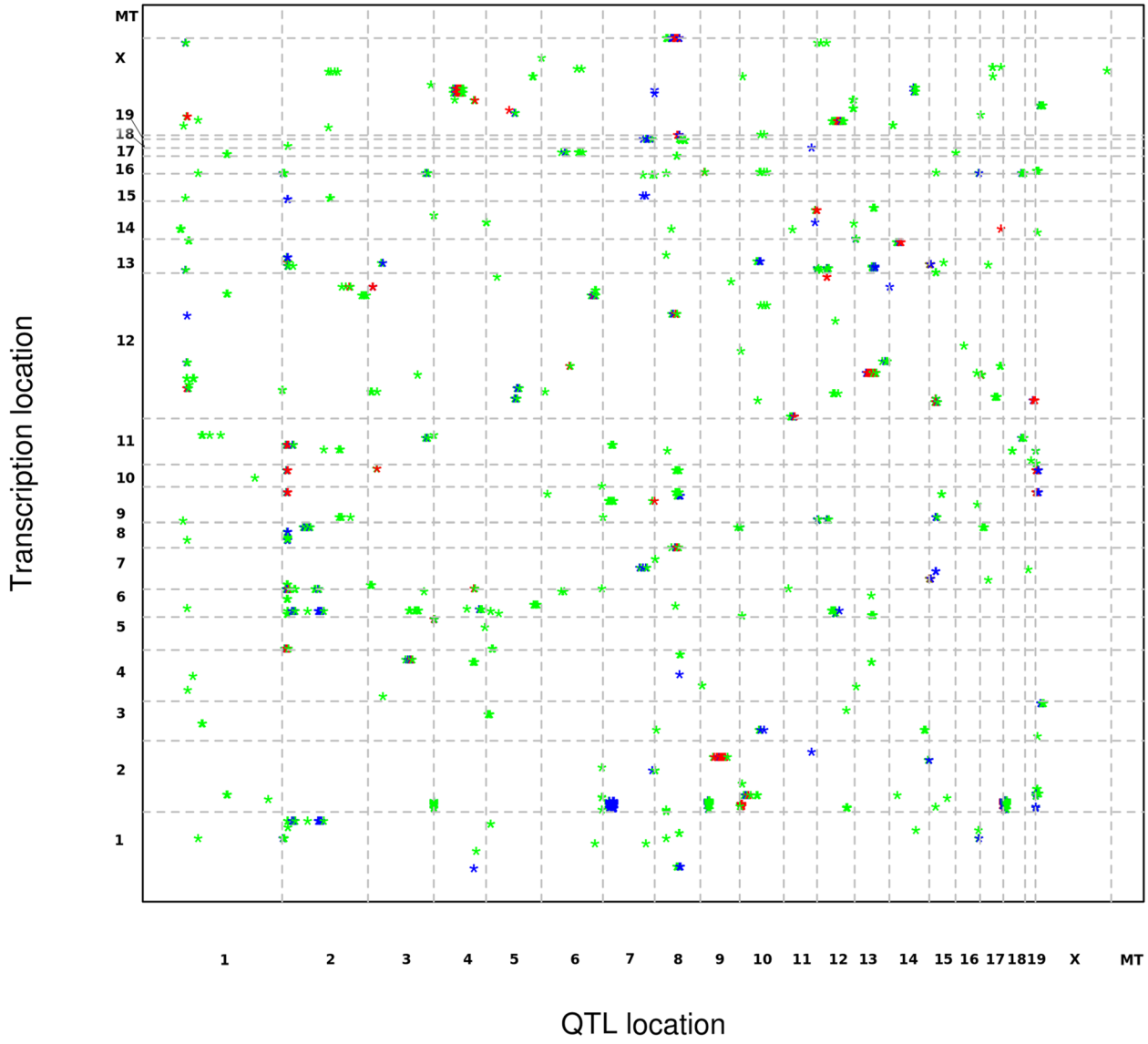


Fig. 1 Transcriptome map for eQTL of miRNA . The x-axis in the figure represents the SNP location in the chromosome, and the y axis represents the transcriptional site of miRNAs. In this figure genome wide significant threshold of $\alpha < 0.25$ is represented in a green. 83 miRNA eQTL were defined as suggestive eQTL for miRNA (genome wide threshold, $\alpha < 0.1$) are represented in blue. 43 miRNA eQTL with $\alpha < 0.05$ were observed as significant eQTL are represented in red

confidence interval of miR-486, miR-487b and miR-501 on chromosome 8. Further, four helicases (*Ddx50*, *Ascc3*, *Ddx21* and *Dna2*) were mapped within the confidence interval of the eQTL controlling miR-126. Genes that are involved in transcriptional processes, such as *Polr3f*, *Polr2a*, *Polr3g* and *Polr2a* were also mapped to the eQTL for miR-409, miR-681, miR-34 and miR-449. Other genes which do not belong to these classes, such as lin28a and its homolog lin28b, have been shown to modulate let-7a [26]. These genes were mapped to the eQTL for miR-290 on chromosome 4 and miR-126 on chromosome 10, suggesting that *Lin28* might modulate the expression of other miRNAs as well.

Some miRNA eQTL were constrained to a specific location in the genome, thereby suggesting potential eQTL hotspots for the regulation of multiple miRNAs. On chromosome 2, five miRNAs (miR-26a, mir-291a, miR-423, miR-671 and miR-23b) were mapped between 28–51 Mb (Fig. 2). Three nearby SNPs (rs6250599 ~ 48.5 Mb, rs6265423 ~ 47 Mb and rs13476472 ~ 45 Mb) showed significant association ($-\log P > 4$) with all five miRNAs that were mapped to this region. A long non coding RNA *1700019E08Rik* was located near SNP rs13476472 (~3 Kb upstream). As for the other two SNPs, the nearest gene to rs6265423 was mapped 6.8 kb apart, coding for the snRNA *U7.39-201*, while the nearest coding gene for SNP

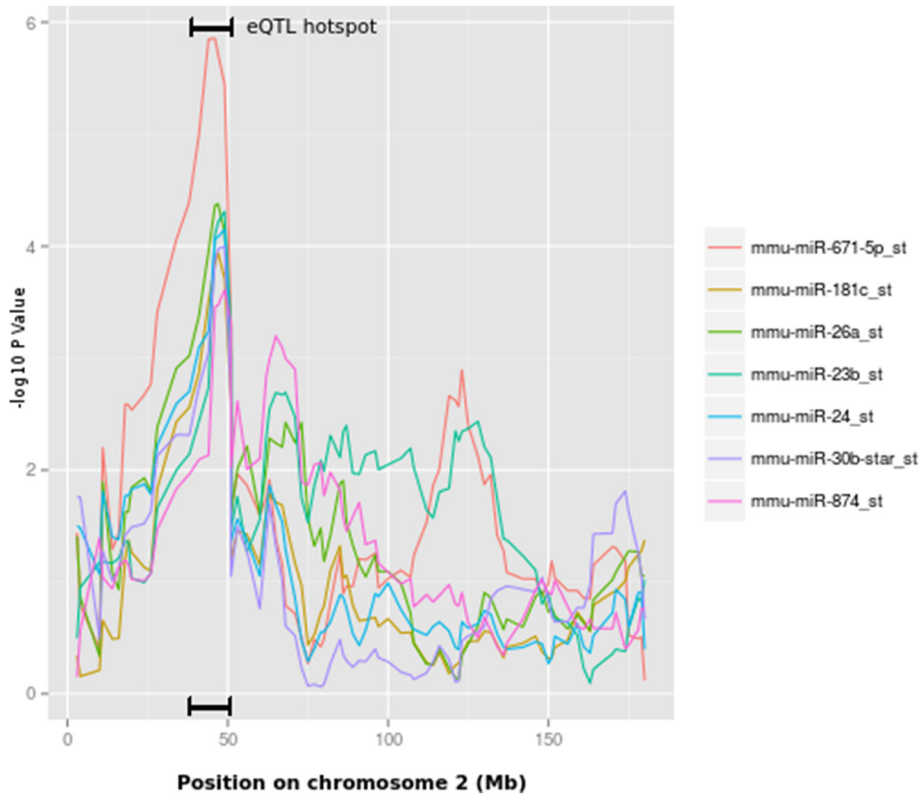


Fig. 2 eQTL hot spots on chromosome 2. Position on x axis represents the coordinates in million base pairs on chromosome 2. The y axis is –log *P* value. The overlapping peaks on y axis represents an eQTL hot spot between 28–51 Mb. miR-181, miR-30b* and miR-874 are additional suggestive eQTL with significant genome-wide (α < 0.1) threshold after permutation

rs6250599 was a pseudogene *Gm13489-001* (~13 kb upstream).

Similarly, on chromosome 8 we identified eQTL for three miRNAs (miR-486, miR487b and miR-501) in confidence interval 72–95 Mb. Three genes *Gm1068*, *Siah1* and *Dnaja2* were located near peak SNPs (rs13479880 and rs6257357).

Furthermore, we found several miRNAs that are associated to more than one locus in the genome. For example miR-7a showed significant association with two loci present on chromosome 15 (3–9 Mb, CEL.15_4222769 and 14–24 Mb, rs13482455). Similarly, miR-466-3c was regulated by two nearby loci on chromosome 10 (14–24 Mb, rs3712394 and 24–38 Mb, rs13480563). Additionally, miR-26a was mapped with two loci on chromosomes 2 and X (28–51 Mb, rs6265423 and 0–49 Mb, gnfX.023.543).

Epistatic control of miRNA expression

A defective of single loci may possibly be compensated by another, or only jointly the effect is the strongest. Therefore, for all the miRNAs significant single-locus eQTL, we analyzed the epistatic effect for each SNP pair. As a result, we identified 200 SNP pairs for 8 miRNAs below the significance level (adjusted *P* value < 0.05) (Fig. 3 and Additional file 1: Table S1). The highest -log *P* value of 10.38 was found between the SNP pairs rs13480360 (chr 10, ~67 Mb, nearest gene: *AK139516*) and rs3689658 (chr 2, ~ 85 Mb, nearest gene: *Olfr1006*) for miR-7a (Table 2). In total, we found 119 SNP pairs for miRNA miR-7a. The hub locus (i.e. SNP with the maximal number of interactions) for miR-7a was observed on chromosome 16 (rs3680665 ~ 84 Mb, nearest gene: *AK04263*). The same SNP (rs3680665) also showed association with miR-542 additionally with another SNP (rs4200124, nearest gene: *Gbe1*) present nearby. For miR-742 we observed 47 SNP pairs with SNP rs3657112 (~148 Mb, nearest gene: *Snora17*) on chromosome 3 showing the highest number of associations (40 SNP pairs). The same SNP i.e. rs3657112 also showed statistical interaction with chromosome 9 loci 67–72 Mb for miR-295. We also found multiple SNPs on chromosome 2 (13–17 Mb) statistical interacting with SNPs on chromosome 1(3–11 Mb) for miR-501. Two miRNAs (miR-136 and miR-337) had only one significant SNP pair: for miR-136 the correspondence was found between SNPs rs37113033 (chr 19, ~5 Mb, nearest gene: *Slc29a2*) and rs13459176 (chr 15, ~3 Mb, nearest gene: *Sepp1*), while

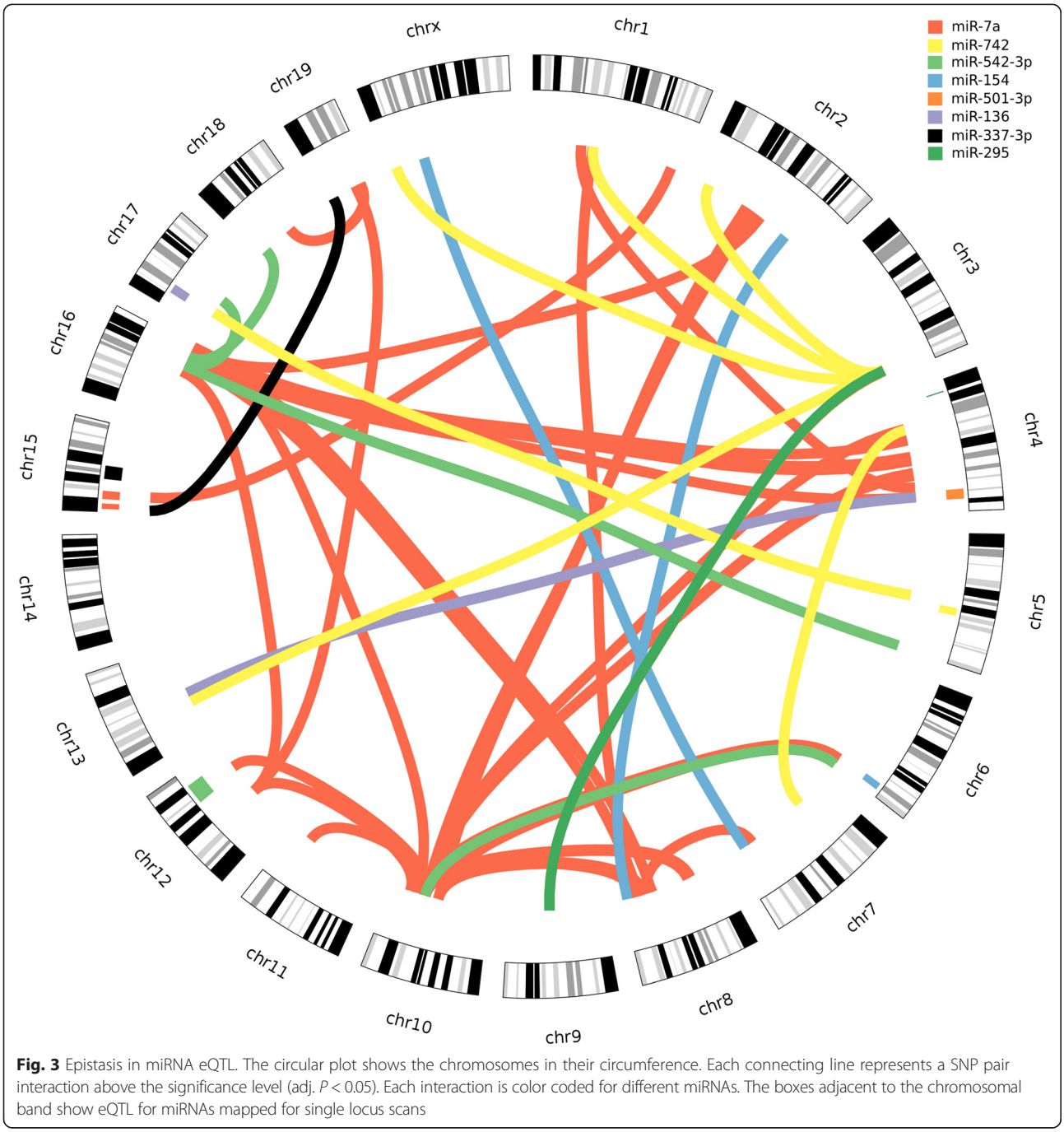


Fig. 3 Epistasis in miRNA eQTL. The circular plot shows the chromosomes in their circumference. Each connecting line represents a SNP pair interaction above the significance level (adj. $P < 0.05$). Each interaction is color coded for different miRNAs. The boxes adjacent to the chromosomal band show eQTL for miRNAs mapped for single locus scans

miR-337 had SNP pair rs3693942 (chr 13, ~55 Mb, nearest gene: *Unc5A*) and rs3663950 (chr4, ~135 Mb, nearest gene: *Il22ra*) (Table 2).

To search the candidate interacting gene pairs we investigated the epistatic control of miR-501-3p. Using the Ingenuity pathway analysis (IPA), we searched for all the possible interacting genes present between loci present on chromosome 1 and chromosome 2 [27]. We found three putative interacting gene pairs; namely *Commd3* with *Cops5*, *Cacnb2* with *Vopip1* and *Commd3-Bmi1* with *Rb1cc*. Interestingly, *Cops5* in the *Nfkb1* pathway possibly be associated with miR-501 via *Tp53* (Additional file 1: Figure S2).

