

Frá arfgerð til svipgerðar:

tengsl breytileika
í genastjórnun við
svipgerðir og sjúkdóma

Hulda Karen Ingvarsdóttir
Arnar Pálsson

Myndun svipgerðar er flókið og margþætt ferli sem byggir á genatjáningu þar sem fjölmargir þættir hafa áhrif á tjáningu gena og þar með breytileika í svipgerð. Erfðabreytileiki, umhverfisáhrif og tilviljanakenndar breytingar spila þar inn í og samspil þessara þátta getur jafnvel stuðlað að myndun sjúkdóma. Með tilkomu lífmengjafræði, sem gerir kleift að greina þúsundir þátta samtímis, hefur skilningur okkar á genatjáningu og undirliggjandi líffræðilegum ferlum aukist verulega. Lífmengjafræðilegar rannsóknir, hvort sem þær beinast að afmörkuðum kerfum eða genatjáningu í heild sinni, geta varpað ljósi á hvernig breytileiki í arfgerð birtist í svipgerð. Í þessari grein verður skoðað hvernig arfgerð mótar svipgerð, með sérstakri áherslu á breytileika á mismunandi stigum genastjórnunar, frá DNA til prótína og hvernig slíkur breytileiki getur haft áhrif á svipgerð og sjúkdóma. Nýjar greiningaraðferðir hafa verið þróaðar til að vinna úr gífurlegu magni ganga sem safnast hefur á síðustu árum. Þær gera m.a. kleift að kortleggja erfðarþætti sem hafa áhrif á mRNA-styrk allra gena lífveru, greina áhættuþætti fyrir sjúkdóma og tengja styrk próteina í blóði við sjúkdóma eins og Alzheimer, sykursýki og hjartasjúkdóma. Með samþættingu gagna af ólíkum stigum genatjáningar má dýpka skilning á samspili sameinda, stjórnkerfi frumna, þroskunarferli og lífeðlisfræði. Slík þekking getur leitt til þróunar persónubundinna lækningaraðferða þar sem ólík líffræðileg gögn úr sama einstaklingi veita upplýsingar um undirgerðir sjúkdóma og leiða til markvissari meðferða. Frekari rannsóknir á breytileika í genatjáningu og áhrifapáttum hennar munu áfram auka þekkingu okkar á sjúkdómum og bæta meðferðarúrræði.

INNGANGUR

Erfðaupplýsingar eru grundvallarþáttur í lífverum þar sem þær stjórna myndun svipgerðar hvers einstaklings. Vörpun arfgerðar yfir í svipgerð (tafla 1) krefst margra mismunandi stjórnunarþátta og sameinda sem leggja sitt að mörkum til þess að byggja eiginleika lífverunar, t.d. form og líffæri. Þessi vörpun er háð umhverfi sem hver lífvera býr við eða velur sér. Til dæmis þroskast fóstur fiska ekki ef þau lenda í of miklum hita eða frosti. Breytileiki í erfðum og umhverfisþáttum, samspil gena og umhverfisþátta sem og tilviljun stuðla að breytileika í eiginleikum svipgerðar eins og meðalhæð fólks eða fjölda bletta á tigrisdýrum. Erfðabreytileiki (tafla 1) er af mörgum gerðum, en algengastur er einkirnabreytileiki (e. single nucleotide polymorphism, SNP).¹ Með tilkomu aðferða eins og sambandsgreininga á erfðamengjum (e. genome wide association study, GWAS hér eftir, tafla 1) hafa fundist margir erfðabreytileikar sem tengjast sjúkdómum.² Einnig hefur RNA-raðgreining verið gífurlega mikilvæg en með henni er hægt að skoða virkni nær allra gena í mörgum sýnum eða einstaklingum.³

