

Elektrophorese und das Paradebeispiel der Klonalität

Smith (Smith et al., 1993) und Caugant (Caugant et al., 1985; Caugant et al., 1987) veröffentlichten elektrophoretische Daten mit Hilfe der MLEE von 15 verschiedenen Loci aus 688 *Neisseria meningitidis*-Isolaten. Sie eine Gesamtzahl von 331 unterschiedlichen elektrophoretischen Typen (ET). Trotz dieser Variabilität wurden 61 ET häufiger als einmal gefunden, von denen 19 vom Kontinent stammten und in einem Zeitraum von bis zu 15 Jahren isoliert worden waren. Dieses Vorkommen gleicher ET trotz einer großen Variabilität wurde auch bei anderen Bakterienspezies nachgewiesen. Diese Ergebnisse bestätigten zunächst die Ansicht, dass trotz des gelegentlichen Transfers von Genen innerhalb einer Spezies (z.B. Flagellin-Gene bei *Salmonella* (Smith and Selander, 1990; Reid et al., 1999; Wang et al., 2000)) die Bakterienpopulationen vornehmlich klonal sind (Selander and Lewin, 1980; Orskov and Orskov, 1983; Ochman and Selander, 1984b; Pupo et al., 2000a; Pupo et al., 2000b).

Bei näherer Betrachtung dieser Daten kamen jedoch Zweifel an der Idee der klonalen Ausbreitung auf. Zunächst wurde eine erneute Analyse dieser elektrophoretischen Daten (Smith et al., 1993) durch einen verhältnismäßig umfangreichen Datensatz von *Neisseria gonorrhoeae* (O'Rourke and Stevens, 1993) angeregt, und zwar wurden neun Loci von 327 weltweit gesammelten Isolaten analysiert. Diese Daten ergaben, dass die einzelnen Allele eine zufällige Verteilung aufwiesen. Dieses Verbindungsgleichgewicht (Linkage Equilibrium) bezeichnet den Umstand, dass das Vorhandensein eines Allels an einer bestimmten Lokalisation im Chromosom unabhängig vom Fehlen oder vom Vorhandensein eines entsprechenden Allels an einem anderen Ort im Chromosom ist. So kam der häufigste ET 35mal vor. Dieses weitverbreitete Vorkommen dieses ET kann jedoch nicht als ein Beweis für Klonalität verwendet werden, da die erwartete Anzahl 32,2 bei zufälliger Anordnung war (berechnet durch Multiplikation der Häufigkeiten eines jeden der Allele von diesem ET innerhalb des Datensatzes).

Ein Verbindungsgleichgewicht oder Linkage Equilibrium ist nicht anders zu erklären als durch die Annahme hoher Rekombinationsraten. Ein Verbindungsungleichgewicht oder Linkage Disequilibrium, also die nicht-zufällige Verbindung von Allelen, kann in einer Probe von Stämmen beobachtet werden, wenn die Population, aus welcher die Probe kam, durch eine hohe Rekombinationsfrequenz gekennzeichnet ist. So besteht in der Tat die Möglichkeit des begrenzten geographischen Auftretens einer Population. Dieses Phänomen wurde übrigens bei Gonokokken nicht nachgewiesen, sondern lediglich bei Meningokokken, also selbst innerhalb einer Gattung von Bakterien bestehen derart gravierende Unterschiede. So wäre es möglich, dass eine ermittelte Verbindung zwischen einzelnen Allelen lediglich daraus resultiert, dass der Genfluss auf zwei oder mehrere Subpopulationen innerhalb einer definierten Gruppe von Bakterien beschränkt ist, weil diese geographisch isoliert sind. Souza (Souza et al., 1992; Souza et al., 1999) belegte die Bedeutung geographischer Strukturen am Beispiel der Nitrogen-fixierenden Bakterien *Rhizobium leguminovarum* biovar *phaseoli*. Diese Bakterien kommen lediglich in Assoziation mit Wildtyp- und Kulturbohnen vor. Diese Autoren verglichen den Umfang einer Verbindung von Allelen in bakteriellen Populationen, die über eine räumliche Skala von der westlichen Hemisphäre bis auf das Niveau einer einzelnen Wirtspflanze reichte. Sie fanden heraus, dass sich der Grad der Verbindung (und der allelen Verschiedenartigkeit) bei Zunahme der räumlichen Skala erhöhte. Daraus folgerten sie eine geringe Migrationsrate und eine signifikante geographische Strukturierung.