Genetic overlap of QTL for clinical scores and miRNA expression

Any genetic variation with an effect on the clinical phenotype and miRNA levels should map to the same chromosomal region. Previously, genetic loci for the clinical phenotype of EBA, an autoimmune skin blistering disease,

Table 2 Epistasis in miRNA. The table includes only the top interacting SNPs for each miRNAs

miRNA	Top interacting SNP 1					Top interacting SNP 2					I.Pval(log)
	SNP ID	chr	Pos(Mb)	Refseq Genes	Pval(log)	SNP ID	chr	pos(Mb)	Refseq Genes	Pval(log)	
miR-7a	rs13480630	10	67.29		0.89	rs3689658	2	85.52	Olfr1006	1.12	10.4
miR-136	rs3713033	19	5.03	Slc29a2	0.073	rs13459176	15	3.23	Ccdc152, Sepp1	1.56	7.2
miR-154	rs3708073	8	124.28	Gm20388,Jph3	0.3	rs8246404	2	136.7	Mkks	0.57	7.67
miR-295	rs13480271	9	72.68	RP23-461P14	0.018	rs3657112	3	148.03		0.32	8.064
miR-337-3p	rs3693942	13	55.05	Unc5a	0.026	rs3663950	4	135.29	Il22ra1	0.2	7.17
miR-501-3p	CZECH-2_15618849	2	15.5	Gm13364	1.33	rs3716083	1	9.01	Sntg1	0.72	8.46
miR-542-3p	rs4200124	16	70.7		1.12	rs3718776	5	150.4	Wdr95	0.14	8.5
miR-742	rs3657112	3	148.028		0.35	CEL.1_49993068	1	49.68		0.33	9.16

Chr Chromosome, *Pos (Mb)* Position in million base pairs, *Pval(log)* Additive log 10 *P* value, *I.Pval(log)* Interacting log 10 *P* value

were studied using a larger cohort of mice from the same breeding scheme. Previously, we mapped the QTL for the onset of EBA to chromosomes 9 (39.5–46.3 Mb), 12 (83.2–109.9 Mb), 14 (49.1–68.9 Mb) and 19 (46.0-end Mb). Additionally, three QTL were mapped for severity of EBA on chromosomes 1 (3.1–27.3 Mb), 15 (61.0–85.7 Mb) and 19 (43.1-end Mb) [21]. In this study, we found 4 eQTL for miRNAs overlapping with the QTL for EBA (Fig. 4). The eQTL for miR130b (Chr 9: 36–44 Mb, −log $P = 4.42$), miR-542-3p (Chr 12: 79–103 Mb, −log $P = 4.73$) and

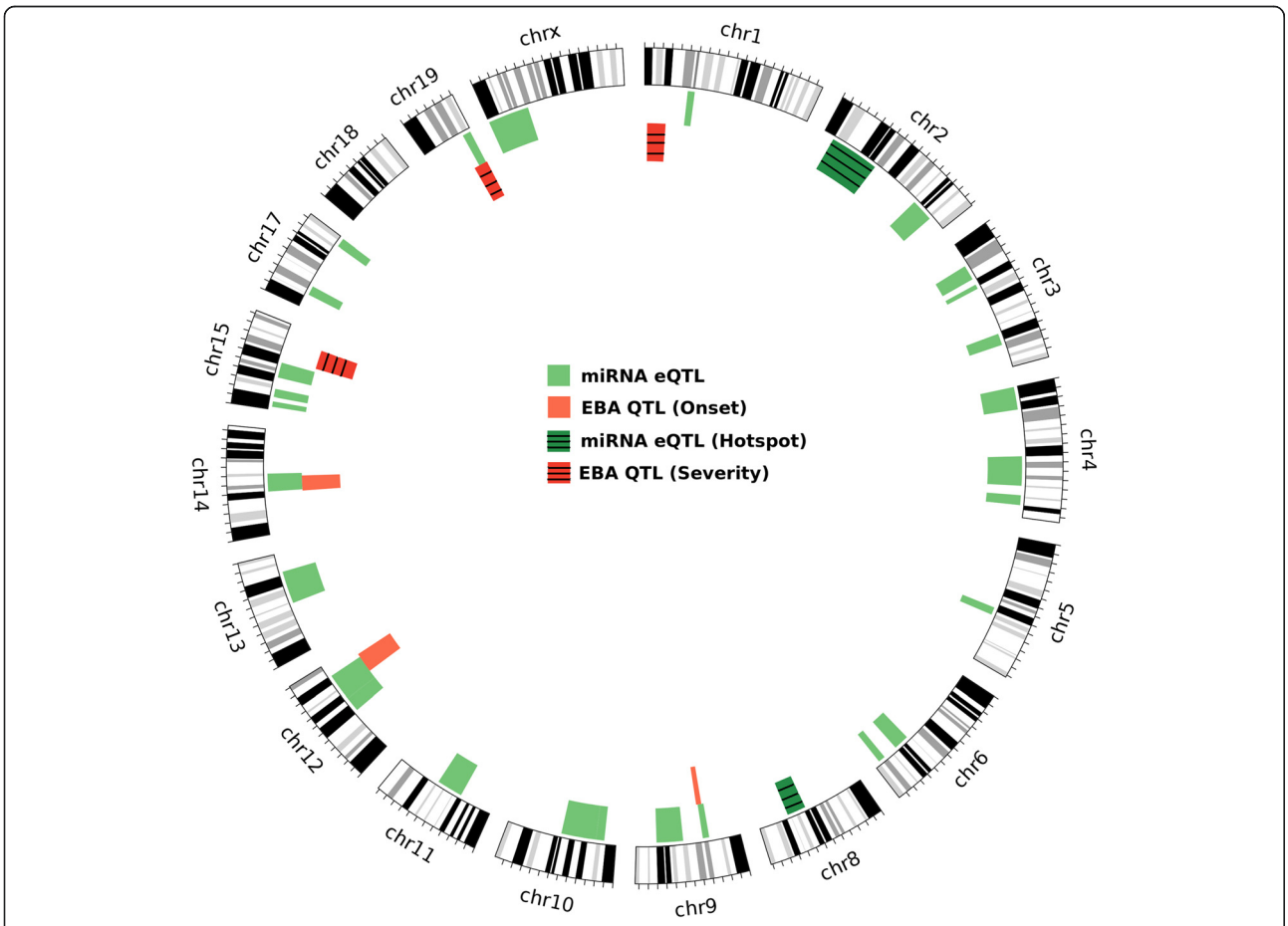


Fig. 4 Overlapping QTL for EBA disease and miRNA eQTL. The circular plot shows all the eQTL for miRNAs (green) and QTL for EBA (red). It also presents the eQTL hot spots (dark green) and EBA QTL for onset (red) and severity (dark red). Each circular band represents a chromosome on which QTL and eQTL are mapped. The region within the chromosome which has either red or dark red and green or dark green bands is overlapping eQTL with EBA QTL

miR-449b (Chr 14: 50–72 Mb, −log $P = 4.49$) were mapped on QTL for disease onset on chromosome 9, 12 and 14. Additionally, we mapped eQTL for miR-493 (50–60 Mb, −log $P = 4.98$) to the QTL for both disease severity and onset on chromosome 19. To further evaluate, if the variation in the genome is associated with the miRNA expression levels and also underlies the clinical phenotype, we investigated the genotype association of significant SNPs of miRNA eQTL with clinical phenotype (Fig. 5). We observed that the most significant SNP (rs6211533) mapped for the eQTL of miR-493 on chromosome 19 also shows variation for EBA severity score and onset. In this eQTL minor allele (BB) derived from MRL/MpJ was associated with higher expression of miR-493, late onset of the disease, and higher clinical score (Fig. 5b). Similar observation was also derived for eQTL mapped (peak SNP, rs6413270) for miR-130b where minor allele (BB) derived for MRL/MpJ and NZM/2410J was associated with lower expression of miR-130b, and late onset of disease phenotype (Fig. 5a).

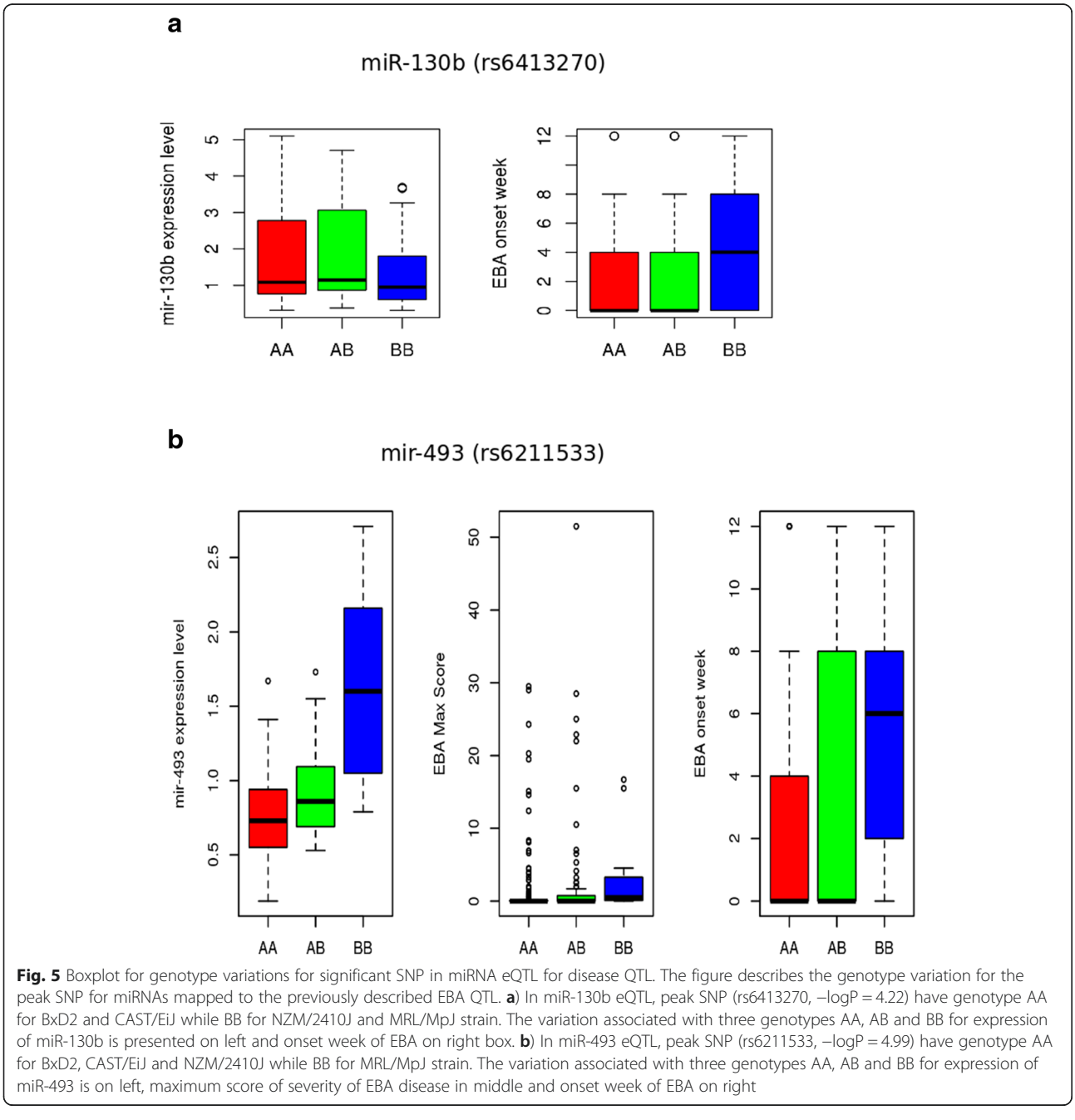


Fig. 5 Boxplot for genotype variations for significant SNP in miRNA eQTL for disease QTL. The figure describes the genotype variation for the peak SNP for miRNAs mapped to the previously described EBA QTL. **a**) In miR-130b eQTL, peak SNP (rs6413270, −logP = 4.22) have genotype AA for BxD2 and CAST/EiJ while BB for NZM/2410J and MRL/MpJ strain. The variation associated with three genotypes AA, AB and BB for expression of miR-130b is presented on left and onset week of EBA on right box. **b**) In miR-493 eQTL, peak SNP (rs6211533, −logP = 4.99) have genotype AA for BxD2, CAST/EiJ and NZM/2410J while BB for MRL/MpJ strain. The variation associated with three genotypes AA, AB and BB for expression of miR-493 is on left, maximum score of severity of EBA disease in middle and onset week of EBA on right

Expression and co-expression of miRNAs

Immunization of mice with recombinant COL7 leads to development of subepidermal blisters and the clinical phenotype of EBA in 1/3 of the immunized mice, while 2/3 remain clinically healthy. To access differential expression of miRNAs, we divided the 4^{th} generation of our mouse cohort into two separate groups: affected and non-affected mice. Comparing the two groups, only two miRNAs were differentially expressed: miR-379 (adj. P value = 0.044) and miR-223 (adj. P value = 0.044) (Fig. 6 and Additional file 1: Table S2). Both miRNAs were significantly over expressed in mice with disease. Additionally, an eQTL for miR-379 (–log P = 4.7, 98–116 Mb) was also mapped on chromosome 6, with its peak at ~104 Mb (rs6208251, nearest gene: *Cntn6*).

The comparison of diseased and non-diseased mice has its drawbacks; the disease severity observed for individual mice might differ depending on the differences in the genome. To investigate miRNA expression affecting disease severity, we performed co-expression analysis among the expression levels of miRNAs (co-expressed miRNAs are in same pathway) in correlation with quantitative scores for severity and onset week of disease. For this purpose, we employed the WGCNA R package, which clusters miRNAs into different modules and further associates them to a phenotypic score, such as EBA severity or disease onset. As a result, we identified 11 clusters (Fig. 7). Using this approach, only the 'black' module, consisting of 23 miRNAs, was significantly associated with EBA (ρ = 0.28, P = 0.005) (Additional file 1: Table S3). Additionally, it was significantly correlating to the onset of EBA (ρ = 0.29, P = 0.005). Due to the fact that the 'black' module stronger correlates with the onset of the disease than with the maximum score, one can speculate that miRNAs from this module is rather involved in the onset than in the severity of the disease. This would relate to our previous observation, in which miRNA eQTL were overlapping with EBA onset QTL. The pathways associated with the 'black' module were predominantly pathways which have been shown to play a crucial role in other autoimmune disorders, such as the MAPK signaling pathway, T-cell receptor and TGF-beta (Table 3).

