

# Genetisches Unterart-Identifizierungs-Tool für europäische Honigbienen

Im Rahmen des europäischen Forschungsprojektes SmartBees, zu welchem auch das ZBF beigetragen hat, wurde, basierend auf einer umfassenden Beprobung der Honigbienendiversität in ganz Europa, ein genetisches Unterart-Identifizierungs-Tool entwickelt. Das Tool verwendet ein maschinelles Lernmodell, um Bienenproben anhand von rund 4000 Variationen in ihrem genetischen Material in 14 verschiedene europäische Unterarten zu klassifizieren.

MELANIE PAREJO<sup>1,2</sup>, JEAN-DANIEL CHARRIÈRE<sup>1</sup>, ANDONE ESTONBA<sup>2</sup> & SMARTBEES WP3 DIVERSITY GROUP

<sup>1</sup> AGROSCOPE, ZENTRUM FÜR BIENENFORSCHUNG, 3003 BERN

<sup>2</sup> LAB. GENETICS, UNIVERSITY OF THE BASQUE COUNTRY (UPV/EHU), LEIOA, SPANIEN

Die Westliche Honigbiene (*Apis mellifera* L.) weist in ihrem natürlichen Verbreitungsgebiet eine beträchtliche Variation auf und besteht aus mindestens 30 beschriebenen Unterarten, die zu verschiedenen evolutionären Linien gehören.

## Honigbienen Vielfalt soll geschützt werden

Europa beherbergt einen grossen Teil dieser Honigbienen Vielfalt mit zahlreichen endemischen Unterarten, die vier evolutionäre Linien repräsentieren, nämlich die afrikanische Linie (A), die mittel- und osteuropäische Linie (C), die west- und nordeuropäische Linie (M) und die nahöstliche und zentralasiatische Linie (O). Diese Vielfalt und das natürliche Verbreitungsgebiet der europäischen Honigbienen wurden jedoch im Laufe der Zeit durch menschliche Aktivitäten beeinflusst. Es wird befürchtet, dass durch intensive Königinnenzucht, Handel von Königinnen und Bienenvölkern und Wanderimkerei die lokal angepassten Honigbienen mit ihren spezifischen Merkmalen verdrängt oder durch Hybridisierung reduziert werden. Es wurde nachgewiesen, dass lokal angepasste Honigbienen eine höhere Überlebensfähigkeit haben, und es daher für den langfristigen Erhalt der Honigbienenpopulationen notwendig ist, die nachhaltige Zucht lokaler Bienen zu fördern.

Zahlreiche Projekte zu Schutzgebieten und zur Zucht einheimischer Honigbienen wurden in ganz Europa

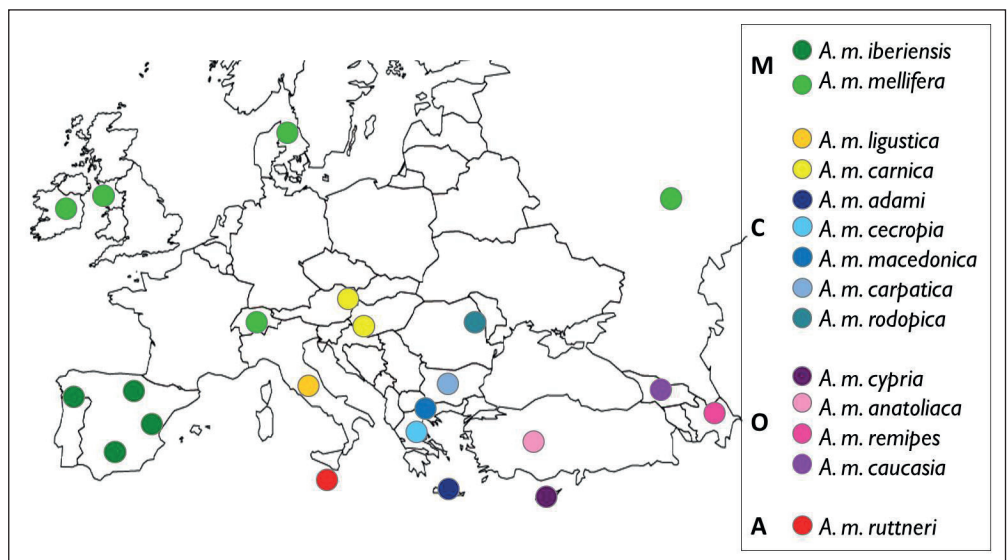


Abbildung 1. Herkunft der 22 beprobten Populationen. Jede Populationsprobe besteht aus ca. 100 Arbeiterinnen, welche aus nicht verwandten Völkern und, wenn möglich, aus verschiedenen Bienenständen der ganzen Region bzw. Nation beprobt wurden. Insgesamt wurden mehr als 2000 Honigbienen beprobt.

