

SAMENVATTING VAN DE LEZING VAN 24 MAART 2012

De moleculaire revolutie in de systematiek: zin en onzin

Edi Gittenberger¹

Het is tegenwoordig niet ongebruikelijk om termen als ‘moleculaire systematiek’ of ‘moleculaire fylogenie’ te gebruiken. Het suggereert dat er een bijzonder soort systematiek bestaat en een apart soort fylogenie. Dat is een vergissing. Er is maar één fylogenie, zoals er ook maar één evolutionaire geschiedenis bestaat, die men zo goed mogelijk wil reconstrueren om vervolgens in de systematische indeling tot uiting te kunnen brengen. De systematicus gebruikt zoveel mogelijk gegevens om tot een classificatie te komen waaruit de natuurlijke verwantschappen kunnen worden afgeleid. Eerst werden bij de molluskensystematiek vooral schelpkenmerken gebruikt. Later kwam daar de anatomie bij. Tegenwoordig levert vooral het DNA nog weer extra gegevens. Daardoor is er meer mogelijk dan vroeger, maar het DNA is geen wondermiddel dat alle problemen oplost.

DNA, nucleotiden, sequenzen

DNA komt voor in de, per afzonderlijke cel, in veelvoud aanwezige mitochondria (mtDNA) en in de enkelvoudige celkern of nucleus (nDNA). Daarom wordt mtDNA het meest gebruikt, ook bijv. als het om zgn. ‘ancient’ DNA (adNA) gaat. Dat is relatief oud, meer of minder sterk gefragmenteerd DNA, dat minder goed bruikbaar is geworden. DNA kan eenvoudig worden omschreven als twee lange ketens van achter elkaar geschakelde bouwstenen, de nucleotiden adenine (A), guanine (G), thymine (T) en cytosine (C). De twee ketens spiraliseren om elkaar heen (de ‘dubbelhelix’), doordat de nucleotiden uit beide ketens volgens vaste regels verbindende bruggen vormen. [Het RNA is opgebouwd als een enkelvoudige keten van nucleotiden, die eveneens bruggen vormen, waardoor er een ingewikkelde ruimtelijke structuur ontstaat. In het RNA is thymine (T) vervangen door uracil (U)].

Bij het moleculair systematisch onderzoek wordt door sequenzen de nucleotidenvolgorde vastgesteld van minstens één bepaald stuk DNA. Zo’n segment wordt meestal een ‘marker’ of, onjuist, ‘gen’ genoemd (het is immers meestal geen compleet gen). Er zijn manieren om homologe (vergelijkbare) stukken DNA, afkomstig uit verschillende soorten, met elkaar te vergelijken (‘alignment’). Dat levert dan bijvoorbeeld voor soort X een reeks van honderden of nog veel meer nucleotiden op, die genummerd zijn op basis van hun positie, bijv. AAGTTTCGATG... voor de posities 1 2 3 4 5 6 7 8 9 10 11 ... Voor soort Y vinden we bijv. AAGTCTCGTTG... en voor soort Z AAGTCTCGTTG... De onderstreepte nucleotiden, op posities 5 en 9, vertellen iets. Als we op basis van voorafgaand onderzoek aannemen dat voor posities 4 5 6 de oorspronkelijke nucleotidenvolgorde TTT is, zien we bij soorten Y en Z de volgorde TCT, dus een mutatie op positie 5. We nemen dan aan dat die twee soorten afstammen van een ge-

zamenlijke voorouder, waarbij die mutatie ooit plaatsvond, zodat destijds, bij die voorouder, op positie 5 in plaats van de oorspronkelijke T een C kwam te staan. Dus zijn Y en Z onderling nauwer verwant dan elk van hen met X. Die verwantschap wordt ondersteund door positie 9, waar Y en Z allebei een T hebben, verwijzend naar de gezamenlijke voorouder met T, en X een A.

Verwantschap

In de praktijk is het helaas lang niet altijd zo eenvoudig als in het voorbeeld met X, Y en Z. In dat voorbeeld suggereren posities 5 en 9 dezelfde verwantschap, maar soms worden tegengestelde verwantschappen aangegeven. Daarom is het nodig om een flink groot stuk DNA te gebruiken. Er zijn immers maar vier alternatieve mogelijkheden, waardoor het mogelijk is om in twee stappen weer bij de oorspronkelijke toestand terug te zijn. In dat geval laten twee mutaties geen sporen in het DNA na. De mutatie van T naar C in ons voorbeeld kan best ook onafhankelijk in soorten Y en Z hebben plaatsgevonden, en we kunnen ook niet uitsluiten dat er bij soort X een zgn. terugmutatie, van C naar T is opgetreden. Het kan allemaal. Bij de computergestuurde analyses van de grote aantallen DNA gegevens gaat het daarom uiteindelijk om waarschijnlijkheden.

Het voorafgaande beschrijft het opstellen van een stamboom voor genetische markers, niet automatisch ook voor de soorten zelf. Het is best mogelijk dat een soort zich door mutaties opsplijt in twee soorten, zonder dat de door de systematicus gebruikte stukken DNA, die mogelijk niets met de soortverschillen te maken hebben, gelijktijdig óók muteren. Daarom gebruikt men tegenwoordig meestal meer dan één marker om de evolutionaire geschiedenis van een taxon te achterhalen, liefst bovendien mt en n.