Individual correlation of miRNA expression levels with the disease severity identified 24 miRNAs to be significant (P < 0.05) (Additional file 1: Table S3), with miR-223 showing the strongest association (ρ = 0.4, P = 4.93e-05). To further validate this result we experimentally verified the expression of miR-223 in EBA skin by qRT-PCR. We found that miR-223 is upregulated in EBA skin in comparison to normal skin (Additional file 1: Figure S3). Another miRNA which was highly correlating with EBA was miR-21 (ρ = 0.36, P = 0.00036). Further, investigations of the co-expression module showed that miRNAs were potentially co-regulated i.e. possible overlapping genetic loci among co-expressed miRNAs. We observed that miRNAs in a given locus were either clustered within the same

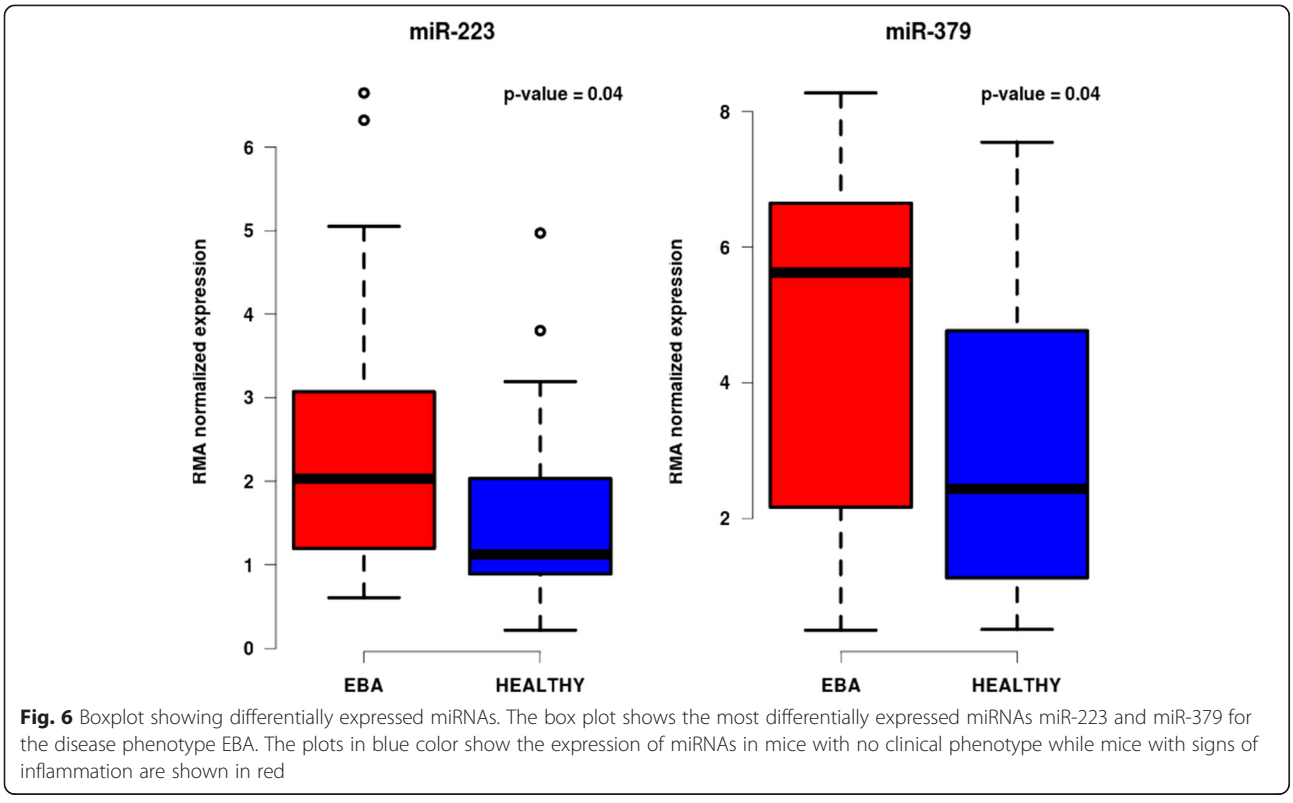


Fig. 6 Boxplot showing differentially expressed miRNAs. The box plot shows the most differentially expressed miRNAs miR-223 and miR-379 for the disease phenotype EBA. The plots in blue color show the expression of miRNAs in mice with no clinical phenotype while mice with signs of inflammation are shown in red

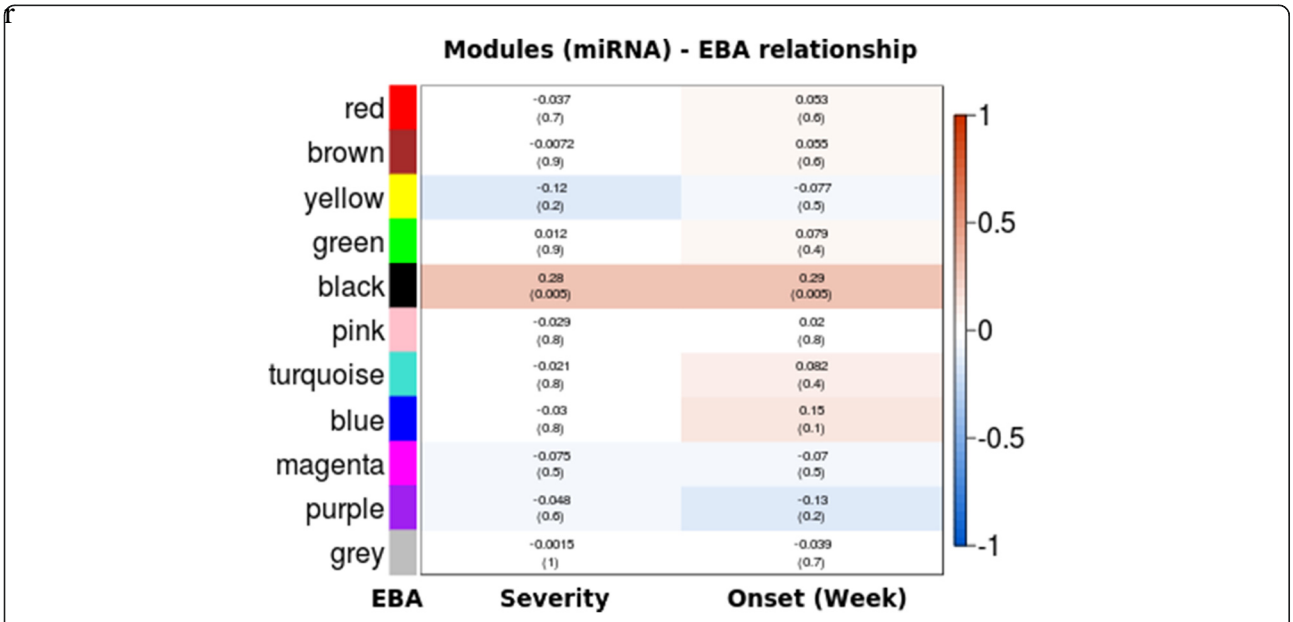


Fig. 7 Module-trait relationship between clusters of miRNAs with EBA severity and onset. The graph is a representation of co-expression analyses of miRNAs. The clusters of miRNAs are called modules which are color coded on the y-axis. The disease phenotypes (traits) are EBA onset and EBA severity given on the x-axis. Blocks represent the correlation of module with phenotype using a Pearson correlation coefficient ranging from −1 to 1. The range is color coded with red representing positive correlation and blue representing negative correlation. *P* values are given in brackets below the correlation coefficients

module or showed stronger inter-module membership for a specific module even if they were assigned to different modules. As an example, miR-322 and miR-431, that were mapped on chromosome 1 (51–59 Mb), are clustered in the 'red' module. In eQTL hot spots as on chromosome 2 (28–51 Mb), miR-423-3p and miR-23b are clustered in the 'yellow' module. Even though other miRNAs in this locus were assigned to a different module, they also show significance for the 'yellow' module. Examples are given by miR-671-5p ($P = 1.224713e-12$), miR-26a ($P = 2.654271e-19$) and miR-291a-3p ($P = 3.530522e-02$) (Additional file 1: Table S3). In line with this observation, for the eQTL mapped on chromosome 8, two miRNAs (miR-501-3p and miR-486) were clustered with the 'brown' module while miR-487 was assigns to the 'red' module but had significant module membership with the brown module as well ($P = 0.004$). Thus, these genomic loci may be considered as confirmed and strengthen the genetic contribution for the control of miRNAs expression levels.

Discussion

Small non-coding RNAs like miRNAs are known to contribute to the onset and severity of various diseases as well as the defense against them [28]. Thus, miRNAs function as tissue-specific key regulators, affecting some of the major pathways towards an aggravation of disease severity when aberrantly expressed [29]. Accordingly, it is not of surprise that miRNAs have been recently recognized as potential therapeutic targets [30, 31].

However, the underlying mechanisms of such dysregulated miRNA expression patterns are not well characterized. Different studies have shown that gene expression alterations in different tissues are genetically derived [29]. Thus, it is plausible that not only the regulation of gene expression is genetically controlled, but also the expression of miRNAs. In this study we explore the diversity of miRNAs in inflamed skin tissue and genetic loci that control variations in miRNA expression levels across a mouse cohort. We provide evidence that miRNA levels in skin tissue are genetically controlled on transcriptional level by helicases and RNA polymerases that are important for the biogenesis of miRNA. Furthermore, we found that some of the miRNA eQTL are restricted to one particular locus in the genome (eQTL hot spots). Deeper investigation revealed that these miRNAs are under multi-locus and/or epistatic control.

Interestingly, eQTL hot spots were predominantly found in genomic regions coding for non-coding RNA. Hence, it is tempting to speculate miRNA expression might not necessarily be solely controlled by protein-coding RNA, but rather by non-coding RNA which would impose an additional level of post-transcriptional regulation. This in turn leads to the tempting hypothesis that non-coding RNAs do at least in part regulate miRNA expression. Such a scenario is supported by the fact that some non-coding RNAs have been shown to

Table 3 List of over-represented KEGG pathways for 'black' module

Total Genes of the Term	Union	Targets in the Term	Union miRNAs in the Term	Score
AXON_GUIDANCE	131	51	17	1.88
PATHWAYS_IN_CANCER	323	98	18	1.788
MAPK_SIGNALING_PATHWAY	271	89	17	1.69
NEUROTROPHIN_SIGNALING_PATHWAY	131	45	16	1.557
T_CELL_RECEPTOR_SIGNALING_PATHWAY	109	42	16	1.424
ENDOCYTOSIS	219	65	16	1.408
TGF-BETA_SIGNALING_PATHWAY	85	30	15	1.334
GLIOMA	65	27	16	1.319
PROSTATE_CANCER	89	32	17	1.314
UBIQUITIN_MEDIATED_PROTEOLYSIS	138	52	16	1.277
COLORECTAL_CANCER	65	25	16	1.271
WNT_SIGNALING_PATHWAY	153	38	18	1.268
MELANOGENESIS	100	34	18	1.256
ERBB_SIGNALING_PATHWAY	87	33	17	1.246
MTOR_SIGNALING_PATHWAY	53	22	13	1.154
INSULIN_SIGNALING_PATHWAY	137	45	16	1.151
B_CELL_RECEPTOR_SIGNALING_PATHWAY	76	23	16	1.137
FOCAL_ADHESION	197	50	18	1.111
ADHERENS_JUNCTION	74	28	17	1.097
CHAGAS_DISEASE	102	34	16	1.094
GNRH_SIGNALING_PATHWAY	99	32	15	1.042
MELANOMA	71	26	16	1.037
PANCREATIC_CANCER	70	27	15	1.028
ENDOMETRIAL_CANCER	52	19	16	1.022

bind to miRNAs at functional level as demonstrated by an interaction of *linc-MD1* with miR-133 and miR-135 [32]. Here, *linc-MD1* works as a sponge and traps these miRNAs preventing the binding to the canonical targets. Moreover, a recent study even shows an interaction network between lncRNAs and miRNAs [32].

Based on our observed overlap between QTL controlling miRNA expression and EBA, a blistering phenotype in autoimmune skin disease, we conclude that there are interconnected pathways to simultaneously regulate both disease development and miRNA expression. This might explain the findings of earlier studies that show a clear correlation of aberrant miRNA expression and autoimmune diseases [33–36]. Accordingly, initiation and/or progression of the disease do not primarily appear to be caused by aberrant miRNA expression. Quite the contrary seems to be true; aberrant miRNA expression could be a consequence of the disease which in turn would lead to a downward spiral. Hence, miRNAs could provide a large and unexplored reservoir of potential biomarkers for EBA and related cutaneous autoimmune skin blistering diseases and an interesting target for therapeutic intervention.

Conclusion

Taken together, our study provides a complex framework of gene-gene and miRNA-gene-interactions, which eventually leads to disease development and progression. Our data provide evidence that miRNAs are important drivers of cutaneous autoimmune diseases by acting on various pathways. Moreover, the study strongly implies there is yet another, so far largely unexplored level of regulatory network, possibly comprised of by non-coding RNAs which on their part affect miRNA expression. In this sense, aberrant miRNA expression would indeed be one the responsible elements for disease progression, however, the driving force behind might be a different one.

Methods

Generation of a 4-way advanced intercross line

The out bred four-way autoimmune-prone advanced intercross line (AIL) was generated (in our group) from the parental mouse strains BXD2/TyJ, MRL/MpJ, NZM2410/J and CAST/EiJ [21, 37, 38]. All strains were purchased from The Jackson Laboratory (Bar Harbor, ME). The four inbred strains were intercrossed following

an equal strain and sex distribution. First generation (G1) offspring mice were then mated based on their parental origin to generate G2 mice in order to maintain an equal distribution of the original strain alleles across the genome. The same procedure was applied for intercrossing G2 and G3 mice. At least 50 breeding pairs were used for each successive generation of mice. Animals were held under pathogen free conditions at a 12-h light/dark cycle with food and water ad libitum. All animal experiments were approved by the Ministerium für Energiewende, Landwirtschaft, Umwelt und ländliche Räume des Landes Schleswig Holstein in Kiel, Germany (Refrence number : V 312–72241. 122–5 (12-2/09)).

Induction of experimental EBA

Experimental EBA was induced by immunization with an immune-dominant peptide within the murine NC1 domain of Collagen type VII (GST-mCOL7C) as previously described [39]. In total, 600 mice of the AIL were immunized out of which 100 mice were randomly selected for miRNA expression profiling. Mice were evaluated every 4th week after immunization regarding their development and extent of skin disease, following an established scoring system for a total period of 12 weeks [40]. Murine skin tissues were obtained for analysis at the end of the experiment.

Expression profiling and statistical analysis

Total RNA was extracted from skin tissue and processed as previously described using the Flash Tag Biotin HSR RNA labeling kit and hybridized with the miRNA expression profiling GeneChip miRNA 2.0 Array according to the manufacturer's protocol [41]. Raw data was preprocessed using R packages and normalized using RMA (Robust Multi-array analysis) to generate expression levels across the samples following instruction of at bioconductors.org [42].

Genotyping and expression QTL analysis

Genomic DNA was isolated from tail clippings and incubated in 500 μl 50 mM NaOH at 95 °C for 2 h. The reaction was neutralized by posterior addition of 50 μl 1 M Tris HCl (pH 8.0). DNA was further processed with DNeasy Blood & Tissue Kit according to manufacturer's instructions. The extracted DNA was quantified using Nanodrop and normalized to 50 ng/μl in TE buffer (10 mM Tris; 1 mM EDTA; pH = 8). Agarose gel electrophoresis was performed for quality control. A total of 1400 SNPs markers evenly spaced across 19 autosomes and the X chromosome were genotyped on 100 generation 4^{th} mice using Illumina mouse medium density array. 200 SNPs were discarded as they were non informative and 1200 markers were retained. The marker order and position in our map is provided in (Additional file 1). The SNP genotype information for each mouse from generation 4 and four founders is provided in (Additional file 1). These informative SNPs were used for performing eQTL analysis using HAPPY 2.3 on Debian Linux for miRNA expression levels [43, 44]. The software infers the haplotype probabilities for each sample using the progenitor SNP information. The miRNA expression levels of mice (600 probes) were considered as quantitative traits and linkage analysis with haplotype probabilities was done using the additive gaussian model. We considered gender as additive covariate. 1000 permutations were performed across all miRNA probes. A significant threshold ($\alpha = 0.05$) was defined genome-wide. Correction for family structure was done using a variance component model: A Kinship matrix was obtained by the EMMA R package to account for relatedness among the individual [45] and interpreted as a random effect with sex as fixed effect to perform single marker association with phenotype.

We defined the confidence interval of the given eQTL based on a -log *P* value drop of 1.5. We considered local regulation if the eQTL for the probe was found on the same chromosome as its genomic location. A trans-regulation was presumed if the QTL controlling the trait was found in a different chromosome than its genomic location. The function "epistasis" from the HAPPY R package was used to find pair wise SNP interactions using F-score for the interaction model for SNPs which were significantly associated with miRNA expression levels in initial genomic scan [43]. All the *P*-values are corrected for multiple testing using the Bonferroni correction. Circos was used for visualization of eQTL and epistasis [46].

Co-expression analysis

The standard WGCNA procedure was used for module detection [47]. For miRNA we considered 97 samples. 3 samples were excluded as these could not phenotype for EBA score. A weighted adjacency matrix of pair-wise connection strengths (correlation coefficients of gene expression levels) was constructed using the soft-threshold approach with a scale independent topological power $\beta = 7$ (miRNA). For each probe, the connectivity was defined as the sum of all connection strengths with all others. Probes were aggregated into modules by hierarchical clustering and refined by the dynamic cut tree algorithm [48]. The Pearson correlation coefficient was determined for each phenotype-module pair. The representative module expression profile, or module eigengene value, is the first principal component of the gene expression profile within a module. The correlation between the module eigengene and the sample trait of interest yields the eigengene significance, as assessed by a correlation test. The modules were assigned by different colors where grey was assigned

to traits that could not be clustered in any other module. To determine the enriched pathways involved in a cluster we used 'miRsystem' which uses the weighted pathway-ranking method for identifying enriched biological functions [49].