initiiert. Der Erfolg solcher Erhaltungsbemühungen einschliesslich genetischer Verbesserungsprogramme (Zucht) hängt von effektiven Massnahmen zur Paarungskontrolle und einer kontinuierlichen Überwachung der genetischen Abstammung der Population ab. Hierfür wird ein schnelles, genaues und einfach zu bedienendes genetisches Testwerkzeug benötigt.

Das EU-Forschungsprojekt SmartBees wurde mit dem Ziel initiiert, neue Werkzeuge zur Beschreibung und Erhaltung der Honigbienen Vielfalt in Europa zu entwickeln. Wir haben ein solches molekulares Werkzeug basierend auf einer umfassenden Stichprobe der europäischen Honigbienen entworfen, das aus hochinformativen SNPs (Einzelnukleotid-Markern, auf Englisch «Single Nucleotide Polymorphisms») besteht,

die geeignet sind, Bienenproben ihrer Herkunftsunterart zuzuordnen. Die Grundlagen der Bestimmung mit der SNP-Technik wurden in einem Artikel in der SBZ 10/2018 beschrieben.<sup>1</sup>

## Vorgehensweise

Insgesamt wurden 22 Populationen, welche die vier europäischen Evolutionslinien und 14 Unterarten repräsentieren, aus ihren heimischen Verbreitungsgebieten in ganz Europa und angrenzenden Regionen beprobt (Abbildung 1), inklusive 100 *Apis mellifera mellifera* Bienen aus der Schweiz. Jede ausgewählte Population umfasste etwa 100 Arbeitsbienen aus nicht verwandten Völkern, insgesamt mehr als 2000 Proben. Dies stellt die bisher umfangreichste Probennahme für die Untersuchung der europäischen Honigbienen dar.

