

Bakteerigenomien tutkimus

Minna Pirhonen

Soveltavan biologian laitos, PL 27, 00014 Helsingin yliopisto, minna.pirhonen@helsinki.fi

Bakteerien ominaisuuksia

Bakteerit ovat mikroskooppisen pieniä yksisoluisia eliöitä. Niitä on kaikkialla, maaperässä, ruuassa, iholla ja suolistossamme. Niiden muoto ja solunrakenne vaihtelee ja suurta vaihtelua on myös niiden vaatimissa kasvuolosuhteissa. Useimmat bakteerit ovat hyödyllisiä. Esimerkiksi ihon ja ruuansulatuksemme terveydelle on välttämätöntä että oikeanlaiset bakteerit pitävät taudinaiheuttajat kurissa. Monet bakteerit tuottavat teollisesti tärkeitä molekyyilejä.

Bakteereiden näkemiseen tarvitaan mikroskooppia. Ensimmäiset havainnot bakteereista tehtiin 1600-luvulla, kun van Leeuwenhoek tarkasteli hampaankolostaan löytämänsä materiaalia kehittämällään mikroskoopilla. Bakteereita tarkasteltiin sen jälkeen useaan otteeseen, vaikka niiden olemassaolo yleisesti kiistettiin aina 1800-luvun puoleenväliin asti. Silloin Pasteur kiistatta osoitti että ilmassa on pieneliöitä ja että kaikkiin pilaantumis- ja käymisilmiöihin tarvitaan mikrobeita. Kun bakteereiden ja muiden mikrobeiden olemassaolo tunnustettiin, huomattiin myös että ne olivat monen tarttuvan taudin aiheuttajia. Monet ihmiskunnan pahimmat vitsaukset, esim. rutto, kolera, spitaali, tuberkuloosi, syfilis, sekä monet viljelykasveja ja kotieläimiä vaivaavat taudit todettiin bakteereiden aiheuttamiksi.

Bakteereilla tehty tutkimus

Suurin osa tutkimuksesta joka käsittelee bakteereita tehdään ihmisellä tauteja aiheuttavilla tai teollisesti tärkeillä bakteereilla. Maatalouden kannalta tärkeitä ovat bakteerit jotka aiheuttavat tauteja kotieläimillä tai kasveilla, sekä myös bakteerit jotka ovat symbioosissa näiden kanssa tai muuten edesauttavat niiden hyvinvointia. Kaikkien näiden eri tavalla tärkeiden bakteereiden tutkimus on edennyt hyvin samaan tapaan kautta vuosien, sillä uudet menetelmät ja keksinnöt on otettu käyttöön suunnilleen samoihin aikoihin huolimatta siitä tehdäänkö työtä taudinaiheuttajalla vai symbiontilla. Eräs yhteinen tavoite kaikelle tutkimukselle on ollut ymmärtää, miten bakteeri aiheuttaa taudin tai miten bakteerin ja sen isännän välinen vuorovaikutus syntyy symbioosissa tai taudinaiheutuksessa. Sen jälkeen tätä tietoa halutaan soveltaa niin että vuorovaikutusta voidaan muuttaa ihmiselle edulliseen suuntaan, joko taudinaiheuttajan määrän vähentämiseen, sen torjuntaan ja diagnostiikkaan tai symbiontin tai kasvua edistävän bakteerin tehokkuuden parantamiseen. Edellä olevat esimerkit käsittelevät kasvien symbiontteja ja bakteeritaudinaiheuttajia, mutta vastaavaa tutkimusta on tehty kaikilla bakteeriryhmillä.

Bakteereilla tehtävä tutkimus tähtäsi alussa solujen muodon, liikkumistavan ja muiden ominaisuuksien kuvaamiseen. Myös niiden aiheuttamat oireet isännässä tai muutokset kasvualustassa kuvattiin tarkasti. Bakteereiden avulla perehdyttiin myös perinnöllisyystieteen perusteisiin, sillä mutanteja syntyy bakteereilla herkästi. Mutanttien avulla saatiin määritetyksi mitkä perintötekijät eli geenit olivat vastuussa bakteerin eri ominaisuuksista, ja kun yhdistelmä-DNA-tekniikka kehitettiin 70- ja 80-luvuilla saatiin monet näistä geneista kloonattua. Näistä kloonatuista geneista selvitettiin niiden DNA:n emäsjärjestys eli sekvenssi. Sekvenssi kertoo, minkälaisen proteiinin geeni tuottaa, ja proteiinista voidaan päätellä, tai ainakin arvailla mikä on proteiinin tehtävä bakteerisolussa. Parin vuosikymmenen työn tuloksena meillä oli tuhansien bakteerigeenien sekvenssi selvillä, ja tämän tiedon avulla voitiin päätellä kuinka bakteerit aiheuttavat esim. taudin tai saavat aikaan symbioosin kasvin kanssa.