Total RNA preparation, cDNA synthesis and qRT-PCR

Perilesional skin samples from mice injected with NC1 domain of Collagen type VII (GST-mCOL7C) or normal skin from mice injected with GST alone were obtained. Total RNA was extracted using TRIzol® reagent (Invitrogen GmbH, Darmstadt, Germany). RNA concentrations were measured on a Nanodrop 2000c spectrophotometer (Thermo Fischer Scientific GmbH, Bremen, Germany). Total RNA was used for cDNA synthesis as previously described [50]. Briefly, 100 ng of total RNA were poly (A) tailed and reverse transcribed in a single reaction tube containing: 1 μl of 10xpoly(A) polymerase buffer, 0.1 mM of ATP, 0.1 mM of dNTPs, 100 units of MuLV reverse transcriptase (New England Biolabs, Frankfurt am Main, Germany), 1 unit of poly(A) polymerase (New England Biolabs, Frankfurt am Main, Germany), and 1 μM of RT primer (5′-CAGGTCCAGTTTTTTTTTTTTTTTGT-3′) in a final volume of 10 μl. Consequently, the tube was incubated at 42 °C for 1 h followed by 5 min at 95 °C. The cDNA was diluted 1:10 before the qPCR reaction. qPCR analysis was performed on a Mastercycler ep Realplex (Eppendorf AG, Hamburg, Germany) using 8 μl of diluted cDNA, 10 μl of SYBR Select Master Mix (Thermo Fischer Scientific GmbH, Bremen, Germany), and 250nM mmu-mir-223 specific primer set. The cycling conditions were 50 °C for 2 min, 95° for 2 min, followed by 40 cycles of 95 °C 15 s and 60 °C for 1 min. The expression levels of mmu-mir-223 were normalized against β-actin. The sequences of the qPCR primers used in this study are: mmu-mir-223 F: 5′-CGCAGTGTCAGTTTGTCA-3′; mmu-mir-223 R:5′-CCAGTTTTTTTTTTTTTTTGGGGTA-3′; mmu-β-actin F:5-CCCCATTGAACATGGCATTG-3′; mmu-β-actin R: 5′- ACGACCAGAGGCATACAGG-3′.

Availability of the supporting data

Data is accessible at NCBI GEO GSE64276.

Additional file

Additional file 1: Figure S1. Flowchart describing the workflow of analysis. The flowchart provides an overview of the analysis performed for understanding regulation of miRNAs and their contribution to the disease phenotype. **Figure S2** Interaction network accessed via IPA software for epistasis of miR-501. The graph depicts the interacting genes identified from epistasis scan of miRNA miR-501 in chromosome 1 and chromosome 2. The graph shows all known gene interactions between the two loci where genes colored in yellow are from locus on chromosome 2 and green are from locus on chromosome 1. The red line shows the possible pathway for the regulation of miR-501. **Figure S3** qRT-PCR validation of miR-223 expression in EBA and normal murine skin. (ZIP 637 kb)

Abbreviations

EBA: epidermolysis bullosa acquisita; eQTL: Expression quantitative trait loci/locus; lncRNA: Long non coding RNA; miRNA: Micro RNA; QTL: Quantitative trait loci/locus; SnoRNA: Small nucleolar RNA; SnRNA: Small nuclear RNA.

Competing interests

The authors have no competing interest.

Authors' contributions

GY, MS, WM, BM, AV, NF, BJ and HR performed the experiment and statistical analysis. Ibrahim S, LR, TR, BJ and ZD designed the experiment. TS, MH and CD. S performed the experimental validation. All the authors have read and approved the final manuscript.

Acknowledgement

The project has been funded by Deutsche Forschungsgemeinschaft: the research training programs Modulation of Autoimmunity [GRK1727/1], Genes, Environment and Inflammation [GRK 1743/1] and Excellence Cluster Inflammation at Interfaces [EXC 306/2]. We thank Miriam Freitag, Illona Klamfuss, Susen Möller and Andreia de Castro Marques for help with the AIL line.

Author details

[1]Lübeck Institute of Experimental Dermatology, University of Lübeck, Lübeck, Germany. [2]Department of Dermatology, University of Lübeck, Lübeck, Germany. [3]Max Planck Institute for Evolutionary Biology, Plön, Germany and Institute for Experimental Medicine, University of Kiel, Kiel, Germany. [4]Institute of Molecular Medicine, University of Lübeck, Lübeck, Germany. [5]IKMB, Molecular Cell Biology, Kiel, Germany.

Received: 18 April 2015 Accepted: 9 February 2016
Published online: 16 February 2016

References

1. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993;75(5):843–54.
2. Chen K, Rajewsky N. The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet. 2007;8(2):93–103.
3. Bueno MJ, Malumbres M. MicroRNAs and the cell cycle. Biochim Biophys Acta. 2011;1812(5):592–601.
4. Rottiers V, Naar AM. MicroRNAs in metabolism and metabolic disorders. Nat Rev Mol Cell Biol. 2012;13(4):239–50.
5. Xiao C, Rajewsky K. MicroRNA control in the immune system: basic principles. Cell. 2009;136(1):26–36.
6. Lee YS, Dutta A. MicroRNAs in cancer. Annu Rev Pathol. 2009;4:199–227.
7. Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, et al. New class of microRNA targets containing simultaneous 5′-UTR and 3′-UTR interaction sites. Genome Res. 2009;19(7):1175–83.
8. Saini HK, Griffiths-Jones S, Enright AJ. Genomic analysis of human microRNA transcripts. Proc Natl Acad Sci U S A. 2007;104(45):17719–24.
9. Lin SL, Miller JD, Ying SY. Intronic microRNA (miRNA). J Biomed Biotechnol. 2006;2006(4):26818.
10. Ozsolak F, Poling LL, Wang Z, Liu H, Liu XS, Roeder RG, et al. Chromatin structure analyses identify miRNA promoters. Genes Dev. 2008;22(22):3172–83.
11. Ghorai A, Ghosh U. miRNA gene counts in chromosomes vary widely in a species and biogenesis of miRNA largely depends on transcription or post-transcriptional processing of coding genes. Front Genet. 2014;5:100.
12. Su WL, Kleinhanz RR, Schadt EE. Characterizing the role of miRNAs within gene regulatory networks using integrative genomics techniques. Mol Syst Biol. 2011;7:490.
13. Zhou Y, Wang J, Lu X, Song X, Ye Y, Zhou J, et al. Evaluation of six SNPs of MicroRNA machinery genes and risk of schizophrenia. J Mol Neurosci. 2013;49(3):594–9.
14. Yi R, Pasolli HA, Landthaler M, Hafner M, Ojo T, Sheridan R, et al. DGCR8-dependent microRNA biogenesis is essential for skin development. Proc Natl Acad Sci U S A. 2009;106(2):498–502.

15. Schneider MR. MicroRNAs as novel players in skin development, homeostasis and disease. Br J Dermatol. 2012;166(1):22–8.
16. Bonazzi VF, Stark MS, Hayward NK. MicroRNA regulation of melanoma progression. Melanoma Res. 2012;22(2):101–13.
17. Ballabio E, Mitchell T, van Kester MS, Taylor S, Dunlop HM, Chi J, et al. MicroRNA expression in Sezary syndrome: identification, function, and diagnostic potential. Blood. 2010;116(7):1105–13.
18. Joyce CE, Zhou X, Xia J, Ryan C, Thrash B, Menter A, et al. Deep sequencing of small RNAs from human skin reveals major alterations in the psoriasis miRNAome. Hum Mol Genet. 2011;20(20):4025–40.
19. Sonkoly E, Janson P, Majuri ML, Savinko T, Fyhrquist N, Eidsmo L, et al. MiR-155 is overexpressed in patients with atopic dermatitis and modulates T-cell proliferative responses by targeting cytotoxic T lymphocyte-associated antigen 4. J Allergy Clin Immunol. 2010;126(3):581–9. e581-520.
20. Li Q, Stram A, Chen C, Kar S, Gayther S, Pharoah P, et al. Expression QTL-based analyses reveal candidate causal genes and loci across five tumor types. Hum Mol Genet. 2014;23(19):5294–302.
21. Ludwig RJ, Muller S, Marques A, Recke A, Schmidt E, Zillikens D, et al. Identification of quantitative trait loci in experimental epidermolysis bullosa acquisita. J Invest Dermatol. 2012;132(5):1409–15.
22. Alberts R, Terpstra P, Li Y, Breitling R, Nap JP, Jansen RC. Sequence polymorphisms cause many false cis eQTLs. PLoS One. 2007;2(7):e622.
23. Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. Nature. 2011;477(7364):289–94.
24. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol. 2009;11(3):228–34.
25. Jankowsky A, Guenther UP, Jankowsky E. The RNA helicase database. Nucleic Acids Res. 2011;39(Database issue):D338–341.
26. Li Y, Liu H, Lai C, Du X, Su Z, Gao S. The Lin28/let-7a/c-Myc pathway plays a role in non-muscle invasive bladder cancer. Cell Tissue Res. 2013;354(2):533–41.
27. Kramer A, Green J, Pollard Jr J, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30(4):523–30.
28. Baumjohann D, Ansel KM. MicroRNA-mediated regulation of T helper cell differentiation and plasticity. Nat Rev Immunol. 2013;13(9):666–78.
29. Liu H, Kohane IS. Tissue and process specific microRNA-mRNA co-expression in mammalian development and malignancy. PLoS One. 2009;4(5):e5436.
30. Schmidt MF. Drug target miRNAs: chances and challenges. Trends Biotechnol. 2014.
31. De Guire V, Robitaille R, Tetreault N, Guerin R, Menard C, Bambace N, et al. Circulating miRNAs as sensitive and specific biomarkers for the diagnosis and monitoring of human diseases: promises and challenges. Clin Biochem. 2013;46(10–11):846–60.
32. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell. 2011;147(2):358–69.
33. Otaegui D, Baranzini SE, Armananzas R, Calvo B, Munoz-Culla M, Khankhanian P, et al. Differential micro RNA expression in PBMC from multiple sclerosis patients. PLoS One. 2009;4(7):e6309.
34. Murata K, Furu M, Yoshitomi H, Ishikawa M, Shibuya H, Hashimoto M, et al. Comprehensive microRNA analysis identifies miR-24 and miR-125a-5p as plasma biomarkers for rheumatoid arthritis. PLoS One. 2013;8(7):e69118.
35. Ma X, Zhou J, Zhong Y, Jiang L, Mu P, Li Y, et al. Expression, Regulation and Function of MicroRNAs in Multiple Sclerosis. Int J Med Sci. 2014;11(8):810–8.
36. Marcet B, Chevalier B, Luxardi G, Coraux C, Zaragosi LE, Cibois M, et al. Control of vertebrate multiciliogenesis by miR-449 through direct repression of the Delta/Notch pathway. Nat Cell Biol. 2011;13(6):693–9.
37. Asghari F, Fitzner B, Holzhuter SA, Nizze H, de Castro MA, Muller S, et al. Identification of quantitative trait loci for murine autoimmune pancreatitis. J Med Genet. 2011;48(8):557–62.
38. Srinivas G, Moller S, Wang J, Kunzel S, Zillikens D, Baines JF, et al. Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering. Nat Commun. 2013;4:2462.
39. Ludwig RJ, Recke A, Bieber K, Muller S, Marques Ade C, Banczyk D, et al. Generation of antibodies of distinct subclasses and specificity is linked to H2s in an active mouse model of epidermolysis bullosa acquisita. J Invest Dermatol. 2011;131(1):167–76.
40. Sitaru C, Chiriac MT, Mihai S, Buning J, Gebert A, Ishiko A, et al. Induction of complement-fixing autoantibodies against type VII collagen results in subepidermal blistering in mice. J Immunol. 2006;177(5):3461–8.
41. Hasler R, Begun A, Freitag-Wolf S, Kerick M, Mah N, Zvirbliene A, et al. Genetic control of global gene expression levels in the intestinal mucosa: a human twin study. Physiol Genomics. 2009;38(1):73–9.
42. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):307–15.
43. Mott R, Talbot CJ, Turri MG, Collins AC, Flint J. A method for fine mapping quantitative trait loci in outbred animal stocks. Proc Natl Acad Sci U S A. 2000;97(23):12649–54.
44. Moller S, Krabbenhoft HN, Tille A, Paleino D, Williams A, Wolstencroft K, et al. Community-driven computational biology with Debian Linux. BMC Bioinformatics. 2010;11 Suppl 12:S5.
45. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. Genetics. 2008;178(3):1709–23.
46. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19(9):1639–45.
47. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
48. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics. 2008;24(5):719–20.
49. Lu TP, Lee CY, Tsai MH, Chiu YC, Hsiao CK, Lai LC, et al. miRSystem: an integrated system for characterizing enriched functions and pathways of microRNA targets. PLoS One. 2012;7(8):e42390.
50. Busk PK. A tool for design of primers for microRNA-specific quantitative RT-qPCR. BMC Bioinformatics. 2014;15:29.

The Journal of Immunology

Radiosensitive Hematopoietic Cells Determine the Extent of Skin Inflammation in Experimental Epidermolysis Bullosa Acquisita

Hiroaki Iwata, Mareike Witte, Unni Krishna S. R. L. Samavedam, Yask Gupta, Atsushi Shimizu, Akira Ishiko, Tobias Schröder, Karsten Seeger, Markus Dahlke, Dirk Rades, Detlef Zillikens and Ralf J. Ludwig

J Immunol 2015; 195:1945-1954; Prepublished online 22 July 2015;
doi: 10.4049/jimmunol.1501003
http://www.jimmunol.org/content/195/5/1945

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
9650 Rockville Pike, Bethesda, MD 20814-3994.

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Radiosensitive Hematopoietic Cells Determine the Extent of Skin Inflammation in Experimental Epidermolysis Bullosa Acquisita

Hiroaki Iwata,*[,1] Mareike Witte,[†] Unni Krishna S. R. L. Samavedam,[†] Yask Gupta,[†] Atsushi Shimizu,[‡] Akira Ishiko,[‡] Tobias Schröder,[†] Karsten Seeger,[§] Markus Dahlke,[¶] Dirk Rades,[¶] Detlef Zillikens,*[,†] and Ralf J. Ludwig*[,†]

Animal models have enhanced our understanding of the pathogenesis of autoimmune diseases. For these models, genetically identical, inbred mice have commonly been used. Different inbred mouse strains, however, show a high variability in disease manifestation. Identifying the factors that influence this disease variability could provide unrecognized insights into pathogenesis. We established a novel Ab transfer-induced model of epidermolysis bullosa acquisita (EBA), an autoimmune disease characterized by (muco)-cutaneous blistering caused by anti-type VII collagen (COL7) autoantibodies. Blistering after anti-COL7 IgG (directed against the von Willebrand factor A–like domain 2) transfer showed clear variability among inbred mouse strains, that is, severe cutaneous blistering and inflammation in C57BL/6J and absence of skin lesions in MRL/MpJ mice. The transfer of anti-COL7 IgG into irradiated, EBA-resistant MRL/MpJ mice, rescued by transplantation with bone marrow from EBA-susceptible B6.AK-H2k mice, induced blistering. To the contrary, irradiated EBA-susceptible B6.AK-H2k mice that were rescued using MRL/MpJ bone marrow were devoid of blistering. In vitro, immune complex activation of neutrophils from C57BL/6J or MRL/MpJ mice showed an impaired reactive oxygen species release from the latter, whereas no differences were observed after PMA activation. This finding was paralleled by divergent expression profiles of immune complex–activated neutrophils from either C57BL/6J or MRL/MpJ mice. Collectively, we demonstrate that radiosensitive cells determine the varying extent of skin inflammation and blistering in the end-stage effector phase of EBA. ***The Journal of Immunology*****, 2015, 195: 1945–1954.**

Autoimmune diseases, such as systemic lupus erythematosus, rheumatoid arthritis, type 1 diabetes, or autoimmune bullous dermatoses, are examples of complex diseases that have become a major health burden worldwide (1). They are characterized by an aberrant immune response against self-Ags (2), which is, for example, reflected by the detection of autoantibodies in patients (3) and the demonstration of autoantibody pathogenicity by transfer into experimental animals (4). Although the current treatment options for patients with autoimmune diseases are far from ideal, they have dramatically improved, as exemplified by the introduction and application of biologicals into clinical practice (5). The present treatment of patients with autoimmune diseases has been shaped, to a great extent, by an increased understanding of the pathogenesis of diseases. In particular, the development and use of animal models have significantly contributed to both an increased understanding of the pathogenesis of autoimmune diseases and the identification of novel treatments (6).

