



Kandidaatintutkielma

CRISPR-Cas9 ja kaksoisjuosteisen
DNA:n virheenkorjaus

Valtteri Peitso

Oulun yliopisto
Biokemian ja molekyyli lääketieteen tiedekunta

2022

Sisällysluettelo

| | |
|---|----|
| 1. Johdanto..... | 4 |
| 2. Historia | 5 |
| 3. CRISPR-Cas | 8 |
| 3.1 TOIMINTAVAIHEET | 8 |
| 3.1.1. <i>Adaptaatiovaihe</i> | 8 |
| 3.1.2. <i>Ekspressiovaihe</i> | 9 |
| 3.1.3. <i>Interferenssivaihe</i> | 10 |
| 3.2 CRISPR-CAS-LUOKITUS | 10 |
| 3.3 CAS9-ENTSYYMIN RAKENNE | 11 |
| 3.3.1. <i>Cas9 on apoentsyymi, joka muuttuu aktiiviseksi konformaatiomuutosten avulla</i> | 12 |
| 3.3.2. <i>Opas-RNA</i> | 13 |
| 3.3.3. <i>Nukleasidomeenit</i> | 14 |
| RuvC:n kaltainen domeeni | 14 |
| HNH-domeeni | 15 |
| 3.3.4. <i>PI-domeeni</i> | 15 |
| 4. DNA:n korjausmekanismit | 15 |
| 4.1 KAKSOISJUOSTEISEN DNA:N VAURIOT..... | 16 |
| 4.1.1. <i>Ei-homologinen päiden yhdistyminen</i> | 16 |
| 4.1.2. <i>Homologinen rekombinaatio</i> | 17 |
| Lyhyen kantaman pään resektio | 17 |
| Pitkän kantaman pään resektio..... | 18 |
| RAD51-filamentin muodostuminen ja juosteeseen hyökkääminen (strand invasion) | 18 |
| DSBR ja SDSA..... | 19 |
| 5. Crispr-Cas9 sovelluksia..... | 19 |
| 5.1 dCAS9 JA nCAS9 | 20 |
| 5.2 ALUKE- JA EMÄSMUOKKAUS | 20 |
| 5.2.1. <i>Emäsmuokkaus</i> | 20 |
| 5.2.2. <i>Alukemuokkaus</i> | 21 |
| 6. Yhteenveto | 22 |
| 7. Kirjallisuusviitteet..... | 23 |

Käytetyt lyhenteet

| | |
|----------|---|
| Cas9 | CRISPR associated protein 9 |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CRISPRa | CRISPR-aktivaatio |
| CRISPRi | CRISPR-inhibitio |
| crRNA | CRISPR-RNA |
| dCas9 | deaktivoitu Cas9-proteiini |
| DSBR | Kaksoisjuosteisen katkoksen korjausmekanismi |
| gRNA | guide RNA – ohjaava RNA |
| HDR | Homologiaohjattu korjaus |
| HNH | Cas9-entsyymin HNH-proteiinin kaltainen nukleasidomeeni |
| HR | Homologinen rekombinaatio |
| nCas9 | Cas9-nikaasi |
| NHEJ | Ei-homologinen päiden yhdistyminen |
| NUC | Cas9-entsyymin nukleasilohko |
| PAM | Protospacer adjacent motif |
| pegRNA | prime editing guide RNA |
| PI | Pam-interacting |
| REC | Cas9-entsyymin tunnistuslohko |
| RPA | Replikaatioproteiini A |
| RuvC | Cas9-entsyymin RuvC-proteiinin kaltainen nukleasidomeeni |
| SDSA | Synteesistä riippuvainen juosteiden yhdistyminen |
| sgRNA | yksittäinen opas-RNA |
| spCas9 | <i>Streptococcus pyogenes</i> –bakteerin Cas9-entsyymi |
| tracrRNA | trans-activating RNA |

1. Johdanto

Usein genomien muokkauksesta puhuttaessa ihmisten mieleen tulevat elokuvista ja uutisista koetut negatiivisävytteiset aiheet, kuten designer-vauvat ja 1990-luvun geeniterapiayritysten optimismin lasku hetkessä vuosituhannen taitteessa. 2010-luku on mahdollistanut geeniterapian imagon parantamisen ja nousun takaisin ajankohtaiseksi vaihtoehdoksi tavallisten terapeuttisten menetelmien rinnalle. Vielä toistaiseksi tutkimukseen riittää rahoitusta ja optimismia varsinkin CRISPR-Cas-menetelmän (eng. Clustered regularly interspaced short palindromic repeats) kehityksen myötä. Jatkuva CRISPR-systeemien tutkimus lisää erilaisia käyttökohteita, jotka eroavat huomattavasti niiden alkuperäisestä tarkoituksesta eli adaptiivisesta immunitetista, tai ylipäätään bakteerien ja arkkien perintöaineksen puolustamisesta (Koonin & Makarova 2022). Kehityksen kannalta mietittäväksi jäävät erityisesti eettiset ongelmat menetelmän käyttökohteita valittaessa ja niitä tutkittaessa.

CRISPR-Cas-koneiston valjastamisessa genomien muokkauksen mahdollistavaksi menetelmäksi bioinformatiikalla on ollut iso rooli. Bioinformatiikka on nykypäivänä läsnä lähes jokaisessa tutkimuksessa ja sen merkitys lisääntyy jatkuvasti. Ennen uuden projektin aloittamista data-analyysin avulla voidaan säästää huomattava määrä käytännön tutkimustyötä ja siten rahaa etsimällä ja käsittelemällä tutkimukseen liittyvää dataa.

Solujensisäiset pienet DNA-vauriot ovat jokapäiväisiä tapahtumia, mitkä aiheuttavat mutaatioita soluissa esimerkiksi happiradikaalien toimesta. Kaksoisjuosteisen DNA:n katkeaminen on kaikista DNA-vaurioista vakavin, mutta silti soluissa luonnollinen tapahtuma esimerkiksi meioosin tekijänvaihdon aikana. Kaksoisjuosteisen katkeamista vastaan solut ovat valjastaneet omia endogeenisiä korjausmekanismeja, mitkä korjaavat katkoksia liittämällä ne toisiinsa tai valmistamalla katkoksen kohtaan luovuttavasta DNA-juosteesta uuden korjatun kaksoisjuosteisen DNA:n. Jälkimmäistä tapaa käytetään hyödyksi CRISPR-Cas9-menetelmässä, jossa Cas9-entsyymi kuljetetaan soluihin katkaistaan kaksoisjuosteista DNA:ta. Katkennut DNA korjataan solujen endogeenisen korjauskoneiston kautta kohde-DNA:han lisäämällä tai poistamalla haluttu geeni. CRISPR-Cas9 on aiempia menetelmiä helpompi suunnitella ja sen valjastamiseksi solujen genomien muokkaukseen tarvitaan ainoastaan suunniteltu yksittäinen opas-RNA ja Cas9-entsyymi. Genomien muokkauksessa, kuin kaikessa muussakin tärkeää on tietää työkalun, tässä tapauksessa CRISPR-Cas9-koneiston historia ja rakenne.

2. Historia

Vuonna 1987 japanilainen työryhmä kloonasi ja sekvensoi iap-geenin, jonka tuottama proteiini toimii entsyyminä kolmen alkaalisen fosfataasin isotsyymien muodostumisen reaktioon *Escherichia coli* -bakteerissa. Isotsyymien muodostuminen on post-translaationaalinen tapahtuma, missä proteolyttisesti iap-geenin tuottama proteiini poistaa isotsyymi 1:n aminoterminaalista arginiinitähteitä. Yhden arginiinitähteen poistaminen muodostaa isotsyymin 2 ja toisen arginiinin poisto muodostaa isotsyymin 3. Työryhmä löysi erikoisen rakenteen toistuvan iap-geenin 3'-pään transkriptioalueen vierestä. Rakenne sisälsi viisi lähes homologista 29 aminohappoa sisältävää sekvenssiä, joiden välissä oli 32:sta heterologisesta aminohaposta koostuvaa DNA-juostetta. Työryhmä ei osannut selittää rakenteiden toimintaa ja oletti niiden stabiloivan RNA:n rakennetta (Ishino et al. 1987). Todellisuudessa he sattuiivat sekvensoimaan *E. coli* CRISPR-lokuksen (eng. Clustered Regularly Interspaced Palindromic Repeats), joka toimii bakteerin koneistona viruksia vastaan. Vuonna 1993 espanjalainen työryhmä löysi ensimmäiset CRISPR:t arkeilta, tarkemmin *Haloferax mediterranei* -arkeilta (Mojica et al. 1993).

Bioinformatiikan avulla vuonna 2002 Cas-geenien (eng. CRISPR-associated) todettiin toimivan yhdessä CRISPR-lokuksen kanssa. Jansen työryhmineen tutki prokaryooteilla toistuvia, sekä eukaryooteilla ja viruksilla esiintymättömiä DNA sekvenssejä. Näitä toistuvia sekvenssejä he nimittivät CRISPR:ksi. Cas-geenien operonin todettiin löytyvän CRISPR-lokuksen vierestä, joten niiden uskottiin olevan toiminnallisesti yhteydessä. CRISPR-lokuksen käyttötarkoitusta ei tiedetty tarkalleen, mutta niiden uskottiin olevan yhteydessä DNA:n aineenvaihduntaan tai geenien ekspressioon (Jansen et al. 2002).

Läpimurto CRISPR-sekvenssien kohdalla tuli vuonna 2005, kun kaksi ryhmää totesivat omissa tutkimuksissaan toistojaksojen välissä olevien spacer-alueiden (ei-koodaava alue geenien välissä) olevan homologisia bakteriofagien, profagien ja plasmidien sekvenssien kanssa. Molemmissa tutkimuksissa huomautettiin, kuinka bakteriofagit ja plasmidit eivät infektoi isäntäkantaa, jossa esiintyy spacer-sekvenssejä CRISPR-lokuksesta. Johtopäätösten avulla tutkimusryhmät esittivät itsenäisesti, että CRISPR-sekvenssien funktio on toimia samanlaisena biologisena puolustusjärjestelmänä prokaryooteilla, kuin RNAi (eng. RNA interference) toimii eukaryooteilla, suojaten soluja vierailta geneettisiltä elementeilä. Molemmat ryhmät ehdottivat CRISPR-sekvenssien käynnistävän tuntemattomalla tavalla vieraiden tunkeutuvien DNA:n osien sieppaamisen ja niiden lisäämisen omaan muis-tiinsa eli genomiin (Mojica et al. 2005, Pourcel et al. 2005). Kolmannessa tutkimuksessa samalta

vuodelta Bolotin työryhmineen varmistivat Mojican ja Pourcelin havainnot todeksi ja he huomasivat korrelaation bakteriofageista peräisin olevien spacer-alueiden lukumäärän ja vastustuskyvyn bakteriofagi-infektioiden välillä. Lisäksi työryhmä teorioi immuniteetin mahdollisesti johtuvan bakteriosolujen tuottaman antisense-RNA:n estävän virusperäisen DNA:n ekspression (Bolotin *et al.* 2005, Barrangou & Horvath 2017).

