

# **Epigenetiikka ja kaksostutkimus**

**Petri Vänni**

**15.6.2016**

**Avainsanat:**

**epigenetiikka, kaksoset, kaksostutkimus, metylaatio, histonimodifikaatiot**

## Sisällysluettelo

- 1. Johdanto**
- 2. Epigeneettiset mekanismit**
  - 2.1 DNA:n metylaatio eläimillä**
  - 2.2 DNA:n metylaatio kasveilla**
  - 2.3 Histonimodifikaatiot**
  - 2.4 Ei-koodaavat RNA:t**
- 3. Epigenetiikan analysointi**
  - 3.1 DNA:n metylaation mittaaminen**
  - 3.2 Histonimodifikaatioiden mittaaminen**
  - 3.3 Ei-koodaavien RNA:den mittaaminen**
- 4. Kaksosten muodostuminen ja kaksoskohortit**
  - 4.1 Kaksosten muodostuminen**
  - 4.2 Kaksoskohortit**
- 5. Kaksosten käyttö epigenetiikan tutkimuksissa**
  - 5.1 Klassinen kaksosmalli tutkimuksessa**
  - 5.2 DNA:n metylaatio tautitapauksissa**
- 6. Yhteenveto**

### **1. Johdanto**

Termin ”epigenetiikka” merkitys on vuosien aikana muuttunut paljon ja vieläkin siltä puuttuu yhtenäinen määritelmä. Conrad Waddington määritteli epigenetiikan vuonna 1942: ”the branch of biology that studies the causal interactions between genes and their products which bring the phenotype into being” (Deans & Maggert 2015, Waddington 2012). Nykyään epigenetiikka voidaan määritellä mahdollisesti palautuvaksi geenien säätelyksi, joka ei toimi DNA-sekvenssin muutosten kautta ja jota ympäristön ärsykkeet voivat muokata (Bell & Spector 2011, Wong et al. 2010).

Kaikki epigeneettiset muutokset, kuten DNA:n metylaatiokohdat, histonimodifikaatiot ja ei-koodaavat RNA:t, muodostavat yhdessä kokonaisuuden, jota sanotaan yksilön epigenomiksi.

Nämä epigeneettiset mekanismit toimivat keskeisessä osassa transkription säätelyssä ja kudosten erilaistumisessa. Tästä johtuen epigeneettiset muutokset ovat usein kudostai solukohtaisia. Häiriöt epigeneettisten mekanismien normaalissa toiminnassa aiheuttavat useita erilaisia sairauksia, kuten syöpää, genomisen leimautumisen häiriöitä ja useita neuropsykiatrisia häiriöitä (Wong et al. 2010). Ympäristön arvellaan vaikuttavan paljon siihen, millainen epigenomi yksilölle muodostuu, DNA:n metylaation erojen on esimerkiksi huomattu lisääntyvän yksilön iän ja ympäristön mukaan (Wong et al. 2010). Epigeneettisten muutosten plastisuus ja kudostai solukohtaisuus vaikeuttavat epigeneettisten mekanismien ja epimutaatioiden tutkimista.

Monotsygoottiset (MZ) kaksoset ovat DNA-sekvenssiltaan identtisiä, poissulkien harvinaiset sattumanvaraiset mutaatiot. Tästä syystä erojen MZ-kaksosten välillä päätellään yleensä johtuvan ympäristöstä. Kaksostutkimuksen on havaittu olevan yksi tehokkaimmista työkaluista epigenetiikan ongelmien tutkimiseen. Kaksosien avulla pystytään erottamaan toisistaan geneettisen, yhteisen ympäristön ja yksilöille ainutlaatuisen ympäristön vaikutus fenotyypin (Bell & Spector 2011). MZ-kaksosia käyttämällä tautitutkimuksissa pystytään kontrolloimaan geneettisiä tekijöitä, ikää, sukupuolta ja ympäristön vaikutusta (kohdunsisäinen ympäristö ja elinympäristö) (Castillo-Fernandez et al. 2014). Kaksostutkimuksella pyritään selvittämään kuinka suuren osan fenotyypin vaihtelusta voidaan selittää epigeneettisillä tekijöillä.

Käsittelen DNA:n metylaatiota myös *Arabidopsis* – kasvien näkökulmasta. *Arabidopsis* – kasvit ovat itsesiittoisia ja niiden jälkeläiset ovat toisiinsa verrattuna hyvinkin homologisia. Geneettinen homologisuus on verrattavissa kaksosiin ja tästä syystä nämä kasvit voivat olla hyödyllisiä monissa erilaisissa epigenetiikan tutkimuksissa. Kasveja on yleensä helpompi ja edullisempi tutkia verrattuna kaksosiin. Useiden epigeneettisten mekanismien tarkoitus solujen toiminnassa on täysi mysteeri. Monet suuret läpimurrot tieteessä on tehty ”Out of the box” – ajattelulla, joten on varmasti hyödyllistä etsiä ratkaisuja näihin kysymyksiin myös muista organismeista.

## 2. Epigeneettiset mekanismit

### 2.1 DNA:n metylaatio eläimillä

Yksi parhaiten tunnetuista epigeneettisistä mekanismeista on DNA:n metylaatio. DNA:n metylaatiota löytyy eniten sentromeeristä, telomeeristä, inaktivoituista X-kromosomista ja toistojaksoista. DNA metylaatiolla tarkoitetaan yleensä sytosiinin metylaatiota 5-metyylisytosiiniksi. Selkärankaisilla eläimillä ja hyönteisillä DNA metylaatiota tapahtuu usein vain CpG-dinukleotideissa, mutta kasveilla se voi tapahtua CHG ja CHH (H on joko adenosiini, tymiini tai sytosiini) trinukleotideissa (Feil & Fraga 2012). Vastinjuosteen CpG:n metyloi mitoosin jälkeen *maintenance methyltransferase* -proteiini, joka tunnistaa hemimetyloidun (vain toinen juoste metyloitu) kohdan kaksoisjuosteessa ja metyloi vastinjuosteen. Metylaatio ei vaikuta emäksen kykyyn sitoutua vastinjuosteeseen. Se voi säädellä geenien ilmenemistä estämällä promoottorialueella transkriptiota aloittavien entsyymien sitoutumisen promoottoriin. DNA metylaatio aiheuttaa kromatiinin pakkautumista ja geenien hiljentymistä ja X-kromosomin inaktivoitumista naaraissa ja suojaa genomia patogeenisiltä retroviruksilta (Wong et al. 2010). Genomissa loisivat retrovirukset voidaan hiljentää metyloimalla retrovirusten toiminnalle tärkeät pitkät terminaaliset toistojaksot (long terminal repeats). Metylaation on havaittu hiljentävän transposoneita *Arabidopsis thaliana*:ssa (Cokus et al. 2008).

Melkein kaikki metyyliryhmät pyyhitään genomista pian hedelmöityksen jälkeen hiljentämällä *maintenance methyltransferase* -proteiini (passiivinen demetylaatio), tai demetyloimalla aktiivisesti entsyymeillä (aktiivinen demetylaatio). Toinen merkitsevä genomien laajuinen demetylaatio tapahtuu ituradan kehityksen aikana (Alberts et al. 2015, Heard & Martienssen 2014). Demetylaation jälkeen genomiin tehdään uusia DNA metylaatiokuvioita *de novo* käyttäen DNA metyylitransferaaseja (DMNT).

Eukaryooteista on löydetty kolmea erilaista DNA metyylitransferaasia: DMNT1, DMNT3a ja DMNT3b. DMNT1 on vastuussa metylaation ylläpidosta (*maintenance methyltransferase*). DMNT3a ja DMNT3b – entsyymit pystyvät myös metyloimaan hemimetyloitunutta DNA:ta, mutta toimivat soluissa lähinnä *de novo*, eli tuottavat uusia metylaatioita.