For several reasons, such as the reproduction rate, housing requirements, standardization, and availability of genetic modifications, inbred mice are most often used in current biomedical research. Translating the experimental outcomes from mouse models into the corresponding human diseases may, however, be hampered by conflicting and contradictory results obtained from different inbred mouse lines, that is, a reduced response to i.v. Ig treatment of immune thrombocytopenia in C57BL/6, as opposed to BALB/c, mice (7). Consistent with this observation of a differential response to treatment, disease induction in mice is strain dependent. This strain dependency has been identified for both immunization-induced (8, 9) and Ab transfer-induced models of autoimmune diseases (10, 11). These observations led to the identification of the C5 locus as an important modulator of joint destruction in the K/B×N serum transfer arthritis model (10). Therefore, novel insights into the pathogenesis of disease and therapeutic responses can be obtained from a detailed understanding of strain differences observed in the majority, if not all, of mouse models of autoimmune diseases.

In this study, we developed a novel Ab transfer-induced mouse model of epidermolysis bullosa acquisita (EBA), which is a rare but prototypical organ-specific autoimmune disease caused by autoantibodies against type VII collagen (COL7) (12, 13). Al-

*Department of Dermatology, University of Lübeck, D-23538 Lübeck, Germany; [†]Lübeck Institute of Experimental Dermatology, University of Lübeck, D-23538 Lübeck, Germany; [‡]First Department of Dermatology, School of Medicine, Faculty of Medicine Toho University, Tokyo 143-8540, Japan; [§]Department of Chemistry, University of Lübeck, D-23538 Lübeck, Germany; and [¶]Department of Radiation Oncology, University of Lübeck, D-23538 Lübeck, Germany

[1]Current address: Department of Dermatology, Hokkaido University Graduate School of Medicine, Sapporo, Japan.

Received for publication April 30, 2015. Accepted for publication June 26, 2015.

This work was supported in part by Deutsche Forschungsgemeinschaft Grants DFG LU877/10-1, GRK 1727/1 (Research Training Group "Modulation of Autoimmunity"), GRK 1743/1 (Research Training Group "Genes, Environment, and Inflammation"), and EXC 306/2 (Excellence Cluster "Inflammation at Interfaces").

The expression data presented in this article have been submitted to Gene Expression Omnibus under accession number GSE70193.

Address correspondence and reprint requests to Prof. Ralf J. Ludwig, Lübeck Institute of Experimental Dermatology, University of Lübeck, Ratzeburger Allee 160, D-23538 Lübeck, Germany. E-mail address: ralf.ludwig@uksh.de

The online version of this article contains supplemental material.

Abbreviations used in this article: COL7, type VII collagen; DEJ, dermal–epidermal junction; EBA, epidermolysis bullosa acquisita; IC, immune complex; IF, immunofluorescent; ROS, reactive oxygen species; vWFA2, von Willebrand factor A–like domain 2.

though the clinical presentation of EBA patients shows considerable variability, that is, EBA patients may present with skin fragility, tense blisters, scarring, milia formation, or widespread cutaneous blistering on inflamed skin, which is often complicated by mucous membrane involvement (14–16), subepidermal blistering is the histopathological hallmark of the disease (17).

In the current study, we first demonstrated the in vitro and in vivo pathogenicity of autoantibodies directed against an immunodominant epitope located within the murine COL7. Secondly, our results confirmed observations of strain dependency in Ab transfer model of arthritis (10) and nephritis (11). By generating bone marrow chimeric mice, we extended these findings, thus demonstrating that the observed strain dependency in Ab transfer-induced EBA is uniquely controlled by radiosensitive cells. Because Gr-1$^+$ myeloid cells are the primary effector cells in Ab transfer-induced EBA (6), we subsequently investigated the transcriptional and functional responses of neutrophils from EBA-resistant and EBA-susceptible mouse strains. Pathway analysis of this data set provided novel insights into the molecular mechanisms of neutrophil activation.

Materials and Methods

Studies involving human material

The approval for studies using biological material from humans was obtained from the Institutional Review Board at the University of Lübeck (Lübeck, Germany), and written informed consent was obtained, as specified in the Declaration of Helsinki.

Mice

Mice from inbred mouse strains (BALB/cJ, C57BL/6J, DBA1/J, and MRL/MpJ,); B6.129P2-Fcer1gtm1Rav/J (γ-chain deficient), B10.D2-Hc0 H2d H2-T18c/oSnJ (C5-deficient), and B10.D2-Hc1 H2d H2-T18c/nSnJ (C5-sufficient); and MHC-congenic B6.C-H2d/bByJ (B6.d), *B6.SJL*-H2s *C3c/1CyJ* (B6.s), and B6.AK-H2k/FlaEgJ (B6.k) were obtained from The Jackson Laboratory (Bar Harbor, ME). Sixth generation mice of a recently described, autoimmune-prone advanced intercross line were made, as previously described (18). All animals were housed in specific pathogen-free conditions, fed standard mouse chow, and provided acidified drinking water ad libitum. All clinical examinations, biopsies, and blood drawings were performed under anesthesia using i.p. administration of a mixture of ketamine (100 μg/g) and xylazine (15 μg/g). The animal experiments were approved by the local authorities of the Animal Care and Use Committee (Kiel, Germany) and conducted by certified personnel.

Generation of bone marrow chimeric mice

Bone marrow from B6.k or MRL/MpJ mouse donors was isolated from femoral bones. The cells (1×10^7; 6.7×10^6/ml) were i.v. injected into irradiated (8 Gy) B6.k and MRL/MpJ female mice. Engraftment of donor cells was evaluated using flow cytometry 6 wk after transplantation. EDTA-anticoagulated blood was stained with PE-CD8a (clone 53-6.7; eBioscience, Frankfurt, Germany) and FITC-CD8b.2 (clone 53-5.8; BioLegend, San Diego, CA) to distinguish the different alloantigens of the MRL/MpJ and B6.k strains (19, 20).

Induction of experimental EBA and clinical evaluation

Experimental EBA was induced by repetitive injections of anti-COL7^{vWFA2} IgG, according to slightly modified established protocols (21). Two New Zealand White rabbits (SA6812 and SA6813) were s.c. immunized with 250 μg murine von Willebrand factor A–like domain 2 (vWFA2) protein (13, 22) that was suspended in CFA. The animals were boosted three times (at 13-d intervals) with the same protein preparation in IFA, and immune sera were characterized using immunofluorescent (IF) microscopy on cryosections of murine skin. IgG from immune and normal rabbit sera was purified using affinity chromatography and protein G affinity, as previously reported (23, 24). The split-inducing capacity of the purified rabbit Abs was evaluated using an in vitro assay, as described elsewhere (25–27). The reactive oxygen species (ROS) release capacity of the purified rabbit Abs was studied using an established in vitro assay (28). In cases of ROS release by mouse neutrophils, the cells were isolated from bone marrow flushed from femurs and tibias, as previously reported (29). In brief, neutrophils were isolated from cell suspension using 62% Percoll (GE Healthcare, Uppsala, Sweden). After centrifugation at $1000 \times g$ for 30 min at room temperature without breaking, the pellet was collected. The cells were subsequently washed with RPMI 1640 medium (Lonza, Basel, Switzerland) and resuspended at 2×10^7 cells/ml in phenol red-free RPMI 1640 medium (Biochrom, Berlin, Germany) supplemented with 25 mM HEPES and 1% FCS. PMA (2 μg/ml) was then used as a positive control. Purified rabbit IgG specific to vWFA2 or normal rabbit IgG was injected s.c. into adult mice every second day for a total of six injections. The mice were examined to determine their general condition and to note evidence of cutaneous lesions (i.e., erythema, blisters, erosions, alopecia, and crusts) every fourth day for 12 or 20 d. Disease severity is expressed as the percentage of body surface area that was affected by skin lesions every fourth day, and total disease severity during the observation period was calculated as the area under the curve of the recorded disease severity. Serum and tail skin samples were collected every fourth day. Serum, ear skin, tail skin, oral mucosa, esophagus, stomach, duodenum, small intestine, and colon samples were obtained on the final day and prepared for examination by histopathology and IF microscopy. Skin samples for electron microscopy were obtained on days 6 and 20.

Direct and indirect IF microscopy, immunoblot analysis, and histopathology

Tissue-bound Abs were detected using direct IF microscopy on 6-μm frozen sections that were prepared from tissue biopsies using 100-fold diluted FITC-labeled Abs specific to rabbit IgG (DakoCytomation) or murine C3 (Cappel Organon-Teknika). The fluorescence intensity at the dermal-epidermal junction (DEJ) was determined using ImageJ software (http://rsbweb.nih.gov/ij/) and dermal fluorescence for background subtraction. NaCl (1 M)-split murine and human skin sections were prepared, as previously described (30). Recombinant vWFA2 proteins or murine dermal extracts were fractionated using 12 and 6% SDS-PAGE, respectively, transferred to nitrocellulose, and analyzed using immunoblotting (31). The dermal extracts were prepared as described (32). H&E staining and direct and indirect IF microscopy were performed as described (33).

Expression profiling of immune complex–activated mouse neutrophils

Gene expression in immune complex (IC)–activated neutrophils of MRL/MpJ and C57BL/6J mice and nonactivated neutrophils from gender-, age-, and strain-matched mice was compared using the Affymetrix GeneChip Mouse Gene 2.0 ST Array (n = 3/group). The raw data were processed using the oligo R/BioConductor package. Thus, the raw intensity values were background corrected, log2 transformed, and then quantile normalized using the Robust Multiarray Average implemented in R. An unpaired t test was performed to identify differential expressed genes. Genes with a fold change of >2 or ≤ 2 and $p < 0.01$ were considered to be statistically significant. The pathway ontology was performed using the Database for Annotation, Visualization, and Integrated Discovery v6.7 as well as the STRING database with Kyoto Encyclopedia of Genes and Genomes as a reference pathway database. Moreover, ingenuity pathway analysis software was used to investigate the unknown putative pathways. For the generation of heat maps, we used the gplots package provided by R.

Statistical analysis

Comparisons or differences in the extent of clinical phenotypes were performed using SigmaPlot (Version 12; Systat Software, Erkrath, Germany). The tests used are presented in the legends of the figures and tables. A p value <0.05 was considered to be statistically significant.

Results

Mice injected with Abs against the vWFA2-like domain of murine COL7 develop skin blisters

In vitro characterization of anti-COL7 (vWFA2) IgG. IgG from immune rabbit sera (SA6812 and SA6813) against the COL7 subdomain vWFA2, but not preimmune sera, bound to the DEJ of murine skin as determined by indirect IF microscopy (Supplemental Fig. 1A, 1B). Immune IgG localized at the dermal side of 1 M NaCl-split murine skin (Supplemental Fig. 1C) with an endpoint serum titer by indirect IF of 1:256,000. Cross-reactivity at lower endpoint titers (1:800) with human skin was also observed (Supplemental Fig. 1D). Indirect immunogold electron microscopy with murine skin as a substrate showed immune IgG deposits at the lamina densa (Supplemental Fig. 1E).

Using immunoblot analysis, immune sera, but not preimmune sera, targeted both recombinant vWFA2 protein and dermal extract of murine skin at 290 kDa (Supplemental Fig. 1F, 1G). To study the pathogenic potential of immune IgG, we first evaluated the impact of immune IgG/vWFA2 IC on neutrophil activation. Incubation of neutrophils with either Ag or immune IgG alone did not lead to their activation assayed by their ROS release. In contrast, incubation of neutrophils with IC dose dependently activated neutrophils, exceeding the ROS release of the positive control (fMLP) (Supplemental Fig. 1H). To study the ability to induce subepidermal splits ex vivo, immune IgG was incubated with cryosections of murine skin. After the subsequent addition of leukocytes, subepidermal splits developed in skin sections incubated with immune IgG (Supplemental Fig. 1I, 1J), whereas no dermal–epidermal separation occurred in sections treated with preimmune IgG (Supplemental Fig. 1K).

Anti-COL7 (vWFA2) IgG dose dependently induces subepidermal blisters in mice. To clarify whether anti-COL7 (vWFA2) IgG is capable of inducing skin blistering in vivo, C57BL/6J male mice were injected with immune IgG. Injecting 0.25 mg/g immune IgG led to the induction of subepidermal blisters in C57BL/6J mice ($n = 16$). Disease severity, determined as the body surface area affected by skin lesions, increased throughout the observation period (Fig. 1A). To confirm these results in an independent experiment and to establish a possible dose–response relationship, mice received 0.25, 0.125, or 0.063 mg/g immune IgG per injection. Again, subepidermal blistering was observed in all mice. The onset of disease and the extent of blistering were, however, dose dependent. The mice injected with high and medium immune IgG doses developed skin lesions within 4 d after the first injection, whereas the mice injected with 0.063 mg/g developed initial skin lesions by day 8 and had an overall reduced cumulative disease severity, expressed as the area under the curve calculated from the plots of affected body surface area against time (Fig. 1B, 1C). In all of the mice injected with immune IgG, skin lesions initially developed on the ears and eyelids. Subsequently, the lesions spread to tails, limbs, and trunk. Most commonly, erosions, crusts, and alopecia were observed (Fig. 1D, 1F). In all mice injected with immune IgG, light-microscopic analysis of skin biopsies revealed extensive subepidermal blisters accompanied by varying degrees of inflammatory infiltrates (Fig. 1G). Transmission electron microscopy showed that blisters developed below the lamina densa (Fig. 1H). Furthermore, deposits of IgG and C3 were observed at the DEJ of mice injected with immune IgG (Fig. 1I, 1J), and the degree of IgG deposits at the DEJ increased over time (Fig. 1K). Regarding other anatomical sites, 7 (46.8%) of 15 oral mucosa and 5 (33.3%) of 15 esophagus specimens showed subepidermal blistering with mild inflammatory infiltrates (Supplemental Fig. 2A). Additionally, failure to gain weight was observed in the diseased mice, but not in the control mice. Control mice gained 13% weight during the 20-d observation period. In contrast, the diseased mice only gained 1% weight (Supplemental Fig. 2B).

Anti-COL7 (vWFA2) IgG-induced blistering depends on activating FcγRs and C5. The γ-chain (Fcer1)–deficient mice were completely resistant to induction of experimental EBA; moreover, doses of 0.5 mg/g immune IgG did not lead to the development of skin lesions (Supplemental Fig. 3). Histopathologically, neither subepidermal blistering nor leukocyte infiltration was observed (data not shown). Direct IF microscopy showed linear deposits of rabbit IgG and murine C3 in all γ-chain–deficient mice injected with immune IgG (Supplemental Fig. 3). In contrast, the control mice in this experiment developed clinical disease (Supplemental Fig. 3). Interestingly, both C5-deficient and C5-sufficient mice ($n = 12/10$) developed initial blisters 4–6 d after the first injection of immune IgG (Supplemental Fig. 3). C5-deficient mice, however, presented with a significantly milder blistering phenotype compared with C5-sufficient mice. Reductions by 76, 78, or 88% were observed in mice injected with 0.125, 0.25, and 0.5 mg/g immune IgG, respectively. Histopathologically, subepidermal blistering was observed in biopsies of lesional skin in both groups (data not shown). Direct IF microscopy of skin showed linear deposits of rabbit IgG and murine C3 (Supplemental Fig. 3), with similar intensities in both groups.