Vuonna 2007 tanskalainen työryhmä todisti kolme tärkeää löydöstä CRISPR:n toiminnan tunnistamisessa. (1) Adaptaatiovaiheen aikana CRISPR-lokukseen on mahdollista lisätä bakteriofagi-DNA:ta toistojaksojen väleihin spacer-sekvensseinä. (2) Lisäämällä, poistamalla tai siirtämällä CRISPR-lokuksen spacer-sekvenssejä resistenssi bakteriofageja vastaan lisääntyy, häviää tai siirtyy vastaavasti muokkauksen mukaan (3) *cas*-geenit toimivat osallisina adaptaatio- ja interferenssivaiheissa. Löydökset vahvistivat CRISPR-järjestelmän olevan bakteerien adaptiivinen immuunisysteemi bakteriofageja vastaan lisäämällä osan bakteriofagin genomia osaksi tunnistuskoneistoa, joka uuden virusinfektion tapahtuessa hiljentää hyökkäävän viruksen spacer-alueiden avulla. Näin vuoden 2005 löydökset, jotka viittasivat RNAi:n kaltaiseen hankittuun puolustusjärjestelmään olivat oikeassa (Barrangou *et al.* 2007, Barrangou & Horvath, 2017). Vuonna 2008 osoitettiin, kuinka CRISPR-sekvenssin transkripti pre-crRNA (eng. precursor CRISPR RNA) muokataan viiden eri Cas-proteiinin toimesta crRNA:ksi (CRISPR RNA). crRNA ohjaa Cas -proteiinikompleksia antiviraaliseen immuunipuolustukseen (Brouns *et al.* 2008). Vuoden 2008 toisessa tutkimuksessa DNA:n osoitettiin olevan CRISPR:n koodaaman immuniteetin kohde. CRISPR:n todettiin muodostavan immuniteetin, joka on erilainen kuin pääosin eukaryoottisolussa RNAi-menetelmällä tapahtuva geenien hiljentäminen, minkä aiemmin oletettiin olevan mahdollinen mekanismi CRISPR-Cas-koneistolle (Marraffini & Sontheimer 2008).

2010 Cas9-endonukleaasin todettiin katkaisevan kohde-DNA:ta kolmen emäksen päästä PAM-sekvenssin (eng. Protospacer Adjacent Motif) 3'-pään reunasta (Garneau *et al.* 2010). 2011 tutkijat löysivät trans-koodaavan crRNA:n (tracrRNA) toimivan pre-crRNA:n prosessoinnissa ja kypsymisessä. Tutkimus tehtiin *Streptococcus pyogenes* -bakteerilla ja tracrRNA:lla todettiin olevan 24 komplementaarista nukleotidia pre-crRNA:n kanssa. Samassa tutkimuksessa esiteltiin myös uusia polkuja crRNA:n kehittymiselle. Esimerkiksi crRNA:n kypsymiseen osallistui tutkimuksen mukaan RNAasi III ja Csn1 (aiempi nimitys Cas9 proteiinille) -proteiineja. Csn1-proteiinin todettiin olevan multidomeeninen proteiini, jonka funktiosta ei tiedetty muuta, kuin että se oli tärkeä osa *Streptococcus thermophilus* -bakteerin CRISPR-johteista immuunipuolustusta. Csn1:n ehdotettiin toimivan

molekulaarisena ankkurina edesauttaen tracrRNA:n ja pre-crRNA:n emäspariutumista, jolloin RNAasi III voisi tunnistaa ja katkaista pre-crRNA:n (Deltcheva *et al.* 2011).

Vuonna 2012 Cas9-entsyymien RuvC- ja HNH-domeenien todettiin katkaisevan vastakkaiset kohde-DNA-juosteet nikaasin tavoin (Gasiunas *et al.* 2012). Saman vuoden *In vitro* -tutkimukset osoittivat valmiin crRNA:n muodostavan erityisen kaksoisjuosteisen RNA-rakenteen yhdessä tracrRNA:n kanssa emästen komplementaarisen pariutumisen mukaisesti. Tämä kaksoisjuosteinen rakenne ohjasi Cas9-proteiinin katkaisemaan tietyn sekvenssin kohde-DNA:sta. Samassa tutkimuksessa kaksoisjuosteisesta tracrRNA:crRNA-dupleksista suunniteltiin yhtenäinen synteettinen yksittäinen opas-RNA (eng. single-guide RNA; sgRNA), joka pystyi ohjaamaan Cas9 entsyymiä luonnollisen dupleksi-RNA:n tavoin katkaisemaan halutun DNA-kaksoisjuosteen RuvC- ja HNH-domeenien avulla (Jinek *et al.* 2012). Vuonna 2013 tyyppi II Cas-entsyymejä ohjelmoitiin katkaisemaan DNA:ta nisäkkäiden soluissa (Cong *et al.* 2013, Mali *et al.* 2013). Samana vuonna Cas9-entsyymistä valmistettiin mutaatioiden avulla dCas9-proteiini (deaktivoitu Cas9), jolla ei ollut aktiivista nukleasidomeenia, ja sen avulla voitiin kontrolloida geenin ekspressiota (Qi *et al.* 2013). Lisäksi erillisissä tutkimuksissa kehitettiin dCas9:n avulla CRISPRa- (CRISPR activation) ja CRISPRi-työkalut (CRISPR inhibition), missä deaktivoituun Cas9-entsyymiin yhdistettiin transkription säätelijöitä, jotka aktivoivat (CRISPRa) tai inhiboivat (CRISPRi) geenin transkriptiota (Gilbert *et al.* 2013, Maeder *et al.* 2013).

Vuonna 2016 Harvardin yliopiston laboratoriossa kehitettiin CRISPR-Cas9-työkalu, missä dCas9-entsyymiin lisättiin APOBEC (sytosiinideaminaasi). Muokkauksen johdosta entsyymien avulla voitiin muokata sytosiiniemäs tymiiniksi tai guaniiniemäs adeniiniksi synteettisen sgRNA:n ohjauksen avulla ilman DNA:n kaksoisjuosteen katkeamista. Työkalulle annettiin nimeksi emäsmuokkausmenetelmä ja sen avulla pistemutaatioiden korjauksesta tuli tehokasta (Komor *et al.* 2016). Myöhemmin seuraavana vuonna samainen työryhmä kehitti adeniini-emäsmuokkausmenetelmän (ABE), minkä avulla A-T emäsparit voitiin vaihtaa C-G emäspareiksi (Gaudelli *et al.* 2017). Kolmas mullistuksellinen työkalu, joka kehitettiin saman työryhmän toimesta, oli vuonna 2019 esitelty prime editing (alukemuokkaus) -menetelmä. Siinä nCas9-entsyymiin (Cas9-nikaasi) lisättiin käänteiskopioijaentsyymi ja pegRNA (eng. prime editing guideRNA) (Anzalone *et al.* 2019).

3. CRISPR-Cas

RNA-välitteisiä CRISPR-Cas-systeemejä esiintyy bakteereilla ja useimmilla arkeilla, missä ne toimivat adaptiivisena osana immuunipuolustusta pääasiassa bakteriofagivälitteisiä viruksia, plasmideja ja transposoneja vastaan. Näistä parhaiten tunnettu ja tutkittu CRISPR-Cas-systeemi on *Streptococcus pyogenes* -bakteerilla esiintyvä tyypin II CRISPR-Cas9. Nämä systeemit käyttävät hyödykseen kohdentavan crRNA:n ja aktivoivan tracrRNA:n muodostamaa kaksois-RNA-rakennetta. Muodostuvaa duplexi-RNA:ta kutsutaan lyhyesti opas-RNA:ksi (eng. guide RNA; gRNA). Opas-RNA ohjaa Cas9-entsyymiä hiljentämään bakteeriin tai arkkiin tunkeutuvia liikkuvia geneettisiä elementtejä (eng. Mobile Genetic Element; MGE) (Jinek *et al.* 2012).

Geneettisen materiaalin vaihdon ansiosta bakteerit saavat uusia ominaisuuksia, jotka voivat aiheuttaa muutoksia metabolisiin reitteihin, patogeneesiin sekä antibioottiresistenttyyteen. Uusien ominaisuuksien vastaanottaminen onnistuu horisontaalisen geeninsiirron avulla, jonka mekanismeja ovat plasmidin vastaanotto toiselta bakteerilta (transformaatio), suora plasmidin siirto, joka vaatii kontaktin luovuttajan ja vastaanottajan välillä (konjugaatio) ja infektoivan viruksen eli bakteriofagin välityksellä siirtyvän geneettisen materiaalin siirto (transduktio). Bakteriofagivälitteinen DNA:n siirto voi aiheuttaa muutoksia geenien ilmenemiseen bakteerisolussa esimerkiksi inaktivoimalla isäntägeenejä, tai muuttamalla viereisten geenien ilmentymistä DNA-liitoksen aikana (Ochman *et al.* 2000, Nozawa *et al.* 2011, le Rhun *et al.* 2019). CRISPR-Cas-järjestelmän läsnäolo voi olla myös haitaksi isäntäbakteerille estäen sen kasvua tai selviytymistä edistävien ominaisuuksien hankkimista. Järjestelmä voi estää esimerkiksi ihmisille patogeenista *Streptococcus pneumoniae* -bakteeria hankkimasta transformaation avulla kapselia koodaavaa geeniä (Nozawa *et al.* 2011).