Mutatiesnelheid

Er zijn nog heel wat andere zaken waar we bij de toepassing van moleculaire methoden rekening mee moeten houden. Het maakt bijvoorbeeld een groot verschil of we de verwantschap tussen individuen (wie is de dader?) of soorten willen onderzoeken, of de fylogenetische relaties tussen veel hogere taxa in het systeem willen kennen, zoals tussen planten, fungi en dieren. Bij al dat soort onderzoek kunnen moleculaire markers worden gebruikt, maar niet steeds dezelfde. Het vinden van de meest geschikte marker is een kunst op zich. Verschillende stukken DNA kunnen namelijk heel sterk verschillen in stabiliteit, dat wil zeggen gevoeligheid voor mutaties, ofwel mutatiesnelheid. Dat hangt nauw samen met de functie van het betreffende stuk DNA. Ook tussen diergroepen kunnen er verschillen in dit opzicht bestaan. Afhankelijk van het taxon niveau, van hoog naar laag, worden markers met een zeer lage tot hoge mutatiesnelheid gebruikt.

De mutatiesnelheid, die binnen ruime grenzen varieert, wordt gebruikt om de ouderdom van taxa te bepalen met behulp van het zogenoemde 'moleculaire-klok model'. Als twee taxa lang geleden uit één voorouder zijn ontstaan, zullen ze sterker in hun DNA verschillen dan wanneer de splitting van veel recenter datum is. Het vinden van de meest geschikte marker, is daarbij alweer niet eenvoudig. De mutatiesnelheid kan te laag of te hoog zijn, waarbij respectievelijk geen verschillen worden gezien, of chaos door 'verzadiging', zoals ook bij 'Verwantschap' werd opgemerkt. In het laatste geval kunnen er op één plaats een onbekend aantal mutaties zijn opgetreden, of geen. Zo kan bijvoorbeeld de waargenomen verandering van A naar C door één mutatie tot stand komen (A-C), maar ook via een omweg (bijv. A-T-C of A-G-A-C). Wat er echt gebeurd is, blijft onzichtbaar. Het kan in dit voorbeeld dus om één (A-C), twee (A-T en T-C) of zelfs drie (A-G, G-A en A-C) mutaties gaan. Die verschillen zijn van belang, want één mutatie impliceert een kortere periode van gescheiden evolutie (nauwere verwantschap) dan twee of zelfs drie mutaties. Soms zijn er mutaties opgetreden, maar is dat niet meer te zien, omdat de begintoestand weer is teruggekeerd, bijv. bij A-G-T-A (dus A-G, G-T en T-A). Er lijkt dan niets gebeurd te zijn, terwijl er in feite al drie keer een mutatie is opgetreden. Kortom, naarmate het muteren langer aanhoudt worden de waargenomen verschillen en overeenkomsten minder goed bruikbaar om er de mate van verwantschap op te baseren.

Aminosuren en eiwitten

Het DNA is essentieel voor de vorming van de eiwitten in een organisme. Eiwitten zijn voor het leven van groot belang. Het zijn lange ketens van aan elkaar gekoppelde aminozuren. Een reeks van drie nucleotiden, een DNA triplet (of codon), is nodig om één bepaald aminozuur te 'vangen'. Het organisme moet vervolgens de aminozuren in de goede volgorde tot een eiwit combineren.

De structuur van sommige stukken DNA is uiterst belangrijk voor een organisme en dergelijke stukken muteren (evolueren) bijzonder langzaam (bijv. de zgn. ribosomale genen). Zgn. 'junk' DNA daarentegen, waarbij geen functionele triplets voorkomen, is het meest variabel. Daarbij is de nucleotidenvolgorde kennelijk veel minder belangrijk en worden mutaties niet of nauwelijks door selectie verwijderd. Of er werkelijk DNA zonder enige functie bestaat, is omstreden. Dat soort DNA zou dan de allerhoogste mutatiesnelheid hebben.

Omdat de derde nucleotide in een triplet niet steeds essentieel is ter bepaling van een aminozuur, wordt die 'derde positie' als minder betrouwbare indicator bij analyses soms weggelaten.

Barcoding

Momenteel wordt wereldwijd gewerkt aan de opbouw van een systeem met diagnostische DNA markers (bar-codes) voor uiteindelijk alle levende soorten van organismen. Een stukje weefsel is dan voldoende voor een determinatie. Er

zitten heel wat haken en ogen aan deze onderneming. Uiteraard zijn het in elk geval de systematici die voorafgaande aan het sequencen de exemplaren op de klassieke manier op naam moeten brengen.

Samenvattend

DNA sequenties kunnen (1) zogeheten tweeling-soorten aan het licht brengen, soorten dus die morfologisch niet of nauwelijks van elkaar lijken te verschillen; (2) beslissend zijn bij alternatieve opvattingen over verwantschappen; (3) een indicatie over de ouderdom van taxa geven (moleculaire-klok model); (4) de classificatie van 'moeilijke' taxa vergemakkelijken; (5) extra gegevens ook over uitgestorven soorten leveren (aDNA); (6) duidelijk maken welke larvale stadia bij volgroeide dieren horen.

DNA sequenties (1) maken morfologisch onderzoek volstrekt niet overbodig; (2) zijn niet maatgevend voor de mate van reproductieve isolatie (soortgrenzen) tussen nauw verwante taxa; (3) geven soms tegenstrijdige informatie en zijn dan in hun toepassing even objectief of subjectief als morfologische gegevens; (4) bieden geen mogelijkheid om de status van taxa, als soort, genus, familie, enz. objectief vast te stellen.

¹Prof. Dr. E. Gittenberger, *Naturalis*,
e-mail: edi.gittenberger@ncbnaturalis.nl