

Elu mitmekesisuse geneetilised ilmingud

Mart Viikmaa

Elu mitmekesisus (biodiversiteet) on evolutsiooni tulemus, selle nüüdisseisundi näitaja ja tuleviku perspektiiv. Nagu teame, seletas C. R. Darwin (1859; 2012) evolutsiooni määratlemata iseloomu ja põhjustega individuaalsete muutuste loodusliku valikuna. Tol ajal valitsenud segupärilikkuse (*blending inheritance*) kontseptsiooni tõttu oli ta kriitikaga tegelikult suures hädas. G. Mendel (1866) avastas oma katsetega hernesortide ristamisel diskreetse olemusega (individuaalsetel püsivatel elementidel, ehk geenidel põhineva) pärilikkuse, seda just nende elementide alternatiivvormide (alleelide) kombineerumise seaduspärasuste mõistmiseks. Mendel aga ei esitanud küsimust – kust need alternatiivid (dominantseid ja retsessiivseid alleelid) tulevad. Selles mõttes oli järgmine avastus, et geenid võivad vahetevahel muutuda (muteeruda), T. H. Morgani (1910) „kärbsetoas“ leitud valgesilmne isaskärbes. Geenmutatsioonide olemus jäi küll kauaks ajaks mõistmata. Peale geenmutatsioonide leiti ka mikroskoopiliselt tuvastatavaid kromosoommutatsioone (osade

ümberpaiknemisi, kordistumisi, kadusid) ja genoommutatsioonide (kromosoomide arvu muutusi). Kromosoomide struktuuris ja arvus ilmnesid ka selged liigilised erinevused, genoomide nähtav mitmekesisus. 1920. aastate alguses tõestati, et inimese ABO-süsteemi vererühmad on määratud ühe geeni kolme alleeliga, mis on kõik enamikus inimpopulatsioonides võrdlemisi suure sagedusega (> 10%), seega n-ö normaalsed. See oli esimene liigisisese polüalleelse polü-morfismi nähtus (vt Stern 1968).

See on eellugu. Nüüd, pärast molekulaargeneetika teket ja genoomika arengut, teame me hoopis rohkem ja ka hoopis midagi muud. Järgnevalt käsitlen mõnda neist ilmingutest.

C-väärtuse mõistatus

Üks esimesi elusolendite suure geneetilise mitmekesisuse näitajaid oli 1940-te lõpus määrama hakatud eukarüootide rakutuuma DNA-sisaldus. Selgus, et see on konstantne ühe liigi isendite kõigis keharakkudes ja haploidses gameedituumas kaks korda väiksem – seda viimast suurust hakati nimetama C-väärtuseks (*C-value*; Swift 1950) mõõdetuna pikogrammides (1 pg = 978 mega-aluspaari). C-väärtus varieerub eri organismirühmade ja mõnikord ka sama rühma liikide vahel väga suures ulatuses ning paradoksaalselt: see väärtus ei korreleeru organismi ehituse ja talitluse keerukusega. Üks esimesi üllatusi oli, et salamandri genoom osutus 40 korda suuremaks kui inimesel. Seletamatu nähtuse kirjeldamisel võeti kasutusele mõiste **C-väärtuse paradoks** (Thomas 1971), nüüd nimetatakse seda enamasti C-väärtuse mõistatuseks (*C-value enigma*). Loomariigis on väikseima ja suurima C-väärtuse erinevus üle 6600 korra, maismaataimedel

üle 1000 korra. Suurim eukariöodigenoom on leitud vabalt elaval amööbil (*Chaos chaos*) – 1400 pg; see on üle 450 korra suurem kui inimese genoom*. Paradoksi põhjus sai arusaadavaks, kui alates 1970. aastatest hakkas saabuma üha rohkem andmeid, et eukariootide genomides on varieerulval hulgal transposoonset kordus- ja muud tüüpi mittekodeerivat DNA-d. Kuid nende suurte erinevuste ja muutuste põhjused on praegugi suuresti seletuseta.