5-metyylisytosiinin lisäksi syntetisoidaan myös muita sytosiinin modifikaatioita (taulukko 1.). Näitä ovat 5-hydroksyylimetyylisytosiini, 5-formyyliytosiini ja 5-karboksyylisytosiini, joita syntetisoi TET – geeniperhe (ten-eleven translocation) (Dawson & Kouzarides 2012). Vaikka näiden muutoksien tarkkaa funktiota ei tiedetä, ne todennäköisesti vaikuttavat passiiviseen ja aktiiviseen demetylaatioon, jotka helpottavat tai vaikeuttavat MBD:n (methylation binding protein) toimintaa (Wu & Zhang 2011).

## 2.2 DNA:n metylaatio kasveilla

Metylaation sijaintia ja ilmenemistapoja on tutkittu *Arabidopsis* – suvun kasveilla (Cokus et al. 2008). *Arabidopsis thaliana*:n ja *Arabidopsis lyrata*:n DNA:n metylaatiot sekvensoitiin käyttämällä koko genomien bisulfaattisekvensointi – tekniikka. Tekniikalla saatiin sekvensoitua 93 % genomien kaikista sytosiineista. CG-dinukleotidien metylaatiota löytyi selvästi enemmän proteiineja koodaavien geenien, pseudogeenien ja transposonien transkriptioalueilta enemmän kuin niitä ympäröiviltä alueilta (Cokus et al. 2008). Myös käänteisten ja tandem – toistojaksojen toistoista löytyi enemmän CG, CHG ja CHH – metylaatiota kuin niitä ympäröivistä alueista. CHG ja CHH – metylaatiokohtia löytyi keskimäärin vähemmän proteiineja koodaavilta transkriptioalueilta, mikä osoittaa näiden metylaatiokohtien säätelevän muita genomien alueita. Telomeerien CCCTAA – toistojaksosta löytyi myös metylaatiota, joista suurin osa havaittiin toistojakson kolmannessa sytosiinissa (Cokus et al. 2008).

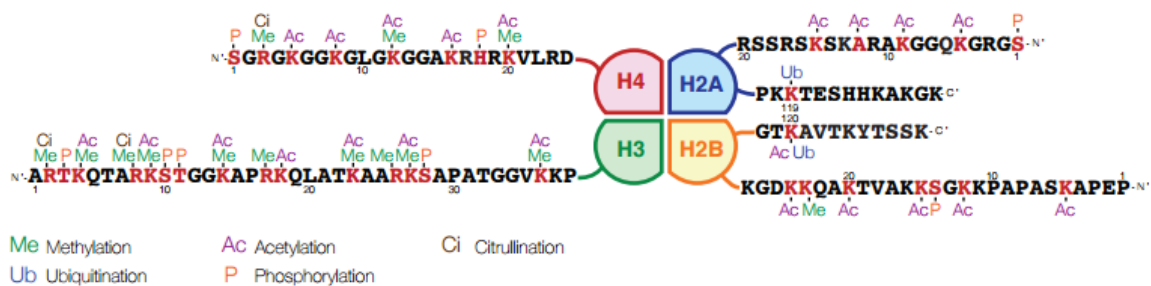
DNA:n metylaation kontekstia tutkittaessa *Arabidopsis* -kasveilla huomattiin, että metyloituneita CHH:ta ja CHG:ta löytyi todennäköisemmin muutama nukleotidi CpG-dinukleotidien jälkeen (Cokus et al. 2008). Metylaatiota havaittiin tapahtuvan säännöllisesti 167 nukleotidien välein, mikä on kasvien nukleosomia (175-185 nukleotidia) hiukan lyhyempi. Näiden metylaatiokohtien arvellaan vaikuttavan kromatiinin rakenteeseen, mahdollisesti lyhentäen histonien välistä DNA:ta tai vähentämällä histonien väliseen DNA:han sitoutuvien proteiinien määrää (Cokus et al. 2008). CpG-dinukleotidien havaittiin myös säännöllisesti olevan kymmenen nukleotidin päässä toisistaan. Tämä voi johtua kahden metyyli transferaasin muodostamasta Dnmt3a-Dnmt3L – entsyymikompleksin toiminnasta. Dnmt3a-Dnmt3L – kompleksilla on kaksi DNA:ta metyloivaa reaktiokohtaa 8-10 nukleotidin päässä toisistaan (Cokus et al. 2008). *Arabidopsis* – kasveilla on myös tutkittu metylaatioprofiilien eroja paikallisen sopeutumisen yhteydessä (Dubin et al. 2015). Tutkimuksessa löydettiin geenien transkriptioalueella (gene body methylation) eroja CG-metylaatioprofiileissa leveyspiirin mukaan. CHH-metylaatiota löydettiin pääasiassa transposoneista ja niiden arvellaan liittyvän transposonien hiljentämiseen (Dubin et al. 2015).

## 2.3 Histonimodifikaatiot

DNA:n täytyy olla helposti saatavilla transkriptiota varten sekä suojattu ulkoisilta tekijöiltä. Genomi sisältää miljoonia emäspareja DNA:ta, joten se on erittäin alttiina sattumanvaraisille

kemiallisille tekijöille, jotka vaurioittavat sitä ja aiheuttavat epäsuotuisia mutaatioita. Genomi on pakattu soluissa kromatiiniksi, joka koostuu miljoonista nukleosomeista. Nukleosomi on DNA:n ja proteiinin muodostama kompleks, joka koostuu kahdeksasta histoniproteiinista (kaksi kappaletta H2A, H2B, H3 ja H4 -proteiineja) ja 147 emäsparista kietoutuneena histonin ympärille. Jokaisella histoniproteiinilla on N-terminaalinen sivuketju, joissa esiintyy kovalenttisia muutoksia, kuten lysiinien asetylaatio ja mono-, di-, ja trimetylaatio, sekä seriinien fosforylaatio. Nämä kovalenttiset modifikaatiot (taulukko 1 ja kuva 1) ovat tärkeitä epigeneettisiä tekijöitä, jotka ovat osa solun epigenomia. Histonimodifikaatiot voivat säädellä transkriptiota muuttamalla DNA:n ja histoniproteiinien sitoutumista toisiinsa. Sääntely voi tapahtua myös värväämällä paikalle transkriptiolle tärkeitä entsyymejä tai muita säätelytekijöitä.

Histonien sivuketjun lysiinien asetylaatio on yksi merkittävimmistä histonien modifikaatioista, joka vaikuttaa kromatiinin rakenteeseen, transkriptioon ja DNA:n korjaukseen. Asetylaatio poistaa lysiinien negatiivisen varauksen, mikä heikentää histonien sitoutumista negatiivisesti varautuneeseen DNA:han, mikä puolestaan aiheuttaa kromatiinin dekontensaatiota. Tästä syystä asetylaatiota esiintyy geenien promoottorien alueella ja joissain tapauksissa koko geenin pituudelta (Dawson & Kouzarides 2012).



**Kuva 1. Histoniproteiinien sivuketjujen kovalenttisiä modifikaatioita.** Erilaisia histonimodifikaatioita neljän histoniproteiinin sivuketjussa. (Mariuswalter 2015, Wikimedia commons)

Histonien asetylaatiota säätelee kaksi geeniperhettä, joista toinen asetyloisi lysiniä (*lysine acetyltransferase* eli KAT) ja toinen deasetyloisi (*histone deacetylases* eli HDAC). KAT-geeniperheen geenit voidaan jakaa kahteen ryhmään; solulimassa vapaina olevia histoneita asetyloivaan ryhmään ja tumassa histoneita asetyloivaan. KAT-geeniperheen entsyymit olivat ensimmäisiä entsyymejä joiden havaittiin muuttavan histonien rakennetta (Dawson & Kouzarides 2012). HDAC-geeniperheen entsyymit poistavat asetylaatiota eli deasetyloivat

N-terminaalien sivuketjuja aiheuttaen positiivisen varauksen histoneissa, joka aiheuttaa histonin voimakkaan sitoutumisen negatiivisesti varautuneeseen DNA:han (Dawson & Kouzarides 2012).