3.1 Toimintavaiheet

CRISPR-Cas9-välitteinen immuniteetti tapahtuu kolmen eri vaiheen kautta, jotka ovat adaptaatiovaihe, ekspressiovaihe eli crRNA:n biogeneesivaihe ja interferenssivaihe.

3.1.1. Adaptaatiovaihe

Adaptaatiovaiheessa bakteriofagien tai plasmidien perimästä lisätään lyhyt tunnistussekvenssi bakteerien ja arkkien perimään. Tätä lyhyttä ulkoista sekvenssiä kutsutaan protospaceriksi ja se lisätään osaksi isäntägenomin CRISPR-lokusta spacer-sekvenssiksi. Isäntäsolu tunnistaa tunkeutuvan nukleinihapposekvenssin, valitsee ja prosessoi prespacer-alueen (sekvenssi, joka prosessoidaan CRISPR-

lokukseen liitettäväksi spacer-sekvenssiksi), tunnistaa CRISPR-lokuksen ja lisää spacer-alueet siihen. Ennen toistojaksoja ja spacer-alueita CRISPR-lokuksessa on leader-sekvenssi, joka sisältää adeniini- ja tyymiinimäksistä rikkaan alueen sekä promoottorialueen (McGinn & Marraffini 2018). Leader-sekvenssin on todettu edistävän CRISPR-lokuksen transkriptiota (Pougach *et al.* 2010). CRISPR-Cas-systeemit käyttävät hyväkseen Cas1- ja Cas2-proteiineja adaptaatiovaiheessa ja ne ovat välttämättömiä CRISPR-lokuksen mukauttamisessa naiivissa immuunipuolustuksessa (Yosef *et al.* 2012, Sternberg *et al.* 2016). Cas1-proteiini on metallista riippuvainen DNA-spesifinen endonukleaasi, kun Cas2-proteiini katkaisee spesifisti yksijuosteista RNA:ta usean eri bakteerin koneistossa (Yosef *et al.* 2012, le Rhun *et al.* 2019).

Sopivien protospacer-alueiden tunnistamiseksi tyyppin I ja II systeemit käyttävät lyhyitä 3-7 nukleotidin pituisia protospacer-adjacent motiiveja (PAM:ejä) (Sternberg *et al.* 2016). Tyyppin I CRISPR-Cas-järjestelmien adaptaatiovaiheella on tutkittu olevan kaksi mahdollista tapaa: naiivi ja pohjustettu. Naiivissa adaptaatiomallissa bakteeri hankkii spacer-sekvenssin vieraasta DNA-lähteestä, kun taas pohjustetussa adaptaatiomallissa spacer-sekvenssin liittämiseen käytetään hyödyksi jo olemassa olevia vanhoja spacer-alueita. Molemmissa tavoissa bakteerisolut tarvitsevat Cas1- ja Cas2-proteiineja. Naiivissa adaptaatiossa tarvitaan ainoastaan Cas1-Cas2-kompleksia, kun taas pohjustetussa adaptaatiossa tarvitaan näiden proteiinien lisäksi kaskadia, Cas3-nukleaasia ja pohjustus-spacer-sekvenssiä, joka tunnistaa olemassa olevan spacer-alueen (Sternberg *et al.* 2016). Ensimmäisen kerran adaptaatiovaihe dokumentoitiin *Streptococcus thermophilus* -bakteerilla, jolla on tyyppin II CRISPR-Cas systeemi (Barrangou *et al.* 2007). Naiivissa adaptaatiossa tyyppin II systeemit tunnistavat PAM-sekvenssin Cas9:n avulla hyödyntäen myös tracrRNA:ta. Ilman Cas9-entsyymiä tyyppin II systeemit eivät pysty lisäämään uusia spacer-alueita CRISPR-lokukseen (Heler *et al.* 2015).

3.1.2. Ekspressiovaihe

Ekspressiovaiheessa tai crRNA:n biogeneesivaiheessa koko CRISPR-lokus transkriptoidaan pre-crRNA:ksi (eng. precursor CRISPR RNA; esiaste-crRNA), joka sisältää kaikki virus- ja plasmiditunnistuksen lyhyet nukleiinihapposekvenssit, sekä lokuksen sisäiset toistojaksot tunnistusekvenssien (spacer-sekvenssien) väleissä. Tyyppin II systeemin transkriptoitu tracrRNA kiinnittyy pre-crRNA:han emäspariperiaatteen mukaan ja kypsä crRNA muokataan pre-crRNA:sta, kun RNAasi III katkaisee spacer-alueen ja toistojaksojen välisen alueen tracrRNA:crRNA-Cas9-kompleksin tunnistuksen avulla (Deltcheva *et al.* 2011). Näin aina yhteen kypsyneeseen crRNA:han jää yksi toistojakso-osa ja

viruksen tai plasmidin sekvenssiä vastaava spacer-sekvenssi. Cas-proteiinien operoni sijaitsee hie-
man ennen CRISPR-lokusta bakteerin perimässä ja Cas-proteiinit valmistuvat samanaikaisesti
crRNA:n kanssa. Yhdessä crRNA:n ja tracrRNA:n kanssa Cas9 muodostaa tracrRNA:crRNA-Cas9-
kompleksin (Jinek *et al.* 2012, le Rhun *et al.* 2019). Tyypin I systeemit hyödyntävät pre-crRNA:n pro-
sessoinnissa Cas6-perheen endoribonukleaaseja. Cas6-proteiinit muodostavat tyypin I systeemeissä
yleensä suuren kompleksin usean eri Cas-proteiinin kanssa, esimerkiksi tyypin I-E CRISPR-Cas-sys-
teemi muodostaa Cas6-entsyymin kanssa kaskadin yhdessä Cse1-, Cse2-, Cas7- ja Cas5-proteiinien
kanssa mahdollistaen crRNA:n kypsymisen (Charpentier *et al.* 2015).

3.1.3. Interferenssivaihe

Interferenssivaiheessa tyypin II systeemeillä muodostunut tracrRNA:crRNA-Cas9-kompleksi (opas-
RNA:n ja Cas9-entsyymin kompleksi) tunnistaa PAM-sekvenssin (eng. protospacer adjacent motif)
avulla crRNA:n spacer-sekvenssiä vastaavan tunkeutuvan kohteen protospacer-sekvenssin. Tunnis-
tuksen jälkeen Cas9-entsyymi katkaisee RuvC- ja HNH-endonukleaasidomeenien avulla kohde-
DNA:n kolmen emäksen päästä PAM-sekvenssistä. *Streptococcus pyogenes* -bakteerilla PAM-tunnis-
tuskohdan sekvenssi on 5'-NGG-3', missä N on mikä tahansa emäs ja G on guaniiniemäs. Katkaisun
jälkeen genomisen DNA korjataan kaksoisjuosteen korjausmekanismeilla (Jinek *et al.* 2012). *Strep-
tococcus pyogenes* -bakteerin Cas9-entsyymin (SpCas9) tunnistus PAM-sekvenssin avulla käynnistää
kaksoisjuosteisen kohde-DNA:n juosteiden eroamisen. Eroamisen jälkeen opas-RNA sitoutuu 20
nukleotidin crRNA:n spacer-alueen avulla kohdejuosteeseen muodostaen R-silmukan (eng. R-loop).
R-silmukassa komplementaarinen juoste on kiinnittyneenä opas-RNA:han, mikä aiheuttaa ei-komp-
lementaarisen juosteen syrjäytymisen. Valmiin R-silmukan muodostumisen jälkeen Cas9 aktivoituu
HNH- ja RuvC-nukleaasidomeenien katalysoidessa kohde-DNA:n katkaisua (Pacesa *et al.* 2022).

CRISPR-Cas9 geeninmuokkaustyökalua voidaan käyttää luomaan pieniä insertioita tai deleeti-
oita (indeleitä) genomiin ei-homologisen päiden yhdistymisen (eng. Non-homological end joining;
NHEJ) tai korkean tarkkuuden homologiaohjatun korjauksen (eng. Homology-directed recombina-
tion; HDR) kautta (Jinek *et al.* 2012).

3.2 CRISPR-Cas-luokitus

Tunnettujen CRISPR-Cas-systeemien monimuotoisuus ja määrät ovat nousseet viime vuosina huo-
mattavasti. Pääsyy monimuotoisuuden kasvulle on tutkimuksen lisääntyminen, mutta myös CRISPR-

Cas-systeemien jatkuva evolutiivinen immuunivaste viruksia vastaan. CRISPR-tutkimuksen edistymisen kannalta systeemien lokeroinnilla ja bioinformatiikalla on suuri vaikutus. Oman suolansa tutkimukseen tuovat jatkuva tutkimuksen kehitys ja universaalien markkereiden puute johdonmukaiselle evoluutioluokitukselle. Makarovan (2019) työryhmineen aiemmin ehdottaman luokitusmenetelmän mukaan CRISPR-Cas systeemit jaotellaan kahteen luokkaan, kuuteen tyyppiin, yli 30:een alatyyppiin ja edelleen alatyypin variantteihin. Luokittelu perustuu useisiin eri tekijöihin sekvenssin ja geenin kontekstin analysoinnissa yhdessä kokeista saatujen tulosten kanssa. Näin tutkimus yhdistää kokeellista tietoa ja bioinformatiikkaa (Makarova *et al.* 2019, Koonin & Makarova 2022).