Þekking á stjórnunarferlum hinna mismunandi stiga genatjáningar hefur knúið framfarir á sviði líf- og læknávisinda. Rannsóknir á líffræðilegum kerfum, bæði afmörkuðum og í heild sinni, geta upplýst okkur um hvað stuðlar að breytileika í svipfari og erfðum. Í grein þessari verður fjallað um nýlegar framfarir í erfðamengjafræði og skyldum greinum. Höfundar rifja upp grundvallaratriði sameindalíffræði með það að markmiði að útskýra hversu miklar framfarir

hafa orðið í þessum geira. Með notkun dæma, m.a. nokkura innlendra verður sagt frá ólíkum aðferðum og hvað má læra af þeim. Greinin hallar töluvert til læknisfræðilegra nota en grundvallatriðin eru öll líffræðileg og eru rætur breytileika í svipfari lífvera í náttúrulegum stofnum sem tengist vistfræði, þróun og verndun lífríkisins.

Fyrst verður útskýrt hvernig breytileiki í gena-stjórnun tengist tilurð breytileika í svipgerðum og líkum á sjúkdómum. Farið verður yfir helstu stig genastjórnunar, frá arfgerð til prótíns, og rakið hvaða mismunandi kerfi stjórna genatjáningu og hvernig aðrir þættir eins og umframerfðir (tafla 1) tengjast genastjórnun. Rætt verður um orsakir svipgerðar-breytileika og í framhaldinu um lífmengjafræði, þ.e.a.s. gögn og aðferðir sem nýtast til að rannsaka genatjáningu á mismunandi stigum stjórnunar. Að greina mögulega orsakaþætti breytileika í stjórnun gena og tengja þá við svipgerð og sjúkdóma er lykilatriði til að dýpka skilning okkar á lífverum og vistkerfum, auk þess sem það getur leitt til þróunar nýrra meðferða við sjúkdómum.

GENASTJÓRNUN

Genatjáning er mikilvægt ferli sem leiðir til stöðugra svipgerða í öllum lífverum. Ferlið samanstendur af mörgum skrefum og kerfum sem í sameiningu ákvarða birtingarmynd svipgerðarinnar en aðrir þættir eins og t.d. mismunandi umhverfisaðstæður, tilviljun og erfðabreytileiki geta haft áhrif á svipgerðina.⁴ Þótt

erfðir séu skilvirkar er genatjáning breytileg á milli einstaklinga og tengist að vissu leyti breytileika í svipgerð þeirra.⁵ Ferlið frá arfgerð yfir í svipgerð er mismunandi eftir eiginleikum og getur verið stutt, langt, flókið eða einfalt (mynd 1).

Erfðamengið (e. genome) inniheldur gen sem geyma erfðaupplýsingar og þurfa að vera umritaðar yfir í RNA. Hér verður ekki fjallað ítarlegar um breytileika í erfðamengjum. Af ólíkum gerðum RNA er mRNA (e. messenger RNA) veigamest þar sem það er þýtt í prótín sem eru mikilvægustu byggingareiningar og efnahvatar frumunnar. Aðalferlin í genatjáningu eru umritun (e. transcription) og þýðing (e. translation). Í umritun er mRNA myndað eftir DNA móti. Í þýðingu er röð mRNAsins túlkað, sem leiðir til myndunar peptíðs með ákveðinni röð aminosýra (keðjur aminosýra myndar prótín).⁶ Eftir þýðingu er í mörgum tilfellum gerðar breytingar á prótínum (e. post-translational modifications). Genastjórn felur í sér flókið samspil margra prótína, RNA sameinda og annarra sameinda í frumum og lífverunni sem í sameiningu undirbyggja svipgerð einstaklings.⁴ Sá hluti verður ekki kannaður frekar hér en það er viðfangsefni þroskunar- og lífeðlisfræði.