Geenide funktsionaalne katkulisus

1970. aastate kõige suurem geneetiline üllatus oli eukariootsete **geenide funktsionaalse katkulisuse** (ekson-intron-struktuuri) ja pre-mRNA splaiissingu (e. kokkupõime) avastamine 1977 (Sharp 1987), mille eest said P. Sharp ja R. Roberts 1993. aastal Nobeli auhinna. Mittekodeerivate intronite sagedus geenides on eukariootide evolutsioonis oluliselt kasvanud. Näiteks pärmseene (*Saccharomyces cerevisiae*) genomis on niisuguseid gene 5-6%, varbussil (*Caenorhabditis elegans*) u. 10%, imetajatel aga ca 95%. Tihti võtavad intronid geeni DNA kogupikkusest kaugelt suurema osa. Näiteks suurim tuntud geen on inimese lihasvalgu düstrofiini geen (*Duchenne muscular dystrophy, DMD*). Selle geeni (Xp21) pikkus on ~2,3 Mb (seega pikem, kui paljude bakterite kogu genoom), ta sisaldab 79 eksonit, splaiissitud mRNA pikkus on aga ainult 14 kb (0,6% pre-mRNA pikkusest). Inimese teise lihasvalgu titiini geen (2p31.2) sisaldab 363 eksonit ning sellelt transleeritud valgu standardvorm koosneb 34 350 aminohappejäägist; hiire vastavas valgus on üle 35 000 aminohappe – see on suurim tuntud polüpeptiidahel. Inimese genoomi

* Wikipedia 2012. C-value. <http://en.wikipedia.org/wiki/C-value>

nukleotiidjärjestustest ainult u. 1,5% kodeerib valke. Palju kõrgemate selgroogsete geene on mosaiikse struktuuriga, see tähendab nad sisaldavad erinevate eellasgeenide domeene. Näiteks düstrofiinigeen koosneb neljast erinevat päritolu domeenist. Niimoodi on saanud tekkida polüfunktsionaalsed valgud.

Ootamatu organismide keerukuse ja geeni arvu vastavus

1960.-70. aastatel oletati, et inimesel peaks olema 2-5 miljonit geeni. See arv tuletati peamiselt genoomi mahtude erinevusest bakterite ja inimese vahel. Mittekodeerivate DNA-järjestuste avastamine geenides ja geenivahemikes vähendas seda hinnangut suurel määral – 1980. aastatel esitati inimese geeni arvuks ümmarguselt 100 000 ja 1990-te keskel langes see 60-80 tuhandeni. Kui 2000. aastal avaldati inimese genoomi järjendamise esmased täisandmed ja esitati hinnangud valgugeenide arvu kohta – ühelt grupilt 31 000 ja teiselt 36 000 – siis oli see suur üllatus ja näis isegi uskumatu (Davies 2001). Hiljem on seda arvu veelgi oluliselt vähendatud, räägitakse umbkaudu 23 000 geenist. Samal ajal on 1 mm pikkusel varbussil hinnatud umbes 13 000 ja äädikakärbsel 15 000 geeni, väikesel rohhtaimel (müürlook, *Arabidopsis thaliana*) 27 000 geeni ja riisil koguni 46 000 geeni. Niisiis, nagu genoomi üldise mahu, nii ka geeni arvu puhul ei ole oodatud kooskõla organismi ehituse ja arengu keerukusega.

Sellele probleemile seletuse otsimisel on leitud, et evolutsioon on leiutanud muutuvate tingimustega kohastumisel geenide arvu suurendamise asemel geenide **alternatiivse splaissingu** (valikkokkupõime) võimaluste laiendamise. On leitud, et inimesel allub alternatiivsele splaissingule enam kui 90% geenide esmastest transkriptidest. See

mehhanism võimaldab toota erinevaid valke palju rohkem, kui on erinevaid geene. Arvatakse, et inimese organismis tema arengu eri etappidel, erinevate kudede ja organite rakkudes ning erinevates välismõjudes sünteesitakse kuni 120 000 vähem või rohkem erinevat valku. Mõnel geenil on alternatiivse splaissingu võimalusi äärmiselt palju. Näiteks inimese ajus avalduvalt neureksiinigeenilt sünteesitakse ca 1000 valgu isovormi. Neureksiinid kuuluvad närvirakkude vaheliste sünapside koostisse, seega need variandid võimaldavad luua erinevate omadustega ühendusi.