Histonien metylaatiota esiintyy N-terminaalien sivuketjujen arginiineissa, lysiineissa ja histidiineissa (Dawson & Kouzarides 2012). Nämä metylaatiot eivät muuta histonien varausta. Sivuketjujen lysiineja metyloivat KMT-geeniperheen (*lysine methyltransferase*) entsyymit. Histonien demetylaatiosta vastaavat *lysiini demetylaasit*. Demetylaasit voidaan jakaa kahteen eri ryhmään, joista ensimmäinen ryhmä demetyloi mono- ja dimetyloituvia lysiineja. Toinen ryhmä, nimeltään Jumonji demetylaasit, pystyvät demetyloimaan mono-, di- ja trimetyloituvia lysiineja (Dawson & Kouzarides 2012).

Histonien sivuketjujen fosforylaatiota tapahtuu seriineissa, treoniineissa ja tyrosiineissa (Dawson & Kouzarides 2012). Histonien sivuketjujen fosforylaatiot vaikuttavat histonin varaukseen, sekä ovat tärkeitä mitooseen, apoptosikseen eli solukuolemaan, DNA:n korjaukseen, replikaation ja transkription säätelijöitä. Sivuketjun fosforylaatiokohdat voidaan jakaa funktioltaan kahteen eri ryhmään; transkriptiota sääteleviin ja kromatiinin kondensaation säätelyyn osallistuviin (Dawson & Kouzarides 2012). Fosforylaatiota säätelevät proteiinikinaasit ja proteiinifosfataasit.

DNA:n modifikaatiot	Termistö	Modifikaation funktio
5-metyylisytyosiini	5mC	Transkriptio
5-hydroksyylimetyylisytyosiini	5hmC	Transkriptio
5-formyylysytyosiini	5fC	-
5-karboksyylisytyosiini	5caC	-
Histonimodifikaatiot		
Asetylaatio	K-ac	Transkriptio, korjaus, replikaatio ja kondensaatio
Metylaatio (lysiini)	K-me1, K-me2, K-me3	Transkriptio, korjaus
Metylaatio (arginiini)	R-me1, R-me2s, R-me2a	Transkriptio
Fosforylaatio (seriini, treoniini)	S-ph, T-ph	Transkriptio, korjaus ja kondensaatio
Fosforylaatio (tyrosiini)	Y-ph	Transkriptio ja korjaus
Ubikitiinilaatio	K-ub	Transkriptio ja korjaus

**Taulukko 1. Luettelo DNA:n modifikaatioista ja histonimodifikaatioista.** Perustuu taulukkoon 1. artikkelista (Dawson & Kouzarides 2012)

## 2.4 Ei-koodaavat RNA:t

Transkriptiota tapahtuu melkein koko genomin alueella ja suurin osa transkriptiotuotteista on ei-koodaavaa RNA:ta (ncRNA) eli ne eivät käy läpi translaatiota (Dawson & Kouzarides

2012). Ei-koodaavat RNA:t voidaan jakaa pieniin, alle 200 nukleotidin pituisiin, sekä pitkiin ei-koodaaviin RNA:hin (lncRNAs).

Pienet ei-koodaavat RNA:t ovat hyvin konservoituneet eliölajien välillä, joten niiden funktio on organismeille tärkeä ja muutokset niiden rakenteeseen yleensä ovat haitallisia. Pieniin ei-koodaaviin RNA:hin kuuluu snoRNA:t (small nucleolar RNAs), piRNA:t (PIWI-interacting RNAs), siRNA:t (small interfering RNA:t) ja miRNA:t (microRNAs) (Dawson & Kouzarides 2012). Niiden arvellaan hiljentävän geenejä transkriptionaalisesti ja posttranskriptionaalisesti. Pitkät ei-koodaavat RNA:t ovat pienempiä ei-koodavia RNA:ta vähemmän konservoitu eliölajien välillä ja ne säätelevät transkriptiota (Dawson & Kouzarides 2012).

### **3. Epigenetiikan analysointi**

#### **3.1 DNA:n metylaation mittaaminen**

DNA:n metylaation tutkimisessa käytetään natrium bisulfaatteja, jotka deaminoivat metyyllittomat sytosiinit urasiiliksi jättäen 5-metyylisytosiinit ennalleen (Mensaert et al. 2014). Bisulfaatti konversio jättää ennalleen myös hydroksyylimetyylisytosiinit, joka voidaan hapettaa formyylysytosiiniksi, jonka jälkeen natrium bisulfaatit deaminoivat formyylysytosiinin urasiiliksi. Tarvittaessa 5-hydroksyylimetyylisytosiineja voidaan tutkia TET-avusteisella bisulfiitti sekvensoinnilla, jossa 5-hydroksyylimetyylisytosiinit glykolisoidaan. Tämän jälkeen TET-entsyymi karboksyloi 5-metyylisytosiinit, mutta ei glykolisoituja 5-hydroksyylimetyylisytosiineja. Bisulfaatit aminoivat muodostuneet karboksyylisytosiinit ja metyyllittomat sytosiinit urasiiliksi.

Natrium bisulfaatilla käsitelty DNA voidaan analysoida genomilaajuisella bisulfaattisekvensoinnilla (whole genome bisulfite sequencing, WGBS), jolla pystytään erottamaan jokainen metyylysytosiini genomissa (Mensaert et al. 2014). Huono puoli DNA:n metylaation tutkimisessa WGBS:llä ja bisulfaattikäsitelyllä on niiden kallis hinta ja se, että PCR-tekniikan käyttö saattaa muuttaa tutkittavan alueen metylaatiota (Mensaert et al. 2014). RRBS-tekniikalla (Reduced representation bisulfite sequencing) pilkotaan DNA pieniksi pätkiksi käyttäen endonukleasi -entsyymeitä. Pilkottujen DNA:n pätkien päissä on CpG:itä. Tutkittavat pätkät valitaan elektroforeesilla ja käsitellään natrium bisulfaatilla, jonka jälkeen DNA:n pätkät ovat valmiita sekvensoitavaksi (Mensaert et al. 2014).



DNA:n metylaatiokohtia voidaan tutkia erilaisilla bead array -koettimilla, jotka voivat tunnistaa noin 500 000 bisulfaatti konversion kautta muodostunutta nukleotidipolymorfismia (Mensaert et al. 2014). Tällä tekniikalla mitataan koettimien avulla tutkittavassa DNA-alueessa metyylisytyosiinit ja metyylittomia sytosiineja edustavat tyymiinit.

Vaihtoehtoinen tapa WGBS:n ja bead array:n käytölle on MSAP-tekniikka (methylation-specific amplified polymorphisms), joka on paljon edullisempi kuin WGBS. MSAP:ssa käytetään hyväksi kahta isoskitsomeerirestriktioentsyymeitä, jotka molemmat sitoutuvat DNA:han samasta tunnistussekvenssistä, mutta sitoutuvat erilaisilla metyylisytyosiinin sisältävään tunnistussekvenssin DNA:han (Baurens et al. 2008). Esimerkiksi *HpaII* ja *MspI* isoschizomeereistä, joiden tunnistussekvenssi on 5' CCGG 3', *HpaII* ei pysty leikkaamaan C<sup>m</sup>CGG metyloituneen tunnistussekvenssin molekyyliä (Baurens et al. 2008). Restriktion jälkeen sekvenssit moninkertaistetaan PCR:llä ja erotellaan elektroforeesilla ja DNA:n metyloituminen tunnistetaan isoskitsomeerien aiheuttamien erojen kautta (Mensaert et al. 2014).

