DeepL

Abonnieren Sie DeepL Pro, um größere Dateien zu übersetzen.
Weitere Informationen finden Sie unter www.DeepL.com/pro

BMC Genomics

Hundsdoerfer *et al. BMC Genomics* (2023) 24:443
https://doi.org/10.1186/s12864-023-09506-y

FORSCHUNG **Freier Zugang**

Hochwertige haploide Genome bestätigen 29 Chromosomen und hochgradig konservierte Gensyntenie bei den Hyles-Schwärmern (Lepidoptera: Sphingidae)

Anna K. Hundsdoerfer[1*], Tilman Schell[2], Franziska Patzold[1], Charlotte J. Wright[3], Atsuo Yoshido[4], František Marec[4], Hana Daneck[1], Sylke Winkler[5], Carola Greve[2], Lars Podsiadlowski[6], Michael Hiller[2] und Martin Pippel[5,7,8]

Abstrakt

Hintergrund Morphologische und traditionelle genetische Untersuchungen der jungen pliozänen Gattung *Hyles* haben zu der Erkenntnis geführt, dass die phänotypische Ähnlichkeit von Flügelmustern trotz ihrer Bedeutung für die Taxonomie nicht korreliert
mit phylogenetischer Verwandtschaft. Um Einblicke in verschiedene Aspekte der Artbildung beim Wolfsmilchschwärmer (*Hyles euphorbiae) zu* gewinnen, haben wir ein Genom auf Chromosomenebene zusammengestellt und einige seiner Merkmale untersucht.

Ergebnisse Das Genom eines männlichen *H. euphorbiae* wurde mit Hilfe von PacBio- und Hi-C-Daten sequenziert. Das Ergebnis war eine 504 Mb große Assemblierung (Scaffold N50 von 18,2 Mb), wobei 99,9 % der Daten von den 29 größten Scaffolds repräsentiert wurden, die den haploiden Chromosomensatz bilden. In Übereinstimmung damit ergab die FISH-Analyse des Karyotyps *n* = 29 Chromosomen und ein WZ/ZZ (weiblich/männlich) Geschlechtschromosomensystem. Schätzungen der Chromosomenlänge auf der Grundlage des Karyotypbildes lieferten ein zusätzliches Qualitätskriterium für die Größe der zusammengesetzten Chromosomen. Die erneute Gerüstbildung des veröffentlichten männlichen *H.* vespertilio-Genoms führte zu einer qualitativ hochwertigen Assemblierung (651 Mb, Gerüst N50 von 22 Mb) mit 98 % der Sequenzdaten in den 29 Chromosomen. Die größere Genomgröße von *H. vespertilio* (durchschnittlicher 1C-DNA-Wert von 562 Mb) ging einher mit einem proportionalen Anstieg
an Wiederholungen von 45 % bei *H. euphorbiae* (gemessen als 472 Mb) auf fast 55 % bei *H. vespertilio*. Mehrere Flügelmustergene wurden bei den beiden Arten auf denselben Chromosomen gefunden, wobei unterschiedliche Mengen und Positionen von repetitiven Elementen und Inversionen ihre Funktion möglicherweise beeinträchtigen.

Schlussfolgerungen Unser zweifacher vergleichender Genomik-Ansatz ergab eine hohe Gensyntenz der Hyles-Genome zu anderen Sphingidae und eine hohe Übereinstimmung mit intakten Merian-Elementen, den angestammten Bindungsgruppen der Lepidoptera, mit Ausnahme von drei einfachen Fusionsereignissen. Wir schlagen einen standardisierten Ansatz für die Genomtaxonomie vor, bei dem die Nukleotidhomologie über die Gerüstverkettung als primäres Werkzeug mit Oxford-Plots auf der Grundlage von Merian-Elementen kombiniert wird, um die Richtung der chromosomalen Umlagerungen abzuleiten und zu visualisieren. Die Identifizierung der Flügelmustergene verspricht ein zukünftiges Verständnis der Evolution der Vorderflügelmuster in der Gattung *Hyles*, obwohl weitere Sequenzierungsdaten von mehr Individuen benötigt werden. Die gewonnenen genomischen Daten bieten zusätzliche zuverlässige Referenzen für weitere vergleichende Studien bei Habichtskäfern (Sphingidae).

*Korrespondenz:

Anna K. Hundsdoerfer
anna.hundsdoerfer@senckenberg.de
Eine vollständige Liste der Autoren finden Sie am Ende des Artikels

© Der/die Autor(en) 2023. **Open Access** Dieser Artikel steht unter einer Creative Commons Attribution 4.0 International License, die die Nutzung, Weitergabe, Anpassung, Verbreitung und Vervielfältigung in jedem Medium oder Format erlaubt, solange Sie den/die ursprünglichen Autor(en) und die Quelle angemessen nennen, einen Link zur Creative Commons-Lizenz angeben und angeben, ob Änderungen vorgenommen wurden. Die Bilder oder anderes Material Dritter in diesem Artikel sind in der Creative-Commons-Lizenz des Artikels enthalten, es sei denn, es wird in einer Kreditlinie zu dem Material anders angegeben. Wenn das Material nicht in der Creative-Commons-Lizenz des Artikels enthalten ist und die von Ihnen beabsichtigte Nutzung nicht durch gesetzliche Bestimmungen erlaubt ist oder über die erlaubte Nutzung hinausgeht, müssen Sie die Erlaubnis direkt beim Urheberrechtsinhaber einholen. Eine Kopie dieser Lizenz finden Sie unter http://creativecommons.org/licenses/by/4.0/. Die Creative Commons Public Domain Dedication-Verzichtserklärung (http://creativeco mmons.org/publicdomain/zero/1.0/) gilt für die in diesem Artikel zur Verfügung gestellten Daten, es sei denn, es wird in einer Kreditlinie zu den Daten anders angegeben.

Schlüsselwörter Karyotyp, Gerüst auf Chromosomenebene, Flügelmustergene, *Optix*, *Wnt*, *Kortex*, *Aristaless*, *Distal-less*, *P*
Supergene

Hintergrund

Die Gattung *Hyles* umfasst etwa 30 Arten, ist weltweit verbreitet und mit einer geschätzten Divergenzzeit von 5,7-9,3 Millionen Jahren (Mya) [1] aus evolutionärer Sicht recht jung [2]. Einige Arten befinden sich noch im Prozess der Speziation, und andere hybridisieren, weil die Fortpflanzungsbarrieren noch niedrig sind. Diese Prozesse sind bei *Hyles euphorbiae*, die früher aus fünf verschiedenen Arten bestand, schwer zu erkennen [3]. Der Wolfsmilchschwärmer *Hyles euphorbiae* ist eine charismatische paläarktische Art mit großen, farbenprächtigen und polymorphen Larven und getarnten, schweren Erwachsenen mit starken Flugfähigkeiten (Abb. 1a). Trotz ihrer aparten Färbung scheiden die Larven die giftigen Wolfsmilch-Diterpenester [4] nicht aus ihren Nahrungspflanzen aus. Stattdessen schützen sie sich, indem sie den noch nicht entgifteten Pflanzenbrei ausspucken. Die Spezialisierung der Larven auf giftige Wirtspflanzen der Gattung *Euphorbia* hat sich innerhalb der Gattung *Hyles* zweimal unabhängig voneinander entwickelt [5, 6]. Die beeindruckend hohe morphologische Variabilität der Larven hat die Taxonomie von *H. euphorbiae* erschwert, da sie zu einer Überschätzung der Artenvielfalt beigetragen hat (Überblick in Hundsdoerfer et al. [3]). In ähnlicher Weise erschwerte die hohe intraspezifische Diversität der mitochondrialen Markergene die Rekonstruktion der molekularen Phylogenie der ehemals fünf Arten (die 2019 mit *H. euphorbiae* synonymisiert wurden [5]), lieferte aber eine wertvolle Auflösung für die Phylogeographie [7]. Obwohl einige Larvenmuster mit der Geografie korreliert sind [7] und daher davon auszugehen ist, dass sie auf einer zugrundeliegenden genetischen Divergenz beruhen, die zu phänotypischer Variabilität führt, scheinen andere durch phänotypische Plastizität umweltbedingt zu sein.

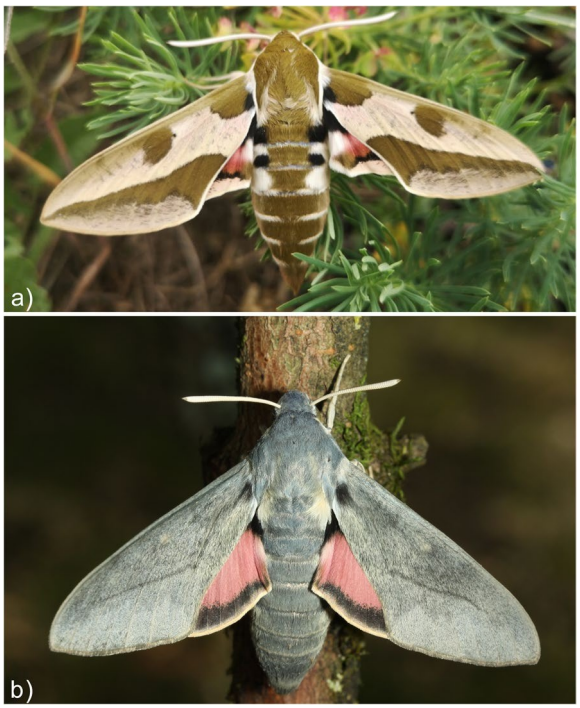

Abb. 1 a) Ein Männchen des Wolfsmilchschwärmers (*Hyles euphorbiae*), gesammelt bei Berbisdorf, Meißen, Deutschland (Foto: AKH); **b)**
Ein Weibchen des Fledermausschwärmers (*Hyles vespertilio*), fotografiert von Jean Haxaire in l'Alpes d'Huez, Frankreich (Copyright und mit freundlicher Genehmigung von Jean Haxaire [8])

Frühe Studien haben gezeigt, dass die Ähnlichkeit der Flügelmuster in dieser Gattung nicht auf phylogenetischer Zugehörigkeit beruht [1, 2]. Bei den zentralpaläarktischen Hyles-Arten werden sieben grundlegende Flügelmuster beobachtet, die weder mit den derzeit definierten Arten korrelieren noch phylogenetische Beziehungen innerhalb der Gattung widerspiegeln. Ein sehr auffälliges Flügelmuster zeigt viele weiße Streifen auf braunen Vorderflügeln, eine Reihe von gemeinsamen Merkmalen [9], die ursprünglich zu der Annahme führten, dass die paläarktische *H. livornica* und die australische *H. livornicoides* mit der Neuweltart *H. lineata* verwandt sind. Nach ihrer Anerkennung als drei Arten wurden sie immer noch als eigenständige Untergattung "Danneria" betrachtet [10], was jedoch durch frühere molekularphylogenetische Arbeiten als Klade widerlegt wurde [2]. Die Arten gruppieren sich in sehr unterschiedlichen Teilen des Stammbaums und bilden keine enge Gruppe von kryptischen Arten. Ein weiteres auffälliges Flügelmuster besteht aus dunkelbraunen Flecken und Streifen auf einem helleren, cremefarbenen Hintergrund, das dem typischen Vorderflügelmuster von *H. euphorbiae* entspricht, und so umfasste die Gruppe zahlreiche Arten mit einem ähnlichen Flügelmuster (oder leichten Variationen davon) [9]. Bei einigen dieser Arten, z. B. *H. centralasiae*, *H. nicaea* und *H. euphorbiae*, wurde jedoch keine enge molekulargenetische Verwandtschaft festgestellt [11]. Die Art *H. vespertilio* (Abb. 1b), deren Genom kürzlich veröffentlicht wurde [12], hat einen Vorderflügel, dem die beschriebenen Musterelemente fehlen. Den Vorderflügeln dieser Art fehlt fast vollständig ein Flügelmuster [12], was die Möglichkeit aufwirft, dass das/die Gen(e), das/die die Bildung des Flügelmusters steuert/steuern, inaktiv ist/sind, was Genomvergleiche

mit dieser Art besonders interessant macht.

Interspezifische Unterschiede in den Vorderflügelmustern innerhalb der Gattung *Hyles* sollten auf nachweisbaren genetischen Unterschieden beruhen, da die Muster innerhalb der Arten in der Gattung Hyles stabil sind.

gut getrennte, älteste neotropische (und nearktische) Taxa [2]. In der Paläarktis erschweren eine unvollständige Stammbaumsortierung und die fortschreitende Hybridisierung solche Erkenntnisse, was die laufenden systematischen, phylogenetischen und taxonomischen Forschungen rechtfertigt (z. B. Patzold et al. [13]).

Genomische Daten versprechen Einblicke in zahlreiche Aspekte der Artbildung, einschließlich des laufenden Genflusses, der Nutzung von Nahrungspflanzen und der genetischen Grundlage für die morphologische Differenzierung. Daher haben wir das Genom von *Hyles euphorbiae* sequenziert und untersucht. Derzeit werden zahlreiche Sphingidae-Genome veröffentlicht (z. B. *Hyles vesper-tilio*, Macroglossinae, von Pippel et al. [12], *Mimas tiliae*, Smerinthinae, von Boyes et al. [14], *Laothoe populi*, Smer- inthinae [15]). Diese Fülle an Daten wird Einblicke in die Evolution der Flügelmuster bei Habichtskäfern ermöglichen. Die Gene, die den Flügelmustern von Lepidoptera zugrunde liegen, wurden an der Familie der Nymphalidae untersucht, was durch ein klassisches Modell erleichtert wurde, das Flügelmusterelemente (z. B. Augenflecken, Bänder) in einem idealisierten "Nymphalid-Grundriss" beschreibt [16, 17]. Die genetische Induktion farbiger Schuppen von Musterelementen auf den Flügeln von Lepi- doptera ist bei mehreren Gattungen der Familie Nymphalidae gut untersucht worden, nämlich bei *Bicyclus* [18, 19], *Heli- conius* [20], *Junonia* [21] und *Vanessa* [22]. Es wurde bereits früher vorgeschlagen, dass die Evolution der Flügelmusterstreifen bei Lepidopteren durch den wiederholten Gewinn, Verlust und die Veränderung von nur einer Handvoll seriell wiederholter Elemente erfolgte, die in einer modularen Architektur mit schmalen Abschnitten des Genoms angeordnet sind, die mit spezifischen Unterschieden in Farbe und Muster in Verbindung stehen [23]. Darüber hinaus wird davon ausgegangen, dass die phänotypische Plastizität bei der Variation von Farben und bestimmten Farbmustern von Arten in erster Linie durch biotische und abiotische Faktoren beeinflusst wird [24], die als Selektionsdruck auf Individuen mit unterschiedlicher Ausprägung der Genotypen wirken [18, 25]. Starker Selektionsdruck kann den Polymorphismus einschränken, so dass das am besten geschützte Flügelmuster erhalten bleibt. Es wurde jedoch gezeigt, dass alternative adaptive Phänotypen durch heterozygote Genotypen des Flügelmuster-P-Supergens in natürlichen Populationen von *Heliconius numata* erhalten werden können [26]. Das P-Supergen stellt einen nicht rekombinanten Teil von Chromosom 15 dar, der aus drei Abschnitten mit koadaptiven Loci (P_1 -P_3 mit einer Größe von 400 kb, 200 kb und 1150 kb) besteht, die durch polymorphe Inversionen in enger Verknüpfung gehalten werden [27].

Optix ist ein Ein-Exon-Gen bei *Heliconius* [20, 28], das für einen Homeobox-Transkriptionsfaktor kodiert, der die Produktion von Ommochrom-Pigmenten steuert [29] und mit der Ausbildung roter und oranger Vorderflügel bei diesen Schmetterlingen in Verbindung gebracht wird. Nahe gelegene nicht-kodierende Regionen kontrollieren seine Expression in der Heliconius-Flügelentwicklung [20] und dies sind die Regionen der Divergenz, während die kodierende Region ein konserviertes Homeobox-Gen ist. Die Vorderflügel von *H. euphorbiae* sind für ihre rosa Färbung bekannt

in einigen Regionen seines Verbreitungsgebiets (siehe Falter in Abb. 3 von Hundsdoerfer, Lee et al. [3]), während die grauen Flügel von *H. vespertilio* nicht rot sind.

Der WntA-Signalligand ist auf Aminosäureebene bei *Heliconius* hoch konserviert [20]) und wird mit dem schwarzen Band der Vorderflügel in Verbindung gebracht. Auch hier ist es wahrscheinlicher, dass die Variation durch die Expression während der Flügelentwicklung beeinflusst wird und daher in der nahen Kontrollregion zu finden ist [20].

Das Flügelmuster-Gen *Cortex* ist für die gelben Muster auf den Heliconius-Flügeln verantwortlich [30], und es wird vermutet, dass es bei Lepidoptera für die Veränderung der Farbmuster verantwortlich ist [31]. So wurde beispielsweise gezeigt, dass die Insertion eines transponierbaren Elements (TE) in ein Intron von *Cortex* zu industriellem Melanismus [32] führt, d. h. zu schwarzen Morphen bei der Pfeffermotte (*Biston betularia*; Geom- etridae), was seine Funktion bei allen Lepidoptera bestätigt, z. B. aufgrund der TE-Insertion von *Cortex*. Diese Insertion [32], die schätzungsweise um 1819 entstanden ist, konnte sich in der Population nur durch einen Selektionsdruck gegen die hellen Falter manifestieren, die auf den dunklen Bäumen auftauchen und somit einer höheren Prädation ausgesetzt sind. Bei *Heliconius numata* liegt der *Kortex* innerhalb des Supergeninversionspolymorphismus P_1 und steuert den Wechsel der Farbmuster zwischen den Mimikry-Morphen [26, 31].