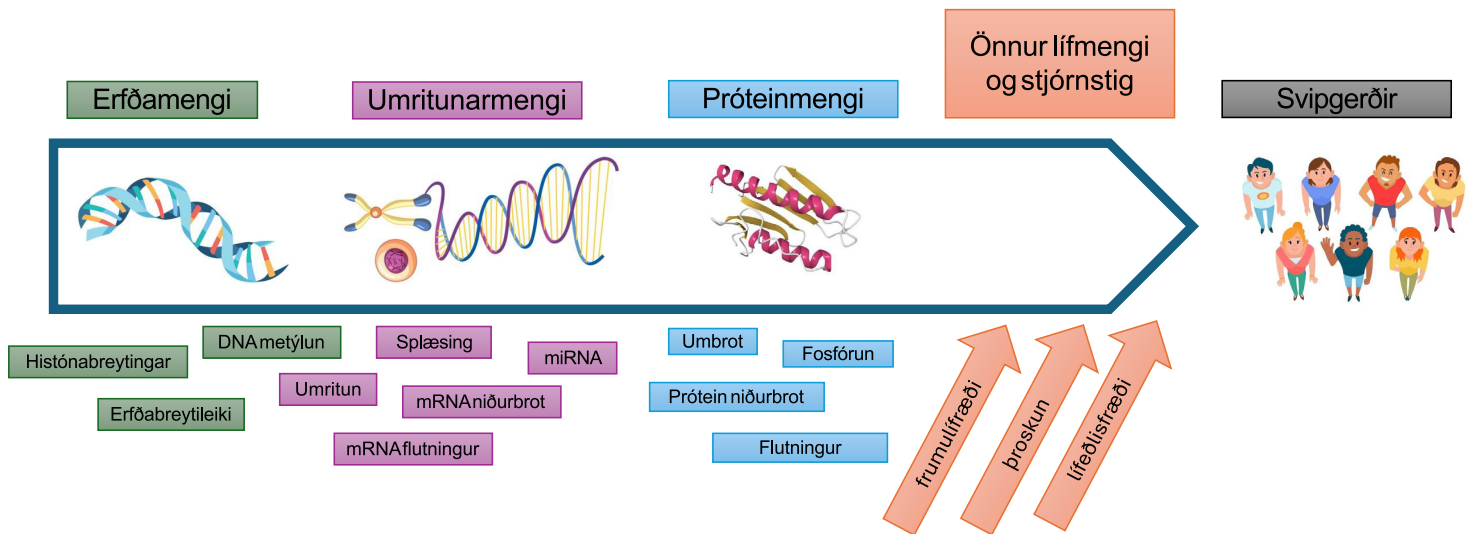
Hver einasta fruma í líkamanum inniheldur sömu erfðaupplýsingar í röðum núkleótíða (með nokkrum undantekningum) sem segja til um hvaða afurðum hvert gen tjáir fyrir. Erfðaupplýsingar gena standa einnig undir kerfum frumunnar sem sinna t.d. umritun og þýðingu en geta einnig framkvæmt efnafræðilegar breytingar á DNA og litni, t.d. metýlun sem leiðbeinir frumunum hvernig á að lesa upplýsingarnar. Þannig eru ólíkar frumur líkamans með sömu DNA röð en með t.d. mismunandi metýlunar mynstur í DNA og á histónum sem mótar líka genatjáninguna.⁷ Mismunur í metýlun DNA og gerðum og skreytingum históna á ólíkum stöðum í erfðamenginu heyrir undir kerfi sem miðla umframerfðum í þroskun fjölfrumunga. Þessi kerfi hjálpa frumum að muna hverjar þær eru og hvaða gen þær eiga að tjá.

STJÓRNUN GENATJÁNINGAR OG ÁHRIF HENNAR Á BREYTILEIKA Í SVIPGERÐ OG MYNDUN SJÚKDÓMA

Genatjáningu er stjórnað af mörgum þáttum á ýmsum stigum.⁴ Stjórn á umritun krefst t.d. erfðaupplýsinga sem byggjast á röð basa í genunum sjálfum sem innihalda bindiset fyrir svokallaða umritunarþætti (e. transcription factors). Umritunarþættir skiptast í almenna og sértæka umritunarþætti. Almennir umritunarþættir hjálpa RNA pólýmerasa að mynda grunnumritunarbúnað (e. basal transcription apparatus) við stýrilsvæði gena, oft nálægt svokölluðu TATA boxi.⁸ Sértækir umritunarþættir binda sértækar stjórnraðir sem geta verið bæði nærri og fjarri stýrlinum og hafa þannig áhrif á hvort umritun fari fram og hversu virk hún er.⁸ Stjórnunin byggist einnig á öðrum stjórnsameindum sem bæla, virkja og/eða

Tafla 1: Skilgreiningar á nokkrum lykilhugtökum erfðamengjafræði (e. genomics) – Definitions of key concepts used in genomics.