Koko genomin metylaation tutkimiseen yritetään jatkuvasti etsiä parempia ja edullisempia keinoja. MethylCap-seq on osoittanut olevan edullinen vaihtoehto WGBS:lle. MethylCap-seq perustuu metyloituja DNA:n pätkiä kaappaaviin GST-MBD – fuusioproteiineihin. GST-MBD koostuu MBD:stä (methyl binding domain), joka on ihmisen MeCP2 –proteiinin metyyliä sitova domeeni ja GST:stä (Glutathione-S-transferase), joihin on kiinnittynyt magneettinen molekyyli (Brinkman et al. 2010). MethylCap-seq:ssä tutkittava genomisen DNA hajotetaan sonikaattorilla pieniksi pätkiksi, jonka jälkeen metyylisytyosiinejä sisältävät pätkät sitoutuvat paikalleen magneettisesta molekyylistä ankkuroituihin GST-MBD –fuusioproteiineihin alhaisessa suolakonsentraatiossa. Pätkät joissa ei ole metylaatiota huuhdellaan pois ja MBD:hen sitoutuneet pätkät kerätään talteen ja sekvensoidaan (Brinkman et al. 2010).

### 3.2 Histonimodifikaatioiden mittaaminen

Nukleosomit koostuvat histoniproteiineista ja noin 147 emäsparista DNA:ta joka on kietoutunut niiden ympärille. Nukleosomien paikkaa genomissa ja histonien modifikaatiota tutkitaan pääasiallisesti vasta-aineilla, jotka sitoutuvat spesifisesti tutkittavaan modifikaatioon tai nukleosomiin, kuten ChIP-seq –tekniikalla (chromatin immunoprecipitation) (Mensaert et al. 2014).

Histonimodifikaatioita tutkittaessa, ChIP-seq aloitetaan yleensä sitomalla tutkittavat histoniproteiinit ja DNA toisiinsa kovalenttisilla sidoksilla (O'Geen et al. 2011). Tämän jälkeen näyte fragmentoidaan ja tutkittavat molekyylit kaapataan spesifisesti tutkittavaan histonimodifikaatioon sitoutuvilla vasta-aineilla. Aikaisemmassa vaiheessa muodostetut kovalenttiset sidokset histoniproteiinien ja DNA:n välillä puretaan. Tämän jälkeen DNA erotetaan histoniproteiineista ja sekvensoidaan (O'Geen et al. 2011). Nukleosomien sijainti genomissa voidaan paikantaa käyttäen sekvensoituja noin 147 emäsparin pätkiä.

Nukleosomiprofiili voidaan luoda MNase-seq -tekniikalla pilkkomalla genomi MNaaseilla (Micrococcal nuclease) nukleosomien välistä. Tämän jälkeen pätkät sekvensoidaan ja muodostetaan nukleosomiprofiili vertaamalla pätkiä referenssigenomiin (Mensaert et al. 2014).

### 3.3 Ei-koodaavan RNA:den mittaaminen

Transkriptomi on kaikkien transkriptiossa muodostuneiden tuotteiden kokonaisuus, jonka tutkimista varten on kehitetty monenlaisia keinoja. Strategioita, jotka perustuvat tutkittavan molekyylin hybridisoitumisesta mikrokoettiin, käytetään paljon tiedossa olevien geenien transkriptiotuotteiden tai mRNA:n vaihtoehtoisten silmukoinnin tutkimisessa (Wang et al. 2009). Näiden tekniikoiden käyttö on suhteellisen edullista, mutta ne nojaavat genomien sekvenssin pohjalta tarkasti suunniteltuihin koettiin. Ei-koodaavan RNA:n tutkimisessa pitää ottaa huomioon, että ne eivät mRNA:n tavoin ole yleensä polyadenoloituja, mikä rajaa pois oligo-dT – alukkeisiin perustuvat lähestymistavat (Mensaert et al. 2014).

RNA-seq – tekniikka käyttämällä muodostetaan solun kaikista RNA-fragmenteista cDNA-kirjasto ja liitetään sekvensointia varten adapterit fragmenttien päihin (Wang et al. 2009). Tämän jälkeen cDNA:t sekvensoidaan ja linjataan referenssigenomia vasten. Transkriptomin tutkiminen cDNA:n avulla saattaa aiheuttaa cDNA-artifakteja ja vääristää transkriptomin eri molekyylien määriä *in vivo* (Yu et al. 2011).

Tiling microarray – tekniikkaa käyttäessä ei aiheudu cDNA-kirjaston käyttöön liittyviä ongelmia. Siinä RNA:n erottamisen jälkeen genomista DNA:sta, käsitellään RNA kemikaaleilla, joka lisää alkyyliryhmän guaniinin, adeniinin ja sytosiinin (Yu et al. 2011). Käsitelyn jälkeen RNA:t puhdistetaan ja hybridisoidaan koettimille, jotka tunnistavat alkyyliryhmät.

## 4. Kaksosten muodostuminen ja kaksoskohortit

### 4.1 Klassinen kaksosmalli tutkimuksissa

Klassinen kaksosmalli pyrkii erottamaan monotsygoottisten- ja ditsygoottisten kaksosten avulla geneettisten tekijöiden ja ympäristötekijöiden vaikutuksen tutkittavaan piirteeseen tai fenotyyppiin. Klassisessa kaksosmallissa tehdään useita oletuksia, jotka eivät välttämättä pidä paikkaansa jokaisessa tutkimusasetelmassa (Rijsdijk & Sham 2002). Oletukset ovat, että MZ- ja DZ-kaksoset jakavat hyvin samanlaisen ympäristön, geenien ja ympäristön väliset korrelaatiot tarkasteltavan piirteen suhteen ovat vähäiset, kaksoset eivät eroa muusta populaatiosta tarkasteltavan piirteen suhteen ja pariutuminen populaatioissa on sattumanvaraista (Rijsdijk & Sham 2002).

Kaksostutkimuksissa totaalinen fenotyyppinen varianssi (Var P) voidaan jakaa seuraaviin neljään komponentteihin:

- 1) Additiivinen geneettinen varianssi (Var A) edustaa vaihtelua fenotyyppissä, joka perustuu alleelien additiiviseen vaikutukseen.
- 2) Dominanssi varianssi (Var D), ei-additiivinen geneettinen tekijä, mikä edustaa alleelien vuorovaikutusten vaihtelun vaikutusta fenotyyppiin.
- 3) Yhteiset ympäristötekijät (Var C), jotka edustavat kaksosten yhteisen ympäristön aiheuttaman vaihtelua fenotyyppissä.
- 4) Ainutlaatuiset yksilölliset ympäristötekijät (Var E), edustavat kaksosten ympäristössä esiintyvien eroavaisuuksien vaihtelua fenotyyppissä.

Fenotyyppinen kokonaisvarienssi piirteen tai fenotyypin suhteen on kaikkien neljän komponentin summa eli:  $(\text{Var P} = \text{Var A} + \text{Var D} + \text{Var C} + \text{Var E})$  (Rijsdijk & Sham 2002).