Die Expressionsmuster der beiden Gene *Distal-less* und *engrailed* [33] wurden als homologer Entwicklungsmechanismus zwischen Schmetterlings- und Mottenaugenflecken postuliert und auch bei saturniiden Motten der Gattung *Antheraea* und *Saturnia* untersucht [19]. Das Gen *aristaless*, von dem bekannt ist, dass es die Variation der Flügelfarbe steuert [34], hat kürzlich gezeigt, dass es eine zusätzliche Rolle bei der Bildung von Flügelanhängseln während der Schmetterlingsentwicklung spielt [35], was zeigt, welch zukunftsträchtige Möglichkeiten dieses Forschungsgebiet für das Verständnis der morphologischen Evolution auf der Ebene der Genfunktion eröffnet. Insgesamt ist davon auszugehen, dass eine Reihe kon- servierter, flexibler Flügelmusterungsgene, die möglicherweise in einem Supergen verbunden sind, auch die schnelle morphologische Diversifizierung in der Gattung *Hyles* vorangetrieben haben.

Im Gegensatz zu *H. vespertilio* [12], bei dem die Vorderflügel ein natürlich auftretendes Knock-out-Erscheinungsbild aufweisen (als ob die Flügelmustergene gestört wären), das aus nahezu gleichmäßig grauen Flügeln ohne kontrastreiche Musterung besteht, zeigt *H. euphorbiae* wichtige Elemente des für die Gattung typischen Grundmusters der Vorderflügel, das als Vorfahren von *Proto-Hyles* rekonstruiert wurde [9]. Durch den Vergleich der Genome von *H. euphorbiae* und *H. vespertilio* auf Chromosomenebene soll die vorliegende Arbeit Referenzgenomdaten für künftige Studien liefern, um den Ursprung der Hyles-Flügelmuster als phänotypische Variabilität zu verstehen, ausgehend von mehreren oben beschriebenen Genen zur Steuerung der Flügelmuster (z. B. [20, 23, 30]).

Ergebnisse

Zusammenbau des Genoms von *Hyles euphorbiae* mit einem Gerüst auf Chromosomenebene

Das *H.* euphorbiae-Genom wurde mit Pac- Bio HiFi-Reads [36] in Kombination mit Hi-C-Daten assembliert. Die Statistiken zur Kontiguität der Assemblierung sind in Tabelle 1 zusammengefasst. Das endgültige Contig-Assembly besteht aus 56 Scaffolds mit einem N50 von 18,2 Mb (321 Contigs mit einem N50 von 2,76 Mb) und einer Größe von 504 Mb und ist beim NCBI verfügbar (SRA-Zugangsnummer SRR17432892, BioProject PRJNA794104, BioSa- mple SAMN24610150 und Genom JALBCW000000000, JALBCX000000000, siehe Erklärung zur Datenverfügbarkeit für aufgeschlüsselte Akzessionsnummern). Das annotierte zirkuläre mitochondriale Genom hat eine Länge von 15.322 bp. Die endgültigen Statistiken der Assemblierung sind in Tabelle 2 zusammengefasst.

Chromosomengerüst mit Hi-C-Daten von *Hyles vespertilio*

HiRise [40] Gerüst von *H. vespertilio* mit mehr als 132 Millionen Lesepaare (NCBI SRA-Zugangsnummer SRX14530528) ergaben ein Gerüst N50 von 22,1 Mb. Insgesamt wurden 146 Verknüpfungen und nur eine Unterbrechung in die Eingabebaugruppe eingeführt. Die Kontaktkarte zeigt deutlich 29 gut unterstützte Gerüste, die Chromosomen darstellen (Abb. 2). Die Größen der Chromosomen beider Hyles-Arten liegen in der

Tabelle 1 Kontiguitätsstatistik des Genoms von *Hyles euphorbiae*

	Contigs	Erste Gerüste	Kuratierte Gerüste	Kuratierte Contigs
#Sequenzen	592	81	56	322
Gesamtlänge	513,190,344	513,292,544	504,259,600	504,323,440
Größte Sequenz	6,062,269	30,760,442	30,347,856	10,621,849
N50	1,441,051	18,456,932	18,182,747	2,758,341

Tabelle 2 Zusammensetzungsstatistiken der vorgestellten und verwandten Arten

	Frau	*Mt.*	*Lp*	*Hf*	*Dp*	*Er*	*Hv*
# Gerüste (> = 10.000 bp)	2822	30	33	34	31	47	380
# Contigs (> = 10.000 bp)	4977	41	135	74	36	313	526
# Gerüste (> = 50.000 bp)	61	29	29	30	29	34	251
# Contigs (> = 50.000 bp)	1129	40	130	65	34	295	396
Gesamtlänge (> = 0 bp) Mb	470.0	478.0	576.4	448.9	402.1	504.3	651.4
Gesamtlänge (> = 1000 bp) Mb	470.0	478.0	576.4	448.9	402.1	504.3	651.4
Gesamtlänge (> = 5000 bp) Mb	468.5	478.0	576.4	448.8	402.1	504.3	651.4
Gesamtlänge (> = 10.000 bp) Mb	463.1	478.0	576.4	448.8	402.1	504.3	651.4
Gesamtlänge (> = 50.000 bp) Mb	410.9	478.0	576.3	448.7	402.0	504.0	647.4
Größtes Gerüst (Mb)	21.11	25.98	29.55	25.57	22.66	30.35	37.22
Größter Kontig (Mb)	3.17	22.90	24.17	24.16	22.66	10.62	13.20
Gesamtlänge (Mb)	470.0	478.0	576.4	448.9	402.1	504.3	651.4
N50 Gerüste (Mb)	14.25	17.90	21.13	16.55	15.07	18.18	22.14
N50-Contigs (Mb)	0.425	17.80	6.97	10.68	14.88	2.76	7.26
N75 (Mb)	12.41	15.70	18.59	14.30	12.99	16.57	20.07
L50-Gerüste (Mb)	14	12	13	12	12	12	13
L50-Contigs (Mb)	297	13	27	14	13	54	34
L75	23	19	20	20	20	19	20
# N's pro 100 kb	230.91	0.65	5.29	3.26	0.4	12.66	2.24
BUSCO (lepidoptera_odb10):							
Vollständig, Einzelexemplar (%)	91.8	98.5	98.5	98.4	98.6	97.9	95.4
Vollständig, dupliziert (%)	6.5	0.3	0.4	0.3	0.2	0.3	2.9
Fragmentiert (%)	0.7	0.3	0.3	0.3	0.6	0.5	0.7
Fehlend (%)	1.0	0.9	0.8	1.0	0.9	1.3	1.0

Ms Manduca sexta [37], *Mt Mimas tiliae* [14], *Lp Laothoe populi* [15], *Hf Hemaris* fuciformis [38], *Dp Deilephila porcellus* [39], *He Hyles euphorbiae*, *Hv Hyles vespertilio*, die neue Zusammenstellung von *H. vespertilio* wurde verwendet (basierend auf PacBio-Sequenzdaten von Pippel et al. [12])

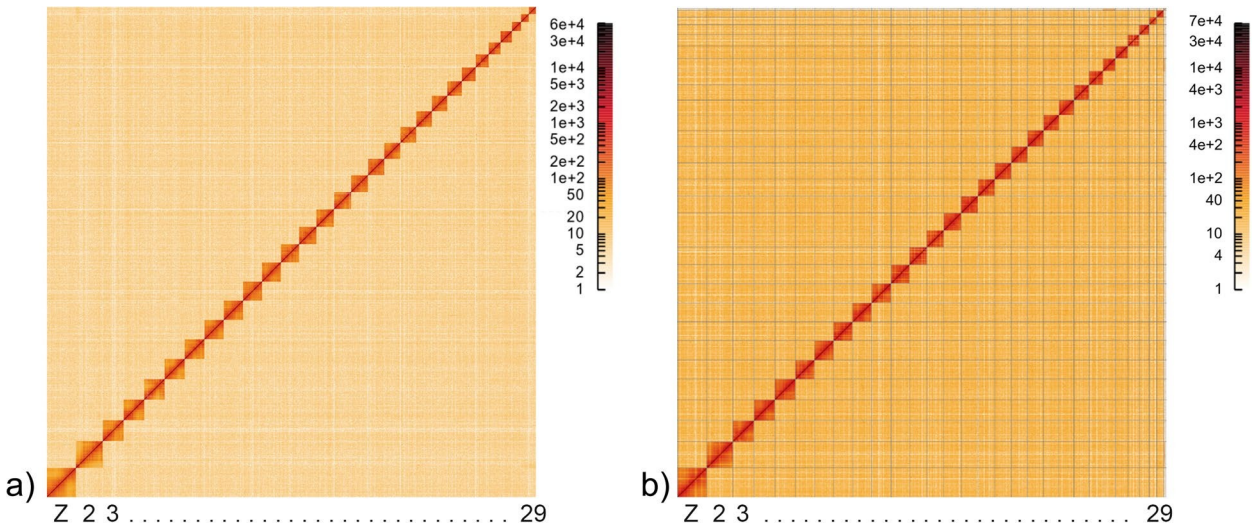

Abb. 2 Mit HiGlass erstellte Kontaktkarte von **a**) *Hyles vespertilio* und **b**) *Hyles euphorbiae*. Die Gerüstbildung auf Chromosomenebene unterstützt eindeutig 29 Gerüste, die Chromosomen darstellen. Die Chromosomen sind nach ihrer Größe von links unten (größte) nach rechts oben (kleinste) geordnet.

in der gleichen Größenordnung, aber nicht identisch (Tabelle 3). Die Baugruppe von *H. vespertilio* ist fast 147 Mb größer als die von *H. euphorbiae*.

Chromosomenstruktur bei den Sphingidae

Der Vergleich der orthologen Einzelexemplare im Genom von *H. euphorbiae* mit denen von *H. vespertilio* zeigt eine starke Erhaltung sowohl der Syntenie als auch der Chromosomenstruktur, definiert als Erhaltung der Genordnung. Dies deutet darauf hin, dass der letzte gemeinsame Vorfahre von *Hyles* 29 Chromosomen hatte, die anschließend intakt blieben (Abb. 3a). Diese 29 Chromosomen entsprechen intakten Merian-Elementen, den angestammten Bindungsgruppen der Lepidoptera [42], mit Ausnahme von drei einfachen Fusionsereignissen zwischen Paaren von Merian-Elementen (M17 + M20, M30 + M9, M24 + M25). Anschließend wurden die Orthologe von *H. euphorbiae* mit denen von *Deilephila porcellus* (Abb. 3b) verglichen, die beide zu den Macroglossinae, einer Unterfamilie der Sphingidae, gehören. Dabei zeigte sich ebenfalls ein Satz von 29 orthologen Chromosomen, was darauf hindeutet, dass die Chromosomenstruktur während der gesamten Evolution dieser Unterfamilie stabil geblieben ist. Um festzustellen, ob sich die Stabilität dieses Satzes von 29 Chromosomen auf die gesamte Unterfamilie der Sphingidae erstreckt, verglichen wir die Orthologe von zwei weiteren Arten, *Manduca sexta* (n = 28; Abb. 3c) aus der Unterfamilie Sphinginae und *Laothoe pop- uli* (n = 28; Abb. 3d), die zu einer dritten Unterfamilie, den Smerinthinae, gehört. Es zeigte sich, dass 27 Chromosomen von *M. sexta* ortholog und hoch syntenisch sind im Vergleich zu *H. euphorbiae*. Das verbleibende Chromosom in *M. sexta* (CM026259.1) ist zwei Autosomen in *H. euphorbiae* zugeordnet (CM041057.1, CM041033.1), was den Unterschied im Karyotyp erklärt. Dieses Chromosom in *M. sexta*

entspricht einer M17 + 20 + M29-Fusion. In ähnlicher Weise ist ein Chromosom von *Laothoe populi* (HG992147.1) ortholo- gous zu zwei Autosomen in *H. euphorbiae* (CM041039.1, CM041057.1). Dies ist auf eine Verschmelzung zwischen M29 und M4 zurückzuführen. In *L. populi* und *M. sexta haben* also zwei unabhängige Fusionsereignisse stattgefunden. Dies würde darauf hindeuten, dass die Sphingidae ursprünglich einen Karyotyp von $n = 29$ hatten, der mit einer begrenzten Anzahl von Fusionsereignissen in den heutigen Arten weitgehend intakt geblieben ist. Außerdem ist die starke Übereinstimmung der Genordnung bei *H. euphorbiae* im Vergleich zu *M. sexta* zeigt begrenzte intrachromosomale Umstellungen seit dem letzten gemeinsamen Vorfahren der Sphingidae.

Genom-Alignments

Die beiden Arten der Gattung *Hyles* haben ein Chromosom mehr als *B. mori* ($n = 28$), aber das Alignment zwischen *H. vespertilio* und *B. mori* kann immer noch illustriert werden (Abb. 4a), um einen Vergleich durch visuelle Inspektion zu ermöglichen. Die Chromosomenanteilswerte dieses Alignments sind in Tabelle S1 aufgeführt.

CIRCOS-Diagramme des Alignments zwischen den beiden Arten zeigen eine gut erhaltene Syntenie, mit Ausnahme einiger Chromosomen, die an Chromosomenumlagerungen beteiligt sein könnten. Chromosom 2 (HvChr2) von *H. ves- pertilio* entspricht den Chromosomen 2 und 26 von *B. mori* (BmChr2, BmChr26). Chromosom 24 von *B. mori*, BmChr24, entspricht Teilen der Chromosomen 22, 24 und 28 *von H. vespertilio* (HvChr22, HvChr24, HvChr28), und Chromosom 11 von *B. mori* (BmChr11) entspricht Teilen der Chromosomen 11 und 26 von *H. vespertilio* (HvChr11, HvChr26). Diese Chromosomenumlagerungen zwischen *B. mori* und *H. vespertilio* stimmen mit denen von

Tabelle 3 Chromosomengrößen der endgültigen Assemblies und Chromosomennummerierung in Homologie zu *Bombyx mori*. Chromosom 1 entspricht dem Z-Chromosom [41].

B. mori **Chr. Nummer**	***H. vespertilio*** **Chr. Nummer**	***H. vespertilio*** **Gerüstname im Browser**	**GenBank-Hinterlegungsnummer**	**Baugruppe Länge (bp) inkl. Lücken**	***H. euphorbiae*** **Chr. Nummer**	***H. euphorbiae*** **Gerüstname im Browser**	**GenBank-Hinterlegungsnummer**	**Baugruppe Länge (bp) inkl. Lücken**
BMChrZ	HvChrZ	ScWp86a_195_HRSCAF_280	CM042833.1	37,218,281	HeChrZ	sc_1	CM041058.1	30,347,856
BmChr2	HvChr2[a]	ScWp86a_159_HRSCAF_226	CM042805.1	24,553,817	HeChr2	sc_7	CM041030.1	19,904,380
BmChr3	HvChr3	ScWp86a_257_HRSCAF_395	CM042806.1	20,065,448	HeChr3	sc_18	CM041031.1	16,855,035
BmChr4	HvChr4	ScWp86a_21_HRSCAF_27	CM042807.1	25,450,662	HeChr4	sc_6	CM041032.1	20,022,273
BmChr5	HvChr5	ScWp86a_119_HRSCAF_165	CM042808.1	26,040,715	HeChr5	sc_4	CM041033.1	21,578,258
BmChr6	HvChr6	ScWp86a_136_HRSCAF_194	CM042809.1	22,136,963	HeChr6	sc_13	CM041034.1	17,936,700
BmChr7	HvChr7	ScWp86a_95_HRSCAF_134	CM042810.1	17,182,831	HeChr7	sc_23	CM041035.1	14,557,976
BmChr8	HvChr8	ScWp86a_214_HRSCAF_310	CM042811.1	21,873,546	HeChr8	sc_14	CM041036.1	17,505,000
BmChr9	HvChr9	ScWp86a_232_HRSCAF_355	CM042812.1	24,188,504	HeChr9	sc_11	CM041037.1	19,134,990
BmChr10	HvChr10	ScWp86a_42_HRSCAF_61	CM042813.1	24,738,224	HeChr10	sc_8	CM041038.1	19,757,600
BmChr11	HvChr11	ScWp86a_178_HRSCAF_255	CM042814.1	22,098,621	HeChr11	sc_15	CM041039.1	17,388,153
BmChr12	HvChr12	ScWp86a_182_HRSCAF_262	CM042815.1	26,852,440	HeChr12	sc_5	CM041040.1	21,432,827
BmChr13	HvChr13	ScWp86a_68_HRSCAF_94	CM042816.1	24,145,169	HeChr13	sc_10	CM041041.1	19,486,524
BmChr14	HvChr14	ScWp86a_111_HRSCAF_155	CM042817.1	20,079,433	HeChr14	sc_21	CM041042.1	15,684,570
BmChr15	HvChr15	ScWp86a_62_HRSCAF_86	CM042818.1	26,014,704	HeChr15	sc_3	CM041043.1	21,768,765
BmChr16	HvChr16	ScWp86a_55_HRSCAF_78	CM042819.1	18,679,767	HeChr16	sc_22	CM041044.1	15,570,054
BmChr17	HvChr17	ScWp86a_210_HRSCAF_302	CM042820.1	24,352,496	HeChr17	sc_9	CM041045.1	19,686,965
BmChr18	HvChr18	ScWp86a_1_HRSCAF_1	CM042821.1	20,909,092	HeChr18	sc_16	CM041046.1	17,373,644
BmChr19	HvChr19	ScWp86a_207_HRSCAF_299	CM042822.1	20,353,206	HeChr19	sc_19	CM041047.1	16,571,156
BmChr20	HvChr20	ScWp86a_378_HRSCAF_523	CM042823.1	16,497,537	HeChr20	sc_24	CM041048.1	12,812,799
BmChr21	HvChr21	ScWp86a_162_HRSCAF_230	CM042824.1	21,114,399	HeChr21	sc_17	CM041049.1	17,181,161
BmChr22	HvChr22[d]	ScWp86a_157_HRSCAF_223	CM042825.1	33,800,775	HeChr22	sc_2	CM041050.1	27,093,441
BmChr23	HvChr23	ScWp86a_31_HRSCAF_41	CM042826.1	22,933,977	HeChr23	sc_12	CM041051.1	18,182,747
BmChr24	HvChr24	ScWp86a_166_HRSCAF_234	CM042827.1	11,894,872	HeChr24	sc_27	CM041052.1	10,049,873
BmChr25	HvChr25	ScWp86a_7_HRSCAF_7	CM042828.1	19,247,757	HeChr25	sc_20	CM041053.1	15,917,517
BmChr26	HvChr26[b]	ScWp86a_81_HRSCAF_111	CM042829.1	9,224,630	HeChr26	sc_29	CM041054.1	7,206,840
BmChr27	HvChr27	ScWp86a_289_HRSCAF_430	CM042830.1	14,304,874	HeChr27	sc_26	CM041055.1	12,048,215

Tabelle 3 (Fortsetzung)

B. mori Chr. Nummer	*H. vespertilio* Chr. Nummer	*H. vespertilio* Gerüstname im Browser	GenBank-Hinterleg ungsnum mer	Baugruppe Länge (bp) inkl. Lücken	*H. euphorbiae* Chr. Nummer	*H. euphorbiae* Gerüstname im Browser	GenBank-Hinterleg ungsnum mer	Baugruppe Länge (bp) inkl. Lücken
BmChr28	HvChr28[d]	ScWp86a_140_HRSCAF_198	CM042831.1	15,210,208	HeChr28	sc_25	CM041056.1	12,771,623
n.a.	HvChr29[c]	ScWp86a_137_HRSCAF_195	CM042832.1	9,587,503	HeChr29	sc_28	CM041057.1	7,786,824

[a] enthält auch Teile von BmChr26,

[b] enthält auch Teile von BmChr11

[c] enthält auch Teile von BmChr23

[d] enthält auch einen Teil von BmChr24

zwischen *B. mori* und *M. sexta* [37, 43], was die starke Syntenie zwischen den beiden Schwärmerarten unterstreicht,

H. vespertilio und *M. sexta* (Abb. 4b). Das Chromosom 23 von *B. mori* (BmChr23) ist in die Chromosomen 23 und 29 von *H. vespertilio* (HvChr23, HvChr29) aufgeteilt, was zu einem zusätzlichen Chromosom in *H. vespertilio* (HvChr29) führt. Bei *M. sexta*, das 28 Chromosomen hat, entsprechen HvChr29 und HvChr15 dem MsChr28 und stellen somit eine andere Kombination dar.