Hugtak	Skilgreining
Arfgerð (e. genotype)	Erfðasamsetning einstaklinga, getur vísað til staks gens eða alls mengisins.
Svipgerð (e. phenotype)	Eiginleikar lífveru (einstaklings) í víðustum skilningi, t.d. útlit, hegðan, þol og efnaskiptageta.
Genatjáning (e. gene expression)	Í þessari grein er þetta skilgreint sem ferli sem breytir erfðaupplýsingum í prótín (eða aðra virkni). Stundum einnig notað í víðari skilningi eins og um myndun svipgerðar frá arfgerð.
Erfðabreytileiki (e. genetic variant)	Breytileiki í röð eða byggingu erfðaefnis innan og milli einstaklinga eða stofna. Getur leitt til breytileika í svipgerð (t.d. sjúkdómum) eða ekki.
Einkirnabreytileiki/Stök basabreyting (e. single nucleotide polymorphism, SNP)	Skipti á einum basa á ákveðnum stað í erfðamenginu (t.d. A yfir T á stöðu 30200 í Egrfr geninu). Slík breyting verður fyrst til í einum einstaklingi en ef hún erfist áfram getur hún orðið algengari og finnst þá í fleiri einstaklingum innan þýðis.
Umframerfðir (e. epigenetics)	Hefur tvennskonar merkingu. Getur fjallað um erfðir eiginleika sem ekki tengjast breytileika á DNA röðinni sjálfri. Einnig notað um byggingu litnis í mismunandi frumum og breytingum á genatjáningu (sem ekki heldur skýrast af frávikum í sjálfri DNA röðinni).
Lífmengjafræði (e. omics)	Safn fagsviða sem fjalla um stór gagnasett um lífrænar sameindir á mismunandi stigum skipulags eins og erfðamengi, umritunarmengi, prótínmengi, efnaskiptamengi. Fjallar um aðferðir og uppgötvanir á þessum sviðum.
Umritunarmengi (e. transcriptome)	Heildarsafn RNA umrita af erfðaefninu í frumu eða sýni á ákveðnum tímapunkti. Til eru fjölbreyttar gerðir af RNA umritum s.s. mRNA, rRNA, tRNA og táknalaus RNA. Greiningar á umritunarmengjum sýna t.d. hvaða gen eru tjáð á ákveðnum tíma eða stað í fóstri og geta gefið innsýn í genastjórnun og starfsemi.
Sambandsgreining á erfðamengi GWAS (e. genome wide association study)	Aðferð sem skannar erfðamengið og prófar fyrir tengslum erfðabreytileika (oftast einkirnabreytileika) og aukinni áhættu á sjúkdómum eða öðrum einkennum. Prófað er fyrir mun á tíðni breytileika meðal einstaklinga með eða án sjúkdóms sjúkdóms. Einnig er hægt að skoða magnbundna eiginleika eins og hæð eða þykkt beina.



1. mynd. Yfirlit yfir ferli genatjáningar, hin mörgu skref á sameindasviðinu frá arfgerð yfir í svipgerð. Stóru þrepin eru úr erfðamengi (e. genome) yfir í umritunarmengi (e. transcriptome) og þaðan yfir í próteinmengi (e. proteome) sem með samspili annarra skrefa þar sem frumurnar, þroskun, og lífeðlisfræði tvinnast saman til að mynda svipgerð einstaklinganna. Mynd byggð á grein Buccitelli & Selbach 2020 4 með vissum breytingum – An overview of the gene expression process and the many molecular steps from genotype to phenotype. The main steps involve the transcription of the genome to the RNA level (transcriptome) and subsequently from mRNA's to the proteome. This process also includes other interacting steps where the cells, development, and physiology intertwine to form the phenotypes of individuals. Picture adapted and reworked from Buccitelli & Selbach 2020.⁴