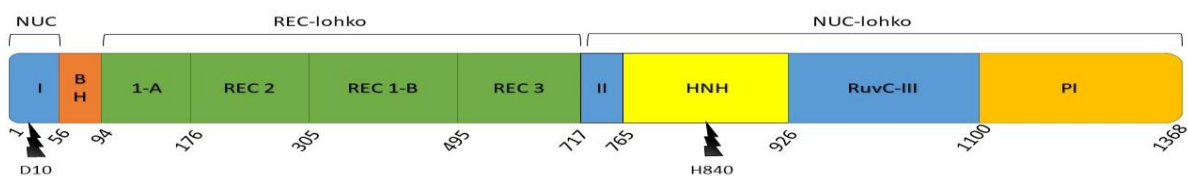
CRISPR-Cas-systeemit on jaoteltu kahteen eri luokkaan Cas entsyymien ja efektorikompleksien mukaan. Luokat on jaoteltu edelleen kuuteen eri tyyppiin. Monimutkaiset luokan I CRISPR-Cas-systeemit (tyypit I, III ja IV) käyttävät hyödykseen usean alayksikön efektorikompleksia liikkuvien geneettisten elementtien hiljentämiseksi, kun luokan II systeemit (tyypit II, V ja VI) käyttävät hiljentämiseen ainoastaan yhtä efektoriproteiinia. Tunnetuin ja tutkituin Cas-proteiini, Cas9 on luokan II tyyppiä II DNA:ta kohdistava endonukleaasi. Tunnetut luokan II variantit ovat tutkimuksen myötä viime vuosina lisääntyneet huomattavasti (Hille *et al.* 2018, Makarova *et al.* 2011, 2015, 2019). Tyypin I, II, V ja VI systeemit tarvitsevat PAM-sekvenssin kohteen tunnistukseen (Barrangou & Horvath 2017.)

cas-geenit voidaan luokka- ja tyyppiluokittelun lisäksi luokitella erilleen toisistaan sen mukaan, osallistuvatko ne adaptaatiovaiheeseen, ekspressiovaiheeseen, häirintävaiheeseen vai signaalin välitykseen. Tästä huolimatta osa geenien tuottamista proteiineista toimii samoissa tehtävissä yhtä aikaa. Esimerkiksi luokan II entsyymit Cas9, Cas12 ja Cas13 toimivat osittain crRNA:n kypsymisessä (ekspressiovaihe), sekä häirintävaiheessa efektoriproteiineina ja kohde-DNA:n nukleaaseina. (Koonin *et al.* 2017, Makarova *et al.* 2019).

3.3 Cas9-entsyymin rakenne

Eniten tutkittu *Streptococcus pyogenes* -bakteerilla esiintyvä Cas9-entsyymin (spCas9) on verrattain suuri 1368 aminohaposta ja useasta eri domeenista koostuva DNA-endonukleaasi. spCas9 katkaisee kaksoisjuosteisen DNA:n kolmen emäksen päästä PAM-sekvenssistä kahden nukleaasidomeenin avulla; HNH-domeeni katkaisee kohdejuosteen, joka on komplementaarinen opas-RNA:n sekvenssille ja RuvC-domeeni katkaisee opas-RNA:lle ei-komplementaarisen juosteen kohde-DNA:sta (Jinek *et al.* 2012, Nishimasu *et al.* 2014, Jiang & Doudna 2017). Cas9 on mukana crRNA:n kypsymisessä ja spacer-sekvenssien hankkimisessa (Heler *et al.* 2015).

spCas9-entsyymin rakenne koostuu kahdesta eri lohkosta, nukleaasi- (NUC) ja tunnistuslohkosta (REC), jotka sitovat yhteen pitkä α -heliksi, jota kutsutaan silta-heliksiksi (eng. bridge helix) (aminohappotähteet 56-93). Tunnistuslohko voidaan jakaa kolmeen tai neljään erilliseen osaan, REC1A-domeeni (aminohappotähteet 94-175) REC2-domeeni (aminohappotähteet 176-304) REC1B-domeeni (aminohappotähteet 305-496) ja REC3-domeeni (aminohappotähteet 496-717). Toisessa esitystavassa REC1B-domeeni ja REC3-domeeni ovat yksi REC3-domeeni (aminohappotähteet 306-716). Nukleasilohko koostuu RuvC- (aminohappotähteet 1-55, 717-764 ja 926-1099), HNH- (aminohappotähteet 765-925) ja PI-domeeneista (aminohappotähteet 1100-1368) (**Kuva 1**) REC-lohkolla ei ole rakenteellisesti samankaltaisuutta muiden tunnettujen proteiinien kanssa (Nishimasu *et al.* 2014, Jiang & Doudna 2017, Pacesa *et al.* 2022).



Kuva 1. *Streptococcus pyogenes* Cas9-endonukleasin domeenien sijainnit: I, RuvC1-domeeni; BH, silta-heliksi; 1-A, REC1A-domeeni; REC2, REC2-domeeni; REC1-B, REC1B-domeeni; REC3, REC3-domeeni; II, RuvC-II-domeeni; HNH, HNH-domeeni, RuvC-III, RuvC-domeeni; PI, PAM-intercating -domeeni. D10 ja H840 esittävät mutaatioita RuvC-I-domeenissa ja HNH-domeenissa. D10- ja H840-aminohappojen mutatoimisella Cas9-entsyymistä voidaan luoda synteettisesti nikaasi (nCas9), (mutatoimalla vain toinen domeeni) tai deaktivoitu Cas9 (dCas9) (mutatoimalla molemmat domeenit). (Kuva muokattu lähteestä (Pacesa *et al.* 2022))

3.3.1. Cas9 on apoentsyymi, joka muuttuu aktiiviseksi konformaatiomuutosten avulla

Cas9-entsyymi on katalyyttisesti inaktiivisessa apoentsyymi-muodossa, kun siihen ei ole sitoutuneena ohjaavaa opas-RNA:ta. Entsyymien kaksi erillistä lohkoa (REC ja NUC) käyvät läpi RNA-johtaisen konformaationmuutoksen muodostaen keskuskanavan, johon kaksijuosteinen kohde-DNA sitoutuu (Jinek *et al.* 2014). Lohkot ovat liittyneet toisesta päästä toisiinsa arginiini-rikkaan silta-heliksien (eng. arginine rich bridge helix) kautta ja toisesta päästä epäjärjestyneellä linkkerillä (eng. disordered linker) (Jiang & Doudna 2017).

Suurin muutos tapahtuu REC-lohkon REC3-domeenissa (kutsutaan myös REC1B- ja helikaali-III-domeeniksi), joka liikkuu kohti HNH-domeenia *tracrRNA*:n sitoutuessa entsyymiin. *tracrRNA*:n tehtävä koaktivaattorina on esitellä entsyymille *crRNA*, joka omaa 20 nukleotidin spacer-sekvenssin DNA:n tunnistamiseksi, jotta Cas9-entsyymi voi toimia DNA:ta katkaisevana endonukleasina. Kohde-DNA:han ja PAM-sekvenssiin sitoutuessaan Cas9 ei käy läpi yhtä suuria muutoksia, kuin RNA:han sitoutuessaan (Doudna & Charpentier 2014). PAM-sekvenssin tunnistuksen avulla Cas9-entsyymi aukaisee kohde-DNA:n kaksoisjuosteen ilman ulkoista energiaa (kaksoisjuosteisen DNA:n alueesta käytetään nimitystä siemen alue, eng. seed region). Opas-RNA:n spacer-sekvenssin ollessa täysin komplementaarinen kohde-DNA:n protospacer-sekvenssille muodostuu stabiili RNA-DNA-hybridi: R-silmukka (Szczelkun *et al.* 2014, Xue & Greene 2021). Kaksoisjuosteisen DNA:n sitoutumisen yhteydessä HNH-domeeni muodostaa vuorovaikutuksen REC2-domeeniin (kutsutaan myös Helikaali-II-domeeniksi), joka lukitsee HNH-domeenin aktiiviseen konformaatioon (Jiang & Doudna 2017). Opas-RNA:n rungon silmukkaan 1 ja toisto-antitoisto-dupleksiin Cas9 sitoutuu suoraan REC1-domeeniin (Helikaali-I-domeeniin), arginiini-rikkaan silta-heliksin ja PI-domeenin avulla, kun rungon silmukkaan 2 tapahtuvat kontaktit pääosin RuvC- ja PI-domeenien kautta ovat vähäisempiä. Synteettiseen *sgRNA*:n rungon silmukkaan 3 ei Cas9-entsyymillä ole kontaktia ollenkaan. Näin ollen rungon 1 silmukka ja toisto-antitoisto-dupleksi ovat ainoat, mitä tarvitaan Cas9-*sgRNA*-kompleksin koaamiseen (Jinek *et al.* 2013, Jiang & Doudna 2017). Lyhyesti tiivistettynä apo-Cas9-entsyymi käy läpi suurimmat muutokset opas-RNA:n sitoutuessa ja lopulliseen aktiiviseen muotoon Cas9-opas-RNA-kompleksi pääsee siemenalueen kaksoisjuosteisen DNA:n PAM-sekvenssin tunnistuksen ja komplementaarisen protospacer-sekvenssin avulla.

3.3.2. Opas-RNA

Opas-RNA (eng. guide RNA; gRNA) johdattaa Cas9 entsyymien katalysoimaan tunkeutuvan liikkuvan kaksoisjuosteisen geneettisen elementin katkaisua. Opas-RNA koostuu kahdesta RNA:sta muodostaen *tracrRNA*:*crRNA*-dupleksin. *sgRNA* (eng. single guide RNA; yksittäinen opas-RNA) on synteettinen yksittäinen RNA, joka toimii samalla tavalla kuin *tracrRNA*:n ja *crRNA*:n yhdistelmä. Se voidaan suunnitella ja siirtää esimerkiksi plasmidissa soluihin, missä se transkriptoidaan ohjaamaan Cas9-entsyymien katalysoimaa kohde-DNA:n pilkkomista opas-RNA:n tapaan. *sgRNA*:n suunnitteluun riittää, että pilkottavan kohteen juoste ja PAM-sekvenssi tunnetaan (Jinek *et al.* 2012).

3.3.3. Nukleaasidomeenit

Cas9-entsyymi pitää sisällään kaksi samankaltaista domeenia, jotka ovat homologisia endonukleaasien HNH ja RuvC kanssa. Domeeneja kutsutaan siis RuvC-domeeniksi ja HNH-domeeniksi. RuvC-domeeni jakautuu kolmeen eri RuvC-motiiviin (RuvC I-III) ja kohtaa PI-domeenin, jonka kanssa se muodostaa positiivisesti varautuneen pinnan. Varautunut proteiinin pinta muodostaa vuorovaikutuksen sgRNA:n 3' hännän kanssa. HNH-domeeni sijoittuu RuvC-domeenin toisen ja kolmannen motiivin väliin, jossa se on muodostaa vain muutamia sidoksia muuhun Cas9-proteiiniin (**Kuva 1**) (Nishimasu *et al.* 2014).