Die Grafik vergleicht die 29 Chromosomensequenzen von

H. vespertilio und *H. euphorbiae* (Abb. 4c) veranschaulicht die hochgradig konservierte Syntenie innerhalb der Gattung *Hyles* in der Definition der Kollinearität, der Erhaltung der Blockordnung in den beiden Chromosomensätzen. Die größere Größe des Genoms von *H. vespertilio,* basierend auf der Schätzung der Genomgröße, und seine längere Assemblierung als die von *H. euphorbiae spiegeln sich* in der größeren Größe der einzelnen Chromosomen wider (Abb. 4c).

Annotation und Methylierungsprofil.

Der Anteil des Genoms, der von den wichtigsten Klassen repetitiver Elemente abgedeckt wird, beträgt bei *H. euphorbiae* 37 % und bei *H. vespertilio* 47 % (Tabelle 4, Abb. 5). Somit hat *H. vespertilio* ein viel größeres Genom (651 Mbp zu 504 Mbp in *H. euphorbiae*) und einen höheren Anteil an repetitiver DNA als *H. euphorbiae*. Ein umfassender Vergleich des Transposon-Gehalts in sieben Arten der Familie Sphingidae (Abb. 5) zeigt die Dynamik der repetitiven Elemente und den Beitrag der jüngsten Ausbrüche in den Familien der transponierbaren Elemente zur Expansion der Genomgröße bei *H. vespertilio* und unabhängig davon bei *Laothoe populi*.

Die Anzahl der Gene pro Chromosom ist in den hier vorgestellten Genomen der beiden Arten sehr ähnlich (Abb. 6), weist aber eine große Bandbreite auf, die von einem Minimum von 169 Genen bei beiden Arten auf Chromosom 29 bis zu 826 / 798 Genen bei *H. vespertilio* / *H. euphorbiae* auf Chromosom 22 reicht.

Das mitochondriale Genom von *H. euphorbiae*, das mit der mitoHifi-Pipeline [45, 46] auf der Grundlage der eng verwandten Art *Reretra oldenlandiae* zusammengestellt wurde, ist in Abb. S1 dargestellt. Es wurde mit MITOS annotiert und

enthält 13 proteinkodierende Gene, zwei ribosomale RNA (rRNA)-Gene sowie 22 Transfer-RNA (tRNA)-Sequenzen und die Kontrollregion.

Das Methylierungsprofil von *H. euphorbiae* (Track "HL Methylation" im Genombrowser; z. B. Abb. S5a) zeigt die Modifikationswahrscheinlichkeiten für jede Stelle. Basenmodifikationen mit einer Modifikationswahrscheinlichkeit von < 50 % (Hypomethylierung) sind blau eingefärbt, und Basenpositionen, die mit hoher Wahrscheinlichkeit methyliert und damit nicht exprimiert werden, sind rot markiert (Basenmodifikation > 50 % Wahrscheinlichkeit, Hypermethylierung).

Flügelmuster-Gene

Die Flügelmustergene *optix, WntA*, *cortex, arista- less* und *Distal-less* (die Hinterlegungsnummern der Referenzsequenzen sind in Tabelle S2 aufgeführt) wurden mit hohem Vertrauen mit BLAT [47, 48] identifiziert; für jedes wurde nur ein Treffer mit einer Identitätsschwelle von über 85 % und einem Bit-Score von über 420 gefunden.

Die Treffer der *Heliconius erato* optix-Proteinsequenz ergaben eine Position auf Chromosom 23 in *H. euphorbiae* (HeChr23, "sc_12: 1.582.862-1.583.680", Größe: 819 bp)
und *H. vespertilio* (HvChr23, "ScWp86a_31_HRSCAF_41: 20.799.551-20.800.375", Größe: 825 bp). Bei beiden Arten wurden keine repetitiven Elemente im Bereich der Codierungssequenz festgestellt (Abb. S2). Bei einer Verkleinerung um das 10-fache oder 100-fache ist im Genombrowser jedoch eine große Anzahl von Indels zu sehen (400 kb stromauf- und stromabwärts der kodierenden Sequenz, Abb. S3).

Die den Hyles-Genomen zugeordnete WntA-Aminosäuresequenz von *Heliconius himera* (Tabelle S2) ergab eine Position auf Chromosom 28 in *H. euphorbiae* (HeChr28, "sc_25: 5.829.059-5.837.439", über sechs Exons und 8.381 bp; Abb. S4a; annotiert durch das Transkript LOC115447391_rna-XM037443222.1.47 von *Manduca*
sexta mit sieben Exons und 17.247 bp) und *H. vespertilio* (HvChr28, "ScWp86a_140_HRSCAF_198: 8.200.403-8.210.485", mit sechs Exons und 10.083 Basenpaaren, Abb. S4b; annotiert von evm07703.91

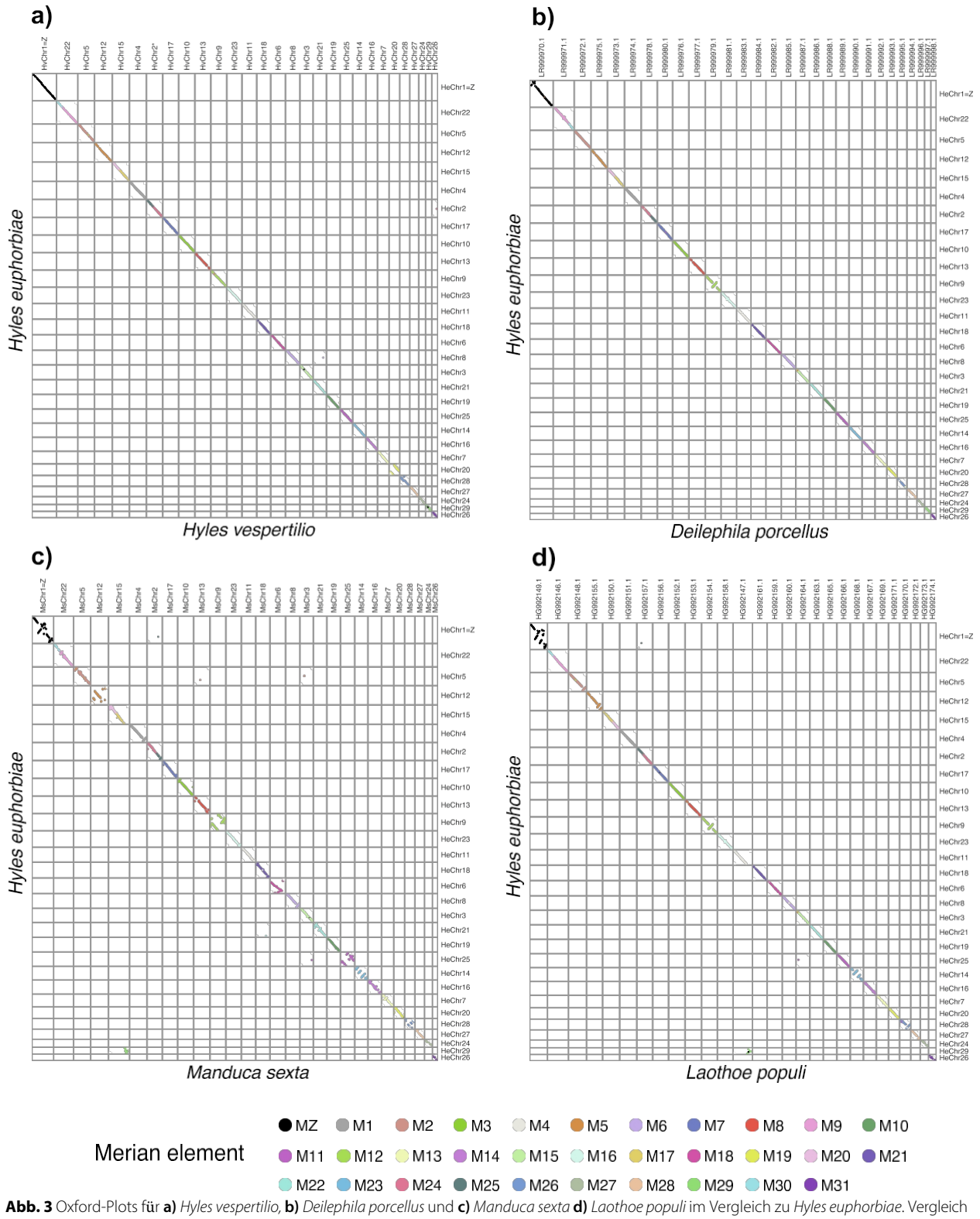

Abb. 3 Oxford-Plots für **a)** *Hyles vespertilio*, **b)** *Deilephila porcellus* und **c)** *Manduca sexta* **d)** *Laothoe populi* im Vergleich zu *Hyles euphorbiae*. Vergleich der Zusammenstellungen zeigen eine starke Erhaltung sowohl der Verknüpfungsgruppen als auch der Genordnung bei den Sphingidae. Orthologe sind nach der Identität der Merian-Elemente eingefärbt. Die Chromosomen sind bei *Hyles euphorbiae* nach abnehmender Größe geordnet

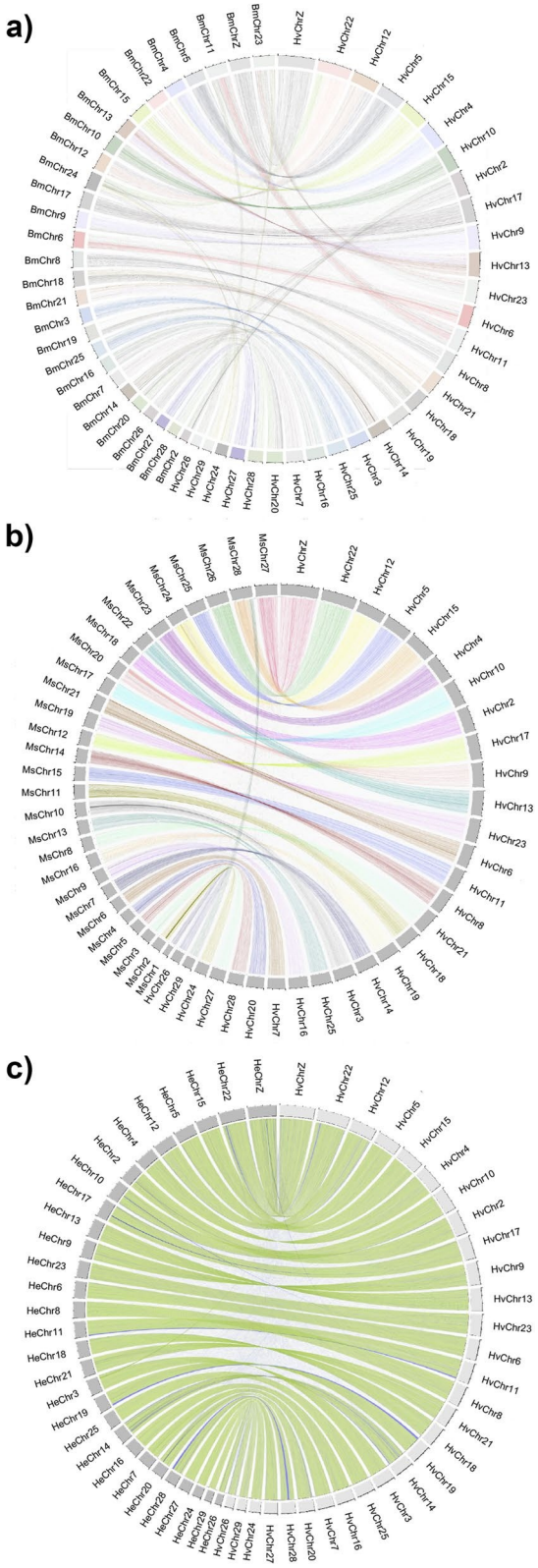

Abb. 4 CIRCOS-Plots von Genom-Alignments auf Chromosomenebene. Das Z-Chromosom entspricht dem Chromosom 1. Die Chromosomen von *H. vespertilio* sind nach Größe geordnet. Die Buchstaben "ScWp86a" im Gerüst werden die Namen von *Hyles vespertilio* aus Gründen der Übersichtlichkeit weggelassen. **a)** *Bombyx mori* und *H. vespertilio* mit automatischer Färbung unter Verwendung von shinyCircos zur Erleichterung der Korrelation. **b)** *Manduca sexta* und *H. vespertilio* mit Färbung nach Chromosomen. **c)** *H. euphorbiae* und *H. vespertilio*, mit den 29 Chromosomen jeder Art (*H. vespertilio* hellgraue Kästchen am äußeren Rand, *H. euphorbiae* dunkelgraue Kästchen). Die Farbe der Verbindungen stellt die Strangorientierung der Chromosomen von *H. euphorbiae* im Vergleich zur Strangorientierung von *H. vespertilio* dar (grün bedeutet + ; blau bedeutet -)

über sieben Exons und 14.395 bp). Ansichten der WntA-BLAT-Ergebnisse sind in Abb. S4 dargestellt. Es gibt zahlreiche Unterschiede in den Wiederholungen zwischen den beiden Hyles-Arten. Zum Beispiel enthält *H. vespertilio* eine 443 bp lange DNA-Repeat-Insertion ("ScWp86a_140_ HRSCAF_198: 8 . 200.940-8.201.382" ; rnd-4_family-526; Familie TcMar-Tc1, Pos. 512-1311 in der Wiederholung) im Intron 2, das *H. euphorbiae* fehlt. Das genomische Interval, das die Exons und Introns von *WntA* überspannt, ist bei *H. euphorbiae* (oder *H. vespertilio*, je nach Polaritätsdefinition) invertiert.

Die BLAT-Suche nach der Kortex-Proteinsequenz von *M. sexta* (8 Exons) ergab eine Übereinstimmung mit 7 Exons auf Chromosom 17 in beiden Arten [Abb. S5; *H. euphor biae* (HeChr17, "sc_9: 7.671.034-7.674.090") und *H. vespertilio* (HvChr17, "ScWp86a_210_HRSCAF_302: 14.798.018-14.801.349")]. Ansichten der BLAT-Ergebnisse des *Kortex* sind in Abb. S5 dargestellt. Das Methylierungsprofil in *H. euphorbiae* (Abb. S5a) deutet auf eine normale Transkriptionsaktivität in diesem Abschnitt der Exons hin, da die Wahrscheinlichkeiten der Basenmodifikationen alle < 50% sind (blau). Die Position, Qualität und Quantität der Wiederholungen in den Introns unterscheiden sich zwischen den beiden Arten. So enthält das erste Intron in *H. euphorbiae* keine Wiederholungen, während es in *H. vespertilio* zwei LINEs zusätzlich zu einem 211 bp langen Abschnitt (ScWp86a_210_HRSCAF_302: 14.798.399-14.798.609; Abb. S5b) mit unbekannten Wiederholungen (rnd-1_family-40 genannt) enthält. Es überschneidet sich stromabwärts mit der LINE der Familie L2, die eine unregelmäßige einfache Wiederholung von 5' TATT aufweist. Sie befindet sich im Intron 1 der *Kortex* X1-Sequenz von *M. sexta* (annotiert mit *M. sexta* tran- script LOC115451467_rna-XM_030179809.2.8). Auch hier ist der Bereich um den *Genkortex* bei *H. euphorbiae* (oder *H. vespertilio*, je nach Polaritätsdefinition) invertiert. Die 100 kb-Ansichten der Kortex-BLAT-Ergebnisse sind in

Abb. S6 dargestellt und zeigen eine hohe Anzahl von Wiederholungen in der Nähe der Gen-Exon-Strecke in beiden Arten.