Bakteerien genomitutkimus

Koko eliön laajamittainen sekvensointi aloitettiin viruksista ja bakteereista. Ensimmäisenä bakteerigenomina julkaistiin aivokalvontulehdusta aiheuttavan *Haemophilus influenzae*-bakteerin sekvenssi vuonna 1995 (Fleischmann et al. 1995). Ensimmäistä kertaa avautui mahdollisuus tietää bakteerin kaikki geenit ja ymmärtää sen biologia kokonaisuudessaan. Tällä hetkellä on julkaistu tai ollaan julkaisemassa noin 145:n bakteerin tai arkkibakteerin genomit. Useimmat niistä aiheuttavat

tauteja ihmisellä, mutta joukossa on myös useita kasveilla tauteja aiheuttavia bakteereita ja symbiontteja. Useat niistä ovat trooppisia kuten esimerkiksi *Xanthomonas axonopodis* ja *Xylella fastidiosa* jotka aiheuttavat tauteja *Citrus*-sukuisilla kasveilla. Suomessa kasvitauteja aiheuttavista bakteereista ollaan sekvenssoimassa perunan vaaleaa rengasmätää (*Clavibacter michiganensis* subsp. *sepedonicus*) ja perunan tyvimätää (*Erwinia carotovora* subsp. *atroseptica*) sekä ruusukasveilla kasvaimia aiheuttavaa *Agrobacterium tumefaciens*-bakteeria. Sekvenssitietoa on saatavissa myös useasta kasvien typensitojabakteerista sekä biologisessa torjunnassa käytettävästä *Pseudomonas fluorescens*-bakteerista. Kaikki valmiit julkiset sekvenssit löytyvät verkosta (<http://www.ncbi.nlm.nih.gov/genomes/MICROBES/Complete.html>) kuten myös monet keskeneräiset (www.tigr.org; www.sanger.ac.uk). Monet yksityiset yritykset ovat sekvenssoineet myös bakteerien genomeja, mutta näitä tuloksia ei ole julkisesti saatavilla.

Genomisen sekvenssin sisältämää tietoa voidaan hyödyntää monella tapaa perustutkimuksessa. Kahden lähisukuisen, eri ympäristöissä elävien tai toiminnaltaan erilaisten bakteerien genomeja voidaan verrata, jolloin selvää mitkä geenit ovat tarpeen kaikille bakteereille. Toisaalta saadaan myös tietoa siitä mitkä geenit tarvitaan tietyissä kasvuolosuhteissa tai esim. eri isäntäkasveissa. Esimerkiksi kun kahden isäntäspesifisyydeltään erilaisen *Xylella fastidiosa*-kannan genomia verrattiin, havaittiin että 95,7 % niiden DNA:n koodaamista proteiineista olivat samanlaisia (Van Sluys et al. 2003). Tämä viittaa siihen että nämä patogeetit, joista toinen aiheuttaa kasvitautin viinirypäleelle ja toinen *Citrus*-sukuisille kasveille, käyttävät lähes samaa strategiaa aiheuttaakseen kasvitautin. Nyt niitä geenejä jotka löytyivät vain toisesta, tutkitaan parhaillaan että ymmärrettäisiin kuinka ne vaikuttavat isäntäspesifisyyteen. Niitä voidaan hyödyntää myös molempien tautien diagnostiikassa.

Vertaamalla kasvitauteja aiheuttavien ja kasvien symbionttien genomeja on löydetty geenejä jotka ovat tyypillisiä kaikille bakteereille jotka kasvavat kasvisolukossa (Van Sluys et al. 2002). Useimmat niistä sijaitsevat bakteerin soluseinässä, ja monien tehtävää ei tiedetä. Tutkimuksessa löytyi myös monia yhtäläisyyksiä kasvitauteja aiheuttavien ja symbionttien bakteerien väliltä, ja näyttääkin siltä että symbiontit ja tauteja aiheuttavat bakteerit käyttävät samoja mekanismeja esimerkiksi estääkseen kasvisolujen puolustusmekanismien käynnistymisen. Hämmästyttävä tulos oli myös se, että kasvitauteja aiheuttavilta bakteereilta löytyi laaja skaala toksiineita ja tarttumiseen tarvittavia adhesiineja, joiden tiedetään olevan tärkeitä ihmissolukkoa infektoivien taudinaiheuttajien infektiassa. Tämä tulos antaa osviittaa myös siitä että mekanismit millä bakteerit tartuttavat aiotumallisten eliöiden soluja ovat paljolti samat kaikilla patogeeneilla.