RuvC:n kaltainen domeeni

RuvC-domeeni koostuu kuudesta säikeisestä ja sekoittuneesta β -laskoksesta, joita reunustaa useampi α -heliksi ja kaksi ylimääräistä kaksisuuntaista ja kaksijuosteista β -laskosta (Nishimasu *et al.* 2014). Domeeni on rakenteellisesti samankaltainen ribonukleasi H-laskoksen omaavien retrovirusintegraasien superperheen jäsenten kuten *Escherichia coli*n RuvC-resolvaasien kanssa. Samankaltaisuutta näiden väliltä löytyy noin 12-14%. RuvC-resolvaasit ovat dimeerisiä endonukleasi-proteiineja, joita löytyy esimerkiksi gram-negatiivisilta bakteereilta. Niistä löytyy neljä katalyyttistä aminohappotähdettä ja ne toimivat osana HJ-migraatiota katkaisten symmetrisesti Holliday-liitoksista (eng. HJ; Holliday junction) kaksi DNA-juostetta kahden metallin mekanismilla (Mn^{2+}) toimien yhdessä Holliday-liitoksia-sitovan RuvA-proteiinin ja helikaasi RuvB-proteiinin kanssa muodostaen RuvABC-kompleksin (Górecka *et al.* 2013). Cas9:n RuvC-domeeni eroaa RuvC-nukleasista siten, että nukleasi muodostaa dimeerisen proteiinin, joka katkaisee neljän katalyyttisen aminohapon (esimerkiksi *Thermos thermophilus* -bakteerin RuvC: Asp7, Glu70, His143 ja Asp146) avulla Holliday-liitoksia, kun RuvC-domeeni katkaisee neljän katalyyttisen aminohapon (Asp10, Glu762, His983 ja Asp986, jotka ovat verrattavissa *T. thermophiluksen* katalyyttisiin aminohappoihin) avulla opas-RNA:ta ei-komplementaarisen juosteen kohde-DNA-kaksoisjuosteesta. Eroavaisuuden takia Cas9:n RuvC-domeenia kutsutaan RuvC:n kaltaiseksi domeeniksi (Jinek *et al.* 2012, Nishimasu *et al.* 2014). Konservoituneen RNAasi-laskoksen lisäksi RuvC-domeeni on rakenteellisesti vuorovaikutuksessa PI-domeenin ja opas-RNA-kohde-DNA-heterodupleksin kanssa (Nishimasu *et al.*, 2014). Mutaatoinnalla RuvC-domeenin aktiivinen aminohappo aspartaatti alaniiniksi (D10A) RuvC-domeenissa saadaan Cas9:n proteiinista muokattua nikaasi (nCas9) (**Kuva 1**). Mutaatio inaktivoi RuvC-domeenin DNA-juosteen katkaisun (Jinek *et al.* 2012, Nishimasu *et al.* 2014).

HNH-domeeni

HNH-domeeni koostuu kaksijuosteisesta β -laskoksesta, joita ympäröi neljä α -heliksiä. Domeeni pitää sisällään samanlaisen $\beta\beta\alpha$ -metallilaskoksen kuin T4-faagin endonukleaasi VII. HNH-nukleaasi toimii kolmen katalyyttisen aminohapon (Asp839, His840 ja Asn863) avulla ja se katkaisee opas-RNA:lle komplementaarisen DNA-juosteen yhden metalli-ionin mekanismilla (Mg^{2+}). Myös HNH-domeenin voi muuttaa nikaasiksi alaniinimutaation avulla, mutatoimalla asparagiinin 839 alaniniksi (**Kuva 1**) (Jinek *et al.* 2012, Nishimasu *et al.* 2014, Jiang & Doudna, 2017).

3.3.4. PI-domeeni

PI-domeeni (eng. PAM-interacting) tai CTD-domeeni (eng. C-terminal domain) sijaitsee nukleaasi-lohkossa koostuen seitsemästä α -heliksistä ja useasta eri β -levystä. REC-lohkon tavoin PI-domeeni esiintyy ainoastaan Cas9-entsyymissä (Nishimasu *et al.* 2014). Apo-muodossa (Ilman opas-RNA:ta) PI-domeeni on epäjärjestyksessä, eikä se voi tunnistaa PAM-sekvenssiä kohde-DNA:sta. Entsyymillä ollessa sitoutuneena kohde-DNA:han PI-domeeni sijoittuu siten, että se tunnistaa opas-RNA:lle ei-komplementaarista DNA-juosteesta PAM-sekvenssin. spCas9 entsyymillä PAM-sekvenssi on muotoa 5'-NGG-3', missä molemmat guaniini-nukleotidit tunnistetaan spesifisti vetysidosten avulla kahden PI-domeenin arginiini-aminohappotähteen avulla (R1333 ja R1335) (Nishimasu *et al.* 2014, Jiang & Doudna 2017).

4. DNA:n korjausmekanismit

Geneettiset mutaatiot toimivat uusien alleelien lähteenä ja niiden avulla populaatiossa yksilöiden välille syntyy geneettisiä eroja. Mutaatiot ovat lisäksi lähde geneettisille muutoksille, jotka voivat johtaa solujen kuolemaan, geneettisiin sairauksiin tai altistaa yksilöä syövän kehittymiseen. CRISPR-Cas-koneistot, joita käytetään genomien editoinnissa, eivät sisällä lyhyitä palindromisia toistojaksoja spacer-sekvenssien väleissä. Ainoastaan sgRNA ohjaa Cas9-entsyymiä katkaisemaan halutun kohde-DNA:n ja varsinaisen DNA:n muokkaus syntyy endogeenisten DNA:n korjausmekanismien ansiosta (Barrangou & Horvath 2017).

4.1 Kaksoisjuosteisen DNA:n vauriot

DNA-vaurioita aiheuttavat tekijät, kuten ionisoiva säteily ja klastogeenit (kemoterapiassa käytetyt lääkkeet) voivat pahimmillaan katkaista kaksoisjuosteen, mikä on DNA-vaurioista vakavin. Eukaryoottisolut ovat kehittäneet DNA-vaurioiden korjaamiselle erilaisia mekanismeja. Solujen luonnollisessa toiminnassa DNA-kaksoisjuoste voi katketa esimerkiksi immunoglobuliinigeenejä muokattaessa tai meioosin tekijänvaihdon aikana (Valerie & Povirk 2003, Chang *et al.* 2017, Al-Zain & Symington 2021). CRISPR-Cas-välitteisessä genomien muokkauksessa Cas-entsyymi katkaisee koodaavan juosteen ja mallijuosteen kohde-DNA:sta kahden katalyyttisen endonukleasidomeenin avustuksella. Katkaistun kaksoisjuosteisen DNA:n korjaus tapahtuu solujen omalla endogeenisellä koneistolla ja katkaistua kohtaa käytetään hyödyksi esimerkiksi lisäämällä tai poistamalla geeni (käytetty nimitys indel; insertion ja deletion). Genomin muokkaus tai kaksoisjuosteen katkaisu voivat johtaa ei-haluttuihin lopputuloksiin endogeenisten korjausmekanismien toimesta (Jiang & Doudna 2017). Yleisin korjausmekanismi on erittäin virhealtis ja templaattivapaa NHEJ (eng. Non-homologous end joining) eli ei-homologinen päiden liittyminen. Toinen hieman harvinaisempi virhevapaa ja templaatti-DNA:ta tarvitseva mekanismi on HR (eng. Homologous recombination) eli homologinen rekombinaatio. Genomin muokkaus onnistuu HDR-mekanismiin (eng. Homology-directed repair) kautta. HDR ei tarkoita yhtä ja tiettyä DNA:n korjausmenetelmää, vaan se kertoo templaatin käytöstä, jolla ohjataan DNA:n korjausta. Homologinen rekombinaatio on yleisin HDR-mekanismi (Xue & Greene 2021).

4.1.1. Ei-homologinen päiden yhdistyminen

Ei-homologisten päiden yhdistämiseen perustuva (eng. Non-homologous end joining; NHEJ) korjaus ei hyödynnä homologista aluetta sisarkromatidista, tai homologista kromosomia DNA:n kaksoisjuosteen korjaukseen. Korjausmenetelmä on nisäkässoluilla ensisijaisesti käytetty katkenneen kaksoisjuosteisen DNA:n korjausmekanismi koko solusyklin ajan. NHEJ-korjauksessa katkenneet kaksoisjuosteiset DNA:t yhdistetään toisiinsa kiinni mahdollisimman nopeasti vähäisillä DNA-päiden muokkauksilla. DNA-kaksoisjuosteen katkeamisen tunnistaa Ku70-Ku80-heterodimeeri lyhyesti Ku (tunnetaan myös nimellä XRCC6-XRCC5), johon muut NHEJ-korjauksen proteiinit sitoutuvat. DNA-riippuvaisella proteiinikinaasi-entsyymillä (eng. DNA-dependent protein kinase catalytic subunit; DNA-PKcs) on voimakas affiniteetti DNA-päihin, joihin Ku-proteiinit ovat sitoutuneet ja yhdessä Ku-proteiinien kanssa ne muodostavat DNA-PK-kompleksin. DNA-PK-kompleksi sitoo itseensä erityisen DNA-ligaasi IV:n eli liittjäentsyymin, joka liittää katkenneet päät toisiinsa. Ligaasi IV voi liittää tylpät

tai yksijuosteiset toisiinsa sopivat kohesiiviset päät yhteen. Ligaation ollessa vaikeaa ylimääräiset entsyymit kuten Artemis-nukleaasi, PNKP (polynukleotidikinaasi 3'-fosfataasi) ja DNA-polymeraasi prosessoivat päät siten, että ne voidaan liittää toisiinsa. Esimerkiksi DNA-PK-kompleksin tai Ku-heterodimeerin suppressio lisää homologisen rekombinaation kautta tapahtuvaa korjausta (Chang *et al.* 2017, Scully *et al.* 2019, Xue & Greene 2021).

Koska osa nukleotidisekvensistä häviää prosessin aikana, on NHEJ-korjaus erittäin virheherkkä. Yksi vaara NHEJ-korjauksessa on samanaikaisesti katkeavien kaksoisjuosteiden mahdollinen liittyminen väärään DNA-kaksoisjuosteeseen, minkä seurauksena voi tapahtua epänormaaleja kromosomirakennelmia tai kromosomistojen tuhoutumista. Geneettisen materiaalin translokaatio voi johtaa syövän kehittymiseen, jolloin geneettinen materiaali vaihtaa kromosomien välillä paikkaa (Kansikas *et al.* 2017). CRISPR-Cas-muokkauksessa NHEJ-korjausta käytetään tekemään deleetioita. Rajoitteina ovat kuitenkin syntyvien deleetioiden laajuuden hallitseminen ja DNA-kohteen ulkopuoliset muokkaukset (eng. off-target edits) (Xue & Greene 2021).