Die BLAT-Suche von *M. sexta aristaless* (in 4 Varianten, Tabelle S2) ergab, dass das Gen auf Chromosom 4 vorkommt

Tabelle 4 Assemblierungslängen und Anteil der Wiederholungen der (am 14.10.2022) verfügbaren Genome im Vergleich (es wurde eine neue Assemblierung von *H. vespertilio* verwendet, die auf PacBio-Sequenzdaten von Pippel et al. basiert [12])

Arten	Gesamtlänge	Maskiert [%]
Manduca sexta [37]	470,036,997	27.8
Mimas tiliae [14]	472,718,480	46.7
Laothoe populi [14]	549,962,493	47.7
Hemaris fuciformis [38]	441,141,813	33.2
Deilephila porcellus [39]	389,730,088	27.5
H. euphorbiae	504,310,614	37.1
H. vespertilio	651,427,907	46.6

bei beiden Arten (HeChr4, "sc_6: 4.437.062-4.474.302", mit einer Länge von 37.241 bp) und *H. vespertilio* (HvChr4, "ScWp86a_21_HRSCAF_27: 5 .766.997-5.884.695", 5
Exons innerhalb eines großen Abschnitts von 117.699 bp). Das erste Intron von *aristaless* enthält in *H. euphorbiae* ein Repeat-Motiv, das in *H. vespertilio* nicht vorkommt (Abb. S7). Abgesehen von diesem Unterschied sind die Positionen der anderen Repeats innerhalb des Exonabschnitts bei beiden Arten sehr ähnlich, aber die Abschnitte der Repeats sind bei *H. vespertilio* länger (Maßstab 2 kb in beiden Ansichten von Abb. S7). Beim Herauszoomen wird die große Vielfalt an Repeats in der Nähe dieses Gens deutlich (Abb. S8), ebenso wie die großen Abschnitte nicht codierender DNA zwischen den zahlreichen Exons.

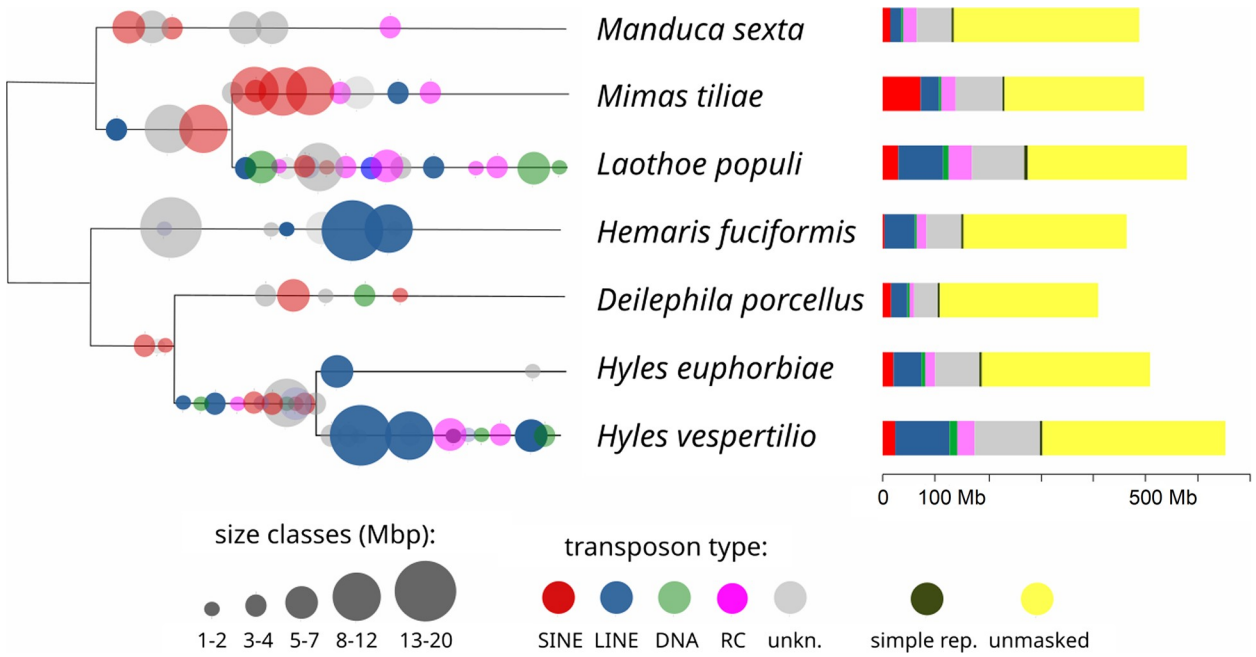

Abb. 5 Transposon-Bursts in einem phylogenetischen Baum, der von Hand nach der Backbone-Topologie in [44] gezeichnet wurde, in relativer Reihenfolge (aufgrund der Sequenzdiversität) und proportionalen Wiederholungsgehalten von sieben Sphingidenarten mit verfügbaren Genomdaten (Transposon-Typen: SINE: short interspersed elements, LINE: long interspersed elements, DNA: DNA-Transposons, RC/Hel: Rolling-Circle-Transposons, z.B. Helitrons)

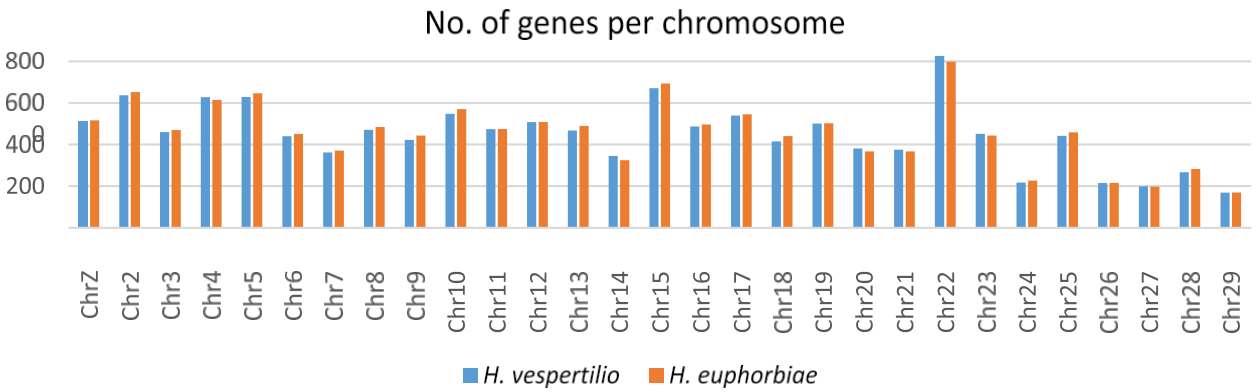

Abb. 6 Anzahl der Gene pro Chromosom bei *H. euphorbiae* und *H. vespertilio* im Vergleich

Das BLAT-Ergebnis von *Distal-less* (Tabelle 2) zeigt eine Position auf Chromosom 2 bei beiden Arten (HeChr2, "sc_7: 16.368.465-16.384.630", 2 Exons innerhalb von 16.166 bp) und *H. vespertilio* (HvChr2, "ScWp86a_159_HRSCAF_226: 4.646.872-4.661.762", 2 Exons innerhalb von 14.891 bp). Das einzelne Intron zwischen den beiden Exons enthält in beiden Arten Wiederholungen, allerdings sehr unterschiedliche Typen in verschiedenen Positionen und Ausmaßen (Abb. S9).

Karyotyp

Die Analyse männlicher mitotischer Chromosomen, die mittels FISH mit einer telomeren Sonde gefärbt wurden (Telomer-FISH), ergab, dass der Karyotyp von *H. euphorbiae* aus $2n = 58$ Chromosomen besteht (Abb. 7a). Wie für Lepidoptera typisch, sind die Chromosomen vom holokinetischen Typ, d. h. sie haben keine primäre Einschnürung (Zentromer) und sind morphologisch einheitlich und unterscheiden sich nur in der Größe. Die Chromosomenzahl wurde durch die Analyse der meiotischen Kerne im Pachytenstadium bestätigt, wo sich homologe Chromosomen paaren und längliche Bivalente bilden. Pachytene Komple- mente, gefärbt durch GISH in Kombination mit Telomer-FISH, zeigten eine haploide Anzahl von 29 Bivalenten in beiden Geschlechtern (Abb. 7b, c). Darüber hinaus identifizierte GISH ein bivalentes WZ-Geschlechtschromosom in pachytenen Oozyten durch Markierung eines großen Teils des W-Chromosoms mit der weiblichen genomischen DNA-Sonde (gDNA) (Abb. 7c), während in pachytenen Spermatozyten kein Bivalent identifiziert wurde (Abb. 7b). Diese Ergebnisse zeigen deutlich, dass *H. euphorbiae* ein WZ/ZZ (weiblich/männlich)-Geschlechtschromosomensystem hat, das bei Lepidoptera üblich ist. Es ist anzumerken, dass das WZ-Bivalent relativ lang ist (Abb. 7c), was darauf hindeutet, dass die W- und Z-Chromosomen zu den größten Chromosomen im Karyotyp von *H. euphorbiae gehören*.

Schätzung der Chromosomengröße anhand von Karyotyp-Bilddaten Die Schätzung der Chromosomengröße basiert auf Abb. 7c, Biva- lenten einer weiblichen Gonadenzelle im Pachytenstadium. Die Schätzungen der Chromosomengröße sind in Abb. 8 dargestellt. Sie bestätigen, dass das WZ-Bivalent das größte Chromosom ist. Das Softwarepaket napari-karyotype [49] basiert auf einer halbautomatischen Bildverarbeitung und verwendet eine schwellenwertbasierte Bildsegmentierung, um chromosomenbezogene Komponenten zu erkennen. Identifizierte chromosomale Objekte sind von roten Rechtecken umgeben und mit den Schätzungen beschriftet.

Schätzung der Genomgröße

Die Genomgröße von *H. vespertilio* und *H. euphorbiae* wurde mit Hilfe der Durchflusszytometrie in drei Wiederholungen gemessen, die jeweils von einem einzigen Individuum stammten. Die durchschnittlichen Ergebnisse von 0,575 pg und 0,484 pg zeigen, dass die 1C-DNA-Werte 562 Mb bzw. 472 Mb betragen (1 pg = 978 Mb).

Diskussion

Qualität der Montage

Die hier vorgestellten chromosomalen Genomassemblies von *H. euphorbiae* und *H. vespertilio* sind in hohem Maße zusammenhängend, wie die N50-Werte der Gerüste zeigen, die über denen der sequenzierten nahen Verwandten *M. sexta* [37] und *B. mori* [41] liegen. Der prozentuale Anteil der Assemblierungslänge, der in Chromosomengerüsten untergebracht werden konnte, ist bei *Hyles euphor biae* am höchsten (Tabelle 2). Beide Metriken unterstreichen die hohe Qualität der beiden neuen Assemblies für die Gattung *Hyles*. Der Gengehalt ist in allen verglichenen Assemblies ähnlich hoch.

Vergleich der beiden neuen Hyles-Genome

Der große Unterschied in der geschätzten Genomgröße zwischen den beiden Arten innerhalb der Gattung *Hyles*, d. h. *H. vespertilio*

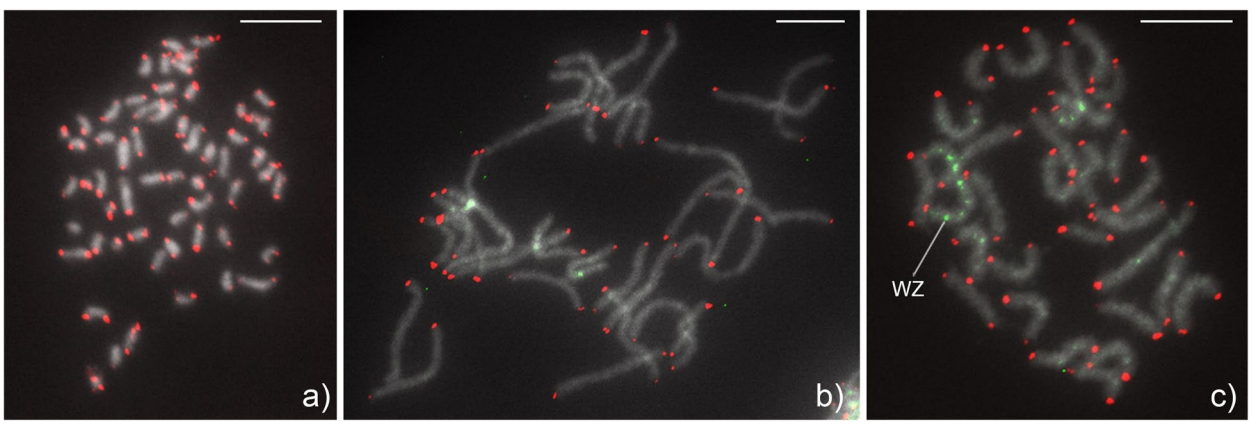

Abb. 7 Telomer- und Analyse der Chromosomen von *Hyles euphorbiae*. Die Hybridisierungssignale der Cy3-markierten (TTAGG)$_n$ telomeren Sonde (rot) zeigen die chromosomalen Enden an (**a-c**), und die Fluorescein-markierte weibliche gDNA-Sonde (grün) identifiziert das Geschlechtschromosomensystem (**b** und **c**) Chromosomen wurden mit DAPI (grau) angefärbt. **a)** Männliche mitotische Prometaphase, angefärbt mit Telomer-FISH, zeigt eine diploide Chromosomenzahl von *2n* = 58. **b)** Männliches pachytenes Komplement, angefärbt mit einer Kombination aus GISH und Telomer-FISH, zeigt 29 Bivalente, aber ohne hervorgehobenes Bivalent, was auf eine ZZ-Geschlechtschromosomenkonstitution hinweist. **c)** Weibliches pachytenes Komplement, angefärbt mit einer Kombination aus GISH und Telomer-FISH, zeigt 29 Bivalente, einschließlich des WZ-Geschlechtschromosomenpaars, das durch das mit der weiblichen gDNA-Sonde markierte W-Chromosom identifiziert wurde. Balken = 10 µm

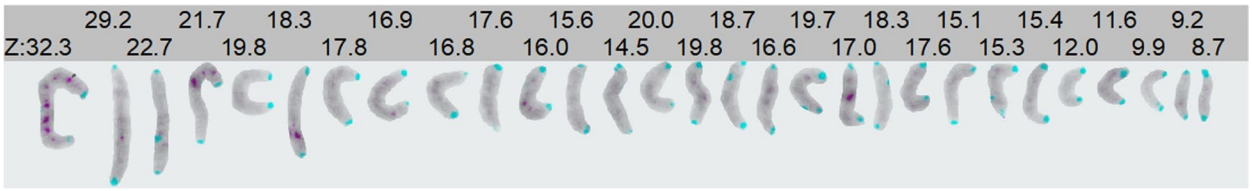

Abb. 8 Annotierte Chromosomenbivalente aus einem weiblichen Pachytenbild von *Hyles euphorbiae* unter Verwendung des Softwarepakets napari-karyotype. Die Chromosomenbivalente wurden manuell nach der geschätzten Größe ausgerichtet. Das größte Chromosomenpaar stellt die Geschlechtschromosomen dar, in diesem Fall ein WZ-Bivalent. Die blauen und roten Punkte entlang der Chromosomen entsprechen der Lage der telomeren Sonde b z w . der weiblichen gDNA-Sonde aus Abb. 7

Die Tatsache, dass das Genom von Hyles lineata mit 562 Mb etwa 20 % größer ist als das von *H. euphor- biae* mit 472 Mb in den durchflusszytometrischen M e s s u n g e n , war aufgrund der kurzen Divergenzzeit zwischen den beiden Arten unerwartet [50]. *Hyles lineata* ist eine der ersten Hyles-Arten, die innerhalb der Gattung divergieren [5] und hat eine noch geringere Genomgröße von 450 Mb (0,46 pg) in der Durchflusszytometrie [51]. Die potenziell geringere Genomgröße dieser dritten Hyles-Art deutet darauf hin, dass das größere Genom von *H. vespertilio* das Ergebnis einer kürzlichen Vermehrung repetitiver Elemente ist und ein abgeleitetes Merkmal darstellt.

Obwohl die Werte der Genomgrößen auf der Grundlage der Assemblierungslängen aufgrund technischer Unterschiede in der Methodik nicht direkt mit den Genomgrößenwerten auf der Grundlage der Durchflusszytometrie vergleichbar sind, liegen die relativen Größen in der gleichen Größenordnung. Die Assemblierungslänge von *H. vespertilio* ist mit 651 MB etwa 30 % länger als die von *H. euphorbiae* mit 504 MB, obwohl die Anzahl der Gene bei beiden Arten sehr ähnlich ist (Abb. 6).

Das Genom von *H. euphorbiae*, das in dieser Arbeit vorgestellt wird, hat viel weniger repetitive Elemente, insbesondere LINEs, als *H. vespertilio*. Der hohe Gehalt an Wiederholungen bei allen vier verglichenen Arten, *H. euphorbiae, H. vespertilio, M. sexta* und *B. mori,* ist typisch für Lepidoptera-Genome [52, 53]. Die Anzahl der Wiederholungen variiert jedoch von Art zu Art. Frühere Untersuchungen haben eine Korrelation zwischen dem repetitiven Anteil des Genoms und der Genomgröße innerhalb und zwischen den Arten gezeigt [54, 55]. In der Tat ist das Genom von *H. vespertilio* das größte in unserem Vergleich und das Genom mit dem größten Anteil an Wiederholungen, während die anderen Sphingiden-Genome eine geringere Anzahl an Wiederholungen aufweisen, was mit ihrer geringeren Genomgröße korreliert. Wie in früheren Forschungsarbeiten beschrieben, sind die hier gefundenen repetitiven Elemente daher wahrscheinlich die treibende Kraft für die Ausdehnung der Genomgröße, möglicherweise aufgrund einer positiven Rückkopplungsschleife, die es diesen Elementen ermöglicht, sich in großen Genomen leichter zu verbreiten [56]. Darüber hinaus wird angenommen, dass das Repeatom eine wichtige Rolle bei der genetischen Innovation und der Speziation spielt [57-61]. Dies könnte erklären, warum *H. vespertilio* als eine der isoliertesten Arten innerhalb der Gattung der Habichtskäfer gilt, da es keine Hybridisierung mit zahlreichen anderen Hyles-Arten gibt. Im Gegensatz dazu ist *H. euphorbiae* bekannt für seine

weite Verbreitung und häufige Hybridisierung selbst mit entfernt verwandten Arten [2, 50, 62].