Liigisisene geneetiline mitmekesisus, polümorfism

Individuaalne ja populatsioonidevaheline erinevus põhineb valdavalt alleelide koosseisul ja eri alleelide sagedustel (Viikmaa 1993). Genoomide järjestusanalüüsid on näidanud, et peamised alleelsed erinevused seisnevad üksiknukleotiidsetes polümorfismides (*single nucleotide polymorphism*, **SNP**). Inimpopulatsioonides esineb üks selline polümorfism keskmiselt 1000 nukleotiidipaari kohta. Seega moodustavad seda tüüpi indiviididevahelised erinevused ümmarguselt 0,1% genoomist. Kuid need polümorfismid ei ole genoomis ühtlaselt jaotunud: on väga konservatiivseid geene, milles polümorfisme ei tunta, ja geene, millel esineb igas populatsioonis 10-40 polümorfismitasemega alleeli (harvima sagedus vähemalt 1%). Selgroogsetel on vist kõige polümorfsemate geenide rühm suur koesobivusgeenide kompleks (*major histocompatibility complex* ehk MHC), inimesel nimetatud ka HLA-geenideks (*human leucocyte antigens*). Need kodeerivad rakupinna valke, mille funktsiooniks on peptiidi-fragmentide esitlemine immuunsüsteemi rakkudele, et need õpiksid eristama oma ja võõrast valku (antigene). Mida

erinevamad need valgud on, seda täpsemalt suudavad nad erinevaid peptiide esile tõsta, järelikult seda paremini võõrast (bakteri, viiruse vm.) valku paljastada. See geenikompleks hõlmab üle 10 polümorfse geeni. Inimesel on neist kõige mitmekesisemate alleelidega geenid HLA-A, -B ja -C. Erinevate MHC-geenide alleelikombinatsioonid on nii palju, et panmiktilises populatsioonis on raske leida ühesuguse genotüübiga indiviide – sellest ka raskused siirdamiseks sobivate organite leidmisel.

Kohastumine parasiitidega on üldiselt üks peamisi populatsioonisisese polümorfismi allikaid. Eriti näitlikud selles mõttes on teatud geneetilise polümorfismi juhud mingi kindla parasiiditüübi leviku piirkondades, mõnikord isegi homosügootsena kahjulike alleelide baasil. Ses suhtes võib mainida malaaria leviku piirkondades (Aafrikas, Vahemere- maades ja Lõuna-Aasias) levinud hemoglobiini- Valkude struktuurseid (HbS, HbC, HbE) või sünteesipuudega (talasseemia- β , $-\delta$ ja α) defekte jmt.

Ulatuslik polümorfism esineb ka mittekodeerivates DNA-järjestustes. Selles osas on hästi tuntud inimese mini- ja mikrosatelliitsetes tandemjärjestustes korduste arvu polümorfism (*variable number tandem repeat*, **VNTR**). Neist enamiku puhul ei teata nende funktsionaalset tähendust organismile, kuigi mõned on inimese omapärase päritavusega haiguste põhjuseks; lisaks on nad leidnud laia rakendust indiviidide tuvastamisel kriminalistikas ja isaduse määramisel.

Selle lühikese ülevaate kokkuvõttena võib konstateerida, et organismide genoomides on avalduvat ja varjatud ning potentsiaalset mitmekesisust kogu elusilma mitmekesisuse ja vormirikkuse seletamiseks.

Kirjandus

- Darwin, C. R. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, John Murray.
- Darwin, C. R. (2012) *Liikide tekkimine loodusliku valiku teel ehk soodustatud rasside säilimine olelusvõitluses*. Tartu, Eesti Loodusuurijate Selts.
- Davies, K. (2001) *Genoome muukides*. Varrak.
- Mendel, G. (1866). Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereins in Brünn*. IV Band.
- Morgan, T. H. (1910) Sex limited inheritance in *Drosophila*. *Science*, **32**, 120–122.
- Stern, C. (1968) *Grundlagen der Humangenetik*. VEB Gustav Fischer Verlag Jena.
- Swift, H. (1950) The constancy of desoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the United States of America*, **36**, 643–654.
- Thomas, C. A. (1971) The genetic organization of chromosomes. *Annual Review of Genetics*, **5**, 237–256.
- Viikmaa, M. (1993) Geneetiline polümorfism populatsioonides ja selle ökoloogiline tähendus. *Schola Biotheoretica*, **19**, 6–12.