Induction of skin blistering is strain dependent

In Ab-transfer mouse models of arthritis and nephritis, a high variability of disease severity among different inbred strains has been observed (10, 11). Previous data in Ab transfer-induced EBA also pointed toward a strain dependency of autoantibody-induced blistering because C57BL/6 mice presented with nearly a 2-fold increase in body surface area affected by subepidermal blistering compared with BALB/c mice (34). To systematically analyze for a potential strain dependency of Ab transfer-induced EBA, C57BL/6J, DBA/1J, BALB/c, and MRL/MpJ ($n = 4$) mice were concurrently injected with 0.125 mg/g anti-COL7 (vWFA2) IgG, which duplicates the immunological, histological, and clinical features of the inflammatory variant of the human disease, as well as of other Ab transfer-induced EBA mouse models (Fig. 1). C57BL/6J and BALB/c mice, but not MRL mice, developed clinical disease, which was most pronounced in C57BL/6J mice (Fig. 2A–F). IgG deposits at DEJ, however, were similar in all of the strains (Fig. 2G–J). Consistent with these observations, we observed a high variation of skin blistering after anti-COL7 (vWFA2) IgG injection into genetically diverse advanced intercross line mice (Fig. 2K–O). Variances in the clinical phenotype in the inbred mouse lines were not associated with an enhanced immune response toward rabbit IgG. Interestingly, anti-rabbit IgG production was detected in all of the strains, except C57BL/6J mice ($n = 11$). Among DBA1/J, BALB/c, and MRL/MpJ ($n = 4$/strain), similar anti-rabbit IgG ODs were measured (Supplemental Table I).

Radiosensitive cells are responsible for the variation in clinical disease presentation in the end-stage effector phase of EBA

To determine whether these differences in susceptibility to develop blisters are controlled by radiosensitive (hematopoietic) or nonradiosensitive cells or both, bone marrow chimeric mice were generated. First, to identify potential MHC-matched donor/recipient pairs and to exclude a possible impact of the MHC haplotype on blister induction, anti-COL7 (vWFA2) IgG was injected into MHC-congenic mouse strains of the B6 genetic background. C57BL/6J (H2b) mice were used as a reference. All MHC-congenic strains developed an identical clinical disease (Fig. 3A, 3B). The MHC-independent blister induction in MHC-congenic B6 mice, the resistance to EBA induction in MRL/MpJ (H2k) mice, and the susceptibility in B6.k mice collectively indicated that genes outside this locus control the susceptibility to autoantibody-induced tissue injury. Furthermore, this finding identified a MHC-matched donor/recipient pair for the generation of chimeric mice. Engraftment of either MRL/MpJ or B6.k bone marrow reached ∼70%. Regardless of the recipient genotype, the recipients of MRL/MpJ bone marrow cells were nearly completely protected from skin blistering after the transfer of anti-COL7 (vWFA2) IgG. Conversely, B6.k bone marrow cells transferred to either recipient genotype did lead to severe clinical EBA manifestation (Fig. 3C–E).

C57BL/6J and MRL/MpJ neutrophils differ in their response after activation with IC

Because Gr-1 myeloid cells are the primary effector cells in experimental EBA (35), we evaluated whether neutrophils from C57BL/6J and MRL/MpJ differ in their response after stimulation.

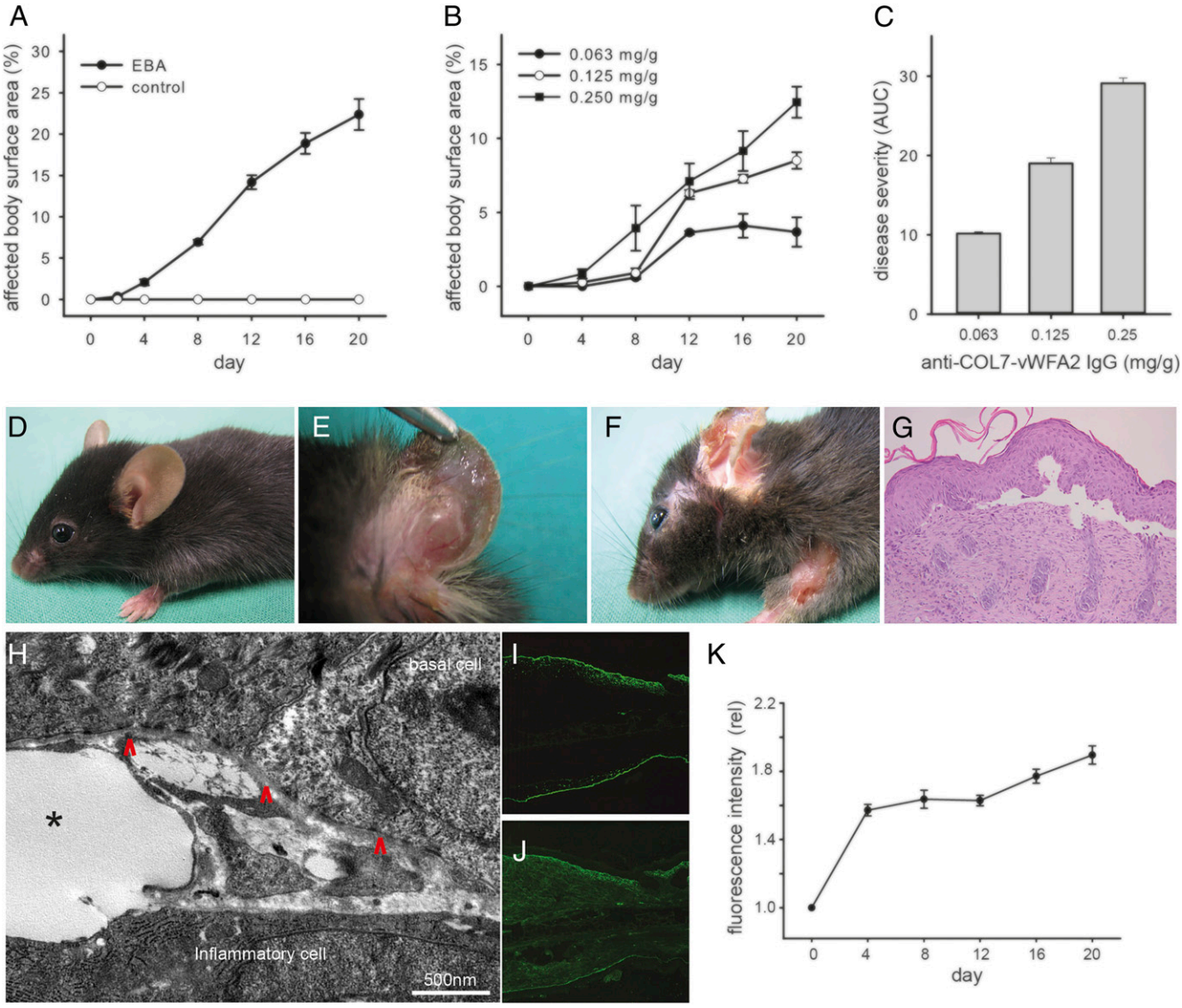


FIGURE 1. Transfer of anti-COL7vWFA2 IgG into C57BL/6 mice induces subepidermal blistering. (**A**) The mean ± SEM of the body surface area (%) affected by skin lesions during the 20-d observation period in anti-COL7vWFA2 (EBA, $n = 16$) or normal rabbit (control, $n = 3$) IgG-injected C57BL/6 mice. (**B**) The induction of skin blistering by anti-COL7vWFA2 IgG into C57BL/6 mice is dose dependent ($n = 3$/group). (**C**) Cumulative disease severity, expressed as the area under the curve (AUC) and calculated from graph b. Clinical images of (**D**) normal rabbit or (**E** and **F**) anti-COL7vWFA2 IgG-injected C57BL/6 that were obtained at the end of the observation period. Whereas (D) normal rabbit IgG-injected mice showed no clinical disease, mice injected with immune IgG developed (E) blisters, (F) erosions, and crusts. (**G**) Histopathologically, subepidermal blistering and dense dermal inflammatory infiltrates were observed (original magnification ×400). (**H**) Using transmission electron microscopy, blister formation was observed below the lamina densa (arrowhead indicates the lamina densa; scale bar, 500 nm). Direct immunofluorescence microscopy showed (**I**) IgG and (**J**) C3 deposition at the DEJ. (**K**) The relative intensity of tissue-bound IgG at the DEJ increased during the observation period; comparison of the relative intensity of day 20 with all other time points showed a significant difference for all comparisons (one-way repeated measures ANOVA with Holm-Sidak posttest).

When stimulated with IC neutrophils derived from MRL/MpJ, bone marrow showed a significant reduced release of ROS compared with C57BL/6J mice (Fig. 4A, 4C). In contrast, when the cells were activated with PMA, no difference was observed between the two strains (Fig. 4B, 4C).

Expression profiling of IC-activated neutrophils differs among EBA-susceptible and -resistant mouse strains

To unravel the molecular basis for the different response to IC stimulation between MRL/MpJ and C57BL/6J mice, the cells were activated using IC, and global mRNA expression was compared between resting or activated neutrophils from MRL/MpJ or C57BL/6J mice. Unexpectedly, fewer genes ($n = 44$) were differentially expressed in C57BL/6J after activation compared with 90 in MRL/MpJ mice (Fig. 5A). Of the differentially expressed genes, 35 were commonly differentially regulated in both strains, whereas 9 genes were only differentially regulated in C57BL/6J mice, in contrast to 55 in MRL/MpJ animals (Fig. 5). The complete expression data have been deposited at National Center for Biotechnology Information Gene Expression Omnibus (accession number GSE70193, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70193). These differentially expressed genes are predominantly allocated to four pathways in MRL/MpJ mice, more specifically, cytokine–cytokine receptor interaction (8 genes, $p = 0.0001$), chemokine signaling (7 genes, $p = 0.0001$), MAPK signaling (7 genes, $p = 0.002$), as well as terpenoid backbone biosynthesis (2 genes, $p = 0.002$). In C57BL/6J mice, only two pathways were found, that is, cytokine–cytokine receptor interaction

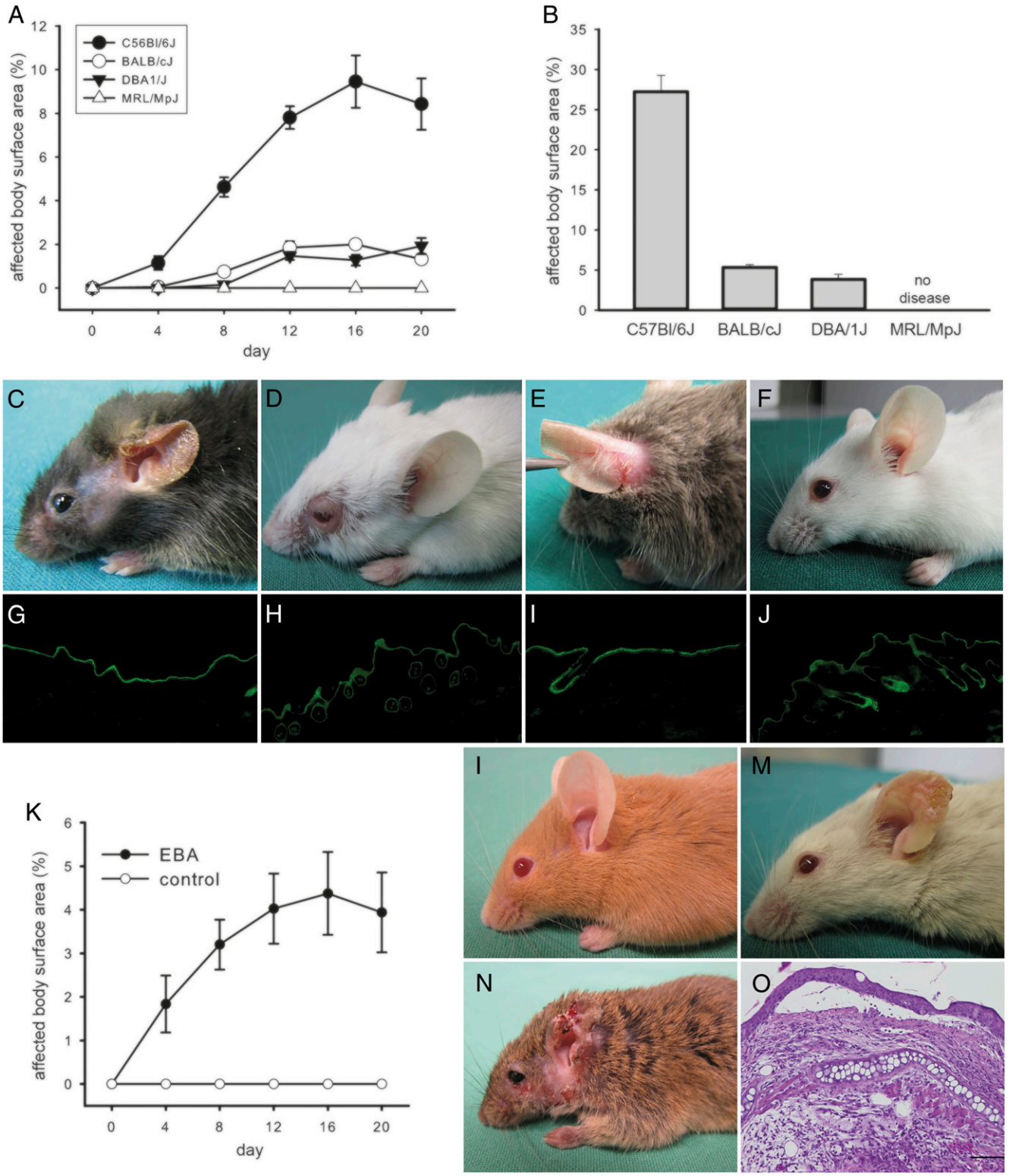


FIGURE 2. Induction of experimental EBA by autoantibody transfer in mice is strain dependent. (**A**) Disease score expressed as the percentage of the body surface area affected by skin lesions during the 20-d observation period in C57BL/6J, BALB/c, and MRL/MpJ mice (n = 4/group) injected with immune IgG showed that C57BL/6J mice exhibited the most severe clinical phenotype, BALB/c mice had a mild phenotype, and MRL/MpJ mice showed no clinical signs. (**B**) The overall disease severity, scored as the area under the curve (calculated from the plots of affected body surface area against time from A), was significantly different between the four groups. (**C**–**F**) Representative clinical pictures of (C) a C57BL/6J mouse with a severe phenotype, (D) BALB/c and (E) DBA1/J mice with mild phenotypes, and (F) a MRL/MpJ mouse with no disease. (**G**–**J**) Direct IF revealed no differences in IgG deposits at the DEJ between the strains. (**K**) Advanced intercross line mice (n = 8) developed clinical disease. (**L**–**N**) Representative clinical pictures of EBA-affected advanced intercross line mice. (**O**) H&E-stained skin section obtained from the lesional skin of an advanced intercross line mouse with evidence of subepidermal blistering and a dense, dermal leukocyte infiltration.