Da die beiden Arten durch eine *p-Distanz* von nur 2,1 % in den neutralen Markergenen *COI* und *COII* getrennt sind [50], was einem Split-Zeitpunkt von nur 4,2-5,6 Mya entspricht (nach den jeweiligen molekularen Uhren von 1,1-1.2 % pro Million Jahre pro Linie, was zu 2,3 % Sequenzdivergenz zwischen Artenpaaren pro Million Jahre Trennung führt [63]), wurde eine hohe Syntenie von Genen zwischen *H. vesper- tilio* und *H. euphorbiae* im Kernchromosomensatz erwartet (Abb. 3, 4). Tatsächlich zeigen die abgebildeten Genom-Alignments (Abb. 3, 4), wie ähnlich sich die beiden Hyles-Genome sind. Das mitochondriale Genom von *H. euphor biae* (Abb. S1) ist dem von *H. vespertilio* ebenfalls sehr ähnlich (siehe [13]). Der Vergleich der orthologen Einzelkopien in jedem Genom zeigt die Erhaltung von 29 orthologen Chromosomen. Diese Chromosomen entsprechen weitgehend den 32 angestammten Chromosomen der Lepidoptera (Merian-Elemente)[42], mit Ausnahme von drei Fusionen, was die Verringerung der Chromosomenzahl erklärt. Darüber hinaus beobachten wir in Übereinstimmung mit anderen Studien [42, 52, 64, 65] ein hohes Maß an Syntenie, d. h. die Erhaltung der Genreihenfolge innerhalb jedes Chromosoms. Die Ausweitung der Vergleiche über *Hyles* hinaus legt nahe, dass diese 29 Chromosomen in der gesamten Sphingidae-Gruppe stabil geblieben sind, mit Ausnahme von zwei unabhängigen Fusionsereignissen in
M. sexta und *L. populi,* die es in *Hyles* und
D. porcellus. Eine solch starke Erhaltung der Chromosomenstruktur wurde bereits bei Nymphalidae beobachtet [66]. Unsere hier vorgestellten Analysen deuten darauf hin, dass auch die Chromosomen der Sph- ingidae über lange Evolutionszeiträume hinweg bemerkenswert syntenisch sind.

Karyotyp, Taxonomie der Chromosomen

Neben der Untersuchung der Erhaltung der Chromosomenstruktur durch den Vergleich von Einzelkopie-Orthologen mit Merian-Elementen in jedem Genom haben wir auch Ganzgenom-Alignments durchgeführt. Ganzgenom-Alignments zwischen jedem Hyles-Genom und dem der gut untersuchten Spezies *B. mori* wurden unter Verwendung der Positionsinformationen von Syntenie-Blöcken (Tabelle S1) durchgeführt, um die Homologie zwischen *B. mori*- und Hyles-Chromosomen zuzuordnen, so dass

Chromosomenbenennung auf der Grundlage von *B. mori*. Dies führte zu einer robusten Chromosomen-Taxonomie auf der Grundlage von *B. mori* (die für die beiden untersuchten Hyles-Arten in Abb. 4c vorgestellt wurde).

Angesichts des jungen Alters der Gattung (der älteste Split innerhalb der Gattung liegt bei 10,8-12,4 Mya, basierend auf der p-Distanz [2]) kann davon ausgegangen werden, dass die Chromosomentaxonomie auch für die gesamte Gattung *Hyles* gilt. Als Goldstandard für die Chromosomentaxonomie schlagen wir unseren Proto- col der Homologie-Inferenz auf der Grundlage von Genom-Alignments vor. Die jüngsten Habichtskraut-Zusammenstellungen [14, 39], einschließlich der von *M. sexta* [37], verwenden nur willkürliche Nummern entsprechend der Chromosomengröße. Eine Tabelle mit der Homologie der Chromosomen, den Zugangsnummern und den Gerüstnummern findet sich in Tabelle S3. Die Verwendung dieses vergleichenden Ansatzes als Standard in der Chromosomentaxonomie der Bombyzoen wird Diskussionen auf der Grundlage von Genomvergleichen für spezifische Forschungsfragen, wie z. B. die Homologien der Flügelmustergene in der vorliegenden Studie, erheblich erleichtern.

In frühen genetischen Studien wurden drei Chromosomen von *B. mori* (BmChr11, BmChr23 und BmChr24) oft so interpretiert, dass sie bei anderen Arten geteilt sind und die Chromosomenzahl im Vergleich zu *B. mori* erhöhen, z. B. [43, 67, 68]. Die jüngste Entdeckung von 32 Merian-Elementen, den angestammten Bindungsgruppen der Lepidoptera, bietet einen Rahmen für das Verständnis der chromosomalen Veränderungen bei *B. mori* im Vergleich zu anderen Arten. Der Urzustand von Ditrysia ist $n = 31$, was auf eine alte Fusion zwischen zwei Merian-Elementen zurückzuführen ist. Daher hat *B. mori* im Vergleich zu anderen Ditrysia-Arten eine reduzierte Chromosomenzahl, die auf drei weitere Fusionen zurückzuführen ist.

Die Ganzgenom-Alignments zwischen *B. mori* und *Hyles* zeigen, dass diese beiden Gruppen unabhängige Fusionen durchlaufen haben, die zu den reduzierten Chromosomenzahlen geführt haben. Dies steht im Einklang mit der früheren Beobachtung, dass drei *B.* mori-Chromosomen bei anderen Arten häufig geteilt sind.

Die erhaltenen Chromosomenbilder von *H. euphorbiae* waren so klar, dass sie mit einer Größenschätzung nach Napari-Karyotyp versehen werden konnten (Tabelle 3) [49]. Es ist jedoch anzumerken, dass dies schwieriger war als erwartet, da sich die Chromosomenbivalente im Pachytenstadium auf dem Bild berührten und daher manuell extrahiert werden mussten. Ohne diesen Schritt wäre die Zuordnung von Objekt und Hintergrund ungenau gewesen. Außerdem sind Chromosomen flexible Strukturen, deren Länge vom Stadium, dem Grad der Kondensation und auch von den Präparationsmethoden abhängt. Daher entspricht ihre gemessene Länge nicht immer ihrer Größe, was insbesondere für meiotische Chromosomen im Pachytenstadium gilt [69]. Um die Chromosomenbilder des Karyotyps mit den Chromosomennummern der *B.* mori-Chromosomentaxonomie in Verbindung zu bringen, wird es in Zukunft notwendig sein, eine weitere in-situ-Hybridisierung mit gen-spezifischen fluorochrom-markierten Sonden nach Yasukochi et al. [43] (deren Abb. 2) durchzuführen.

Flügelmuster-Gene

Die Lage zahlreicher Flügelmustergene ist bei Nymphalidae gut untersucht worden. Hier leiten wir die wahrscheinlichen Standorte von *optix*, *WntA*, *cortex, aristaless* und *Distal-less* in *Hyles* auf der Grundlage von Heliconius-Aminosäuren oder der homologen *M.* sexta-Nukleotid-Ähnlichkeit ab und charakterisieren ihren genomischen Kontext.

Optix

Optix und die umliegenden Gene sind bei Lepi- doptera hoch konserviert [70]. Zhang et al. [28] haben gezeigt, dass Optix-Knock-Outs einen vollständigen Ersatz der Farbpigmente durch Melanine aufweisen, was zu schwarzen und grauen Schmetterlingen führt. Im Gegensatz dazu könnte der hohe Prozentsatz an Indels, der im Alignment vor und nach *optix* zwischen den Genomen von *H. vespertilio* und *H. euphorbiae* gefunden wurde, die zuvor vorgeschlagene Hypothese [71] unterstützen, dass die Flügelmuster tatsächlich durch cis-regulatorische Elemente in der Nähe der Position von *optix* gesteuert werden. Wenn mehr Sequenzdaten von mehr Individuen zur Verfügung stehen, können F_{ST} Plots berechnet werden, um Korrelationen zwischen bestimmten Kandidatenregionen zu bestätigen. Obwohl *Man- duca sexta* das für den hier vorgestellten Vergleichsrahmen am besten geeignete Mottengenom ist, besteht das Flügelmuster dieser Art ebenfalls aus Schwarz und Grau, ohne farbige Streifen oder Flecken. Daher ist die Inversion von *optix* auf Chromo- einige 23 in *H. euphorbiae* im Vergleich zu *H. vespertilio* und
M. sexta, könnte aber auch andersherum sein, nämlich dass *H. euphorbiae* mit seinen cremefarbenen, olivbraunen Vorderflügeln, die manchmal rosa sind, besitzt das Gen mit normaler Funktion, und
H. vespertilio und *M. sexta* haben eine nicht-funktionale Inversion. Die Hypothese, ob *optix* Teil eines Supergens ist, soll durch gezielte Resequenzierung von Phänotypen im Vergleich genauer untersucht werden. Der kausale Zusammenhang zwischen grauen Flügeln bei *H. vespertilio* und *optix* und den umliegenden Genregionen konnte jedoch nur durch die Analyse von Knock-outs beantwortet werden, die mit dem CRISPR/Cas9-System erzeugt wurden.

WntA

Die sehr unterschiedlichen Mengen, Klassen und Familien von Wiederholungen zwischen den beiden sequenzierten Hyles-Arten lassen vermuten, dass die Funktion dieses Gens bei dem grauen *H. vespertilio* gestört sein könnte, da diese Art das größere Genom und mehr Wiederholungen aufweist. Es werden Sequenzdaten von mehr Individuen benötigt, um eine Kandidatenregion zu bestimmen und eine robuste Korrelation herzustellen.

Kortex

Ein weiterer auffälliger Unterschied zwischen den beiden Arten in der Nähe des Abschnitts von 8 Kortex-Exons ist die Insertion von 229 bp, die als unbekannte Wiederholung (Abb. S5b) auf Chromosom 17 in *H. vespertilio* markiert wurde. Die Zuordnung dieses Fragments zum Genom von *H. euphorbiae* mit BLAT ergab 200 Treffer zwischen 89,8-96,7 % Identität auf jedem Chromosom

und 203 Treffer im neuen Genom von *H. vespertilio* (Daten nicht gezeigt), was auf die weite Verbreitung dieser Wiederholungsregion in den Genomen hinweist. Sie wurde auch bei der sehr eng verwandten Makroglossinenart *Deilephila porcellus* gefunden [39] und *Hemaris fuciformis* [38], sowie mehrere Treffer im Genom der Smerinthenart *Mimas tiliae* [14], einer phylo- genetisch etwas weiter entfernten Art. Auch bei *H. euphorbiae* liegt der *Kortex* auf einer Inversion (obwohl es in diesem Vergleichsrahmen möglicherweise umgekehrt ist), was auf eine mögliche Verbindung innerhalb eines zweiten Supergens hindeutet (da es auf einem anderen Chromosom liegt als *optix*). Die beiden neuen Hyles-Referenzgenome eröffnen die Möglichkeit, diese Möglichkeiten zu untersuchen. Weitere Analysen und Vergleiche zwischen den hier vorgestellten Genomen und anderen Genomen, die jetzt verfügbar werden, dürften zu einem besseren Verständnis der Mechanismen führen, die die Flügelmuster in der Familie der Sphingidae steuern.

Aristallos & Distallos

Die beiden Gene scheinen in Bezug auf die Anzahl und Art der Wiederholungen zwischen den Arten ähnlich unterschiedlich zu sein. Die beiden untersuchten Hyles-Falter weisen keine Augenflecken auf ihren Flügeln auf, ebenso wenig wie andere Arten der Gattung. Weitere Untersuchungen könnten jedoch klären, ob diese Gene die einfachen braunen Flecken auf den Vorderflügeln bei *H. euphorbiae* und vielen anderen Arten dieser Gattung steuern, indem sie klären, ob die Gene bei *H. vespertilio* nicht exprimiert oder möglicherweise zum Schweigen gebracht werden.

Schlussfolgerungen

Frühere Studien hatten bereits gezeigt, dass die Ähnlichkeit von Flügelmustern nicht mit phylogenetischen Beziehungen in der Gattung *Hyles* korreliert [1, 2, 5], und zwar weder nach der derzeitigen morphologischen Definition noch nach der molekularen Phylogenie. Morphologen haben lange Zeit argumentiert, dass sie sich in Ermangelung molekularer Daten auf phänotypische Merkmale verlassen müssen, und haben oft auffällige Unterschiede in den Flügelmustern in ihre Artbeschreibungen aufgenommen. Bei *Hyles* jedoch spiegelt die Entwicklung der Flügelmustermerkmale nicht die Entwicklung der Art wider. Natürlich sind Genbäume keine Artbäume [72], so dass selbst traditionelle genetische Analysen, z. B. [1, 2, 5], nicht unbedingt den wahren Baum widerspiegeln.

In dieser Studie stellen wir zwei hochwertige annotierte Chromosomenzusammenstellungen vor und berichten über das Vorhandensein, die Sequenz und die Lage von Flügelmustergenen. Anhand von Referenzgenomen von *H. euphorbiae* und *H. vespertilio*, die sehr unterschiedliche Vorderflügelmuster und -zusammensetzungen aufweisen, eröffnen wir Möglichkeiten zur Untersuchung der Evolution von Flügelmustern auf der Grundlage einer numerisch auswertbaren, objektiven Datenquelle. Die Gene *optix* (auf Chromosom 23), *WntA* (auf Chromosom 28), *cortex* (auf Chromosom 17), *aristaless* (auf Chromosom 4) und *Distal-less* (auf

Chromosom 2) versprechen, für eine vergleichende Studie zahlreicher Hyles-Individuen von Interesse zu sein, da die genomischen Regionen um diese Gene eine hohe Anzahl von Indels aufweisen, die ihre Funktion beeinträchtigen könnten, und Inversionen zwischen den beiden Hyles-Arten festgestellt wurden. Da die fünf Gene, die in dieser Studie verglichen wurden, alle auf unterschiedlichen Chromosomen liegen, ist noch viel Arbeit nötig, bis wir die Grundlage der Veränderungen z. B. zwischen den beiden in der Einleitung beschriebenen auffälligen Hyles-Vorderflügelmustern verstehen, ganz zu schweigen von den zusätzlichen, oft einzigartigen Vorderflügelmustern bei anderen Arten. Die Ergebnisse unserer vergleichenden Repeatomanalyse bestätigen, dass repetitive Elemente wahrscheinlich die treibende Kraft für die Ausdehnung der Genomgröße bei *Hyles vespertilio* im Vergleich zu *H. euphorbiae* sind. Die in dieser Studie für diese beiden Arten bereitgestellten Genomdaten auf Chromosomenebene bieten zuverlässige Referenzen in der Familie der Sphingidae für künftige Studien, die so viele Arten wie möglich einbeziehen, um die Evolution der Vorderflügelmuster in dieser Gruppe der Lepidoptera zu erforschen.

Wir empfehlen nachdrücklich die Anwendung eines standardisierten Chromosomennennamens für jedes neu sequenzierte Genom, um die Vergleichbarkeit zu erleichtern, und schlagen die Anwendung eines zweifachen vergleichenden Genomikansatzes vor. Die Kombination von CIRCOS-Diagrammen auf der Grundlage paarweiser Alignments in einem bestimmten System (z. B. *B. mori* für Bombycoidea) für die Chromosomentaxonomie mit Oxford-Diagrammen auf der Grundlage von Merian-Elementen zur Ableitung und Visualisierung der Richtung chromosomaler Umlagerungen ermöglicht präzisere evolutionäre Ableitungen.

Material und Methoden

Insekten

Für die Karyotyp-Analyse wurden Exemplare der *H.* euphorbiae-Population aus Griechenland (leg. P. Mazzei, Serifos) im Labor aufgezogen (Sommer 2019). Mehrere Larven und junge Puppen wurden für Chromosomenpräparationen verwendet (siehe unten). Für die Genomsequenzierung wurde ein männliches Exemplar von *H. euphorbiae* (Abb. 1a) am 27. Juli 2021 in der Nähe von Berbisdorf (Deutschland) gesammelt und erhielt im Molekularlabor der Senckenberg Naturhistorischen Sammlungen Dresden, Museum für Zoologie, die Gewebebelegnummer MTD-TW- 13387. Ein zweiter, ähnlicher männlicher Falter vom selben Fundort wurde als Belegexemplar in die Hauptsammlung der SNSD Entomologie aufgenommen. Die Hi-C-Daten für *H. vespertilio* (Abb. 1b) wurden aus gefrorenem (-80 °C) Gewebe desselben Individuums aus Vallonina, Italien (gesammelt 2018), gewonnen, das für die ursprüngliche Genompublikation in Pippel et al. verwendet wurde [12]. Das Gewebe wurde vom Max-Planck-Institut für Molekulare Zellbiologie und Genetik (Dresden), wo es die Zugangsnummer LG2117 hatte, in die SNSD MTD-Gewebesammlung überführt und erhielt die Belegnummer MTD-TW-13386.

Karyotyp-Analyse

Spreizchromosomenpräparate wurden wie von Yoshido et al. beschrieben hergestellt [73]. Mitotische Chromosomen wurden aus Flügelimaginalscheiben oder Hoden von Larven des letzten Stadiums gewonnen. Meiotische Chromosomen im Pachytenstadium der Prophase I wurden entweder aus den Hoden von Larven im letzten Instadium oder aus den Eierstöcken von 3-5 Tage alten Puppen gewonnen. Die Gewebe wurden in Kochsalzlösung präpariert, in hypotoner Lösung (75 mM KCl) entweder 5 Minuten (Eizellen) oder 15 Minuten (Hoden und Flügelimaginalscheiben) gequollen und dann 10 Minuten lang in Carnoy's Fixiermittel (Ethanol, Chloroform, Essigsäure, 6:3:1) fixiert. Die in 60%iger Essigsäure dissoziierten Zellen wurden auf einer Heizplatte bei 45 °C ausgebreitet. Alle Chromosomenpräparate wurden durch eine abgestufte Ethanolreihe (70 %, 80 % und 100 %, jeweils 30 s) geleitet und bei -80 °C gelagert.

Die Fluoreszenz-in-situ-Hybridisierung (FISH) mit der telomeren Sonde $(TTAGG)_n$ und die genomische In-situ-Hybridisierung (GISH) wurden nach dem von Yoshido et al. [74] beschriebenen Verfahren durchgeführt. $(TTAGG)_n$ telomere Sequenzen wurden durch Nicht-Template-PCR nach dem Protokoll von Sahara et al. erzeugt [75]. Männliche und weibliche genomische DNAs (gDNAs) von *H. euphorbiae* wurden separat aus Larven des letzten Instars durch Standard-Phenol-Chloroform-Extraktion gewonnen. Die DNA-Sonden wurden durch Nick-Translation mit einer Mischung aus DNase I und DNA-Polymerase I (beide Thermo Fisher Scientific, Waltham, MA, USA) mit entweder Aminoallyl-dUTP-Cy3 oder Fluorescein-12-dUTP (beide Jena Bioscience, Jena, Deutschland) markiert.