Klassinen kaksosmalli mahdollistaa fenotyypin kokonaisvarienssin komponenttien arvioinnin. MZ-kaksosilla ei ole geneettistä vaihtelua, koska pari on DNA-sekvenssiltään identtisiä ( $\text{Var A} = 1$  ja  $\text{Var D} = 1$ ). DZ-kaksosten Var A-komponentti on  $\frac{1}{2}$  ja Var D-komponentti  $\frac{1}{4}$ , sillä he ovat geneettisesti samanlaisia kuin normaalit sisarukset (Rijsdijk & Sham 2002). MZ- ja DZ-kaksosilla ei ole vaihtelua yhteisen ympäristön suhteen, sillä he

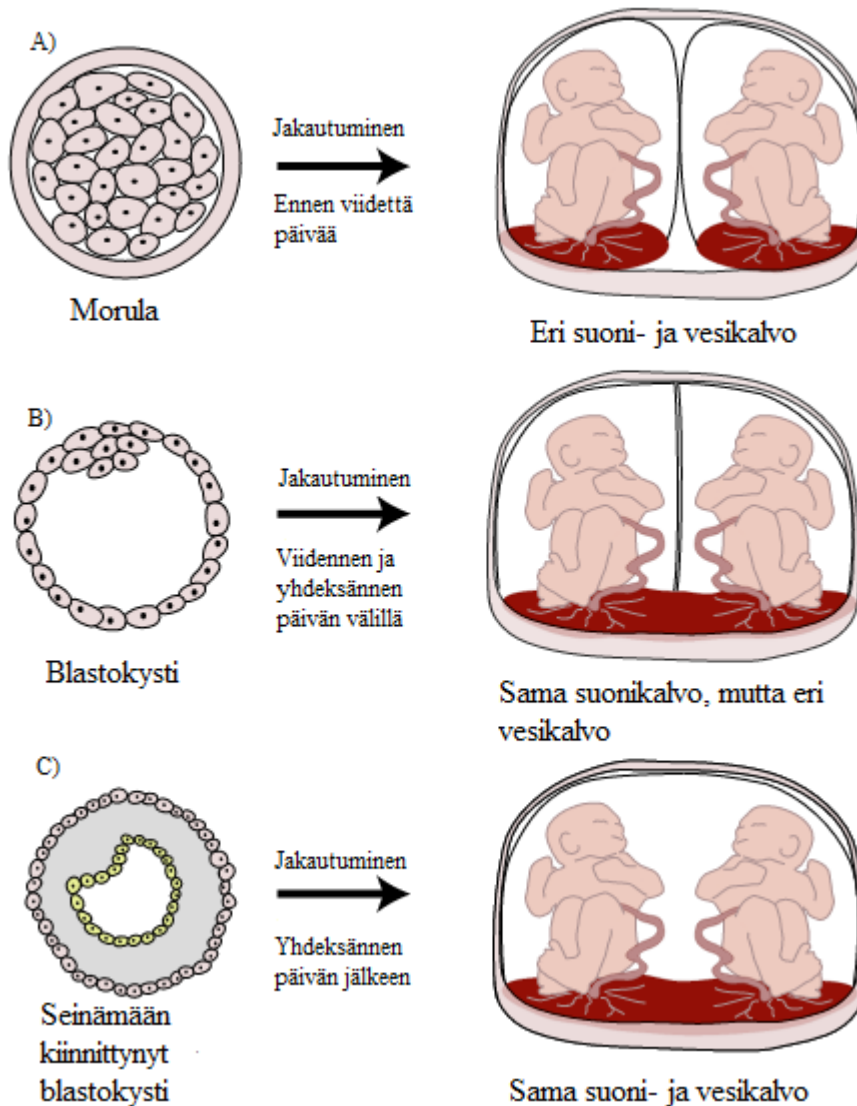
jakavat oletuksen mukaan ympäristötekijät. E-komponentti tarkoittaa yksilökohtaisia ympäristötekijöitä, joten Var E on nolla yksilöiden välillä MZ- tai DZ-kaksosissa. Klassisessa kaksosmallissa heritabiliteetti tarkoittaa geneettisten tekijöiden vaikutuksen osuutta fenotyypisesta kokonaisvarianssista (Rijsdijk & Sham 2002).

#### 4.1 Kaksosten muodostuminen

Kaksoset ovat tärkeä tiedonlähde sairauksien ja muiden mielenkiintoisten ihmisen genetiikkaan liittyvien kysymyksen tutkimiseen. Kaksoset luokitellaan joko monotsygoottisiin tai ditsygoottisiin sen mukaan, että ovatko ne muodostuneet yhdestä vai kahdesta munasolusta. Ditsygoottiset kaksoset ovat muodostuneet kahdesta erillisestä hedelmöityneestä munasolusta. Monotsygoottiset kaksoset muodostuvat samasta munasolusta ja erkanevat toisistaan noin ensimmäisen 14 päivän aikana.

Monotsygoottiset kaksoset voidaan jakaa edelleen ryhmiin sen mukaan, että jakavatko sikiöt korionin eli sikiön suonikalvon ja vesikalvon (Kuva 2). Suonikalvojen ja vesikalvojen määrä määräytyy alkion jakautumisesta ajankohdasta. Jakautumisen tapahtuessa ennen trofoblastin muodostumisesta ensimmäisen viiden päivän aikana alkioille kehittyvät omat suoni- ja vesikalvot. Viidennen ja yhdeksännen päivän välillä tapahtuneen jakautumisen seurauksena alkioille kehittyvät sama suonikalvo, mutta eri vesikalvo. Tämä ryhmä sisältää 2/3 kaikista monotsygoottisista kaksosista. Kaikista harvinaisin ryhmä on saman vesikalvon ja suonikalvon monotsygoottiset kaksoset, jotka muodostuvat alkion jakautuessa yhdeksännen päivän jälkeen.

Kaksosia vuonna 2014 Yhdysvalloissa syntyi 135 336, jonka osuus oli noin 3.5% kaikista synnytyksistä (Hamilton et al. 2015). Identtisten kaksosten syntyvyys on noin 0.25% luokkaa ja ne ovat harvinaisempia kuin ditsygoottiset kaksoset. Epigenetiikan tutkimuksiin halutaan usein kaksosia, jotka ovat identtisiä ja fenotyypiltään eroavia tutkittavan ominaisuuden suhteen. Tällaisissa tutkimuksissa halutaan, että tutkittavilla yksilöillä on kaksi erittäin harvinaista fenotyyppiä, joka usein rajoittaa vakavasti tutkimuksessa käytettävien kaksosten määrää. On tärkeää muodostaa kansallisia ja kansainvälisiä kaksoskohortteja, jotka pitävät kirjaa kaksosista tutkimuksia varten. Useissa tutkimuksissa käytetään myös muita tutkimuksia varten muodostettuja kaksoskohortteja.



**Kuva 2. Monosygoottisten kaksosten muodostuminen ja munasolun jakautumisen ajankohdan vaikutus hedelmöityksen jälkeen.** A) Jakautuminen ensimmäisen viiden päivän aikana johtaa kahteen suonikalvoon ja vesikalvoon. B) 5-9 päivän aikana jakautuminen johtaa samaan suonikalvoon, mutta eri vesikalvoon. C) Yhdeksännen päivän jälkeen jakautuneille kaksosille kehittyä sama suoni- ja vesikalvo. (Muokattu kuvasta: Kevin Dufendach (2008), Wikimedia Commons)

#### 4.2 Suomen kaksoskohortti ja Ruotsin kaksosrekisteri

Tauteja aiheuttavien geenien tai epigeneettisten tekijöiden löytäminen on erittäin vaikeaa. Joissakin tapauksissa genomista löydetään tautitapauksissa esiintyviä modifikaatioita, mutta ei voida todistaa sen kausaalisuutta, eli aiheuttaako modifikaatio sairautta vai onko se syntynyt sairauden takia. Kaksosten käyttö tutkimuksissa poistaa monta tunnettua ja

tuntematonta sekoittavaa tekijää, joten tästä syystä fenotyypiltään erilaisten (diskordanttien) kaksosten käyttö epigeneettisten tekijöiden tutkimisessa on eräs tehokkaimmista menetelmistä. Taudin suhteen diskordantti MZ-kaksospari yleensä osoittaa taudin pohjautuvan epigeneettisten mekanismien häiriöihin.