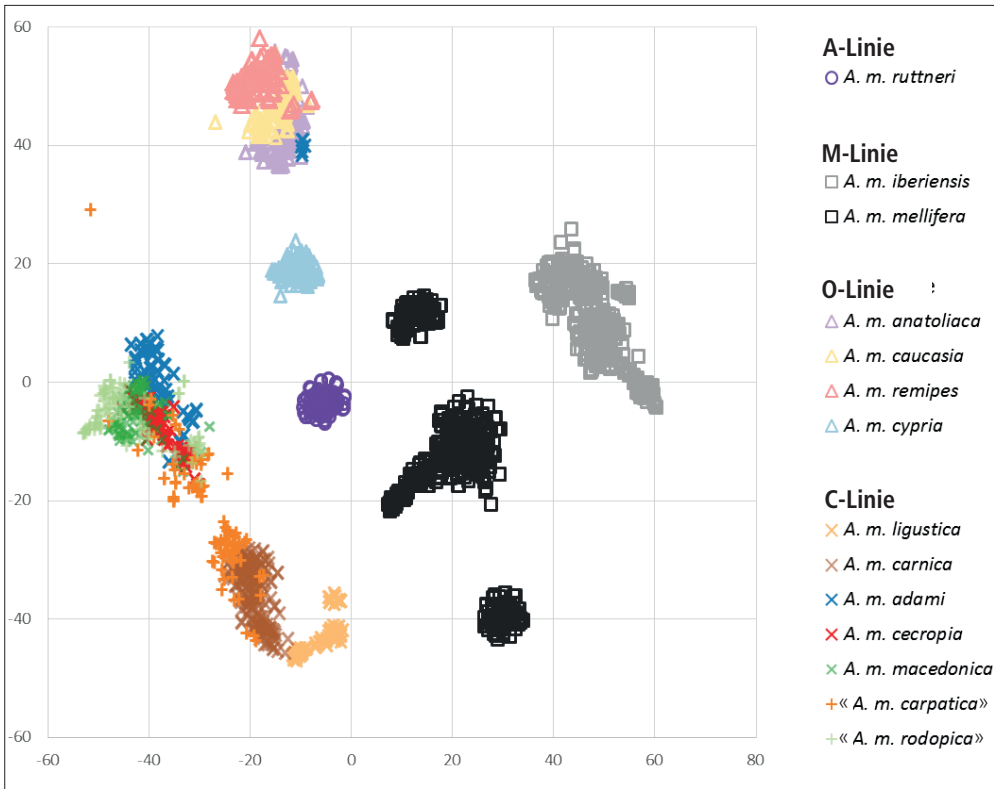


Abbildung 2: Visualisierung (t-SNE Manifold Plot) der Genotypisierungs-Daten der Proben in der Abbildung 1, die individuell für mehr als 4000 SNPs genotypisiert wurden. Es ist wichtig zu beachten, dass diese Art von Plot versucht, die Proben so eng wie möglich zu gruppieren, um ein optimales Clustering zu erhalten. Relative Abstände zwischen Gruppen und einzelnen Proben stellen jedoch keine evolutionäre Differenzierung oder genetische Verwandtschaft dar, z. B. sind *Apis mellifera cyprica* und die oberste Untergruppe von *Apis mellifera mellifera* genetisch weit entfernt verwandt, obwohl sie auf diesem t-SNE-Plot vergleichsweise nah erscheinen.

Die Proben aus jeder Population wurden zusammgelegt, ihre DNA extrahiert und dann sequenziert. Basierend auf den resultierenden Sequenzierungsdaten wählten wir die für die Abstammungsbestimmung informativsten genetischen Marker aus, d. h. genetische Marker, die am besten in der Lage sind, die 14 verschiedenen europäischen Honigbienen-Unterarten zu unterscheiden. Insgesamt waren dies mehr als 4000 SNPs. Anschließend wurden alle 100 Arbeiterinnen für jede der 22 Populationen einzeln genotypisiert, insgesamt mehr als 2000 Arbeiterinnenproben. Diese Genotypisierungs-Daten werden in einem sogenannten «t-SNE manifold plot» visualisiert, der die Proben so eng wie möglich auf einer zweidimensionalen Grafik abbildet, wobei jede Unterart durch ein anderes Symbol dargestellt wird (Abbildung 2). Es zeigt sich, dass sich die Proben in mehreren getrennten Clustern entsprechend ihrer evolutionären Abstammung oder Unterarten gruppieren:

die einzige Unterart der A-Linie in unserer Studie, *Apis mellifera ruttneri*, ist in der Mitte zwischen den anderen Clustern platziert. In der O-Linie bilden die *A. m. cyprica*-Bienen ein separates Cluster, während *A. m. anatoliaca*, *A. m. caucasica* und *A. m. remipes* nah beieinander gruppiert sind, welche auch geografisch nahe beieinander liegen. Die beiden Unterarten der M-Linie sind gut differenziert, wobei sich die *A. m. mellifera*-Populationen in drei Subclustern gruppieren, welche die weit entfernten (Burzyan-Region, Russland, oberes *A. m. mellifera*-Cluster in Abbildung 2) oder isolierten (Insel Læsø, Dänemark, unteres *A. m. mellifera*-Cluster) Probenregionen trennen. Die Schweizer *A. m. mellifera*-Proben befinden sich im mittleren Cluster zusammen mit den *A. m. mellifera*-Proben aus Irland und Isle of Man. Die Proben der C-Linie sind in drei Subcluster gruppiert: (i) *A. m. ligustica*, (ii) *A. m. carnica*-Bienen einschliesslich eines Teils der *A. m. carpatica*-Proben und (iii) ein

heterogenes Subcluster aus *A. m. macedonica*, *A. m. cecropia*, *A. m. adami*, *A. m. rodopica* und dem Rest der *A. m. carpatica*-Bienen.