Die Chromosomenpräparate wurden aus dem Gefrierschrank entnommen, durch die abgestufte Ethanolreihe geleitet, an der Luft getrocknet und dann in 70 %igem Formamid in 2 × SSC für 3,5 Minuten bei 70 °C denaturiert. Für ein Präparat enthielt der Sondencocktail 500 ng Fluorescein-markierte weibliche gDNA, 100 ng Cy3-markierte telomere Sonde, 3 µg unmarkierte männliche sonifizierte
gDNA, und 25 µg beschallte Lachssperma-DNA (Sigma-Aldrich, St. Louis, MO, USA) in 10 µl Hybridisierungspuffer
(50% Formamid, 10% Dextransulfat in 2 × SSC). Die Denaturierung des Sondencocktails erfolgte für 5 Minuten bei 90 °C. Die Präparate wurden unter einem Zeiss Axioplan 2 Mikroskop (Carl Zeiss, Jena, Deutschland) untersucht. Die digitalen Bilder wurden mit einer Olympus CCD-Monochromkamera XM10 und der Software cellSens 1.9 (Olympus Europa Holding, Hamburg, Deutschland) aufgenommen und mit Adobe Photoshop CS4 bearbeitet.

Karyotyp-basierte automatische Chromosomenbeschriftung und Größenschätzung

Das Bild der meiotischen Chromosomen wurde mit der Bildbearbeitungssoftware GIMP (Version 2.10) [76] vorverarbeitet, um einzelne Chromosomenbivalente manuell auszuschneiden. Das bearbeitete Bild wurde in das Tool napari-kar- yotype (Version c41103e) [49] geladen. Der Schwellenwert für die Bildsegmentierung, der Unschärfefaktor und die Genomgröße wurden auf 0,13 festgelegt,
0,5 bzw. 504 Mb.

Schätzung der Genomgröße

Die Genomgrößen von zwei Schwärmerarten wurden nach dem Durchflusszytometrie-Protokoll mit Propidiumjodid-gefärbten Zellkernen geschätzt, wie zuvor beschrieben [77]. Neuralgewebe von gefrorenen (- 80 °C) adulten Proben von *H. vespertilio* und *H. euphorbiae* (von je einem Kopf; dieselben Individuen, die für die PacBio-Genom- und Hi-C-Sequenzierung verwendet wurden) und Neuralgewebe vom internen Referenzstandard *Acheta domesticus* (Weibchen, 1C = 2 Gb) wurden jeweils mit einer Rasierklinge in einer Petrischale mit 2 ml eiskaltem Galbraith-Puffer zerkleinert. Die Suspension wurde durch ein 42-µm-Nylonsieb gefiltert, dann mit dem interkalierenden Fluorochrom Propidiumjodid (PI, Thermo Fisher Scientific) gefärbt und mit Rnase A (Sigma-Aldrich) behandelt, jeweils in einer Endkonzentration von 25 µg/ml. Die mittlere rote PI-Fluoreszenz der gefärbten Zellkerne wurde mit einem Beckman-Coulter CytoFLEX-Durchflusszytometer mit einem Festkörperlaser, der bei 488 nm emittiert, quantifiziert. Die Fluoreszenzintensitäten von 10.000 Kernen pro Probe wurden aufgezeichnet. Anschließend wurden die Kernsuspensionen von *H. vespertilio* und *H. euphorbiae* jeweils mit der Kernsuspension des inter- nationalen Referenzstandards (siehe oben) gemischt und wiederum die Fluoreszenzintensitäten von 10.000 Kernen pro Mischprobe aufgezeichnet. Für die Histogramm-Analysen verwendeten wir CytExpert 2.3-Software. Die Gesamt-DNA-Menge in jeder Probe der beiden Hyles-Arten wurde berechnet als das Verhältnis des mittleren Fluoreszenzsignals des 2C-Peaks der gefärbten Zellkerne der jeweiligen Art geteilt durch das mittlere Fluoreszenzsignal des 2C-Peaks der gefärbten Zellkerne des Referenzstandards mal der 1C-DNA-Menge im Referenzstandard. Drei Wiederholungen, jeweils von demselben Individuum von *H. vespertilio* und *H. euphorbiae*, wurden an drei verschiedenen Tagen gemessen, um mögliche zufällige instrumentelle Fehler zu minimieren. Die Genomgröße wurde als 1C angegeben, die mittlere DNA-Menge in Mb in einem haploiden Zellkern.

PacBio-Genom-DNA und Sequenzierung

Kopfgewebe (38 mg) von einem einzelnen männlichen *Hyles* euphorbiae-Exemplar (dasselbe Einzelindividuum, das für die Schätzung der Genomgröße verwendet wurde) wurde für die Extraktion von DNA mit hohem Molekulargewicht unter Verwendung einer Anpassung des Protokolls von Miller et al. verwendet [78]. Die endgültige Reinheit und die Konzentrationen der DNA wurden mit dem NanoPhotometer (Implen GmbH, München, Deutschland) und dem Qubit Fluorometer (Thermo Fisher Scientific, Waltham, MA) gemessen. Eine SMRTbell-Bibliothek wurde gemäß den Anweisungen des SMRT- bell Express Prep kit v2.0 with Low DNA Input Protocol (Pacific Biosciences, Menlo Park, CA) erstellt. Die gesamte Input-DNA für die Bibliothek betrug 3 µg. Die Bibliothek wurde mit einer On-Plate-Konzentration von 80 pM mittels Diffusionsladung geladen. Ein SMRT-Zellsequenzierungslauf wurde auf dem Sequel System II im CCS-Modus unter Verwendung von 30 Stunden Filmzeit mit 2 Stunden Vorextension und Sequenzierungschemie V2.0 durchgeführt.

Genomaufbau von *Hyles euphorbiae*

Wir erstellten PacBio CCS-Reads aus der Datei *Hyles euphor- biae* subreads.bam mit dem ccs-Kommandozeilen-Tool von PacBio (Version 6.3.0, -all-Flag wurde verwendet). Wir erhielten 7,9 Gb hochwertiger CCS-Reads (HiFi-Reads, rq > 0,99) mit einer N50 von 11,74 kb. Um die Leseabdeckung weiter zu erhöhen, wendeten wir DeepConsensus (v0.2 mit Standardeinstellungen, actc v0.1.1) [79, 80] auf alle Reads des vorangegangenen CCS-Schrittes an und erhielten eine Gesamtausbeute von 8,8 Gb (N50: 11,83 kb). Wir führten HiFiasm (Version 0. 16.1-r375) [36] aus, um das Contig-Assembly zu erstellen. Verbleibende haplotypische Duplikationen im primären Contig-Set wurden mit den Standardparametern der purge-dups Pipeline (v.1.2.3) [81, 82] entfernt. Die Zusammenstellung wurde mit HiC-Daten unter Verwendung von yahs (v 1.1a) [83] gerüstet und manuell mit HiGlass [84] kuratiert. Verbleibende Lücken in den Gerüsten wurden durch Mapping der rohen PacBio-Subreads mit pbmm2 (Version 1.7.0) [85] gefüllt, und für Alignment-Piles, die die Lückenregionen mit 1000 Ankerbasen auf beiden Seiten vollständig überspannten, wurde mit gcpp (Version 2.0.2) [86] eine Konsensussequenz erzeugt. Die Konsensussequenz wurde nur dann zum Füllen einer Lücke verwendet, wenn: 1) die gesamte Konsensussequenz mindestens 5x abgedeckt war und 2) das Abdeckungsprofil der geschlossenen Lücken die Konsensussequenz vollständig unterstützte (d. h. es traten keine Alignment-Brüche oder große Wiederholungs-Alignment-Piles auf). Es wurden zwei Runden der Fehlerkorrektur durchgeführt, indem die DeepConsensus-Reads mit pbmm2 an der Assembly ausgerichtet, Varianten mit DeepVariant (Version 1.3.0) [87] aufgerufen und homozygote Fehler mit bcftools [88, 89] consensus korrigiert wurden. Die Assemblierung wurde mit Blobtoolkit (Version 1.1) [90] und einer internen Pipeline, die mehrere Blast-Datenbanken durchsucht, auf Kontaminationen überprüft. Die Bewertung der Vollständigkeit in Bezug auf Einzelkopie-Orthologe erfolgte mit Benchmarking Using Single Copy Ortho- logues (BUSCO) (Version 5.2.2) [91] zusammen mit dem lepidoptera_odb10-Set und den Optionen "-long -offline". und merqury (Version 1.3) [92]. Die QV-Werte wurden für die endgültigen Gerüste generiert (*H. euphorbiae* QV = 58,1).

Das mitochondriale Genom wurde mithilfe der mitoHifi-Pipeline (Version 2) [45, 46] auf der Grundlage von CCS-Reads und dem eng verwandten mitochondrialen Referenzgenom von *Reretra oldenlandiae* (NCBI-Zugang: MN885801.1) erstellt.

Hi-C Sequenzdaten

Schwalbenschwanz-Hi-C-Bibliotheken für *H. vespertilio* und *H. euphor biae* wurden aus Kopfgeweben (52,2 mg und 40,6 mg; dieselben Einzelindividuen, die für die PacBio-Genomsequenzierung und die Schätzung der Genomgröße verwendet wurden) unter Verwendung des Dovetail Hi-C-Kits (Dovetail Genomics, Scotts Valley, CA, USA) gemäß dem Protokoll des Herstellers (Version 1.4) für Insektenproben. Kurz gesagt, Chromatin wurde fixiert mit

Formaldehyd fixiert und dann extrahiert. Das fixierte Chromatin wurde mit DpnII verdaut, die 5'-Überhänge wurden mit biotinylierten Nukleotiden aufgefüllt und die freien stumpfen Enden wurden ligiert. Nach der Ligation wurden die Vernetzungen rückgängig gemacht, die assoziierten Proteine abgebaut und die DNA gereinigt. Die DNA wurde dann auf eine mittlere Fragmentgröße von ca. 350 bp geschert, und mit Illumina-kompatiblen Adaptern wurden Sequenzierungsbibliotheken erstellt. Biotinylierte Fragmente wurden vor der PCR-Amplifikation mit Streptavidin-Perlen eingefangen.

Die Fragmentgrößenverteilung und die Konzentration der endgültigen PacBio- und Dovetail-Hi-C-Bibliotheken wurden mit der TapeStation (Agilent Technologies) bzw. dem Qubit Fluo- rometer (Thermo Fisher Scientific, Waltham, MA) bewertet. Die Hi-C-Bibliotheken wurden auf einer NovaSeq 6000-Plattform bei Novogene (UK) sequenziert, wobei jeweils 100 Millionen 2×150 bp Paired-End-Reads mit einem Gesamtvolumen von 30 Gb erzeugt wurden.

Gerüstbau für den Zusammenbau von *H. vespertilio* mit HiRise

Die *H.* vespertilio-Eingangsassemblierung von Pippel et al. [12] und die *H. vespertilio* Dovetail Hi-C Library Reads wurden als Eingabedaten für HiRise verwendet, eine Software-Pipeline, die speziell für die Verwendung von Proximity-Ligationsdaten zum Scaf- fen von Genomassemblierungen entwickelt wurde [40]. Die Sequenzen der Dovetail-Hi-C-Bibliothek wurden mit einem modifizierten SNAP-Read-Mapper (http://snap.cs.berkeley. edu) an den Entwurf der Eingangsbaugruppe angeglichen. Die Trennungen der Hi-C-Lesepaare, die innerhalb der Entwurfsgerüste kartiert wurden, wurden mit HiRise Version 2.1.7 analysiert, um ein Likelihood-Modell für den genomischen Abstand zwischen den Lesepaaren zu erstellen. Das Modell wurde verwendet, um mutmaßliche Misjoins zu identifizieren und aufzulösen, potenzielle Joins zu bewerten und Joins oberhalb eines Schwellenwerts zu erstellen.

Anmerkung

Die strukturelle Annotation von proteinkodierenden Genen wurde mit TOGA [93] durchgeführt, einer Methode, die paarweise Genom-Alignment-Ketten zwischen einem annotierten Referenzgenom (hier *Manduca sexta*) und anderen Abfragespezies (hier *Hyles euphorbiae* und *H. vespertilio*) verwendet. Kurz gesagt, TOGA nutzt maschinelles Lernen, um orthologe Loci für jedes Referenztranskript abzuleiten, wobei das Konzept verwendet wird, dass orthologe Gene mehr Alignments zwischen intronischen und flankierenden intergenen Regionen aufweisen. TOGA projiziert dann jedes Referenztranskript zu seinem orthologen Abfragelokus unter Verwendung von CESAR 2.0 [94], einer Hidden-Markov-Modell-Methode, die Leseraster und Spleißstellenannotationen der Referenzexons berücksichtigt. CESAR vermeidet unerwünschte Frameshifts und ist in der Lage, evolutionäre Spleißstellenverschiebungen und präzise Intron-Deletionen zu erkennen [94, 95]. Anhand des CESAR-Alignments bestimmt TOGA, ob das Transkript inaktivierende Mutationen aufweist (Frameshifting-Mutationen, vorzeitige Stoppcodons, Mutationen, die die Spleißstelle stören, Deletionen ganzer kodierender Exons).

Um ganze Genom-Alignments als Input für TOGA zu generieren, haben wir die Assemblies von *H. euphorbiae* und *H. vespertilio* an das *Manduca sexta* (GCF_014839805.1) [37] Assembly angeglichen. Zum Vergleich der Kontiguität zwischen *H. euphorbiae*, *H. vespertilio*, *L. populi*, *Hemaris fuciformis*, *D. porcellus*, *Mimas tiliae* und *Manduca sexta*, wurde Quast 5.0.2 [96] verwendet. Die Genome wurden mit LASTZ (Version 1.04.03) aligniert [97] mit den Parametern (K = 2400, L = 3000, Y = 9400, H = 2000 und der Standard-Scoring-Matrix von Lastz). Dann verwendeten wir axtChain [47] (Standardparameter außer lin- earGap = loose), um kolineare Alignment-Ketten zu berechnen, RepeatFiller [98] (Standardparameter), um zuvor verpasste Alignments zwischen repetitiven Regionen zu erfassen, und chainCleaner [99] (Standardparameter außer minBro- kenChainScore = 75.000 und -doPairs), um die Alignment-Spezifität zu verbessern. *H. vespertilio* wurde als Referenz verwendet und *H. euphorbiae*, *B. mori* und *M. sexta* als Abfragen. Wir verwendeten die NCBI-Annotation (GCF_014839805.1_JHU_Msex_ v1.0_genomic.gff.gz) als Eingabe für TOGA.

Das mitochondriale Genom von *H. euphorbiae* wurde mit dem MITOS WebServer [100] annotiert und das Ergebnis mit shinyCircos [101] dargestellt. Diese Software wurde auch für die CIRCOS-Plots der alignierten Genome verwendet.

Genom-Alignments

Mit dem Ziel, die Chromosomen nach ihrer Homologie zu *B. mori* zu benennen, wurde die Nukleotidhomologie auf der Grundlage des Sequenzalignments von *H. vespertilio*, *H. euphorbiae* und *M. sexta* Contigs mittels Scaffold Chaining als primärem Werkzeug zur Entdeckung von Ähnlichkeiten unter Verwendung von LASTZ bewertet [97]. Die Visualisierung erfolgte mithilfe von CIRCOS-Plots (Abb. 4) unter Verwendung der Positionsinformationen der Syntenie-Blöcke (in den bed-Dateien) der Genom-Alignments als Input (siehe Tabelle S3).

Die Genomzusammensetzungen von *Hyles euphorbiae* und *H. vespertilio* (mit Anmerkungen und Wahrscheinlichkeiten für Modifizierungs-/Methylierungsstellen sowie Genom-Alignments, siehe unten) sind im Senckenberg Genome Browser zugänglich (https:// genome.senckenberg.de/cgi-bin/hgTracks?db=HLhylEup1; https://genome.senckenberg.de/cgi-bin/hgTracks?db= HLhylVes2). Ein Genom-Alignment von *H. vespertilio* mit *B. mori* (GCF_014905235.1 [41]) wurde wie oben beschrieben erstellt und verwendet, um Chromosomen-Homologien zu postulieren und Chromosomen entsprechend zu benennen. Detaillierte Werte für die Anteile homolo- gischer Regionen pro Chromosom sind in Tabelle S1 aufgeführt.

Die Genome von *Mimas tiliae* [14], *Laothoe populi* [15], *Hemaris fuciformis* [38] und *Deilephila porcellus* [39] (dies sind die zusätzlichen Genome der Sphingidae, die am 14.10.2022 verfügbar waren) wurden für statistische Vergleiche und einen phylogenetischen Vergleich der Wiederholungsentwicklung herangezogen.

Analysen des Repeat- und Transposongehalts

Vergleichende Annotation von Wiederholungen bei *H. vespertilio*, *H. euphorbiae*, *D. porcellus*, *Mimas tiliae*, *L. populi*, *H.*

fuciformis und *Manduca sexta* wurde mit RepeatModeler (Version 2.0.3) [102] und RepeatMasker (Version 4.1.2-p1) [103], wie in Dfam TETools Docker-Container (Version 1.85) [104], in Kombination mit rmblastn (Version 2.11.0 +). Die Repeat-Bibliothek für RepeatMasker enthält eine Kombination aller mit Repeat-Modeler identifizierten Repeat-Familien, die mithilfe von Cdhit (Version 4.8.1) [105] und manueller Kuration der Modelle gemäß den in [106] beschriebenen Richtlinien kombiniert und nicht redundant gemacht wurden. Das Vorkommen verschiedener Familien transposabler Elemente wurde auf den jeweiligen Ästen eines Baumes dargestellt, der von Hand auf der Grundlage von Kawahara und Barber [44] gezeichnet wurde, und zwar in relativer Reihenfolge entsprechend der Assemblierungsdivergenz und der Sequenzmenge, wie sie von einem Hilfsskript von repeatmasker (calcDivergenceFromAlign.pl) bereitgestellt wurden. TE-Familien mit höherer Sequenzdiversität (vermutlich ältere Transposon-Bursts) werden weiter links auf den Zweigen eingezeichnet, TE-Familien, die zwischen Taxa geteilt werden (und eine hohe Sequenzdiversität aufweisen), werden auf dem Zweig ihrer gemeinsamen Vorfahren eingezeichnet.