viðhalda umritun. Stjórnþættir eru t.d. skilgreindir eftir staðsetningu miðað við umritaða genið, þ.e.a.s. annars vegar cis-stjórnraðir (e. cis regulatory elements) eða trans-þættir (e. trans-factors). Til þeirra fyrri flokkast t.a.m. stýrlar (e. promoter), bæliðir (e. suppressors) og efliraðir (e. enhancers)⁵ sem eru staðsettar á sama eintaki litnings og umritaða genið, oftast í námunda við komandi gen. Stýrlar eru staðsettir fyrir framan genin (miðað við stefnu umritunar, þ.e. 5' megin við genið). Aðrar stjórnraðir geta hins vegar verið staðsettar innan gensins sjálfs (t.d. í innröðum), fyrir aftan (þ.e. 3' megin við genið) eða jafnvel skarast við táknröðina. Þannig að samkvæmt skilgreiningu virka cis-raðir bara á gen á sama litningi. Þessar stjórnraðir samanstanda af sérstökum kinnaröðum (e. nucleotide sequences) sem umritunarþættir þekkja og bindast. Slík binding getur haft áhrif á hvort gen í nágrenni eru tjáð eða þöggud.⁹ Á hinn bóginn hafa svokallaðir trans-stjórnþættir áhrif á umritun gena t.d. ef þeir bindast DNA röðum nálægt þeim. Þessir þættir eru afurðir gena og geta virkað víða í erfðamenginu og þannig virkjað/hindrað mörg gen. Ólíkt cis-röðum geta trans-stjórnþættir (samkvæmt skilgreiningu) virkað á gen sem eru staðsett á öðrum litningum og hafa því víðtækari stjórnunargetu.¹⁰ Mikilvægastir fyrir ákveðnar svipgerðir eru umritunarþættir sem hafa mikla sértækni í ákveðin bindiset og eru mismunandi tjáðir eftir frumgerðum og þroskun og lífeðlisfræðilegu ástandi lífverunnar.

Breytileiki í cis-röðum eða trans-stjórnþáttum getur haft veruleg áhrif á mRNA styrk tiltekins gens og þar af leiðandi valdið mun á svipgerðum innan og milli tegunda og tengist þannig þróun og

aðlögun lífvera.¹⁰ Stökkbreytingar í cis-röðum s.s. efliröðum og stýrlum geta haft áhrif á magn mRNA ákveðins gens, staðsetningu þess eða tímasetningu í þroskun og stuðlað að fjölbreytileika í svipgerð innan og milli tegunda. Stökkbreytingar í stýrlum og stjórnstöðum manna tengjast margar sjúkdómum.⁵ Þar sem cis-raðir hafa áhrif á eitt eða nokkur gen en virka sjaldnast í öllum vefjum lífverunnar umber þróunin fleiri stökkbreytingar í þeim en í táknröðum gena. Því finnst meiri breytileiki á þessum svæðum á milli tegunda en innan táknaðanna. Áður ræddum við í Náttúrufræðingnum hvernig breytileiki í genastjórn tengist þróun nýrra forma.¹¹ Spyrja má hvort breytileiki í cis-röðum eða trans-þáttum hafi meiri áhrif á breytileika í genatjáningu á milli einstaklinga? Það er auðveldara að greina áhrif cis-breytileika en breytileiki í trans-þáttum getur haft áhrif á mörg gen eða víðtækari ferla sem getur leitt til svipfarsbreytileika á milli tegunda.¹² CAP próteinið (umritunarþáttur) er dæmi um trans-stjórnþátt sem hefur áhrif á tjáningu margra gena, þar á meðal hóp gena sem kóðar fyrir prótínum sem sjá um að brjóta niður (e. catabolize) galaktósa og arabínósa. Ef stökkbreyting verður í geninu sem framleiðir CAP (kallað crp) getur það leitt til skerðingar á tjáningu margra gena sem taka þátt í niðurbroti sykra eins og galaktósa og arabínósa og þannig haft víðtæk áhrif á síðari skref í efnaskiptaferlinu.¹³ Cis-raðir og trans-þættir, þrátt fyrir að vera bæði stjórnþættir, hafa ólík áhrif á genatjáningu og mismiklar afleiðingar. Breytileiki í fleiri þáttum síðar í ferlinu getur haft áhrif á breytileika í svipfari.