Tutkimukselle on eduksi jos näytteitä otetaan useiden vuosien ajan yleensä mahdollisimman varhaisesta kehitysvaiheesta, sekä ennen taudin ilmenemistä. Ympäri maailmaa on olemassa kaksoskohortteja, joissa vaihtelevin tavoin pidetään kirjaa kaksosten elämästä kyselyillä tai näytteitä ottamalla. Nämä kaksoskohortit ovat tärkeitä apuvälineitä luotettavien perusasetelmien ja kehikkopopulaatioiden luomisessa kaksostutkimuksissa.

Suomalainen kaksoskohortti perustettiin vuonna 1974 ja siihen kuuluu nykyään 12 966 MZ- ja DZ-kaksosta. Kyseisessä kohortissa tutkitaan kaksosia postikyselyillä ja kliinistä tutkimusta varten on kerätty näytteitä seerumista ja muista kudoksista, sekä kohortissa on yli 15 000 DNA näytettä saatavilla tutkijoille (Kaprio 2006). Kohorttia on käytetty syömishäiriöiden muodostumisen tutkimiseen ja niiden esiintyvyyden arvioimiseen suomalaisissa kaksosissa. Kohortilla on myös tutkittu seurantakyselyllä alkoholin käytön vaikutuksista aivojen toimintaan (Kaprio 2006).

Ruotsalainen kaksosrekisteri, joka perustettiin 1950-luvulla, muodostuu kolmen kohortin yhdistelmästä ja on yksi maailman suurimpia kaksosrekistereitä. Se on muodostettu useasta kaksoskohortista, joista varhaisin sijoittuu 1800-luvun loppupuolelle (Lichtenstein et al. 2002).

Ruotsalaisesta kaksosrekisteristä on muodostettu CATSS (The Child and Adolescent Twin Study in Sweden) – kohortti, joka sisälsi yhteensä 25 306 MZ- ja DZ-kaksosta, joita oli seurattu kyselyillä ja näytteenotoilla kolmen vuoden ajan yhdeksännen ja kahdennentoista ikävuoden aikana (Ullemer et al. 2016). CATSS – kohortissa oli saatavilla 12 388 kaksosen sylkinäytteet. Kohortin avulla on tutkittu lapsuusajan astman ja siihen liittyvien fenotyyppien heritabiliteettiä (Ullemer et al. 2016). Tutkimuksessa pyrittiin myös toistamaan aikaisempia GWAS (genome wide association studies) tuloksia. Tutkimuksessa mainitut astmaan liittyvät fenotyypit olivat hengityksen vinkuminen, heinäkuume, ruoka-allergiat ja atooppinen ekseema. Tutkimuksessa valittiin 18 lapsuusajan astmaan liittyvää nukleotidipolymorfismia (SNP, eli single nucleotide polymorphism) assosiaatioanalyysia varten. Tutkimuksessa löydettiin selvä yhteys kuuden nukleotidipolymorfismin ja astmaa sairastavien kaksosten

välillä ja tulosten perusteella pääteltiin, että astma on erittäin heritabiliteetti sairaus (Ullemar et al. 2016). Vaikka tutkimuksen asettelussa ei hyödynnetty kaksosmallia täyteen potentiaaliin, osoittaa tutkimus silti kaksoskohortien hyödyllisyyden tietopankkeina.

Vuonna 2014 julkaistussa tutkimuksessa käytettiin TwinGene -kohortin ruotsalaisilta kaksosilta kerättyjä verinäytteitä (Viktorin et al. 2014). Tutkimuksen kaksoset olivat syntyneet 1911–1958 aikavälillä ja kerättyjä verinäytteitä oli yhteensä 12 591 kappaletta, joista 3128 oli MZ-kaksosia ja 9402 DZ-kaksosia. Tutkimuksen tarkoitus oli arvioida veren IgA-tasojen heritabiliteettiä hyödyntäen klassista kaksosmallia tutkimuksessa, jossa fenotyypin varianssia arvioidaan A-, C- ja E-komponenttien (kappale 4.1) avulla. Genomin laajuudessa assosiaatioanalyysissä (GWAS) kaksi SNP-lokusta ylitti analyysin merkitsevyystason. Tutkimuksessa havaittiin merkittävää assosiaatiota veren IgA tasoihin rs6928791 ja rs13300483 lokuksissa (Viktorin et al. 2014).

## **5. Kaksosten käyttö epigenetiikan tutkimuksissa**

Ensimmäisessä suuren mittakaavan kaksosia hyödyntävässä epigenetiikan tutkimuksessa tutkittiin DNA:n metylaatiota ja histonien asetylaatiota useissa eri kohdissa kahdenkymmenen 3-vuotiaan ja kahdenkymmenen 50-vuotiaan genomeissa (Fraga et al. 2005). Tutkimuksessa huomattiin MZ-kaksosilla olevan hyvin samanlaiset epigeneettiset profiilit, mutta erot profiilissa lisääntyivät iän ja elinympäristön erojen myötä (Fraga et al. 2005). Tutkimus oli tyypiltään poikkileikkaava, eli samalta yksilöltä ei otettu näytettä useampaa kertaa eri ajankohtina.

### **5.1 DNA:n metylaatio tautitapauksissa**

Toisessa tutkimuksessa seurattiin yksilön elämän aikana DNA:n metylaatiota kaksosilla kahdessa ikävaiheessa. Näytteitä otettiin 46 MZ-kaksosesta ja 45 DZ-kaksosesta 5-vuotiaana ja uudelleen 10-vuotiaana (Wong et al. 2010). DNA:n metylaatiota mitattiin sylkirauhasten soluista käyttäen bisulfaatti konversioon perustuvaa menetelmää (Wong et al. 2010). Tässä tutkimuksessa pyrittiin selvittämään DNA:n metylaation heritabiliteettiä. Tutkimuksessa yksi geeni osoitti heritabiliteettiä ja eroja DNA:n metylaatioissa oli havaittavissa kummassakin vaiheessa, mutta niiden määrä kasvoi iän myötä.

Vuonna 2010 tutkittiin DNA:n metylaatiota vastasyntyneissä vauvoissa (Ollikainen et al. 2010). Tutkimuskohteiksi valittiin neljä erilailla metyloitunutta aluetta (DMR) jotka ovat

*IGF2* ja *H19* -geeneissä. Tutkimuksessa kerättiin näytteitä vastasyntyneeltä 56 MZ-kaksoselta ja 35 DZ-kaksoselta. Tutkimuksessa käytettiin Perinatal Epigenetic Twin Study:n kaksosilta kerätyjä näytteitä. Kaksoset syntyivät Melbournessa, Australiassa ja näytteet kerättiin napanuorasta istukan kudoksesta ja buccal:n soluja.

*H19* on geeni jonka transkriptiotuote on lncRNA, eli pitkä ei-koodaava RNA ja se sisältää DMR:n, nimeltä H19 CTCF6 DMR, joka on tärkeä *IGF2*:n ja *H19*:n genomisen leimautumisen säätelyssä (Ollikainen et al. 2010). H19 CTCF6 DMR sisältää CTCF-insulaattorin, joka inhiboi *IGF2* ja sallii *H19*:n ekspression. Metylaatioprofiilin muutokset H19 CTCF6 DMR:ssä on yhdistetty syöpään (Ollikainen et al. 2010). Kolme muuta tutkittavaa aluetta oli *H19* promoottorin DMR ja kaksi *IGF2*:n DMR:ää (Ollikainen et al. 2010). Tutkimuksessa käytettiin DNA:n metylaation mittaamisessa bisulfaatti konversioon perustuvia menetelmiä.