4.1.2. Homologinen rekombinaatio

Toisin kuin NHEJ-korjausmenetelmässä HR-korjausta voidaan käyttää ainoastaan S- ja G1-vaiheessa jakautuvissa soluissa kuten kantasoluissa ja sen tehokkuus riippuu solutyypistä, mikä rajaa CRISPR:n avulla hoidettavien sairauksien määrää (Saleh-Gohari & Helleday 2004). Korjausmenetelmä perustuu sisarkromatidin tai ulkoisen templaatti-DNA:n kautta tapahtuvaan korjaukseen (Scully *et al.* 2019). Kaksoisjuosteen korjaus somaattisissa soluissa suosii sisarkromosomia korjaustemplaattina homologisen kromosomin sijasta (Kadyk & Hartwell 1992). Homologisessa rekombinaatiossa katkaistun DNA-kaksoisjuosteen päitä voidaan lyhentää kaksivaiheisesti. Ensimmäisessä vaiheessa (lyhyen kantaman päiden resektio) käytetään lyhyen kantaman nukleaaseja ja toisessa vaiheessa (pitkän kantaman päiden resektio) pitkän kantaman nukleaaseja (Bonetti *et al.* 2018). Lyhentämisen tarkoituksena on valmistaa yksijuosteisia 3'-päihin päättyviä DNA-juosteita katkaisemalla 5'-päähän yksijuosteiset DNA-rihmat molemmilta puolilta katkaistua DNA:ta.

Lyhyen kantaman päiden resektio

Lyhyen kantaman päiden lyhentämisessä katkenneen DNA-kaksoisjuosteen 5'-päihin kiinnittyy erilliset MRN-kompleksit (MRE11-RAD50-NBS1-kompleksit), jotka lopulta liittyvät toisiinsa RAD-50:n (DNA:n korjausproteiini) korjausprosessin myötä. Ennen liittymistä katkenneita kaksoisjuosteisia päitä on lyhennettävä nukleolyttisellä degradaatiolla (Gobbini *et al.* 2020). Päiden lyhentäminen

käynnistyy MRE11-endonukleaasin (eng. double-strand break repair protein) kautta, missä MRE11 katkaisee 5'-pääteiset juosteet jopa 300 nukleotidin päästä kaksoisjuosteen katkoksesta edistään 3'-5'-eksonukleasiaktiivisuutta. MRN-kompleksin NBS1-proteiiniin (eng. Nijmegen breakage syndrome 1) kiinnittyy CtIP-eksonukleasi (eng. C-terminal-binding protein interacting protein), jonka seriinit S233, S276 ja S347 fosforyloituvat solusykliä säätelevien CDK-proteiinien (sykliinistä riippuvaisten kinaasien) ja ATM-proteini-kinaasin (eng. protein kinase ataxia-telangiectasia mutated) avulla edistään MRE11-proteiinin aktiivisuutta pilkkoa loput 5'-pääteisen juosteen nukleotidit aina kaksoisjuosteen katkokseen asti (Shiloh & Ziv 2013, Wang *et al.* 2013, Carusillo & Mussolino 2020, Gobbini *et al.* 2020). CtIP:n on oltava fosforyloituna CDK:n kautta ennen kuin se voidaan fosforyloida ATM:n toimesta. Tämän takia katkenneen kaksoisjuosteisen DNA:n homologinen rekombinaatio on solusykleistä riippuvainen tapahtuma (Wang *et al.* 2013).

Pitkän kantaman pään resektio

Pitkän kantaman päiden resektiossa MRN-kompleksin tuottamiin lyhyisiin yksijuosteisiin 3'-pään ulokkeisiin liittyy EXO1- ja/tai BLM/DNA2-entsyymejä. EXO1-entsyymi (eksonukleasi 1) poistaa nukleotideja 5'-3'-suuntaan muodostaen pitkän, yksijuosteisen 3'-pään ulokkeen (Xue & Greene 2021). BLM/DNA2-välitteisessä juosteen katkaisemisessa BLM (eng. Bloom syndrome protein) toimii helikaasina erottaen katkaistavan 5'-juosteen ja pysyvän 3'-juosteen. DNA2 (ATP-riippuvainen nukleasi/helikaasi DNA2) toimii nukleasina katkaisten erotettua 5'-juostetta (Daley *et al.* 2017).

RAD51-filamentin muodostuminen ja juosteeseen hyökkääminen (strand invasion)

Päiden resektiossa syntyneeseen yksijuosteiseen DNA:han sitoutuu heti heterotrimeerinen RPA-kompleksi (replikaatioproteiini A), joka koostuu RPA1-, RPA2- ja RPA3-proteiineista. Kompleksi estää yksijuosteista DNA:ta sitomasta muita tekijöitä ja suojaa DNA:ta nukleaseilta (Yates *et al.* 2018). DNA:n korjausproteiini RAD52:n avulla RPA-kompleksi poistetaan yksijuosteisesta DNA:sta, mikä mahdollistaa BRCA2-proteiinin (eng. Breast cancer type 2 susceptibility protein) avulla DNA:n korjausproteiini RAD51-rekombinaasin sitoutumisen yksijuosteiseen DNA:han ja RAD51-filamentin muodostumisen (Ma *et al.* 2017). Resektoitujen 3'-päiden täytyy löytää homologinen alue templaatti-DNA:sta (esimerkiksi sisarkromatidi tai synteettinen DNA-kaksoisjuoste) ja hyökätä (eng. strand invasion) sitoutuen emäspariperiaatteen mukaisesti templaattiin muodostaen D-silmukan. Sitoutumisen jälkeen katkennut juoste korjataan templaatin avulla (Wright *et al.* 2018).

DSBR ja SDSA

Somaattisten solujen homologisessa rekombinaatiossa on tunnistettu useampia RAD51-välitteisiä reittejä. Yksi reiteistä on synteestistä riippuvainen juosteiden yhdistyminen (eng. synthesis-dependent strand annealing; SDSA), missä toinen RAD51-filamentin peittämä lyhennetty juoste hyökkää ja sitoutuu emäspariperiaatteen mukaisesti homologiseen templaatti-DNA:han toisen pään ollessa passiivinen. Toisin sanoen SDSA-korjauksessa DNA-syntaasi korjaa katkenneen kaksoisjuosteesta toisen 3'-pään homologisen juosteen mukaisesti ja toinen katkaistu 3'-pää korjataan uuden homologisesta templaattista valmistetun DNA:n mukaisesti (Wright *et al.* 2018, Scully *et al.* 2019). SDSA ei muodosta Holliday-liitoksia, joten meiosisin aikana tekijänvaihtoa ei pääse tapahtumaan (Scully *et al.* 2019).

Toinen yleinen reitti on kaksoisjuosteisen katkoksen korjausmekanismi (eng. Double-stranded break repair; DSBR). Resektion ja juosteeseen hyökkäämisen jälkeen DSBR ja SDSA ovat erotettavissa toisistaan uuden DNA:n valmistuksen myötä (Scully *et al.* 2019). DSBR voi tapahtua kahdella eri tavalla, ensimmäisessä tapahtumassa molemmat resektoidut 3'-päät hyökkäävät templaatti-DNA:han ja molemmat päät syntetoidaan templaatin juosteiden mukaisesti. Toisessa tapahtumassa 3'-pään hyökkäyksessä muodostuneeseen D-silmukan vapaaseen juosteeseen hyökkää toinen resektoitu vapaa 3'-pää. Molemmissa tapahtumissa muodostuu kaksois-Holliday-liitos (eng. double Holliday-junction; dHJ), mikä voi aiheuttaa tekijänvaihtoa (Wright *et al.* 2018, Scully *et al.* 2019).

5. Crispr-Cas9 sovelluksia

Cas9-entsyymistä ja CRISPR-Cas9-tekniikasta on vuoden 2012 sgRNA:n kehityksen jälkeen tehty lukuisia tutkimuksia, joista usean tavoitteena on etsiä ratkaisuja erilaisten geneettisten mutaatioiden aiheuttamien sairauksien hoitoon. Tutkimusten määrä erilaisten Cas-entsyymien käytöstä on kasvanut valtavasti ja uusia erilaisia mitä mielenkiintoisempia sovelluksia syntyy päivittäin. Sovellusten avulla Cas9-entsyymi voidaan valjastaa esimerkiksi tutkijoiden diagnostiseksi työkaluksi lisäämällä siihen fluoresoivia proteiineja. Entsyymiin voidaan lisätä esimerkiksi käänteiskopijoijaentsyymi, mikä mahdollistaa entistä luovemman tavan muokata jopa ihmisen genomia.

5.1 dCas9 ja nCas9

Cas9-entsyymi toimii toisen DNA-juosteen katkaisevana nikaasina (eng. Cas9 nickase; nCas9), jolloin toinen domeeni on aktiivinen ja toinen inaktivoitu. Jinek työryhmineen (2012) inkuboi natiivia plasmidi-DNA:ta Cas9-proteiinin kanssa, josta oli poistettu käytöstä toinen domeeneista HNH tai RuvC. Tuloksena kaksois-RNA-johdetut mutantti-Cas9-proteiinit olivat nikaasin tavoin katkaisseet vain toisen juosteen kaksoisjuosteisesta plasmidi DNA:sta. Lopputuotteena syntyi avoin pyöreä plasmidi, josta toinen juoste oli katkennut. Kokeessa villityypin Cas9-proteiini-tracrRNA:crRNA-kompleksi tuotti lineaarisen kaksoisjuosteisen DNA-tuotteen. nCas9:n tuottama yhden juosteen katkaisu lisää homologisen rekombinaation kautta tapahtuvaa DNA:n korjausta ja voi vähentää indel-mutaatioiden määrää, kun kahden DNA-juosteen sijaan katkaistaan vain yksi juoste (Ran *et al.* 2013). Kahden DNA-juosteen katkaisu on aina altis virheisiin ja kromosomin toiminnan hiljenemiseen.