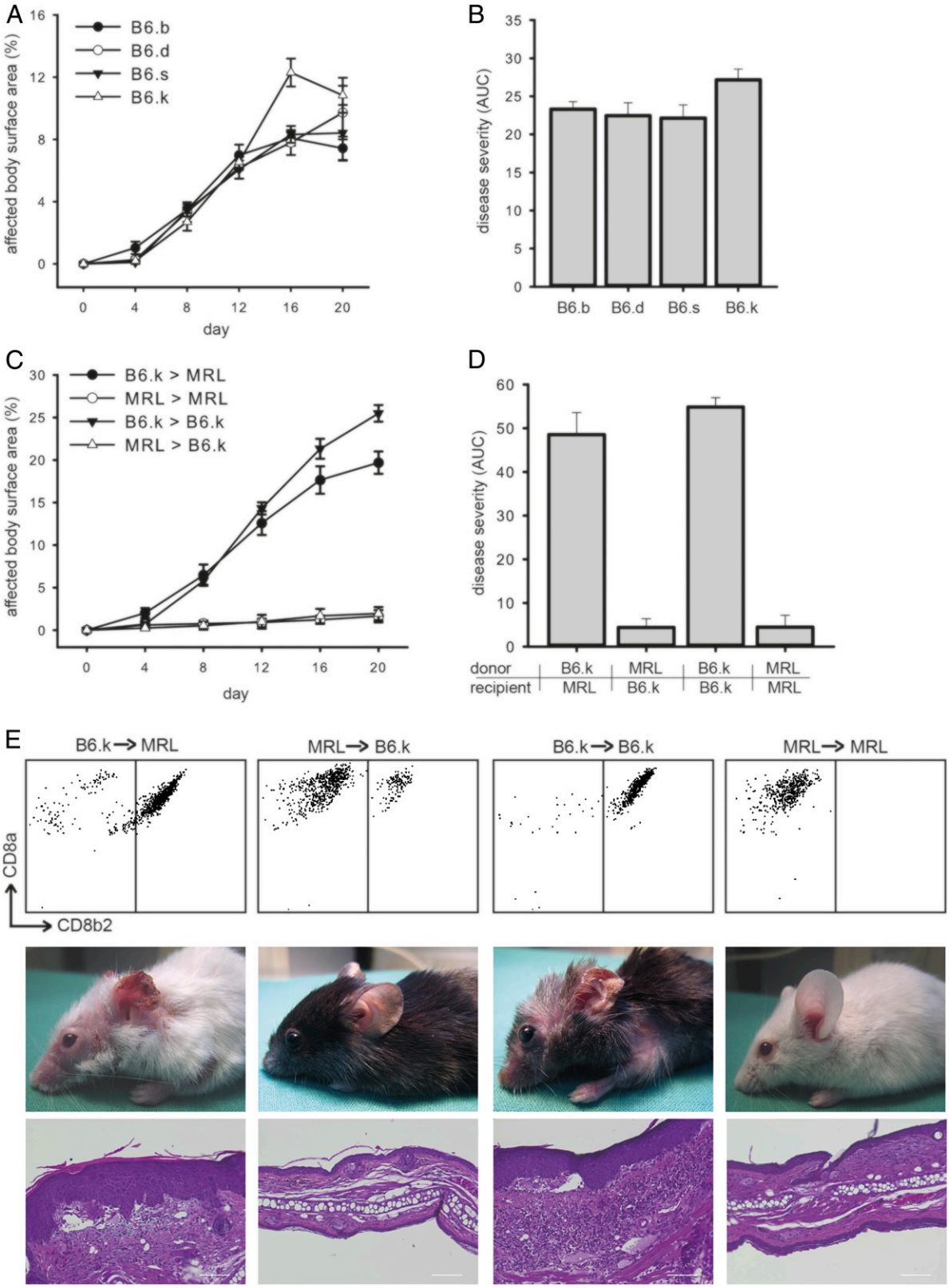


FIGURE 3. Radiosensitive cells are responsible for the varying clinical disease manifestations in the end-stage effector phase of EBA. (**A**) Disease scores, which were expressed as the percentage of the body surface area affected by skin lesions, in B6 mice carrying different MHC-H2 haplotypes (H2-b, -d, -s, and -k) showed similar disease severity in all groups (n = 6/group). (**B**) Overall disease severity, scored as the area under the curve (AUC), was not significantly different between the groups. (**C**) Disease score over the duration of the experiment and (**D**) overall disease severity (AUC) in bone marrow chimeric mice. Independent of the donor strain, mice reconstituted with MRL/MpJ bone marrow were resistant to blister induction after anti-COL7 IgG transfer. In contrast, mice reconstituted with B6.k bone marrow developed severe clinical EBA. (**E**) Representative results from flow cytometry (*upper panel*), clinical images (*middle panel*), and H&E-stained sections from the ears of bone marrow chimeric mice (*lower panel*).

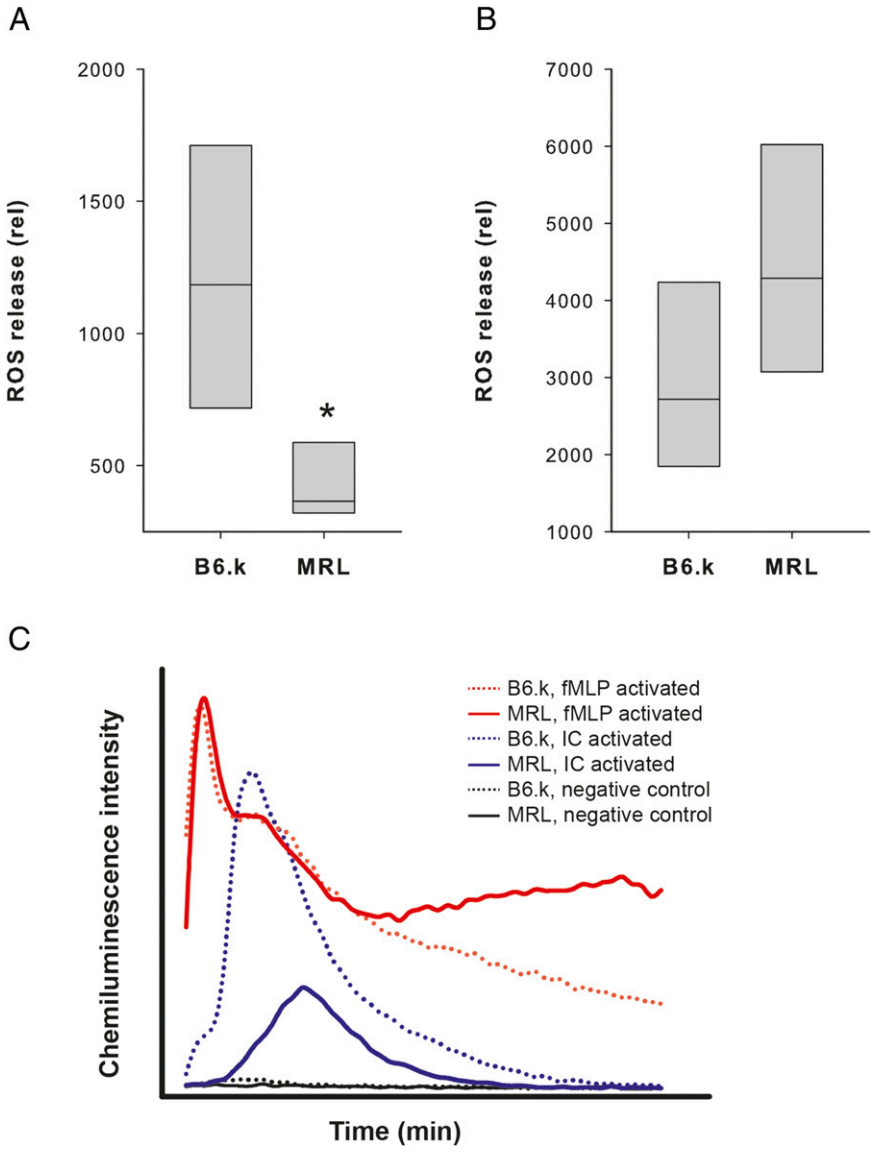


FIGURE 4. Distinct responses of neutrophils derived from C57BL/6J and MRL/MpJ mice are induced after stimulation with IC or PMA. Bone marrow–derived neutrophils from either C57BL/6J or MRL/MpJ mice were stimulated with (**A**) IC or (**B**) PMA, and their activation was assessed by measurement of ROS, which was normalized to resting cells of each strain (n = 4/group). When stimulated with IC, the neutrophils from MRL mice showed a significantly lower degree of activation, whereas the stimulation using PMA showed no differences among the two strains. Due to the nonparametric distribution, data are presented as median (black line) and 75 percentiles (boxes). *$p < 0.05$ (rank-sum test, n = 4 mice/group). (**C**) Representative ROS release (chemiluminescence intensity) over time (0–60 min) in PMA (red)- or IC (blue)-stimulated neutrophils compared with resting cells (black) in MRL/MpJ (straight lines) and C57BL/6J mice (dotted lines).

(4 genes, p = 0.006) and MAPK signaling (4 genes, p = 0.009). This finding has been visualized by integrating the differential gene expression in the pathway analysis (Fig. 6).

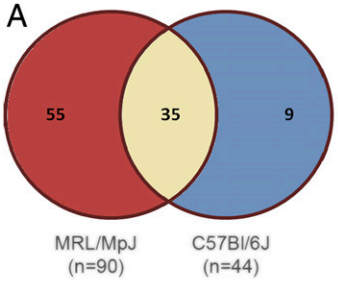
Discussion

Our data document a strain dependency in a novel Ab transfer-induced mouse model of EBA, a prototypical, organ-specific autoimmune disease. Using bone marrow chimeric mice, we were able to link this observation to radiosensitive cells. Our results from the different functional and transcriptional response of IC-activated neutrophils from EBA-resistant and EBA-susceptible mouse strains demonstrate that neutrophils of different inbred mouse lines have distinct responses to neutrophil agonists.

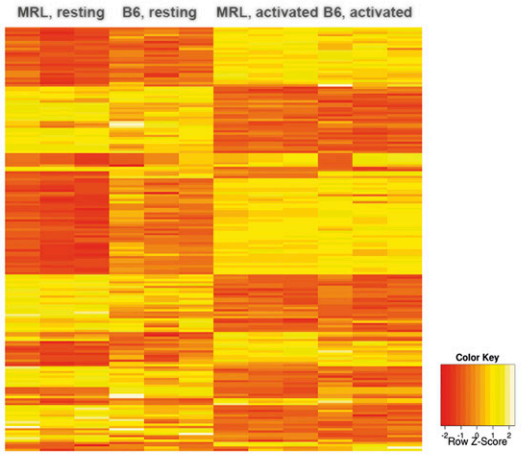
We had previously established a well-characterized Ab transfer-induced model reproducing the inflammatory variant of EBA (33). In an attempt to reproduce also the findings in the mechano-bullous variant of EBA (14, 36), we immunized rabbits of the vWFA2 domain. This domain had been shown to be important for the interaction of COL7 with type I collagen (37), with the fibronectin 3-like domain 9 of COL7 (22) and the cysteine-knot located between the NC-1 and collagenous domain of COL7 (38). In this study, we show that Abs to the vWFA2 domain induce frank blistering when injected into wild-type mice. Interestingly, however, blister formation in this model was also inflammation dependent, because, like in our previously described Ab transfer-induced EBA model (34), clinical disease manifestation was mediated by FcγR-induced inflammatory events. We were also intrigued by the distinct differences regarding the clinical phenotype in this novel Ab transfer-induced EBA model, which ranged from widespread blistering in C57BL/6J mice, in which up to 25% of the body surface area was affected by blistering and inflammatory skin lesions, to a complete absence of blistering in MRL/MpJ mice. Such strain differences have previously also been described in the K/B×N serum transfer arthritis mouse model (10) and in an immune-mediated nephritis model, induced by injection of anti-mouse glomerular basement membrane rabbit serum (11). In the latter model, the differences in nephritis were not due to an increased immune response to rabbit IgG, which we also observed in this study. Interestingly, in all three model systems, MRL/MpJ mice showed the least susceptibility to disease induction.

Intrigued by the striking differences in clinically evident blistering disease in the different inbred mouse lines, we aimed at excluding an impact of the MHC locus on Ab-induced blistering because this locus is a major susceptibility locus in immunization-induced EBA (9). The similar extent of blistering in MHC congenic B6 mice (H2d, H2k, H2s, and H2b as a control) in Ab

A



B



C

Common genes	Only MRL/MpJ	Only C57Bl/6J
CCL2	ABCG1	CASS4
CXCL2	APLNR	CCNO
DDIT3	ARMC7	FRAT1
DPH1 (OVCA2)	ATF3	P2RY13
DUSP2	CCDC85B	PLA1A
EGR1	CCL3 (CCL3L1, CCL3L3, LOC101060267)	SLCO4C1
EGR2	CCL4	THBD
GADD45G	CCNL1	TNFAIP2
GPR65	CCR1	UBTD1
H6PD	CCR5	
HILPDA	CCRL2	
HMGCS1	CD83	
HYOU1	COQ10B	
IER2	CSTB	
IER5	CXCL3	
IMPACT	DUSP4	
INSIG1	DUSP5	
KCTD7 (RABGEF1)	EGR3	
KLF10	F3	
MAFB	FABP5 (LOC101060453)	
MAFK	FFAR2	
MTUS1	G0S2	
MYC	GJB4	
NAB2	GPRC5A	
OSM	GZMB	
PLK3	HAS2	
PTGER2	HMGCR	
PTGIR	HOMER1	
PTGS2	HSPBAP1	
RGS1	IFRD1	
SPP1	ITPKC	
TNFAIP6	KCNJ2	
XCL1	LFNG	
XCL1 (XCL2)	MKL1	
ZBTB39	NPC1	
	NR4A1	
	NR4A3	
	OLFML3	
	OSGIN2	
	P2RY14	
	PER1	
	PLK2	
	PMAIP1	
	PRICKLE3	
	RAB20	
	RARA	
	RIN2	
	SH3BP2	
	SQLE	
	TBX6	
	TPD52L2	
	TRIB1	
	VPS37B	
	WBP5	
	ZFP36	

FIGURE 5. Distinct mRNA expression profile of IC-activated neutrophils from C57BL/6J and MRL/MpJ mice. (**A**) Venn diagram of differentially activated genes in IC-activated versus resting neutrophils in C57BL/6J and MRL/MpJ mice ($n = 3$ mice per group and experimental condition). (**B**) Heat map of the differential expressed genes in the indicated mouse strains. (**C**) List of differentially expressed genes in MRL/MpJ and C57BL/6J mice after activation with IC. The data have been deposited at National Center for Biotechnology Information Gene Expression Omnibus (accession number GSE70193, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70193).

transfer-induced EBA, however, argues against a major contribution of the MHC locus in controlling the blistering phenotype. Additionally, these data identified a suitable donor/recipient pair to investigate the contribution of radiosensitive and nonradiosensitive cells to the observation of a strain dependency in autoantibody transfer-induced blistering in, more specifically, EBA-resistant MRK/MpJ (H2k) and EBA-susceptible B6.AK-H2k/FlaEgJ (B6.k) mice. We assumed that both cell types contribute to blistering because radiosensitive cells [i.e., Gr-1$^+$ myeloid cells (35)] and nonradiosensitive cells [i.e., complement proteins produced resident tissue cells and hepatocytes (33, 39)] have been shown to be required for blister induction. Interestingly, when hematopoietic cells from EBA-susceptible mice were used to rescue irradiated recipient mice, blistering was induced independent of the donor genotype. Conversely, when the cells from EBA-resistant mice were transplanted, none of the recipient mice developed blistering. Collectively, this finding indicates that radiosensitive, hematopoietic cells solely determine the strain dependency observed in this work. This is in contrast to observations noted in the K/B×N arthritis model, in which C5, which is predominantly produced in the liver, was identified as a major susceptibility locus (10).