Meðal heilkjörnunga með innraðir (e. introns)

er for-mRNA (e. pre-mRNA) verkað til að mynda mRNA. Á þessu stigi eru raðir innraðir fjarlægðar úr for-mRNA-inu með ferli sem kallast splæsing (e. mRNA splicing). Splæsing getur leitt til ólíkra mRNA afbrigða úr for-mRNA hvers geni og þannig leitt til þess að mismunandi prótín-ísóform myndast, jafnvel hvert með ólíka virkni.¹⁴ Þetta er talið stuðla að mun á frumgerðum og jafnvel fjölbreytileika í svipgerð. Á svipaðan hátt og í umritunarstjórn ræðst splæsing af bæði cis-röðum innan umritsins og trans-þáttum (bæði RNA og prótín). Þetta ferli er mikilvægur partur af genajáningu flestra gena í manningum og öðrum dýrum.¹⁴ Breytileiki í mRNA splæsingu hefur verið tengdur við svipgerðarbreytileika í mönnum en u.þ.b. 10% af öllum stökkbreytingum í mönnum sem tengjast sjúkdómum virðast hafa áhrif á splæsingu.¹⁵ Annað mat er að um þriðjungur allra stakra basabreytinga sem hafa verið tengdir við sjúkdóma geti truflað splæsingu.¹⁴

Auk mRNA finnast í erfðamengjum dýra aðrar tegundir af RNA sem kóða ekki fyrir prótínum og kallast einu nafni táknaless RNA (e. non-coding RNA). Þetta er fjölbreyttur hópur með misvel skilgreind hlutverk. Tegundir af táknalessum RNA sameindum sem hafa mest áhrif á genatjáningu í umritun og seinni skrefum eru: miRNA (e. microRNA), siRNA (e. small interfering RNA), piRNA (e. piwi-interacting RNA) og löng táknaless RNA eða lncRNA (e. long non-coding RNA).¹⁶ Fyrst verður sagt frá s.k. löngum táknalessum RNA-sameindum en þau eru tiltölulega ný-uppgötvað með umrit lengri en 200 núkleótíð en afurðir þeirra skrá ekki fyrir prótínum. Löng táknaless RNA geta verkað óbeint sem umframerfða stjórnþáttur og mótað genastjórnun og tjáningu.¹⁷ Tjáning lncRNA er oft vefja- eða frumusérhæfð og koma þau að stjórnun genatjáningar á öllum stigum s.s. umritun, umframerfðum, prótein nýmyndun, RNA-verkun og stjórnun eftir-þýðingu.^{17,18} Slík lncRNA eru umrituð frá mörgum svæðum í erfðamengjum dýra og nýlegar rannsóknir ýja að mikilvægi þeirra fyrir mörg líffræðileg ferli þó margt sé ennþá óljóst. Upplýsingar úr GENCODE verkefninu benda til að í erfðamengi mannsins séu meira en 16.000 lncRNA gen.¹⁹ Til samanburðar eru í erfðamengi okkar um 20.000 gen sem skrá fyrir prótínum. Rannsóknir hafa sýnt að nokkrir stakir erfðabreytileikar innan lncRNA tengjast sjúkdómum sem gefur til kynna að þessir SNP (borið fram snippar) geti breytt virkni lncRNA og þannig truflað stjórnun genatjáningar og ýtt undir framvindu sjúkdóma.¹⁸