Kun tiedetään mitkä sekvenssit geenien säätelyalueilla tai itse geenissä tarvitaan tietyissä oloissa ilmentymiseen tai erittymiseen, voidaan uusia samoissa oloissa tarvittavia geenejä löytää tietokonehauilla genomisesta sekvenssistä. Tomaatilla tautia aiheuttavalla *Pseudomonas syringae*-bakteerilla oleellinen virulenssimekanismi on sen kyky kuljettaa taudinaiheutukseen tarvittavia proteiineja suoraan kasvisolun sisään, missä ne todennäköisesti lamaavat isäntäkasvin puolustusvasteen. Näitä proteiineja tuottavien geenien säätelyalueet ja proteiinien alkupään eritykseen tarvittavat sekvenssit ovat samankaltaisia. Tietokonehauilla voitiin löytää *Pseudomonas*-bakteerista yli kolmekymmentä proteiinia, jotka bakteeri todennäköisesti siirtää isäntäsolun sisään (Petnicki-Ocwieja et al. 2002; Zwiesler-Vollick et al. 2002). Nyt tutkitaan näiden proteiinien merkitystä taudin syntymisessä.

Myös bakteereilla voidaan geenien ilmentymisen aktiivisuutta tutkia käyttäen hyväksi mikrosirutekniikkaa. Vihanneksia ja koristekasveja mädättävän *Erwinia chrysanthemi*-bakteerin genomista valittiin sattumanvaraisesti 5000 geeniä joiden ilmentymistasoa vertailtiin bakteereissa jotka olivat kasvaneet rikkaalla laboratorioalustalla ja kasvissa (Okinaka et al. 2002). Tutkimuksessa todettiin että useat bakteerin perusaineenvaihdunnan geenit suljettiin kun taas monet kasvimolekyylien hajottamiseen ja/tai kuljettamiseen tarvittavat geenit sekä geenit jotka liittyvät bakteerin puolustusmekanismeihin käynnistyivät infektion aikana. Tässä kuten monessa muussakin tutkimuksessa löytyi kymmeniä tehtävältään tuntemattomia geenejä, jotka indusoituvat taudinaiheuttajan kasvaessa isäntäkasvissa.

Bakteerien proteomiikkatutkimus

Genomitutkimuksissa löytyvät uudet tuntemattomat geenit ovat toisaalta paljon odotettu tutkimustulos, onhan tarkoitus saada uutta tietoa, mutta toisaalta niiden tehtävän selvittäminen tulee todennäköisesti

viemään paljon voimavaroja tulevaisuudessa. Ensimmäinen haaste on saada selville, mitkä niistä todella tuottavat toimivan proteiinin ja kuinka ne toimivat yhdessä.

Proteomiikkatutkimuksella voidaan tunnistaa solun tietyissä oloissa tuottamat proteiinit. Proteomiikka-analyysillä tarkoitetaan kaksisuuntaisen (2D) geelielektroforeesin ja ultraherkän massaspektrometrian yhdistelmää. 2D-geeleissä proteiinit eritellään ensin proteiinin varauksen mukaan ja sitten niiden massan perusteella, ja näin sadatkin näytteen sisältämät proteiinit saadaan toisistaan erilleen ns. proteiinkartaksi. Vertailemalla eri olosuhteissa kasvaneista soluista valmistettuja proteiininäytteitä saadaan selville, mitä proteiineja tuotetaan eri oloissa. Viimeisessä vaiheessa, kun kiinnostavat proteiinit on valittu, pilkotaan ne entsyymaattisesti peptideiksi, joiden massat tunnistetaan massaspektrometrialla. Vertaamalla peptidien massoja tietokoneen laskemiin genomisesta sekvenssistä laskemiin massoihin voidaan proteiinit tunnistaa. Tämä tunnistus on mahdollista tehdä kymmenille ja sadoille proteiineille vain jos eliön genominen sekvenssi on tunnettu. Vasta proteiinien tutkiminen kertoo lopullisesti mitä proteiineja perimän valtava informaatiomäärä todella koodaa ja miten proteiineja muokataan synteessin jälkeen. Tällä menetelmällä saadaan myös tietoa siitä mitkä proteiinit toimivat kompleksina tai samoissa olosuhteissa. Se missä proteiiniympäristössä proteiini toimii, kertoo todennäköisesti mihin toiminnalliseen kokonaisuuteen se liittyy, ja tämä taas voi helpottaa proteiinin funktion selvittämistä.