Tutkimuksessa havaittiin DMR:n metylaatioissa eroja MZ-kaksosten välillä, mikä tukee ideaa siitä, että epigeneettisiä eroja muodostuu jo varhaisessa lapsuudessa. Tutkimus osoittaa, että *in utero* – ympäristö on tärkeää aikaa yksilön epigeneettisten erojen muodostumisessa (Ollikainen et al. 2010).

Monotsygoottisia kaksosia käytettiin vuonna 2010 julkaistussa tutkimuksessa, jossa selvitettiin kolmen autoimmuunisairauden liittyvyyttä DNA:n metylaatioon (Javierre et al. 2010). Tutkimuksessa käytettiin sairauksien suhteen eroavaisia MZ-kaksosia ja tutkittavat sairaudet olivat systeeminen lupus erythematosus (SLE), nivelreuma (RA) ja dermatomyosiitti (DM). Jokaista tautia kohden käytettiin kymmentä MZ-kaksosparia ja tutkittu DNA kerättiin valkosoluista (Javierre et al. 2010).

Tutkimuksessa muodostettiin metylaatioprofiilit 807 CpG-dinukleotidista, joiden on todistettu olevan kytköksissä useaan signaalireittiin (signaling pathway), DNA:n korjaukseen, solusyklin säätelyyn ja metastaasikseen (syövän leviäminen toiseen kudokseen) (Javierre et al. 2010). Valituissa 807 CpG:ssä ei ollut merkittäviä metylaatio eroja nivelreuman ja dermatomyosiitin suhteen, mutta systeeminen lupus erythematosus MZ-kaksoset osoittivat 49 geenin suhteen tilastollisesti merkitsevää eroavaisuutta. Geenien ontologisessa analyysissä selvisi geenien liittyvän immuunijärjestelmän toimintaan, mikä osoittaa epigeneettisten tekijöiden vaikuttavan systeemiseen lupus erythematosus:een (Javierre et al. 2010).



Norjalaisen kaksosrekisterin monotsygoottisilla kaksosilla vuonna 2012 tutkittiin DNA:n metylaation yhteyttä psoriasis:een (Gervin et al. 2012). Psoriasis on iho- ja niveloireita aiheuttava krooninen autoimmuunisairaus. Tutkimuksessa käytettiin psoriasis:n suhteen fenotyypiltään eroavaa 27 MZ-kaksosparia. Metylaatioprofiilin eroavaisuuksia tutkittiin kaksosten valkosoluissa. Metylaatiota tutkittiin genomilaajuisesti CD4<sup>+</sup>- ja CD8<sup>+</sup>-valkosoluissa käyttäen bead array-koetinta, joka pystyy arvioimaan 27 578 CpG-dinukleotidin metylaation, joista 26 690 valittiin tutkimuksessa (Gervin et al. 2012). Näytteet deaminoitiin bisulfaatti konversiolla ennen koettimen käyttöä. Myös geenien ekspressiota tutkittiin koettimella, mikä pystyy tunnistamaan yli 25 000 geenin transkriptiotuotteita (Gervin et al. 2012).

Tutkituista 26 690 CpG-dinukleotidista 1288 osoitti tilastollisesti merkitseviä metylaation eroja ja 37 856 transkriptiotuotteesta 2126 osoitti geeniekspression eroja kaksosten välillä (Gervin et al. 2012). Geeniekspression ja metylaatioprofiilien erojen pohjalta tehty geenien ontologinen analyysi osoitti, että DNA:n metylaatiossa havaitut erot vaikuttavat geenien toimintaan, jotka ovat yhteyksissä immuunivasteeseen ja keskeisiin immuunijärjestelmän signaalireitteihin (Gervin et al. 2012). Tutkimuksessa pystyttiin siis osoittamaan DNA:n metylaatioiden olevan kytköksissä psoriasis:een.

## 5.2 Histonimodifikaatioiden tutkimukset

DNA:n metylaatio on tunnetuin ja tutkituin epigeneettinen tekijä ja kaksosia käytetään jatkuvasti alan johtavissa tutkimuksissa. Kaksosia on käytetty myös histonimodifikaatioiden vaikutuksien tutkimiseen tautitapauksissa. Fenotyypiltään taudin suhteen eroavia kaksospareja voidaan käyttää hyödyksi näissä tutkimuksissa.

MLL-assosioitunut akuutti myeloinen leukemia (AML) on syöpä, joka aiheuttaa valkosoluissa epätavallista kasvua (Krivtsov & Armstrong 2007). MLL (mixed lineage leukaemia) on geeni jonka koodaama H3K4 metyltransferaasi metyloi H3 histonin sivuketjun lysiinissä 4. MLL-assosioituneissa leukemioissa on tapahtunut translokaatio kromosomistossa kohdassa 11q23, jonka seurauksena H3K4 metyltransferaasin toiminta on lakannut. On kuitenkin huomattu, että MLL-assosioitunut AML tarvitsee 11q23-translokaation lisäksi myös muita mekanismeja, kuten epimutaatioita, puhjetakseen sairaudeksi (Krivtsov & Armstrong 2007, Zhu et al. 2014).

Vuonna 2014 tutkittiin leukemiaa aiheuttavia mutaatioita kolmevuotiaalla kaksosparilla, joista toinen sairasti MLL-assosioitunutta AML:ä (Zhu et al. 2014). Tutkimuksessa löydettiin koko genomin sekvensoinnin perusteella somaattisista leukemiasoluista kaksi ei-synonyymista, eli aminohappoa muuttavaa, hiljentävää mutaatiota *SETD2*-geenissä, joka koodaa histonimodifikaatioita aiheuttavaa entsyymiä. Tämä entsyymi aiheuttaa H3K36me3 modifikaatiota eli histoniproteiinin H3 sivuketjussa lysiinissä 36 kovalenttista kolmen metyyliryhmän lisäämistä. Metylaatiot H3K36:ssa on yhdistetty säätelevän transkription elongaatiovaihetta (Zhu et al. 2014). Jatkotutkimuksissa huomattiin *SETD2* – geenin toimintaa hiljentävien mutaatioiden osallistuvan leukemioiden syntyyn ja ylläpitoon (Zhu et al. 2014).

H3K4-lysiinin metyylaatiota on tutkittu myös Downin oireyhtymän yhteydessä vuonna 2014 julkaistussa tutkimuksessa (Letourneau et al. 2014). Tutkimuksessa käytettiin yhtä kaksosparia, joista toisella oli Downin oireyhtymä. Tutkimuksessa analysoitiin molempien kaksosen sikiöllisten fibroblastien transkriptomeerin eroja. Kaksosista eristetyt mRNA:t analysoitiin ja huomattiin merkitsevästi 182 erilailla expressoitua geeniä, joista 42 transkriptiotuotteet olivat lncRNA:ta (Letourneau et al. 2014). Myös koko genomin alueelta arvioitiin expressiotasojen muutoksia. Yhteensä 337 GEDD:tä (gene expression dysregulation domain) löytyi Downin oireyhtymä kaksosessa. H3K4me3 –metylaatioprofiilit luotiin ja verrattiin kaksosten välillä. H3K4me3-metylaatioprofiilin, eli trimetylaatio 4 lysiinissä histonissa 3, huomattiin korreloivan GEDD-alueiden kanssa. (Letourneau et al. 2014). Nämä epigeneettisten profiilien erot voivat olla syynä eroihin kaksosten välillä.