Önnur mikilvæg gerð af táknalessum RNA sem getur mótað svipgerðir og sjúkdóma er miRNA. Þessi hópur sameinda var uppgötvaður af m.a. Garry Ruvkun og Victor Ambros sem fengu Nóbelsverðlaunin í læknisfræði árið 2024.²⁰ miRNA eru litlar sameindir, yfirleitt í kringum 21 núkleótíð, sem gegna mikilvægu hlutverki við stjórnun genatjáningar eftir umritun²¹ með því að bindast við mRNA til að bæla tjáningu.²²

Þau gegna hlutverki í mörgum líffræðilegum ferlum og tjáning miRNA sameinda gerist á mismunandi stigum þroskunar dýra.²¹ Til dæmis eru ákveðin miRNA nauðsynleg fyrir myndun fósturvísis á meðan önnur eru mikilvæg á síðari stigum þroskunar eins og við aðgreiningu vefjaforvera eða myndun líffæra. Rannsóknir hafa afhjúpað að breytileiki í tjáningu ákveðinna miRNA gena tengist sjúkdómum í mönnum eins og krabbameinum og taugahrörunarsjúkdómum.²¹ Einnig hefur sjónin beinst að miRNA sameindum sem mögulegum lífmörkum (e. biomarkers) fyrir ýmsa sjúkdóma þar sem breytingar á styrk þeirra í blóði eru vísbendingar um sjúkdóma eða forstíg þeirra.