Deaktivoidussa Cas9-entsyymissä (eng. deactivated Cas9; dCas9) endonukelaasidomeenit HNH ja RuvC ovat inaktivoituja mutaatioiden kautta (esimerkiksi spCas9:n H840A- ja D10A-mutaatioilla) (**Kuva 1**). Entsyymi sitoutuu kohdejuosteeseen sgRNA:n tai tracrRNA:crRNA-dupleksin ja PAM-sekvenssin tunnistuksen avulla, mutta ei katkaise sitä vaan pysyy sitoutuneena. Menetellä esimerkiksi kohdesolun transkriptioita voidaan aktivoida (eng. CRISPR activation; CRISPRa) fuusioimalla transkription aktivaattoridomeeneja effektoriproteiineista kuten VP64 tai p65 dCas9-entsyymiin. Menetelmää voidaan lisäksi käyttää kohdesolun transkription inhibitioon (eng. CRISPR inhibition; CRISPRi) fuusioimalla dCas9-entsyymiin transkription repressoridomeeneja kuten KRAB tai SID (Brocken *et al.* 2017). dCas9-entsyymiä voidaan käyttää diagnostiikassa lisäämällä siihen fluoresoivia molekyylejä kuten GFP (vihreää fluoresoiva proteiini; eng. green fluorescent protein) (Knott & Doudna 2018).

5.2 Aluke- ja emäsmuokkaus

DNA:n emäs- ja alukkeidenmuokkaustekniikat ovat työkaluja, joiden avulla ongelmat kaksoisjuosteisen DNA:n korjauksessa perinteisellä CRISPR-Cas9 menetelmällä verrattuna voidaan kiertää katkaisemalla vain toinen juoste kohde-DNA:sta. Näin soluille toksista kaksoisjuosteen katkeamista ei tapahdu ja genomin muokkaus voi olla turvallisempaa.

5.2.1. Emäsmuokkaus

Vuonna 2016 kehiteltiin täysin uusi tapa muokata genomia CRISPR-Cas9-koneiston avulla. Menetelmän nimeksi tuli emäsmuokkausmenetelmä (eng. base editing), missä kohde-DNA:n emäs pystytettiin

vaihtamaan toiseen ilman kaksoisjuosteen katkaisua tai ulkoista DNA-templaattia käyttämällä hyödyksi Cas9-nikaasia, johon lisättiin deaminaasientsyymi ja emäksenkorjauksen (eng. Base excision repair; BER) estäjäentsyymi (Komor *et al.* 2016).

DNA-emäsmuokkaajat pitävät sisällään kaksi pääasiallista rakenteellista tehtävää, Cas-entsyymi toimii apuna DNA:n kohdennetulle sitoutumiselle ja yksijuosteisen DNA:n muokkaajaentsyyminä kohdennetun nukleotidin muokkaukseen. Emäsmuokkaajia on kahdenlaisia: sytosiiniemäsmuokkaajia ja adeniiniemäsmuokkaajia, mitkä pystyvät kollektiivisesti muokkaamaan kaikkia neljää eri transitiomutaatiota (C->T, T->C, A->G ja G->A) (Kantor *et al.* 2020). Kahdeksan eri transversiomutaation (A->C, C->A, A->T, T->A, T->G, G->T, G->C ja C->G) muokkaaminen on kuitenkin mahdotonta menetelmän avulla ilman kaksoisjuosteen katkaisua. Tähän pystyy vielä uudempi menetelmä: alukkeidenmuokkausmenetelmä (eng. prime editing) (Anzalone *et al.* 2019).

Emäsmuokkausmenetelmän ensimmäiset ihmiskokeet aloitettiin vastikään Verve Technologies -nimisen yhtiön yrittäessä parantaa esimerkiksi familiaalista hyperkolesterolemiaa (FH-tauti) estäen PCSK9-geenin (eng. proprotein convertase subtilisin/kexin 9) toimintaa maksan soluissa. PCSK9-geenin tuottaman proteiinin pitoisuuden lasku verenkierrrossa laskee LDL-kolesterolipitoisuutta (eng. low density lipoprotein) lisäämällä LDL-reseptorien määrää verenkierrrossa. Reseptorien määrän nousu lisää LDL-kolesterolin viemistä soluihin (ClinicalTrials.Gov 2022).

5.2.2. Alukemuokkaus

Kaksoisjuosteisen DNA:n katkeamisen aiheuttamien hankaluuksien takia vuonna 2019 kehitetty alukkeisiin perustuva alukemuokkausmenetelmä (eng. prime editing) mahdollistaa genomien muokkauksen ilman DNA:n kaksoisjuosteen katkaisua, minkä myötä DNA:n korjausmekanismeja ei tarvita. Menetelmässä käytetään heikennettyä nCas9-entsyymiä, johon on sitoutuneena käänteiskopioijaentsyymi. Tätä nCas9-käänteiskopioijaentsyymi-kompleksia ohjaa sgRNA:sta edelleen kehitetyllä prime editing guide RNA:lla (pegRNA). Kokonaisuudessaan tätä entsyymi-RNA-kompleksia kutsutaan prime editoriksi. pegRNA määrittää halutun kohteen ja koodaa halutun genomien muokkauksen (Anzalone *et al.* 2019). Menetelmän avulla on mahdollista muokata kohde-DNA:han lähes mikä tahansa paikallinen mutaatio, minkä lisäksi useiden kymmenien emäsparien korvaaminen, lisääminen ja poistaminen on mahdollista (Nelson *et al.* 2021).

6. Yhteenveto

Vaikka CRISPR-Cas-välitteisiä menetelmiä on kehitetty kymmenen vuotta, ovat yhden pistoksen tai lääkkeen hoidot vielä monien mutkien takana. Uuden hyväksyvän ja jatkuvasti kehittyvän sektorin perusteita voivat ravistella yksittäiset tapaukset, kuten 1999 Yhdysvalloissa tapahtunut adenovirusvektorin aiheuttama kuolema. Potilaan kuolema aiheutui solujen immuunireaktiosta vektorina toimivaa adenovirusta vastaan (Stolberg 1999). Uusilla menetelmillä on kuitenkin saatu positiivisia tuloksia, ehkä 1990-luvun esimerkistä valveutuneina paljon kontrolloidumpien tutkimusten, kuten sirppisoluanemian hoidossa. Oman haasteensa sektorille tuo nykyinen nopea kehitys, kuten uusien lääkkeiden yhä nopeampi hyväksymisprosessi ja sijoittajien mielenkiinnon hiipuminen tuloksia odotellessa. Ihmisten perimää muokattaessa eettiset kysymykset ovat kaikkien huulilla ja lainsäädäntö kulkee vielä menetelmien jäljessä. Tulevaisuudessa muokkaavan entsyymin rakenteen, historian, DNA:n kaksoisjuosteen korjauksen ja erilaisten menetelmien tunteminen voivat auttaa CRISPR-Cas-välitteisten menetelmien kehittymistä.

7. Kirjallisuusviitteet

- Al-Zain, A. M., & Symington, L. S. (2021). The dark side of homology-directed repair. *DNA Repair*, 106, 103181. <https://doi.org/10.1016/J.DNAREP.2021.103181>
- Anzalone, A. v., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., Chen, P. J., Wilson, C., Newby, G. A., Raguram, A., & Liu, D. R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 576(7785), 149–157. <https://doi.org/10.1038/S41586-019-1711-4>
- A Study of VERVE-101 in Patients With Familial Hypercholesterolemia and Cardiovascular Disease - Full Text View - ClinicalTrials.gov.* (n.d.). Retrieved 25 August 2022, from <https://clinicaltrials.gov/ct2/show/NCT05398029?cond=base+editing&draw=2&rank=1>
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science (New York, N.Y.)*, 315(5819), 1709–1712. <https://doi.org/10.1126/SCIENCE.1138140>
- Barrangou, R., & Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. *Nature Microbiology* 2017 2:7, 2(7), 1–9. <https://doi.org/10.1038/NMICROBIOL.2017.92>
- Bolotin, A., Quinquis, B., Sorokin, A., & Dusko Ehrlich, S. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading, England)*, 151(Pt 8), 2551–2561. <https://doi.org/10.1099/MIC.0.28048-0>
- Bonetti, D., Colombo, C. V., Clerici, M., & Longhese, M. P. (2018). Processing of DNA ends in the maintenance of genome stability. *Frontiers in Genetics*, 9(SEP), 390. <https://doi.org/10.3389/FGENE.2018.00390/BIBTEX>
- Brocken, D. J. W., Tark-Dame, M., & Dame, R. T. (2017). dCas9: A Versatile Tool for Epigenome Editing. *Current Issues in Molecular Biology* 2018, Vol. 26, Pages 15-32, 26(1), 15–32. <https://doi.org/10.21775/CIMB.026.015>
- Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. v., & van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science (New York, N.Y.)*, 321(5891), 960–964. <https://doi.org/10.1126/SCIENCE.1159689>
- Carusillo, A., & Mussolino, C. (2020). DNA Damage: From Threat to Treatment. *Cells* 2020, Vol. 9, Page 1665, 9(7), 1665. <https://doi.org/10.3390/CELLS9071665>

- Chang, H. H. Y., Pannunzio, N. R., Adachi, N., & Lieber, M. R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* 2017 18:8, 18(8), 495–506. <https://doi.org/10.1038/NRM.2017.48>
- Charpentier, E., Richter, H., van der Oost, J., & White, M. F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiology Reviews*, 39(3), 428. <https://doi.org/10.1093/FEMSRE/FUV023>
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science (New York, N.Y.)*, 339(6121), 819–823. <https://doi.org/10.1126/SCIENCE.1231143>
- Daley, J. M., Jimenez-Sainz, J., Wang, W., Miller, A. S., Xue, X., Nguyen, K. A., Jensen, R. B., & Sung, P. (2017). Enhancement of BLM-DNA2-Mediated Long-Range DNA End Resection by CtIP. *Cell Reports*, 21(2), 324–332. <https://doi.org/10.1016/J.CELREP.2017.09.048>
- Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 471(7340), 602–607. <https://doi.org/10.1038/NATURE09886>
- Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346(6213). <https://doi.org/10.1126/SCIENCE.1258096>
- Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A. H., & Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*, 468(7320), 67–71. <https://doi.org/10.1038/NATURE09523>
- Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 109(39). <https://doi.org/10.1073/PNAS.1208507109/-/DCSUPPLEMENTAL/PNAS.201208507SI.PDF>
- Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., & Liu, D. R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 551(7681), 464–471. <https://doi.org/10.1038/NATURE24644>
- Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*, 154(2), 442. <https://doi.org/10.1016/J.CELL.2013.06.044>