Methylierungsprofil

Für das Methylierungsprofil wurden die PacBio HiFi-Reads von *H. euphorbiae*, die Tags für die Modifikation von 5mC-Basen enthalten, mit pb-CpG-tools (Version 1.1) [107] in CpG/5mC-Daten umgewandelt und in den Genombrowser hochgeladen (s.o.). Stellen mit einer Änderungswahrscheinlichkeit von < 50 % sind blau und > 50 % rot markiert.

Analyse der Chromosomenstruktur

Um die Chromosomenstruktur zwischen den Arten zu vergleichen, wurden mit Hilfe von BUCSOs (Version 5.4) ("lepidoptera odb 10"-Datensatz, Metaeuk-Modus) [108] Orthologe mit einfacher Kopie abgeleitet. Orthologe wurden gefiltert, um diejenigen zu erhalten, die einem Merian-Element zugeordnet werden können, wobei die Tabelle der Orthologenzuordnungen zu Merian-Elementen von Wright et al. verwendet wurde [42]. Anschließend wurden nicht lokalisierte Gerüste entfernt, indem nur Gerüste mit drei oder mehr Einzelkopie-Orthologen beibehalten wurden. Schließlich wurde die Syntenie zwischen den Genompaaren mit Hilfe von Oxford-Plots visualisiert, die mit Hilfe eines eigenen Codes [109] die relative Position jedes Orthologs entlang der Chromosomen eines jeden Genoms darstellen. In den sich ergebenden Oxford-Plots wurden die Orthologe nach der Identität des Merian-Elements eingefärbt, um Chromosomen zu identifizieren, die aus Fusionsereignissen resultieren. Chromosomen, die aus Fusionsereignissen resultieren, wurden auch überprüft, indem https://github.com/charlottewright/lep_fusion_fission_finder mit einer Fenstergröße von 17 ausgeführt wurde. Dies identifiziert Chromosomen, die das Produkt von Fusionsereignissen sind, basierend auf dem Vorhandensein von Orthologen, die mehr als einem Merian-Element innerhalb eines einzelnen Chromosoms entsprechen.

Flügelmuster-Gene

Die Positionen der Flügelmustergene *optix, WntA*, *cortex, aristaless* und *Distal-less* (siehe Tabelle S2 für die Hinterlegungsnummern der Referenzen) wurden in den beiden Hyles-Genomen mit Hilfe des Online-BLAT-Tools [47, 48] mit den im Senckenberg-Genombrowser implementierten Standardoptionen ermittelt. Die BLAT-Ergebnisse werden nach Alignement-Länge sortiert dargestellt, wobei für jedes Gen das längste ausgewählt wurde.

Für das Cortex-Gen haben wir die Hyles-Daten zusätzlich mit den Sequenzen von *Biston betularia* (Geom- etridae; KT182637) verglichen, bei dem die übliche helle (*typica*) Form während der industriellen Revolution durch eine bisher unbekannte schwarze (*car- bonaria*) Form ersetzt wurde, die durch die Wechselwirkung zwischen Vogelprädation und Rauchverschmutzung [110] und durch eine Transposon-Insertion in einem Cortex-Intron verursacht wurde [32, 111].

Ergänzende Informationen

Die Online-Version enthält zusätzliches Material, das unter https://doi.org/10.1186/s12864-023-09506-y verfügbar ist.

A dditional file 1: Abb. S1. Das annotierte mitochondriale Genom von *Hyles euphorbiae*. **Abb. S2.** BLAT-Ergebnis der *Heliconius* optix-Proteinsequenz (KC469894) aus dem Genombrowser auf Chromosom 23. das Hyles-Genom-Alignment, die Annotation und das Vorkommen von Wiederholungen in der Nähe bei a) *H. euphorbiae* (invertiert) und b) *H. vespertilio*. **Abb. S3.** BLAT-Ergebnis der *Heliconius* optix-Proteinsequenz (KC469894) aus dem Genombrowser auf Chromosom 23 in a) *Hyles euphorbiae* (umgekehrt) und b) *Hyles vespertilio* in einer 700 kb-Ansicht. **Abb. S4.** BLAT-Ergebnis der WntA-Protein-Sequenz (AFC75686) von *Heliconius* aus dem Genom-Browser, das das Hyles-Genom-Alignment, die Annotation und das Vorkommen von Wiederholungen in der Nähe in a) *H. euphorbiae* (Rückseite) und b) *H. vespertilio*. **Abb. S5.** BLAT-Ergebnis aus dem Genombrowser, das die Exon/Intron-Strukturen des Cortex-Proteins von *Manduca sexta* und *Biston betularia* (Tabelle S2) auf Chromosom 17 in
a) *Hyles euphorbiae* und b) *H. vespertilio*. **Abb. S6.** BLAT-Ergebnis der *Manduca sexta* mRNA der *Rinde* (Tabelle S2) auf Chromosom 17 in a) *Hyles euphorbiae* und b) *Hyles vespertilio* in einer 100 kb-Ansicht. **Abb. S7.** BLAT-Ergebnis von *Manduca sexta* mRNAs von *aristaless* (Tabelle 2) auf Chromosom 4 in a) *H. euphorbiae* und b) *H. vespertilio*. **Abb. S8.** BLAT-Ergebnis von *Manduca sexta* mRNAs von *aristaless (Tabelle 2*) auf Chromosom 4 in a) *H. euphorbiae* 100 kb Ansicht, b)
H. vespertilio 150 kb Ansicht (beachten Sie die unterschiedlichen Maßstäbe). **Abb. S9.** BLAT-Ergebnis der *Manduca sexta* mRNA von *Distal-less* (AY616435.1) auf Chromosom 2 in a) *Hyles euphorbiae* (invertiert) und b) *Hyles vespertilio*. **Abb. S10.** BLAT-Ergebnis von *Manduca sexta* mRNA von *Distal-less* (AY616435.1) auf Chromosom 2 in a) *Hyles euphorbiae* (invertiert) und b) *Hyles vespertilio* in einer 100 kb-Ansicht.

Zusätzliche Datei 2: Tabelle S1. Chromosomenanteilswerte von der *B. mori*
- *H. vespertilio* Ausrichtung. Chr. = Chromosom.

Zusätzliche Datei 3: Tabelle S2. Zugriffsnummern von Referenz-Flügelmuster-Sequenzen und Datentyp, die für blat in dem Genombrowser verwendet wurden. *Manduca sexta WntA* und *optix* werden im NCBI-Webportal nicht angezeigt.

Zusätzliche Datei 4: Tabelle S3. Chromosomenhomologie, Gerüstnamen und Hinterlegungsnummern von *B. mori*, *M. sexta*, *H. euphorbiae* und *H. vesper- tilio*, mit Chromosomengrößen basierend auf bioinformatisch ermittelten Assemblierungslängen.

Danksagung

Diese Studie profitierte von der gemeinsamen Nutzung von Fachwissen im Rahmen des DFG-Schwerpunktprogramms SPP 1991 Taxon-Omics. Dovetail Genomics stellte den HiRise Gerüstservice zur Verfügung. Unser Dank geht an Alexander Ben Hamadou (LOEWE-TBG) für die DNA-Isolierung und die Bibliotheksvorbereitung. Wir danken Ian J. Kitching für sprachliche Korrekturen und hilfreiche Vorschläge sowie Markward H. Fischer für die Verbesserung zahlreicher Abbildungen. Wir danken Paolo Mazzei für die Bereitstellung lebender H. euphorbiae aus Griechenland.

Beiträge der Autoren
AKH hat die Studie und die Partnerschaften konzipiert, die Feldarbeit durchgeführt, die
Analysen, erstellte und bearbeitete Abbildungen und schrieb den Text. TS unterstützte den Zusammenbau, das Scaffolding, die Kommentierung, erstellte Abbildungen und schrieb und redigierte den Text. FP, CW & LP schlugen Analysen vor und führten sie durch, erstellten Abbildungen, und schrieb den Text. AY & FM steuerten Karyotyp-Analysen mit Abbildungen bei
und Text (mit zusätzlicher Bearbeitung des gesamten T e x t e s). HD trug zur Feld- und Laborarbeit bei und SW & CG übernahmen die Verantwortung für die DNA-Isolierung.
tion, Sequenzierung und steuerte Text bei. MH lieferte Bioinformatik Infrastruktur und Überwachung, Skripte, Zwischenergebnisse, Beratung, und durchgeführte Alignments, MP durchgeführte Montage, Gerüstbau, Numerik Analysen und lieferte die Grundlage für die Abbildungen. Alle Autoren haben das endgültige Manuskript gelesen und genehmigt.

Finanzierung
Open-Access-Förderung ermöglicht und organisiert durch Projekt DEAL.
Diese Studie wurde durch Fördermittel der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Schwerpunktprogramms SPP 1991: Taxon-OMICS (HU 1561/5-2) unterstützt.
MP wurde teilweise vom BMBF finanziert (Zuschuss 01IS18026C). AY und FM danken für die Unterstützung durch das Stipendium 20-13784S der Czech Science Foundation. CJW wurde vom Wellcome Trust finanziert (Zuschüsse 206.194 und 218.328). Der HiRise-Gerüstservice von Dovetail Genomics für *H. vespertilio* wurde von LOEWE-TBG finanziert.

Verfügbarkeit von Daten und Materialien
Die im Rahmen der aktuellen Studie erzeugten und analysierten Daten sind im NCBI-Repository verfügbar
[https://www.ncbi.nlm.nih.gov/genome/?term=hyles]. Die SRA-Zugangsnummer für *H. euphorbiae* PacBio Sequel HiFi ist SRX13604162, und für die Dovetail Hi-C Daten ist SRX14310646. Das assemblierte Genom ist unter GCA_023078785.2 zugänglich, und die Zugriffsnummern sind im F o l g e n d e n aufgeführt. Mitochondrium: CM041059, Chromosom Z: CM041058 und Chromosomen 2-29: CM041030-CM041057).
Die SRA-Zugangsnummer für die Hi-C-Rohdaten von Dovetail *Hyles vespertilio* lautet SRX14530528. Die neue Zusammenstellung, die in dieser Arbeit vorgestellt wird, ist unter GCA_009982885.2 zugänglich, und die Zugangsnummern sind im Einzelnen aufgeführt
im Folgenden. Chromosom Z: CM042833 und Chromosomen 2-29: CM042805-CM042832.

Erklärungen

Ethische Genehmigung und Zustimmung zur Teilnahme
Das Individuum der geschützten Art ***H. euphorbiae*** wurde mit einer vom Landratsamt Meißen, Kreisumweltamt, Untere Naturschutzbehörde, erteilten Ausnahmegenehmigung (AZ 364.621-377/2009-4026/2017-45577/2021) zum Zwecke der wissenschaftlichen Untersuchung gesammelt.

Zustimmung zur Veröffentlichung
Nicht anwendbar.

Konkurrierende Interessen ests
Die Autoren erklären, dass sie keine konkurrierenden Interessen haben.

Angaben zum Autor
[1]Senckenberg Naturhistorische Sammlungen Dresden, Königsbrücker Landstr. 159, 01109 Dresden, Deutschland. [2]LOEWE-Zentrum für Translationale Biodiversitätsgenomik (LOEWE-TBG), Frankfurt am Main, Deutschland. [3]Tree of Life, Wellcome Sanger Institute, Cambridge CB10 1SA, Großbritannien. [4]Biologiezentrum der Tschechischen Akademie der Wissenschaften, Institut für Entomologie, Branišovská 31, 370 05 České Budějovice, Tschechische Republik. [5]Max-Planck-Institut für Molekulare Zellbiologie und Genetik, Pfotenhauerstraße 108, 01307 Dresden, Deutschland. [6]Zentrum
für Molekulare Biodiversitätsforschung, Leibniz-Institut zur Analyse des Biodiversitätswandels, Adenauerallee 127, 53113 Bonn, Deutschland. [7]Zentrum für Systembiologie Dresden, Pfotenhauerstr. 108, 01307 Dresden, Deutschland. [8]Abteilung für Zell- und Molekularbiologie, Universität Uppsala, Husargatan 3, Uppsala 751 23, Schweden.

Eingegangen: 15. März 2023 Angenommen: 5. Juli 2023
Published online: 07 August 2023

Referenzen

1. Hundsdoerfer AK, Päckert M, Kehlmaier C, Strutzenberger P, Kitching IJ. Museum Archive revisited: Zentralasiatische Habichtskäfer zeigen eine außergewöhnlich hohe Artenvielfalt im späten Pliozän (Lepidoptera, Sphingidae). Zoolog Scr. 2017;46(5):552-70.
2. Hundsdoerfer AK, Kitching IJ, Wink M. A molecular phylogeny of the hawkmoth genus Hyles (Lepidoptera: Sphingidae, Macroglossinae). Mol Phylogenet Evol. 2005;35:442-58.
3. Hundsdoerfer AK, Lee KM, Kitching IJ, Mutanen M. Genome-wide SNP data reveal an overestimation of species diversity in a group of hawkmoths. Genome Biol Evol. 2019;11(8):2136-50.
4. Hundsdoerfer AK, Tshibangu JN, Wetterauer B, Wink M. Sequestrierung von Phorbolestern durch aposematische Larven von Hyles euphorbiae (Lepidoptera: Sphingidae)? Chemoökologie. 2005;15(4):261-7.
5. Hundsdoerfer AK, Rubinoff D, Attié M, Kitching IJ, Wink M. A revised molecular phylogeny of the globally distributed hawkmoth genus Hyles (Lepidoptera: Sphingidae), based on mitochondrial and nuclear DNA sequences. Mol Phylogenet Evol. 2009;52:852-65.
6. Hundsdoerfer AK, Buchwalder K, O'Neill MA, Dobler S. Chemical ecology traits in an adaptive radiation: TPA-Empfindlichkeit und Entgiftung bei Larven von Hyles und Hippotion (Sphingidae, Lepidoptera). Chemoecology. 2019;29(1):35-47.
7. Mende MB, Bartel M, Hundsdoerfer AK. A comprehensive phylogeography of the Hyles euphorbiae complex (Lepidoptera: Sphingidae) indicates a "glacial refuge belt." Sci Rep. 2016;6:29527.
8. Les Sphingidae de France [http://sphingidae-haxaire.com/index.php/ macroglossinae-2/hyles-vespertilio/].
9. Hundsdoerfer AK, Kitching IJ. Morphologische Entwicklung bei Hyles Hübner, 1819 Habichtskäfern (Lepidoptera, Sphingidae): Rekonstruktion des ursprünglichen Hyles-Habitus tral. Nota Lepidopterologica. 2020;43:181.
10. Danner F, Eitschberger U, Surholt B. Die Schwärmer der westlichen Palaearktis. Bausteine zu einer Revision (Lepidoptera: Sphingidae). Herbipoliana. 1998;4:1-368 (Textband), 361-720 (Tafelband).
11. Hundsdoerfer AK, Kitching IJ. Ancient incomplete lineage sorting of Hyles and Rhodafra (Lepidoptera: Sphingidae). Org Divers Evol. 2020;20(3):527-36.
12. Pippel M, Jebb D, Patzold F, Winkler S, Vogel H, Myers G, Hiller M, Hundsdoerfer AK. Ein hochgradig zusammenhängendes Genom der Fledermausschwärmer Hyles vespertilio (Lepidoptera: Sphingidae). GigaScience. 2020;9(1):giaa001.
13. Patzold F, Zilli A, Hundsdoerfer AK. Vorteile einer einfach anzuwendenden DNA-Extraktionsmethode für die minimaldestruktive Analyse von Sammlungsspezi- menten. PLoS ONE. 2020;15(7): e0235222.
14. Boyes D, Holland P, University of Oxford and Wytham Woods Genome Acquisition Lab, et al. The genome sequence of the lime hawk-moth, Mimas tiliae (Linnaeus, 1758). Wellcome Open Res. 2021;6:357.
15. Boyes D, Holland PW, Konsortium DToL. Die Genomsequenz des Pappelschwärmers, Laothoe populi (Linnaeus, 1758). Wellcome Open Res. 2021;6:237 [version 1 peer review: 1 approved, 1 approved with reservations].
16. Beldade P, Brakefield PM. Die Genetik und Evo-Devo von Schmetterlingsflügeln patterns. Nat Rev Genet. 2002;3(6):442-52.
17. Nijhout HF. Elemente der Flügelmuster von Schmetterlingen. J Exp Zool. 2001;291(3):213-25.
18. Beldade P, Monteiro A. Öko-Evo-Devo-Fortschritte mit Schmetterlingsaugenflecken. Curr Opin Genet Dev. 2021;69:6-13.
19. Monteiro A, Glaser G, Stockslager S, Glansdorp N, Ramos D. Comparative insights into questions of lepidopteran wing pattern homology. BMC Dev Biol. 2006;6:52.
20. Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Sup- ple MA, Hanly JJ, Mallet J, Lewis JJ, Hines HM, et al. Complex modular architecture around a simple toolkit of wing pattern genes. Nat Ecol Evol. 2017;1(3):52.
21. Dhungel B, Ohno Y, Matayoshi R, Iwasaki M, Taira W, Adhikari K, Gurung R, Otaki JM. distal-less induziert elementare Farbmuster in Junonia butterfly Flügeln. Zool Lett. 2016;2:4.
22. Connahs H, Rhen T, Simmons RB. Transkriptomanalyse des Schmetterlings Vanessa cardui während der Entwicklung des Flügelmusters. BMC Genomics. 2016;17:270.
23. Martin A, Reed RD. wingless und aristaless2 definieren einen Entwicklungsplan für die Evolution der Flügelmuster von Motten und Schmetterlingen. Mol Biol Evol. 2010;27(12):2864-78.
24. Stevens M, Merilaita S. Animal camouflage: current issues and new perspectives. Philos Trans R Soc B Biol Sci. 2009;364(1516):423-7.
25. Hiyama A, Taira W, Otaki JM. Farbmusterevolution als Reaktion auf Umweltstress bei Schmetterlingen. Front Genet. 2012;3:15.
26. Bastide H, Saenko SV, Chouteau M, Joron M, Llaurens V. Dominanzmechanismen in Supergen-Allelen, die die Variation des Flügelmusters von Schmetterlingen steuern: Erkenntnisse aus der Genexpression bei Heliconius numata. Heredity. 2023;130(2):92-8.
27. Jay P, Leroy M, Le Poul Y, Whibley A, Arias M, Chouteau M, Joron M. Asso- ciation mapping of colour variation in a butterfly provides evidence that a supergene locks together a cluster of adaptive loci. Philos Trans R Soc Biol Sci. 1856;2022:377.
28. Zhang L, Mazo-Vargas A, Reed RD. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. Proc Natl Acad Sci U S A. 2017;114(40):10707-12.
29. Livraghi L, Martin A, Gibbs M, Braak N, Arif S, Breuker CJ. CRISPR/Cas9 as the key to unlocking the secrets of butterfly wing pattern development and its evolution. Adv Insect Physiol. 2018;54:85-115 Elsevier.
30. Jiggins CD, Wallbank RW, Hanly JJ. Waiting in the wings: what can we learn about gene co-option from the diversification of butterfly wing patterns? Philos Trans R Soc B Biol Sci. 2017;372(1713):20150485.
31. Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko SV, Wallbank RW, Wu GC, Maroja L, Ferguson L, Hanly JJ, et al. The gene cortex controls mim- icry and crypsis in butterflies and moths. Nature. 2016;534(7605):106–10.
32. van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. Die industrielle Melanismus-Mutation in britischen Nachtfaltern ist ein transponierbares Element. Nature. 2016;534(7605):102–5.
33. Brunetti CR, Selegue JE, Monteiro A, French V, Brakefield PM, Carroll SB. Die Entstehung und Diversifizierung der Farbmuster von Schmetterlingsaugenflecken. Curr Biol. 2001;11(20):1578-85.
34. Westerman EL, VanKuren NW, Massardo D, Tenger-Trolander A, Zhang W, Hill RI, Perry M, Bayala E, Barr K, Chamberlain N, et al. aristaless con- trols butterfly wing color variation used in mimicry and mate choice. Curr Biol. 2018;28(21):3469.
35. Bayala EX, VanKuren N, Massardo D, Kronforst MR. aristaless1 has a dual role in appendage formation and wing color specification during but- terfly development. BMC Biol. 2023;21:1-19.
36. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18(2):170-5.
37. Gershman A, Romer TG, Fan Y, Razaghi R, Smith WA, Timp W. De novo genome assembly of the tobacco hornworm moth (Manduca sexta). G3. 2021;11(1):1-9.
38. Mulhair PO, Holland PW. Die Evolution des Hox-Genclusters bei Insekten: V e r g l e i c h e n d e Analyse über 243 Arten. Semin Cell Dev Biol. 2022. https:// doi.org/10.1016/j.semcdb.2022.11.010.
39. Boyes D, University of Oxford and Wytham Woods Genome Acquisition Lab LDarwin Tree of Life Barcoding collective, et al. The genome sequence of the small elephant hawk moth, Deilephila porcellus (Linnaeus, 1758). Wellcome Open Res. 2022;7:80.
40. Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW. Chromosome scale shotgun assembly using an in vitro method for long-range linkage. Genome Res. 2016;26(3):342-50.
41. Kawamoto M, Jouraku A, Toyoda A,Yokoi K, MinakuchiY, Katsuma S, Fujiyama A, Kiuchi T, Yamamoto K, Shimada T. High quality genome assembly of the silkworm, Bombyx mori. Insect Biochem Mol Biol. 2019;107:53-62.
42. Wright CJ, Stevens L, Mackintosh A, Lawniczack M, Blaxter M. Chromosome evolution in Lepidoptera. bioRxiv. 2023. https://doi.org/10.1101/ 2023.05.12.540473.
43. Yasukochi Y, Tanaka-Okuyama M, Shibata F, Yoshido A, Marec F, Wu C, Zhang H, Goldsmith MR, Sahara K. Extensive conserved synteny of genes between the karyotypes of Manduca sexta and Bombyx mori revealed by BAC-FISH mapping. PLoS ONE. 2009;4(10): e7465.
44. Kawahara AY, Barber JR. Tempo und Art der Ultraschall-Antikörperbildung und Sonar-Störungen bei den verschiedenen Habichtskäferarten. Proc Natl Acad Sci USA. 2015;112(20):6407–12.
45. Allio R, Schomaker-Bastos A, Romiguier J, Prosdocimi F, Nabholz B, Del- suc F. MitoFinder: Effiziente, automatisierte, groß angelegte Extraktion von mitog- enomischen Daten in der Zielanreicherung der Phylogenomik. Mol Ecol Resour. 2020;20(4):892-905. https://github.com/marcelauliano/MitoHiFi.