Because ROS production by neutrophils is a key prerequisite for blister formation in experimental EBA (35), we compared the response to stimulation with IC of neutrophils from EBA-susceptible C57BL/6 to EBA-resistant MRL/MpJ mice. The results showed that the production of ROS after IC, but not after PMA, stimulation is significantly reduced in the EBA-resistant MRL/MpJ mice. This difference in the neutrophil response to stimulation with IC may, in part, explain the resistance of MRL/MpJ mice to the induction of experimental EBA by autoantibody transfer. This difference is also

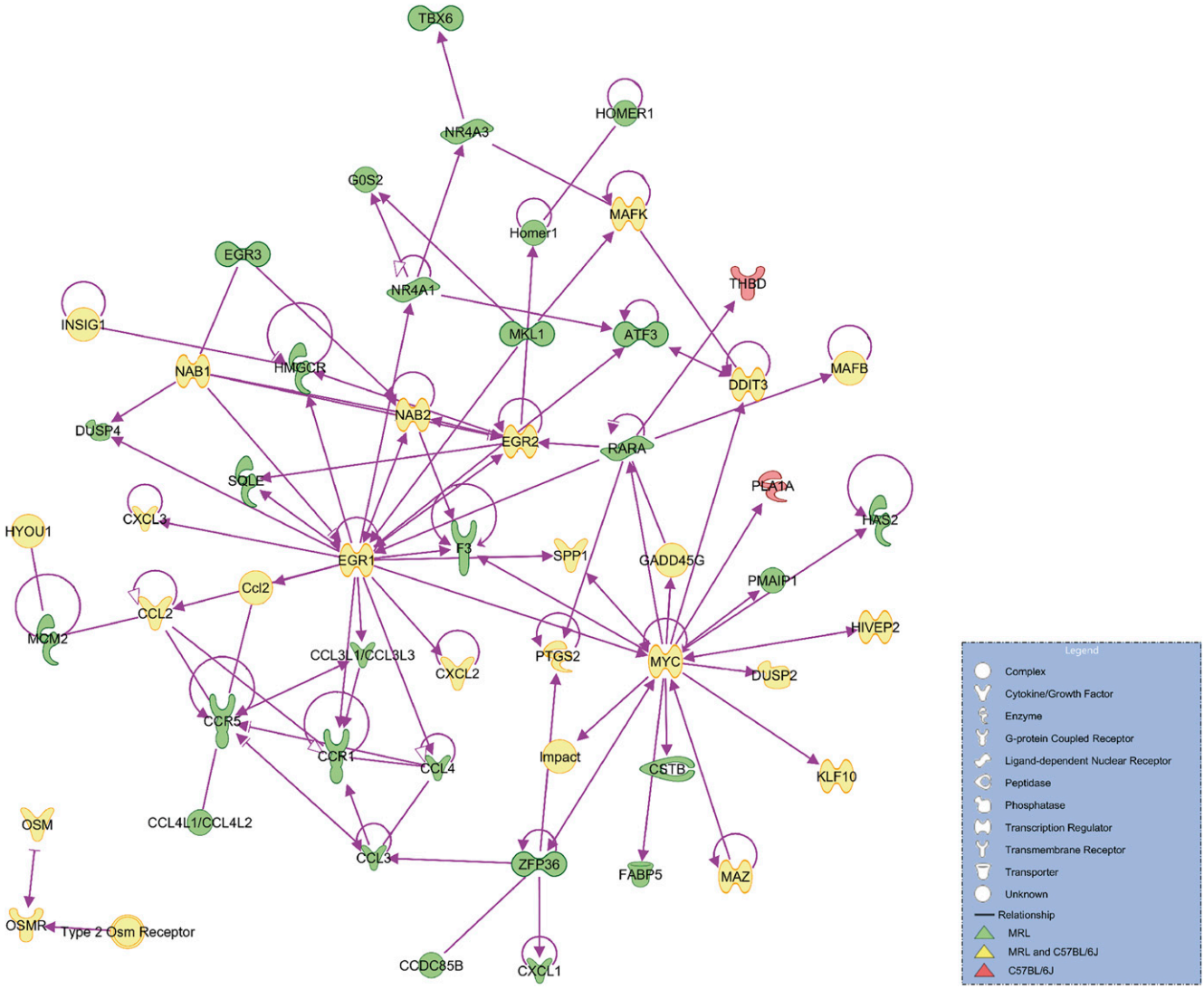


FIGURE 6. Network of differentially expressed genes of IC-activated neutrophils. The data shown in Fig. 5B were used for pathway analysis. This diagram illustrates the network by which the greater number of differential expressed genes in MRL/MpJ mice may modulate IC-induced activation.

reflected by the differential transcriptome response of neutrophils from these two strains after stimulation with IC. Based on the observed difference regarding IC-induced ROS release, we expected that more genes are differentially regulated in the EBA-susceptible C57Bl6/J mice. Contrary to our assumptions, we found more differentially regulated genes in EBA-resistant MRL/MpJ mice. This may point toward the fact that, after IC binding, both inhibitory and activating signals are induced, that is, inhibitory signaling through the FcγRIIB or through dectin-1 (40). Therefore, it is tempting to speculate that the chemokine signaling and terpenoid backbone biosynthesis pathways can dampen the proinflammatory pathways, which is reflected by the cytokine–cytokine receptor interaction and MAPK signaling pathways.

These insights into the pathogenesis of experimental EBA and the transcriptome of IC-activated neutrophils from mice with a diverse response to neutrophil activation may aid the development of neutrophil-modulating strategies, which could provide therapeutic benefits in EBA, as well as in other neutrophil-dependent diseases.

Acknowledgments

We thank Claudia Kauderer for providing excellent technical assistance.

Disclosures

The authors have no financial conflicts of interest.

References

1. Bach, J. F. 2002. The effect of infections on susceptibility to autoimmune and allergic diseases. *N. Engl. J. Med.* 347: 911–920.
2. Rose, N. R., and C. Bona. 1993. Defining criteria for autoimmune diseases (Witebsky's postulates revisited). *Immunol. Today* 14: 426–430.
3. Beutner, E. H., and R. E. Jordon. 1964. Demonstration of skin antibodies in sera of pemphigus vulgaris patients by indirect immunofluorescent staining. *Proc. Soc. Exp. Biol. Med.* 117: 505–510.
4. Rock, B., C. R. Martins, A. N. Theofilopoulos, R. S. Balderas, G. J. Anhalt, R. S. Labib, S. Futamura, E. A. Rivitti, and L. A. Diaz. 1989. The pathogenic effect of IgG4 autoantibodies in endemic pemphigus foliaceus (fogo selvagem). *N. Engl. J. Med.* 320: 1463–1469.
5. Reichert, J. M. 2012. Marketed therapeutic antibodies compendium. *MAbs* 4: 413–415.
6. Ludwig, R. J., K. Kalies, J. Köhl, D. Zillikens, and E. Schmidt. 2013. Emerging treatments for pemphigoid diseases. *Trends Mol. Med.* 19: 501–512.
7. Leontyev, D., Y. Katsman, and D. R. Branch. 2012. Mouse background and IVIG dosage are critical in establishing the role of inhibitory Fcγ receptor for the amelioration of experimental ITP. *Blood* 119: 5261–5264.
8. Watson, W. C., and A. S. Townes. 1985. Genetic susceptibility to murine collagen II autoimmune arthritis: proposed relationship to the IgG2 autoantibody subclass response, complement C5, major histocompatibility complex (MHC) and non-MHC loci. *J. Exp. Med.* 162: 1878–1891.
9. Ludwig, R. J., A. Recke, K. Bieber, S. Müller, Ade. C. Marques, D. Banczyk, M. Hirose, M. Kasperkiewicz, N. Ishii, E. Schmidt, et al. 2011. Generation of

antibodies of distinct subclasses and specificity is linked to H2s in an active mouse model of epidermolysis bullosa acquisita. *J. Invest. Dermatol.* 131: 167–176.
10. Ji, H., D. Gauguier, K. Ohmura, A. Gonzalez, V. Duchatelle, P. Danoy, H. J. Garchon, C. Degott, M. Lathrop, C. Benoist, and D. Mathis. 2001. Genetic influences on the end-stage effector phase of arthritis. *J. Exp. Med.* 194: 321–330.
11. Xie, C., R. Sharma, H. Wang, X. J. Zhou, and C. Mohan. 2004. Strain distribution pattern of susceptibility to immune-mediated nephritis. *J. Immunol.* 172: 5047–5055.
12. Woodley, D. T., R. E. Burgeson, G. Lunstrum, L. Bruckner-Tuderman, M. J. Reese, and R. A. Briggaman. 1988. Epidermolysis bullosa acquisita antigen is the globular carboxyl terminus of type VII procollagen. *J. Clin. Invest.* 81: 683–687.
13. Iwata, H., K. Bieber, B. Tiburzy, N. Chrobok, K. Kalies, A. Shimizu, S. Leineweber, A. Ishiko, A. Vorobyev, D. Zillikens, et al. 2013. B cells, dendritic cells, and macrophages are required to induce an autoreactive CD4 helper T cell response in experimental epidermolysis bullosa acquisita. *J. Immunol.* 191: 2978–2988.
14. Buijsrogge, J. J., G. F. Diercks, H. H. Pas, and M. F. Jonkman. 2011. The many faces of epidermolysis bullosa acquisita after serration pattern analysis by direct immunofluorescence microscopy. *Br. J. Dermatol.* 165: 92–98.
15. Zumelzu, C., C. Le Roux-Villet, P. Loiseau, M. Busson, M. Heller, F. Aucouturier, V. Pendaries, N. Lièvre, F. Pascal, M. D. Brette, et al. 2011. Black patients of African descent and HLA-DRB1*15:03 frequency overrepresented in epidermolysis bullosa acquisita. *J. Invest. Dermatol.* 131: 2386–2393.
16. Kim, J. H., Y. H. Kim, and S. C. Kim. 2011. Epidermolysis bullosa acquisita: a retrospective clinical analysis of 30 cases. *Acta Derm. Venereol.* 91: 307–312.
17. Schmidt, E., and D. Zillikens. 2013. Pemphigoid diseases. *Lancet* 381: 320–332.
18. Srinivas, G., S. Möller, J. Wang, S. Künzel, D. Zillikens, J. F. Baines, and S. M. Ibrahim. 2013. Genome-wide mapping of gene-microbiota interactions in susceptibility to autoimmune skin blistering. *Nat. Commun.* 4: 2462.
19. Ledbetter, J. A., and L. A. Herzenberg. 1979. Xenogeneic monoclonal antibodies to mouse lymphoid differentiation antigens. *Immunol. Rev.* 47: 63–90.
20. Ledbetter, J. A., R. V. Rouse, H. S. Micklem, and L. A. Herzenberg. 1980. T cell subsets defined by expression of Lyt-1,2,3 and Thy-1 antigens: two-parameter immunofluorescence and cytotoxicity analysis with monoclonal antibodies modifies current views. *J. Exp. Med.* 152: 280–295.
21. Vorobyev, A., H. Ujiie, A. Recke, J. J. Buijsrogge, M. F. Jonkman, H. H. Pas, H. Iwata, T. Hashimoto, S. C. Kim, J. Hoon Kim, et al. 2015. Autoantibodies to multiple epitopes on the non-collagenous-1 domain of type VII collagen induce blisters. *J. Invest. Dermatol.* 135: 1565-1573.
22. Leineweber, S., S. Schönig, and K. Seeger. 2011. Insight into interactions of the von-Willebrand-factor-A-like domain 2 with the FNIII-like domain 9 of collagen VII by NMR and SPR. *FEBS Lett.* 585: 1748–1752.
23. Sadeghi, H., A. Lockmann, A. C. Hund, U. K. Samavedam, E. Pipi, K. Vafia, E. Hauenschild, K. Kalies, H. H. Pas, M. F. Jonkman, et al. 2015. Caspase-1-independent IL-1 release mediates blister formation in autoantibody-induced tissue injury through modulation of endothelial adhesion molecules. *J. Immunol.* 194: 3656–3663.
24. Kasprick, A., X. Yu, J. Scholten, K. Hartmann, H. H. Pas, D. Zillikens, R. J. Ludwig, and F. Petersen. 2015. Conditional depletion of mast cells has no impact on the severity of experimental epidermolysis bullosa acquisita. *Eur. J. Immunol.* 45: 1462–1470.
25. Iwata, H., E. Pipi, N. Möckel, P. Sondermann, A. Vorobyev, N. van Beek, D. Zillikens, and R. J. Ludwig. 2015. Recombinant soluble CD32 suppresses disease progression in experimental epidermolysis bullosa acquisita. *J. Invest. Dermatol.* 135: 916–919.
26. Hirose, M., B. Tiburzy, N. Ishii, E. Pipi, S. Wende, E. Rentz, F. Nimmerjahn, D. Zillikens, R. A. Manz, R. J. Ludwig, and M. Kasperkiewicz. 2015. Effects of intravenous immunoglobulins on mice with experimental epidermolysis bullosa acquisita. *J. Invest. Dermatol.* 135: 768–775.
27. Recke, A., L. M. Trog, H. H. Pas, A. Vorobyev, A. Abadpour, M. F. Jonkman, G. van Zandbergen, C. Kauderer, D. Zillikens, G. Vidarsson, and R. J. Ludwig. 2014. Recombinant human IgA1 and IgA2 autoantibodies to type VII collagen induce subepidermal blistering ex vivo. *J. Immunol.* 193: 1600–1608.
28. Yu, X., K. Holdorf, B. Kasper, D. Zillikens, R. J. Ludwig, and F. Petersen. 2010. FcγRIIA and FcγRIIIB are required for autoantibody-induced tissue damage in experimental human models of bullous pemphigoid. *J. Invest. Dermatol.* 130: 2841–2844.
29. Hirahashi, J., D. Mekala, J. Van Ziffle, L. Xiao, S. Saffaripour, D. D. Wagner, S. D. Shapiro, C. Lowell, and T. N. Mayadas. 2006. Mac-1 signaling via Src-family and Syk kinases results in elastase-dependent thrombohemorrhagic vasculopathy. *Immunity* 25: 271–283.
30. Hirose, M., A. Recke, T. Beckmann, A. Shimizu, A. Ishiko, K. Bieber, J. Westermann, D. Zillikens, E. Schmidt, and R. J. Ludwig. 2011. Repetitive immunization breaks tolerance to type XVII collagen and leads to bullous pemphigoid in mice. *J. Immunol.* 187: 1176–1183.
31. Zillikens, D., K. Herzele, M. Georgi, E. Schmidt, I. Chimanovitch, H. Schumann, J. M. J. Mascaro, Jr., L. A. Diaz, L. Bruckner-Tuderman, E. B. Bröcker, and G. J. Giudice. 1999. Autoantibodies in a subgroup of patients with linear IgA disease react with the NC16A domain of BP1801. *J. Invest. Dermatol.* 113: 947–953.
32. Zillikens, D., Y. Kawahara, A. Ishiko, H. Shimizu, J. Mayer, C. V. Rank, Z. Liu, G. J. Giudice, H. H. Tran, M. P. Marinkovich, et al. 1996. A novel subepidermal blistering disease with autoantibodies to a 200-kDa antigen of the basement membrane zone. *J. Invest. Dermatol.* 106: 1333–1338.
33. Sitaru, C., S. Mihai, C. Otto, M. T. Chiriac, I. Hausser, B. Dotterweich, H. Saito, C. Rose, A. Ishiko, and D. Zillikens. 2005. Induction of dermal-epidermal separation in mice by passive transfer of antibodies specific to type VII collagen. *J. Clin. Invest.* 115: 870–878.
34. Kasperkiewicz, M., F. Nimmerjahn, S. Wende, M. Hirose, H. Iwata, M. F. Jonkman, U. Samavedam, Y. Gupta, S. Möller, E. Rentz, et al. 2012. Genetic identification and functional validation of FcγRIV as key molecule in autoantibody-induced tissue injury. *J. Pathol.* 228: 8–19.
35. Chiriac, M. T., J. Roesler, A. Sindrilaru, K. Scharffetter-Kochanek, D. Zillikens, and C. Sitaru. 2007. NADPH oxidase is required for neutrophil-dependent autoantibody-induced tissue damage. *J. Pathol.* 212: 56–65.
36. Roenigk, H. H. J., Jr., J. G. Ryan, and W. F. Bergfeld. 1971. Epidermolysis bullosa acquisita: report of three cases and review of all published cases. *Arch. Dermatol.* 103: 1–10.
37. Wegener, H., S. Leineweber, and K. Seeger. 2013. The vWFA2 domain of type VII collagen is responsible for collagen binding. *Biochem. Biophys. Res. Commun.* 430: 449–453.
38. Wegener, H., H. Paulsen, and K. Seeger. 2014. The cysteine-rich region of type VII collagen is a cystine knot with a new topology. *J. Biol. Chem.* 289: 4861–4869.
39. Mihai, S., M. T. Chiriac, K. Takahashi, J. M. Thurman, V. M. Holers, D. Zillikens, M. Botto, and C. Sitaru. 2007. The alternative pathway of complement activation is critical for blister induction in experimental epidermolysis bullosa acquisita. *J. Immunol.* 178: 6514–6521.
40. Karsten, C. M., M. K. Pandey, J. Figge, R. Kilchenstein, P. R. Taylor, M. Rosas, J. U. McDonald, S. J. Orr, M. Berger, D. Petzold, et al. 2012. Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. *Nat. Med.* 18: 1401–1406.