- Gobbini, E., Casari, E., Colombo, C. V., Bonetti, D., & Longhese, M. P. (2020). The 9-1-1 Complex Controls Mre11 Nuclease and Checkpoint Activation during Short-Range Resection of DNA Double-Strand Breaks. *Cell Reports*, 33(3), 108287. <https://doi.org/10.1016/J.CELREP.2020.108287>
- Górecka, K. M., Komorowska, W., & Nowotny, M. (2013). Crystal structure of RuvC resolvase in complex with Holliday junction substrate. *Nucleic Acids Research*, 41(21), 9945–9955. <https://doi.org/10.1093/NAR/GKT769>
- Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., & Marraffini, L. A. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature*, 519(7542), 199–202. <https://doi.org/10.1038/NATURE14245>
- Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., & Charpentier, E. (2018). The Biology of CRISPR-Cas: Backward and Forward. *Cell*, 172(6), 1239–1259. <https://doi.org/10.1016/J.CELL.2017.11.032>
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakamura, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *Journal of Bacteriology*, 169(12), 5429–5433. <https://doi.org/10.1128/JB.169.12.5429-5433.1987>
- Jansen, R., van Embden, J. D. A., Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43(6), 1565–1575. <https://doi.org/10.1046/J.1365-2958.2002.02839.X>
- Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 Structures and Mechanisms. *Https://Doi-Org.Pc124152.Oulu.Fi:9443/10.1146/Annurev-Biophys-062215-010822*, 46, 505–529. <https://doi.org/10.1146/ANNUREV-BIOPHYS-062215-010822>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096), 816–821. <https://doi.org/10.1126/SCIENCE.1225829>
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., & Doudna, J. (2013). RNA-programmed genome editing in human cells. *ELife*, 2(2). <https://doi.org/10.7554/ELIFE.00471>
- Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., & Doudna, J. A. (2014). Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science (New York, N.Y.)*, 343(6176), 1247997. <https://doi.org/10.1126/SCIENCE.1247997>

- Kadyk, L. C., & Hartwell, L. H. (1992). Sister chromatids are preferred over homologs as substrates for re-combinational repair in *Saccharomyces cerevisiae*. *Genetics*, *132*(2), 387–402.
<https://doi.org/10.1093/GENETICS/132.2.387>
- Kansikas, M., Nyström, M., & Peltomäki, P. (2017). *DNA:n korjausmekanismien häiriöt ja niiden lääketieteellisen merkitys*. <https://helda.helsinki.fi/handle/10138/237123>
- Kantor, A., McClements, M. E., & Maclaren, R. E. (2020). CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *International Journal of Molecular Sciences*, *21*(17), 1–22. <https://doi.org/10.3390/IJMS21176240>
- Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. *Science (New York, N.Y.)*, *361*(6405), 866. <https://doi.org/10.1126/SCIENCE.AAT5011>
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* *2015* *533*:7603, *533*(7603), 420–424.
<https://doi.org/10.1038/NATURE17946>
- Koonin, E. v., & Makarova, K. S. (2022). Evolutionary plasticity and functional versatility of CRISPR systems. *PLOS Biology*, *20*(1), e3001481. <https://doi.org/10.1371/JOURNAL.PBIO.3001481>
- Koonin, E. v., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, *37*, 67–78. <https://doi.org/10.1016/J.MIB.2017.05.008>
- le Rhun, A., Escalera-Maurer, A., Bratovič, M., & Charpentier, E. (2019). CRISPR-Cas in *Streptococcus pyogenes*. *RNA Biology*, *16*(4), 380–389. <https://doi.org/10.1080/15476286.2019.1582974>
- Ma, C. J., Gibb, B., Kwon, Y., Sung, P., & Greene, E. C. (2017). Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Research*, *45*(2), 749–761. <https://doi.org/10.1093/NAR/GKW1125>
- Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H., & Joung, J. K. (2013). CRISPR RNA-guided activation of endogenous human genes. *Nature Methods*, *10*(10), 977–979.
<https://doi.org/10.1038/NMETH.2598>
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J. M., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. v. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* *2011* *9*:6, *9*(6), 467–477.
<https://doi.org/10.1038/NRMICRO2577>
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P.,

- White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., ... Koonin, E. v. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* 2015 13:11, 13(11), 722–736. <https://doi.org/10.1038/NRMICRO3569>
- Makarova, K. S., Wolf, Y. I., Iranzo, J., Shmakov, S. A., Alkhnbashi, O. S., Brouns, S. J. J., Charpentier, E., Cheng, D., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Scott, D., Shah, S. A., Siksny, V., Terns, M. P., Venclovas, Č., White, M. F., Yakunin, A. F., ... Koonin, E. v. (2019). Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology* 2019 18:2, 18(2), 67–83. <https://doi.org/10.1038/S41579-019-0299-X>
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science (New York, N.Y.)*, 339(6121), 823–826. <https://doi.org/10.1126/SCIENCE.1232033>
- Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science (New York, N.Y.)*, 322(5909), 1843–1845. <https://doi.org/10.1126/SCIENCE.1165771>
- McGinn, J., & Marraffini, L. A. (2018). Molecular mechanisms of CRISPR–Cas spacer acquisition. *Nature Reviews Microbiology* 2018 17:1, 17(1), 7–12. <https://doi.org/10.1038/S41579-018-0071-7>
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, 60(2), 174–182. <https://doi.org/10.1007/S00239-004-0046-3>
- Mojica, F. J. M., Juez, G., & Rodriguez-Valera, F. (1993). Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Molecular Microbiology*, 9(3), 613–621. <https://doi.org/10.1111/J.1365-2958.1993.TB01721.X>
- Nelson, J. W., Randolph, P. B., Shen, S. P., Everette, K. A., Chen, P. J., Anzalone, A. v., An, M., Newby, G. A., Chen, J. C., Hsu, A., & Liu, D. R. (2021). Engineered pegRNAs improve prime editing efficiency. *Nature Biotechnology* 2021 40:3, 40(3), 402–410. <https://doi.org/10.1038/s41587-021-01039-7>
- Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. (2014). Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*, 156(5), 935–949. <https://doi.org/10.1016/J.CELL.2014.02.001>
- Nozawa, T., Furukawa, N., Aikawa, C., Watanabe, T., Haobam, B., Kurokawa, K., Maruyama, F., & Nakagawa, I. (2011). CRISPR Inhibition of Prophage Acquisition in *Streptococcus pyogenes*. *PLOS ONE*, 6(5), e19543. <https://doi.org/10.1371/JOURNAL.PONE.0019543>

- Ochman, H., Lawrence, J. G., & Grolsman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 2000 405:6784, 405(6784), 299–304. <https://doi.org/10.1038/35012500>
- Pacesa, M., Loeff, L., Querques, I., Muckenfuss, L. M., Sawicka, M., & Jinek, M. (2022). R-loop formation and conformational activation mechanisms of Cas9. *Nature* 2022, 1–6. <https://doi.org/10.1038/s41586-022-05114-0>
- Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L., & Severinov, K. (2010). Transcription, processing and function of CRISPR cassettes in Escherichia coli. *Molecular Microbiology*, 77(6), 1367–1379. <https://doi.org/10.1111/J.1365-2958.2010.07265.X>
- Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading, England)*, 151(Pt 3), 653–663. <https://doi.org/10.1099/MIC.0.27437-0>
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 152(5), 1173–1183. <https://doi.org/10.1016/J.CELL.2013.02.022>
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 2013 8:11, 8(11), 2281–2308. <https://doi.org/10.1038/NPROT.2013.143>
- Saleh-Gohari, N., & Helleday, T. (2004). Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Research*, 32(12), 3683. <https://doi.org/10.1093/NAR/GKH703>
- Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology* 2019 20:11, 20(11), 698–714. <https://doi.org/10.1038/s41580-019-0152-0>
- Shiloh, Y., & Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nature Reviews Molecular Cell Biology* 2013 14:4, 14(4), 197–210. <https://doi.org/10.1038/nrm3546>
- Sternberg, S. H., Richter, H., Charpentier, E., & Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Molecular Cell*, 61(6), 797–808. <https://doi.org/10.1016/J.MOLCEL.2016.01.030>
- Stolberg, S. G. (1999). *The Biotech Death of Jesse Gelsinger - The New York Times*. <https://www.nytimes.com/1999/11/28/magazine/the-biotech-death-of-jesse-gelsinger.html>

- Szczelkun, M. D., Tikhomirova, M. S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., & Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(27), 9798–9803. https://doi.org/10.1073/PNAS.1402597111/SUPPL_FILE/PNAS.1402597111.SAPP.PDF
- Valerie, K., & Povirk, L. F. (2003). Regulation and mechanisms of mammalian double-strand break repair. *Oncogene* *2003* 22:37, *22*(37), 5792–5812. <https://doi.org/10.1038/SJ.ONC.1206679>
- Wang, H., Shi, L. Z., Wong, C. C. L., Han, X., Hwang, P. Y. H., Truong, L. N., Zhu, Q., Shao, Z., Chen, D. J., Berns, M. W., Yates, J. R., Chen, L., & Wu, X. (2013). The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR-Mediated Double-Strand Break Repair. *PLOS Genetics*, *9*(2), e1003277. <https://doi.org/10.1371/JOURNAL.PGEN.1003277>
- Wright, W. D., Shah, S. S., & Heyer, W. D. (2018). Homologous recombination and the repair of DNA double-strand breaks. *Journal of Biological Chemistry*, *293*(27), 10524–10535. <https://doi.org/10.1074/JBC.TM118.000372>
- Xue, C., & Greene, E. C. (2021). DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends in Genetics*, *37*(7), 639–656. <https://doi.org/10.1016/J.TIG.2021.02.008>
- Yates, L. A., Aramayo, R. J., Pokhrel, N., Caldwell, C. C., Kaplan, J. A., Perera, R. L., Spies, M., Antony, E., & Zhang, X. (2018). A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nature Communications* *2018* 9:1, *9*(1), 1–14. <https://doi.org/10.1038/s41467-018-07883-7>
- Yosef, I., Goren, M. G., & Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Research*, *40*(12), 5569. <https://doi.org/10.1093/NAR/GKS216>