Im Gegensatz dazu sind bei Spezies wie z.B. *E. coli* die Migrationsraten sehr hoch, so dass ein großes Ausmaß der globalen Diversität lokal bestimmt werden kann (Whittam et al., 1983a; Whittam et al., 1983b). Bei anderen nicht-pathogenen Spezies, wie z.B. bei *Bacillus subtilis*-Stämmen, sichern effiziente Verbreitungsmechanismen wie die Sporenbildung eine hohe Migrationsrate und helfen, einen einzelnen gemeinsamen Genpool aufrecht zu erhalten. Eine geographische Strukturierung scheint außerdem bei *N. gonorrhoeae* und anderen pathogenen Spezies wie *N. meningitidis* und *S. pneumoniae* vorzuliegen. Große DNA-Datensätze für die beiden letztgenannten Spezies belegen dies. Allel-Austausche beziehen die importierten Allele mit ein, so dass auch die nicht-verwandten Stämme im Datensatz nachweisbar sind. Diese Daten globaler Stammsammlungen sind zudem ein guter Beweis für hohe Mutationsraten. Zum Beispiel kann ein Allel, das zuerst in einem holländischen Stamm bemerkt wird, mittels Rekombination durch ein Allel z.B. eines Stammes aus Gambia ersetzt werden.

Obschon bei einigen Bakterien eine hohe Rekombinationsrate bewiesen ist, variiert die Rate von Rekombinationen zwischen verschiedenen Bakterienpopulationen. Ein nützlicher statistischer Test, um einen Maßstab für das Maß von Klonalität innerhalb einer Population zur Verfügung zu stellen, ist der „Assoziationsindex“ (Brown et al., 1980). Unter Anwendung der Assoziationsindices gelingt in der Multilocus-Analyse eine Berechnung des Verbindungsungleichgewichts. Ist der Wert des Index signifikant größer als Null, ist dies ein Beleg für eine stattgefundene Rekombination. Ein Indexwert von 1 würde eine freie Rekombination bedeuten.

Der Assoziationsindex ist jedoch für einen Vergleich der Häufigkeit von Rekombination bei unterschiedlichen Bakterien wenig effektiv, weil er selbst einen Wert für eine gegebene Häufigkeit von abhängigen Rekombinationen bei einer Anzahl analysierter Loci erwartet. Eine Messung, die diesen Nachteil vermeidet, wurde von Koufopanou (Koufopanou et al., 2001) vorgestellt¹. Ein zweites Problem liegt in der Benennung der Gründe für dieses Ungleichgewicht. Es könnte sein, dass eine Rekombination selten oder gar nicht auftritt, oder dass dort eine geographische Struktur innerhalb der Population vorliegt. Alternativ könnte es sein, dass eine Population viel häufiger Rekombinationen aufweist, aber dass möglicherweise ein Genomtyp auftaucht, der stark durch Selektion favorisiert wird. Dieser Genomtyp wird dann schnell in der Häufigkeit ansteigen, ein Ungleichgewicht herstellen, und erst dann haben Rekombinationen Zeit, den genetischen Hintergrund zu randomisieren. Solch eine „epidemische Struktur“ wurde anhand der elektrophoretischen Daten für *N. meningitidis* unterstellt und verdeutlicht um so mehr den überzeugenden Nutzen von MLST Daten (Smith et al., 2000).

¹ (zum Zwecke der Bestimmung, ob eine Adaptation an einen neuen Lebensstil eines Pathogens aufgetreten ist, wurde eine Anzahl von stillen und exprimierten Änderungen innerhalb und zwischen reproduktiven Taxa isoliert. Das Verhältnis $Q = (D/P)_E / (D/P)_S$ vom „zwischen-Taxa Unterschied“ (D) zum „innerhalb-Taxon Polymorphismus“ (P) für exprimierte und stille Substitutionen (E resp. S) ist eine allgemein akzeptierte Messung zur Bestimmung positiver, adaptiver Evolution, mit der neutralen Erwartung, dass $Q=1$ ist. Auf der anderen Seite würde ein mit Erfolg exprimierter „zwischen-Taxa Unterschied“ ($Q > 1$) zur schnellen Fixation von vorteilhaften Aminosäureänderungen und Adaptation an eine neue Umwelt führen)