## 6. Yhteenveto

Kaksosten käyttö epigeneettisten mekanismien käytössä on korvaamaton työkalu tutkijoille. Kyky kontrolloida tutkimuksen kohdepopulaation ikää, sukupuolta, perimää ja ympäristöä on tuottanut useissa tautitapauksissa luotettavaa tietoa epimutaatioiden vaikutuksesta yksilöihin. Kaksoskohorttien muodostaminen ja ylläpito on mahdollistanut pitkäaikaisen kaksosten seurannan ja tuottanut suuren määrän tietoa tutkijoiden käytettäväksi. Tulevaisuudessa olisi hyvä nähdä kansallisia ja kansainvälisiä tietopankkeja, joissa kerättäisiin kaksosista toistuvasti syntymästä lähtien useista eri kudoksista näytteitä. Tämä helpottaisi suuresti esimerkiksi tautien muodostumiseen liittyvien tekijöiden löytymistä.

- Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, Walter P, Wilson J & Hunt T (2015) *Molecular biology of the cell*. New York, NY, Garland Science, Taylor and Francis Group.
- Baurens F, Causse S & Legavre T (2008) Methylation-sensitive amplification polymorphism (MSAP) protocol to assess CpG and CpNpG methylation in the banana genome. *Fruits* 63(2): 117-123.
- Bell JT & Spector TD (2011) A twin approach to unraveling epigenetics. *Trends Genet* 27(3): 116-125.
- Brinkman AB, Simmer F, Ma K, Kaan A, Zhu J & Stunnenberg HG (2010) Whole-genome DNA methylation profiling using MethylCap-seq. *Methods* 52(3): 232-236.
- Castillo-Fernandez J, Spector TD & Bell JT (2014) Epigenetics of discordant monozygotic twins: implications for disease. *Genome Medicine* 6(7): 60.
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M & Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452(7184): 215-219.
- Dawson MA & Kouzarides T (2012) *Cancer Epigenetics: From Mechanism to Therapy*. *Cell* 150(1): 12-27.
- Deans C & Maggert KA (2015) What Do You Mean, "Epigenetic"? *Genetics* 199(4): 887-896.
- Dubin MJ, Zhang P, Meng D, Remigereau M, Osborne EJ, Casale FP, Drewe P, Kahles A, Jean G, Vilhjalmsson B, Jagoda J, Irez S, Voronin V, Song Q, Long Q, Raetsch G, Stegle O, Clark RM & Nordborg M (2015) DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife* 4: e05255.
- Feil R & Fraga MF (2012) Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics* 13(2): 97-109.
- Fraga M, Ballestar E, Paz M, Ropero S, Setien F, Ballestar M, Heine-Suner D, Cigudosa J, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector T, Wu Y, Plass C & Esteller M (2005) Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 102(30): 10604-10609.
- Gervin K, Vigeland MD, Mattingsdal M, Hammero M, Nygard H, Olsen AO, Brandt I, Harris JR, Undlien DE & Lyle R (2012) DNA Methylation and Gene Expression Changes in Monozygotic Twins Discordant for Psoriasis: Identification of Epigenetically Dysregulated Genes. *PLoS Genet* 8(1): e1002454.
- Hamilton BE, Martin JA, Osterman MJK, Curtin SC & Matthews TJ (2015) Births: Final Data for 2014. *National vital statistics reports : from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 64(12): 1-64.
- Heard E & Martienssen RA (2014) Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* 157(1): 95-109.
- Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, Rodriguez-Ubrevia J, Berdasco M, Fraga MF, O'Hanlon TP, Rider LG, Jacinto FV, Javier Lopez-Longo F, Dopazo J, Forn M, Peinado MA, Carreno L, Sawalha AH, Harley JB, Siebert R, Esteller M, Miller FW &

- Ballestar E (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. *Genome Res* 20(2): 170-179.
- Kaprio J (2006) Twin studies in Finland 2006. *Twin Res Hum Genet* 9(6): 772-777.
- Krivtsov AV & Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* 7(11): 823-833.
- Letourneau A, Santoni FA, Bonilla X, Sailani MR, Gonzalez D, Kind J, Chevalier C, Thurman R, Sandstrom RS, Hibaoui Y, Garieri M, Popadin K, Falconnet E, Gagnebin M, Gehrig C, Vannier A, Guipponi M, Farinelli L, Robyr D, Migliavacca E, Borel C, Deutsch S, Feki A, Stamatoyannopoulos JA, Herault Y, van Steensel B, Guigo R & Antonarakis SE (2014) Domains of genome-wide gene expression dysregulation in Down's syndrome. *Nature* 508(7496): 345-+.
- Lichtenstein P, De Faire U, Floderus B, Svartengren M, Svedberg P & Pedersen N (2002) The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. *J Intern Med* 252(3): 184-205.
- Mensaert K, Denil S, Trooskens G, Van Criekinge W, Thas O & De Meyer T (2014) Next-generation technologies and data analytical approaches for epigenomics. *Environ Mol Mutagen* 55(3): 155-170.
- O'Geen H, Echipare L & Farnham PJ (2011) Using ChIP-Seq Technology to Generate High-Resolution Profiles of Histone Modifications. *Methods Mol Biol* 791: 265-286.
- Ollikainen M, Smith KR, Joo EJ, Ng HK, Andronikos R, Novakovic B, Aziz NKA, Carlin JB, Morley R, Saffery R & Craig JM (2010) DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Hum Mol Genet* 19(21): 4176-4188.
- Rijsdijk FV & Sham PC (2002) Analytic approaches to twin data using structural equation models. *Briefings in bioinformatics* 3(2): 119-133.
- Ullemar V, Magnusson PKE, Lundholm C, Zettergren A, Melen E, Lichtenstein P & Almqvist C (2016) Heritability and confirmation of genetic association studies for childhood asthma in twins. *Allergy* 71(2): 230-238.
- Viktorin A, Frankowiack M, Padyukov L, Chang Z, Melen E, Saaf A, Kull I, Klareskog L, Hammarstrom L & Magnusson PKE (2014) IgA measurements in over 12 000 Swedish twins reveal sex differential heritability and regulatory locus near CD30L. *Hum Mol Genet* 23(15): 4177-4184.
- Waddington CH (2012) The epigenotype. 1942. *Int J Epidemiol* 41(1): 10-13.
- Wang Z, Gerstein M & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10(1): 57-63.
- Wong CCY, Caspi A, Williams B, Craig IW, Houts R, Ambler A, Moffitt TE & Mill J (2010) A longitudinal study of epigenetic variation in twins. *Epigenetics* 5(6): 516-526.
- Wu H & Zhang Y (2011) Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev* 25(23): 2436-2452.

Yu W, Hovik H, Olsen I & Chen T (2011) Strand-specific transcriptome profiling with directly labeled RNA on genomic tiling microarrays. *BMC Mol Biol* 12: 3.

Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H, Hua C, Zhang Y, Yang X, Lu X, Cao L, Hao L, Dong L, Zou W, Wu J, Li X, Zheng S, Yan J, Zhou J, Zhang L, Mi S, Wang X, Zhang L, Zou Y, Chen Y, Geng Z, Wang J, Zhou J, Liu X, Wang J, Yuan W, Huang G, Cheng T & Wang Q (2014) Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet* 46(3): 287-+.

Kuva 1:

Mariuswalter 2015, Wikimedia Commons, lisenssi CC-BY-SA-4.0. URI:  
[https://commons.wikimedia.org/wiki/File:Histone\\_modifications.png](https://commons.wikimedia.org/wiki/File:Histone_modifications.png)

Kuva 2 perustuu:

Kevin Dufendach 2008, Wikimedia Commons, lisenssi CC BY 3.0. URI:  
<https://commons.wikimedia.org/wiki/File:Placentation.svg>