46. Uliano-Silva M, Gabriel JR, Ferreira N, Krasheninnikova K. Darwin Tree of Life Consortium, Giulio Formenti, Linelle Abueg, James Torrance, Eugene W. Myers, Richard Durbin, Mark Blaxter, Shane A. McCarthy bioRxiv. 2022.12.23.521667. https://doi.org/10.1101/2022.12.23.521667. Zugriff am 1. Februar 2023.
47. Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. Proc Natl Acad Sci USA. 2003;100(20):11484–9.
48. Lee BT, Barber GP, Benet-Pages A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee CM, et al. The UCSC Genome Browser database: 2022 update. Nucleic Acids Res. 2022;50(D1):D1115-22.
49. Sofroniew N, Lambert T, Evans K, Nunez-Iglesias J, Bokota G, Winston P, Bragantini J. napari: a multi-dimensional image viewer for Python. Zenodo. 2022. https://doi.org/10.5281/zenodo. 3555620.
50. Hundsdoerfer AK, Kitching IJ, Wink M. The phylogeny of the Hyles euphorbiae-complex (Lepidoptera: Sphingidae): molecular evidence from sequence data and ISSR-PCR fingerprints. Org Divers Evol. 2005;5:173-98.
51. Hanrahan SJ, Johnston JS. Neue Schätzungen der Genomgröße von 134 Arten von Arthropoden. Chromosome Res. 2011;19(6):809-23.
52. d'Alencon E, Sezutsu H, Legeai F, Permal E, Bernard-Samain S, Gimenez S, Gagneur C, Cousserans F, Shimomura M, Brun-Barale A, et al. Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. Proc Natl Acad Sci U S A. 2010;107(17):7680-5.
53. Lavoie CA, Platt RN, Novick PA, Counterman BA, Ray DA. Die Entwicklung transponierbarer Elemente bei Heliconius deutet auf eine Genomvielfalt innerhalb der Lepidoptera hin. Mob DNA. 2013;4(1):21.
54. Lynch M, Walsh B. Die Ursprünge der Genomarchitektur, Bd. 98. Sunder- land, MA: Sinauer Associates; 2007.
55. Shah A, Hoffman JI, Schielzeth H. Comparative analysis of genomic repeat content in Gomphocerine grasshoppers reveals expansion of satellite DNA and helitrons in species with unusually large genomes. Genome Biol Evol. 2020;12(7):1180-93.
56. Hollister JD, Gaut BS. Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. 2009;19(8):1419-28.
57. Charlesworth B, Sniegowski P, Stephan W. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature. 1994;371(6494):215–20.
58. Talla V, Suh A, Kalsoom F, Dinca V, Vila R, Friberg M, Wiklund C, Backstrom N. Rapid increase in genome size as a consequence of transposable element hyperactivity in Wood-White (Leptidea) Butterflies. Genome Biol Evol. 2017;9(10):2491-505.
59. Ellegren H, Smeds L, Burri R, Olason PI, Backstrom N, Kawakami T, Kunstner A, Makinen H, Nadachowska-Brzyska K, Qvarnstrom A, et al. The genomic landscape of species divergence in Ficedula flycatchers. Nature. 2012;491(7426):756–60.
60. Feliciello I, Akrap I, Brajkovic J, Zlatar I, Ugarkovic D. Satellite DNA as a driver of population divergence in the red flour beetle Tribolium castaneum. Genome Biol Evol. 2014;7(1):228-39.
61. Maumus F, Fiston-Lavier AS, Quesneville H. Impact of transposable elements on insect genomes and biology. Curr Opin Insect Sci. 2015;7:30-6.
62. Pittaway AR. Die Habichtskäfer der westlichen Paläarktis. Colchester: Harley Books; 1993.
63. Brower AVZ. Rapid morphological radiation and convergence among races of the butterfly Heliconius erato inferred from patterns of mito- chondrial DNA evolution. Proc Natl Acad Sci U S A. 1994;91:6491-5.
64. Pringle EG, Baxter SW, Webster CL, Papanicolaou A, Lee SF, Jiggins CD. Syntenie und Chromosomenevolution bei den Lepidoptera: Beweise aus der Kartierung bei *Heliconius melpomene*. Genetics. 2007;177(1):417-26.
65. Yasukochi Y, Ashakumary LA, Baba K, Yoshido A, Sahara K. A second-generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. Genetics. 2006;173(3):1319-28.
66. Ahola VLR, Somervuo P, Salmela L, Koskinen P, Rastas P, et al. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. Nat Commun. 2014;5:4737.
67. Sahara K, Yoshido A, Shibata F, Fujikawa-Kojima N, Okabe T, Tanaka-Okuyama M, Yasukochi Y. FISH-Identifizierung von Helicoverpa armigera und Mamestra brassicae Chromosomen durch BAC- und Fosmid-Sonden. Insect Biochem Mol Biol. 2013;43(8):644-53.
68. van't Hof AE, Nguyen P, Dalíková M, Edmonds N, Marec F, Saccheri IJ. Linkage map of the peppered moth, Biston betularia (Lepi- doptera, Geometridae): a model of industrial melanism. Heredity. 2013;110(3):283-95.
69. Yoshido A, Bando H, Yasukochi Y, Sahara K. The *Bombyx mori* karyotype and the assignment of linkage group. Genetics. 2005;170:675-85.
70. Jiggins CD. Was können wir aus der Genetik der Flügelmuster von Heliconius-Schmetterlingen über Anpassung lernen? In: Diversity and evolution of but- terfly wing patterns. Singapore: Springer; 2017. p. 173-88.
71. Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, Jiggins CD, Chamberlain NL, Kronforst MR, Chen R, et al. optix treibt die wiederholte konvergente Evolution der Schmetterlingsflügelmimikry an. Science. 2011;333(6046):1137–41.
72. Maddison WP. Genbäume in Artenbäumen. Syst Biol. 1997;46(3):523-36.
73. Yoshido A, Sahara K, Yasukochi Y. Seidenspinner (Lepidoptera). In: Sharakhov IV, editor. Protocols for cytogenetic mapping of arthropod genomes. Boca Ranton, FL, USA: CRC Press; 2015. p. 219-56.
74. Yoshido A, Marec F, Sahara K. Resolution of sex chromosome constitution by genomic in situ hybridization and fluorescence in situ hybridization with (TTAGG)(n) telomeric probe in some species of Lepidoptera. Chromosoma. 2005;114(3):193-202.
75. Sahara K, Marec F, Traut W. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. Chromosome Res. 1999;7(6):449-60.
76. Team TGD: GIMP. Abgerufen von https://www.gimp.org. In:; 2019.
77. Hare EE, Johnston JS. Bestimmung der Genomgröße mittels Durchflusszytometrie von Propidiumjodid-gefärbten Zellkernen. In: Molecular methods for evolu- tionary genetics. Springer; 2012. p. 3-12.
78. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215.
79. Ausrichten von Subreads an ccs-Reads [https://github.com/PacificBiosciences/actc].
80. Baid G, Cook DE, Shafin K, Yun T, Llinares-Lopez F, Berthet Q, Wenger AM, Rowell WJ, Nattestad M, Yang H. DeepConsensus: gap- aware sequence transformers for sequence correction. bioRxiv. 2021;41:232.
81. Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R. Identifizierung und Entfernung haplotypischer Duplikationen in primären Genomassemblies. Bioinformatics. 2020;36(9):2896-8.
82. Werkzeug zur Identifizierung haplotypischer Duplikationen [https://github.com/dfguan/ purge_dups#step-1-run-minimap2-to-align-pacbio-data-and-generate-paf-files-then-calculate-read-depth-histogram-and-base-level-read-depth-commands-are-as-follows].
83. Zhou C, McCarthy SA, Durbin R. YaHS: yet another Hi-C scaffolding tool. bioRxiv. 2022;39:btac808.
84. Kerpedjiev P, Abdennur N. Lekschas Fea: HiGlass: webbasierte visuelle Exploration und Analyse von Genom-Interaktionskarten. Genome Biol. 2018;19:125.
85. Ein minimap2-Frontend für native PacBio-Datenformate [https://github. com/PacificBiosciences/pbmm2].
86. GCpp Generieren Sie hochpräzise Referenz-Contigs.
87. Poplin R, Chang P-C, Alexander D, Schwartz S, Colthurst T, Ku A, Newburger D, Dijamco J, Nguyen N, Afshar PT. Ein universeller SNP- und Small-Insel-Varianten-Caller mit tiefen neuronalen Netzen. Nat Biotechnol. 2018;36(10):983-7.
88. bcftools - Dienstprogramme für den Variantenaufruf und die Bearbeitung von VCFs und BCFs [https://samtools.github.io/bcftools/bcftools.html].
89. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abe- casis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
90. Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit-interaktive Qualitätsbewertung von Genomassemblies. G3 Genes Genomes Genet. 2020;10:1361-74.
91. Manni M, Berkeley MR, Seppey M, Simao FA, Zdobnov EM. BUSCO Update: Neue und optimierte Arbeitsabläufe sowie eine breitere und tiefere phylogenetische Abdeckung für die Auswertung von eukaryotischen, prokaryotischen, und viralen Genomen. Mol Biol Evol. 2021;38(10):4647-54.
92. Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: referenzfreie Qualitäts-, Vollständigkeits- und Phasenbewertung für Genomassemblies. Genome Biol. 2020;21(1):1-27.
93. Kirilenko B, Munegowda C, Osipova E, Jebb D, Sharma V, Blumer

Mo, Ellen et al., Mitchell A, Montopoulos J, Rogers L, et al. Integrat- ing gene annotation with orthology inference at scale. Wissenschaft. 2023;380(6643):eabn3107.

94. Sharma V, Schwede P, Hiller M. CESAR 2.0 verbessert die Geschwindigkeit und Genauigkeit der vergleichenden Genannotation erheblich. Bioinformatics. 2017;33(24):3985-7.
95. Sharma V, Elghafari A, Hiller M. Coding exon-structure aware realigner (CESAR) utilizes genome alignments for accurate comparative gene annotation. Nucleic Acids Res. 2016;44(11): e103.
96. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5.
97. Harris RS. Verbesserte paarweise Ausrichtung genomischer DNA. Eine Dissertation in Informatik und Ingenieurwissenschaften. The Pennsylvania State University; 2007.
98. Osipova E, Hecker N, Hiller M. RepeatFiller identifiziert neu Megabasen von ausgerichteten repetitiven Sequenzen und verbessert Annotationen von kon- servierten nicht-exonischen Elementen. bioRxiv. 2019;8:132.
99. Suarez HG, Langer BE, Ladde P, Hiller M. chainCleaner verbessert die Spezifität und Empfindlichkeit des Genom-Alignments. Bioinformatics. 2017;33(11):1596–603.
100. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M, Stadler PF. MITOS: Verbesserte de novo Metazoen mito- chondriale Genom Annotation. Mol Phylogenet Evol. 2013;69(2):313-9.
101. Yu Y, Ouyang Y, Yao W. shinyCircos: an R/Shiny application for interactive creation of Circos plot. Bioinformatics. 2018;34(7):1229-31.
102. Smit A, Hubley R, Green P: RepeatModeler Open-1.0. 2008-2015. Institute for Systems Biology, Seattle, USA Verfügbar unter: http://www.repea tmaskerorg 2015.
103. Smit A, Hubley R, Green P: RepeatMasker Open-4.0. 2013-2015. Available from: http://www.repeatmasker.org. In.; 2015.
104. TETools [https://github.com/Dfam-consortium/TETools].
105. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: Beschleunigt für das Clustering der Next-Generation-Sequencing-Daten. Bioinformatics. 2012;28:3150-2.
106. Goubert C, Craig RJ, Bilat AF, Peona V, Vogan AA, Protasio AV. Ein Leitfaden für A n f ä n g e r zur manuellen Kuratierung von transposablen Elementen. Mob DNA. 2022;13(1):7.
107. pb-CpG-Tools [https://github.com/PacificBiosciences/pb-CpG-tools].
108. Simao Neto F, Waterhouse R, Ioannidis P, Kriventseva E, Zdobnov E. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31:3210-2.
109. Skripte im Zusammenhang mit Hundsdoerfer et al, 2023 [https://github.com/ charlottewright/hyles_manuscript].
110. Koch LM. Aufstieg und Fall der Carbonaria-Form der gepfefferten Motte. Q Rev Biol. 2003;78(4):399-417.
111. van't Hof AE, Edmonds N, Dalíková M, Marec F, Saccheri IJ. Industrieller Melanismus bei britischen Pfeffermotten hat einen singulären und rezenten mutationsbedingten Ursprung. Science. 2011;332(6032):958–60.

Anmerkung des Herausgebers

Springer Nature verhält sich neutral in Bezug auf juristische Ansprüche in v e r ö f f e n t l i c h t e n Karten und institutionelle Zugehörigkeiten.

Sind Sie bereit, Ihren Forschungsbericht einzureichen? C\oose BMC und profitieren Sie davon:

- schnelle, bequeme Online-Einreichung
- gründliche Begutachtung durch erfahrene Forscher auf Ihrem Gebiet
- schnelle Veröffentlichung bei Annahme
- Unterstützung für Forschungsdaten, einschließlich großer und komplexer Datentypen
- goldener Open Access, der eine breitere Zusammenarbeit und mehr Zitate begünstigt
- Maximale Sichtbarkeit für Ihre Forschung: über 100 Millionen Website-Aufrufe pro Jahr

Bei BMC ist die Forschung immer im Gange.

Mehr erfahren

biomedcentral.com/einreichungen