### Zusätzliche Testproben und maschinelles Lernmodell

Wir haben zusätzlich 1900 Testproben genotypisiert, die von verschiedenen SmartBees-Zuchtbienenständen aus ganz Europa stammen. Diese 1900 Proben wurden zusammen mit den 2000 bereits genotypisierten Proben (insgesamt mehr als 3900 Proben) verwendet, um ein statistisches Modell mit Algorithmen des maschinellen Lernens für die Klassifizierung der europäischen Honigbienen zu erstellen. Das Modell berechnet die Vorhersagewahrscheinlichkeit für jede neue Probe, zu einer der 14 verschiedenen Unterarten zu gehören. Der Vorteil der Verwendung eines maschinellen Lernansatzes ist, dass er nicht von vorherigen Annahmen abhängig ist. Auf diese Weise können subtile Unterschiede aufgedeckt werden, was in unserer Studie aufgrund der hohen Anzahl eng verwandter Unterarten, die wir unterscheiden wollten, besonders relevant war. Weitere Details zu den spezifischen Methoden für die Klassifizierung finden sich in unserer wissenschaftlichen Veröffentlichung (Momeni et al. 2021).<sup>2</sup>

### Genauigkeit der Bestimmung der Unterart

Das Modell hat die Proben mit einer mittleren Genauigkeit von 96,2 % richtig zugeordnet. Dies lässt sich am besten in einer sogenannten Konfusionsmatrix visualisieren, die die Prozentsätze der vom Modell richtig oder falsch vorhergesagten Proben anzeigt (Abbildung 3). Die Zahlen auf der Diagonale stellen den Prozentsatz der korrekt klassifizierten Proben dar, während im oberen und unteren Dreieck die falsch klassifizierten Proben gekennzeichnet sind. Fehlklassifizierungen, d. h. Proben, die vom Modell mit einer anderen Unterart als der gekennzeichneten vorhergesagt wurden, sind jedoch nicht unbedingt ein «Fehler» des Modells, es kann auch sein, dass die Testproben falsch gekennzeichnet waren.

Wenn die Vorhersagewahrscheinlichkeit für eine beliebige Probe unter dem Schwellenwert von 90 % liegt, wird sie als «nicht zugeordnet» klassifiziert. Wir haben diesen Schwellenwert so definiert, damit das Modell in der praktischen Erhaltung und Zucht eingesetzt werden kann.

### Die Herausforderungen der Vielfalt und die nächsten Schritte

Die beträchtliche Honigbienen Vielfalt in ganz Europa stellte eine Herausforderung dar, als es darum ging, ein Werkzeug zur Unterartbestimmung zu entwickeln. Die Evolutionslinien (M, C, O, A) liessen sich genetisch leicht mittels nur weniger SNPs unterscheiden, da sie stark divergent sind; auf niedrigeren Differenzierungsebenen war es jedoch kniffliger. Unterarten innerhalb der gleichen Evolutionslinie entwickelten sich evolutionsbiologisch gesehen erst vor Kurzem und sind daher genetisch sehr nahe verwandt. Ausserdem gibt es in Europa einige Gebiete, in denen die Variation der *A. mellifera*-Unterarten noch nicht vollständig beschrieben wurde, während in anderen Gebieten die vom Menschen verursachte Hybridisierung dazu beiträgt, die natürlichen Grenzen zwischen den Unterarten zu verwischen. Auch nationale Zuchtprogramme können den natürlichen Genfluss stören und dazu beigetragen haben, den genetischen Hintergrund der ursprünglichen Unterarten zu verändern. Dies sind Gründe, warum einige Unterarten leichter zu klassifizieren waren als andere, und warum einige Proben mit unserem Identifizierungstool nicht zugeordnet werden konnten. Es handelt sich jedoch um ein dynamisches Werkzeug, das aktualisiert werden kann, um den Referenzdatensatz zu verfeinern und/oder weitere Unterarten mit einzubeziehen. Laufende Untersuchungen deuten darauf hin, dass dieser Ansatz auf die Unterart *A. m. siciliana* aus Sizilien erfolgreich anwendbar ist.