Proteomiikkatutkimuksella on todettu *Xylella fastidiosa*-bakteerin tuottavan hyvin monenlaisia tarttumiseen tarvittavia molekyyliä, jotka muodostavat todennäköisesti proteiini-kompleksin bakteerin pinnalle (Smolka et al. 2003). *Sinorhizobium meliloti*, sinimailasen juurinysträbakteerin genomista on löydetty proteomiikka-analyysillä uusia symbioosin aikana toimivia proteiineja, joiden osuutta typensitomiseen tarvittaviin toimintoihin ei olisi pelkän sekvenssin perusteella voinut arvata (Djordjevic et al. 2003). Samalla todettiin tässä niin kuin monessa muussakin tutkimuksessa että monet geenit tuottavat enemmän kuin yhden toisistaan poikkeavan proteiini-muodon, tulos mitä ei myöskään voi luotettavasti todeta pelkän sekvenssin perusteella.

Tulosten hyödyntäminen

Bakteerien genomitutkimus on tuottanut lyhyessä ajassa suuren määrän uutta tietoa jota voidaan käyttää soveltavassa tutkimuksessa. DNA-sekvenssejä tai proteiineja jotka löytyvät vain taudinaiheuttajasta mutta eivät lähisukulaisista vaarattomista bakteereista voidaan hyödyntää diagnostiikassa. Kun bakteereista löydetään uusia, niiden taudinaiheuttamiskyvylle oleellisia proteiineja, voidaan niiden toimintaa pyrkiä estämään. Monet bakteerit tuottavat muille bakteereille myrkyllisiä yhdisteitä joita voidaan ehkä hyödyntää kehitettäessä antibiootteja korvaavia hoitomuotoja. Bakteereiden tuottamista uusista proteiineista paljastuu varmasti myös uusia, teollisuutta hyödyttäviä entsyymeitä. Genomitutkimuksen antamat tulokset ovat kuitenkin niin uusia, että niiden todellinen hyödyntäminen on vielä edessä.

Referenssit:

Djordjevic MA, Chen HC, Natera S, Van Noorden G, Menzel C, Taylor S, Renard C, Geiger O, Weiller GF, 2003. A global analysis of protein expression profiles in *Sinorhizobium meliloti*: discovery of new genes for nodule occupancy and stress adaptation. MPMI 16(6): 508-524.

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269(5223): 496-512.

Petnicki-Ocwieja T, Schneider DJ, Tam VC, Chancey ST, Shan L, Jamir Y, Schechter LM, Janes MD, Buell CR, Tang X, Collmer A, Alfano JR. 2002. Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. *tomato* DC3000. PNAS 99(11): 7652-7657.

Okinaka Y, Yang CH, Perna NT, Keen NT, 2002. Microarray profiling of *Erwinia chrysanthemi* 3937 genes that are regulated during plant infection. MPMI 15(7): 619-629.

Van Sluys MA, Monteiro-Vitorello CB, Camargo LE, Menck CF, Da Silva AC, Ferro JA, Oliveira MC, Setubal JC, Kitajima JP, Simpson AJ, 2002. Comparative genomic analysis of plant-associated bacteria. Annu. Rev.

Phytopathol. 40: 169-189.

Van Sluys MA, de Oliveira MC, Monteiro-Vitorello CB, Miyaki CY, Furlan LR, Camargo LE, da Silva AC et al., 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J. Bacteriol. 185(3): 1018-1026.

Smolka MB, Martins D, Winck FV, Santoro CE, Castellari RR, Ferrari F, Brum IJ, Galembeck E, Della Coletta Filho H, Machado MA, Marangoni S, Novello JC, 2003. Proteome analysis of the plant pathogen *Xylella fastidiosa* reveals major cellular and extracellular proteins and a peculiar codon bias distribution. Proteomics 3(2): 224-37.

Zwiesler-Vollick J, Plovianich-Jones AE, Nomura K, Bandyopadhyay S, Joardar V, Kunkel BN, He SY, 2002. Identification of novel hrp-regulated genes through functional genomic analysis of the *Pseudomonas syringae* pv. *tomato* DC3000 genome. Mol. Microbiol. 45(5): 1207-1218.