### Anwendungsbereich

Mit diesem Werkzeug ist es also möglich, neue Bienenproben zu testen und ihren genetischen Ursprung zu


Bestimmung mit 90% Schwellenwert		Vorausgesagte Unterart													# Nicht zugeordnet		
		A			M			O			C						
Beschriftete Unterart	A	Ruttneri	Iberiensis	Mellifera	Cyprina	Caucasia	Anatoliaca	Remipes	Carnica	Ligustica	Cecropia	Macedonica	Rodopica	Carpatica	Adami		
	A	Ruttneri	100	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	M	Iberiensis	0	100	0	0	0	0	0	0	0	0	0	0	0	0	6
		Mellifera	0	0	100	0	0	0	0	0	0	0	0	0	0	0	48
	O	Cyprina	0	0	0	100	0	0	0	0	0	0	0	0	0	0	3
		Caucasia	0	0	0	0	100	0	0	0	0	0	0	0	0	0	11
		Anatoliaca	0	0	0	0	0	100	0	0	0	0	0	0	0	0	6
		Remipes	0	0	0	0	0	0	100	0	0	0	0	0	0	0	4
		Carnica	0	0	0	0	0	0	0	99	0	0	0	0	0	0	70
	C	Ligustica	0	0	0	0	0	0	0	0	100	0	0	0	0	0	45
		Cecropia	0	0	4	0	0	0	0	0	0	92	4	0	0	0	55
		Macedonica	0	0	0	0	0	0	0	0	0	0	100	0	0	0	125
		Rodopica	0	0	0	0	0	0	0	0	0	0	0	100	0	0	9
Carpatica		0	0	0	0	0	0	0	0	0	0	0	0	100	0	15	
Adami		0	0	0	0	0	0	0	0	0	0	0	0	0	100	8	
															407		

Abbildung 3. Konfusionsmatrix für die getesteten Proben unter Verwendung eines 90%-Wahrscheinlichkeits-schwellenwerts, die die Prozentsätze (gerundet) der tatsächlich zugeordneten Individuen (Diagonale) und die Prozentsätze der Individuen zeigt, die einer anderen Unterart zugeordnet wurden; einige Proben werden als «nicht zugeordnet» betrachtet und aus der Konfusionsmatrix ausgeschlossen. Zum Beispiel wurden 100 % der Proben als Ruttneri klassifiziert, mit zwei nicht zugeordneten Proben. Für Carnica wurden 99 % korrekt zu Carnica zugeordnet, mit 70 nicht zugeordneten. Andererseits wurden z. B. bei Cecropia-Proben 92 % der Proben richtig zugeordnet, während jeweils 4 % falsch als Macedonica oder Mellifera klassifiziert wurden.

untersuchen, was unter anderem in folgenden Anwendungen nützlich sein kann: Für Imkerinnen und Imker, welche die Unterarten ihrer Bienen kennen wollen, oder wissen möchten, ob sie gekreuzt/hybridisiert sind; für Naturschutzmanager in Europa, wo die Unterartbestimmung wesentlich ist, um die Hybridisierungsrate von Bienenvölkern in Schutzgebieten zu überwachen; für Tierärzt/-innen, um den Handel mit Königinnen zu kontrollieren; für Bienenzüchter und -züchterinnen, um die Unterartabstammung ihrer Königinnen zu zertifizieren; und für Imkerinnen und Imker, um zu dokumentieren, dass ihr Honig von einer bestimmten Unterart produziert wurde.

### Dank

Diese Arbeit war nur aufgrund der wertvollen Proben möglich, die von Forscherkollegen, Imkern, Bienenzüchtern und anderen Mitwirkenden zur Verfügung gestellt wurden, wofür wir besonders dankbar sind.

Für die Schweizer Proben im Spezifischen danken wir herzlich Padruot Fried und mellifera.ch. Das SmartBees-Projekt wurde von der Europäischen Kommission im Rahmen ihres FP7 KBBE-Programms (2013.1.3-02, Grant nr. 613960) gefördert. 

### Literatur

1. Parejo, M. (2018) Entwicklung einer neuen Methode für die Hybridanalyse bei Honigbienen. *Schweizerische Bienen-Zeitung* 10: 13–15.
2. Die wissenschaftliche Veröffentlichung kann hier aufgerufen werden: Momeni, J.; Parejo, M.; Nielsen, R. O.; Langa, J.; Montes, I.; Papoutsis, L.; ... & Estonba, A. (2021) Authoritative subspecies diagnosis tool for European honey bees based on ancestry informative SNPs. *BMC Genomics* 22(1): 1–12. (<https://doi.org/10.1186/s12864-021-07379-7>)