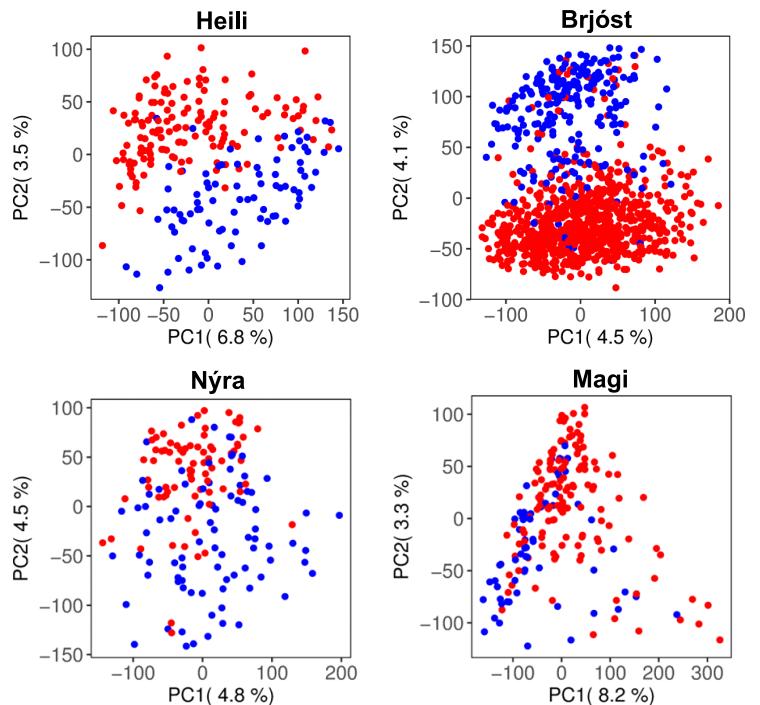
ÞROSKUN LÍFFÆRAKERFA OG GENATJÁNING

Líkt og flest dýr með flókna líkamsgerð hefur hefur maðurinn fjölbreytt líffærakerfi þar sem hvert þeirra hefur flókna lífeðlisfræðilega stjórnun. Líffærakerfi vinna einnig saman og hjálpast að við að byggja upp líkamann og viðhalda góðri heilsu.²³ Öll líffæri mótast vegna genatjáningar²⁴ en á vissum tímasteiðum þroskunar verða miklar breytingar á umritunarmengi (tafla 1) fruma samfara „ákvörðunum“ sem frumurnar taka.²⁵ Nýleg rannsókn kannaði genatjáningu á mismunandi stigum þroskunar í sex tegundum spendýra. Í ljós kom að þróun líffæra spendýranna er tengd þroskun þeirra. Tegundir spendýra hafa svipaða genatjáningu í líffærum snemma í þroskun en munurinn eykst síðar og er mestur hjá tegundum sem eru fjarskyldastar. Rannsóknin sýndi að snemma í þroskun hafa virk gen mörg hlutverk í mismunandi líffærum og á ólíkum stigum en eftir því sem líffæri aðgreinast og þroskast fá virk gen sértækara hlutverk. Þetta getur verið vegna breytinga í fjölvirkni (e. pleiotropy).²⁵ Fjölvirkni þýðir að eitt gen hefur áhrif á margar svipgerðir eða einkenni. Svona rannsóknir hjálpa okkur að skilja betur hvernig genatjáning og stjórnun hennar hefur áhrif á þroskun og þróun líffæra og veitir innsýn inn í það hvernig breytileiki í erfðum og/eða umhverfi hefur áhrif á svipfar.

UMFRAMERFÐIR

Aðrir þættir sem geta haft áhrif á breytileika í svipgerðum eða líkurnar á sjúkdómum eru umframerfðir (e. epigenetics). Orðið umframerfðir er notað um tvennt. Í fyrsta lagi um breytingar á genatjáningu eða frumustarfsemi sem erfast milli kynslóða og tengjast ekki mun í DNA röðinni sjálfri. Nokkrar rannsóknir hafa afhjúpað svokölluð foreldraáhrif eða kynslóða áhrif þar sem ástand foreldra eða annara fyrri kynslóða mótast ástand einstaklinga. Leitin að orsakaþáttum sem bera slík áhrif milli kynslóða bendir til kerfa umframerfða en margt er enn óvitað um þessi ferli. Einnig er lítið vitað hvort og hvernig þessi áhrif komist í gegnum endurstillingu umframerfða (e. epigenetic reprogramming) hjá fósturvísnum.²⁶ Hug-

2. mynd. Myndin er tekin úr rannsókn Vidman o.fl 2019⁴⁷ sem mat genatjáníngu í fjórum gerðum krabbameina. Á myndinni er notast við PCA aðferð til þess að bera kennsl á nýjar undirgerðir krabbameina. Hver punktur er eitt krabbameinssýni sem var RNA raðgreint. Á myndinni eru undirgerðirnar táknaðar með mismunandi lit (rautt og blátt) fyrir krabbamein úr fjórum ólíkum vefjum. Meginþættir (e. principal components) eru nýjar breytur sem byggðar eru á upprunarlegu breytunum í gagnasettinu þar sem fyrsti höfuðþátturinn (PC1) útskýrir mestan breytileika í gögnunum, næsti þáttur (PC2) næst mestan breytileika o.s.frv. Þetta minnkar þannig fjölda breyta sem þarf að vinna með án þess að mikilvægar upplýsingar tapist. Sýndir eru tveir fyrstu þættirnir (PC1 og PC2) og tilgreint á ásunum hversu mikinn hluta breytileikans hver þeirra útskýrir. Skýrt sjást undirgerðir krabbameinanna í heila (e. brain) og brjósti (e. breast) aðskiljast á fyrstu tveimur ásunum, en á sýnum úr nýra (e. kidney) og maga (e. stomach), er aðgreiningin ekki eins skýr. – The plot is from a study by Vidman et al. 2019⁴⁷ that investigated gene expression in four types of cancer. A PCA method was used to identify new cancer subtypes, with each dot representing an RNA-sequenced cancer sample. The subtypes are represented by different colors (red and blue), and each plot corresponds to a different tissue type. Principal components are new variables derived from the original dataset, where the first principal component (PC1) explains the largest proportion of variability in the data, the second principal component (PC2) explains the second largest proportion, and so on. This method reduces the number of dimensions to analyze while retaining essential information. The figure shows the first two principal components (PC1 and PC2), with the proportion of variability explained by each component indicated on the axes. A clear separation of cancer subtypes can be observed for brain and breast tissues along the first two principal components. However, for kidney and stomach tissues, the separation of subtypes is less distinct.



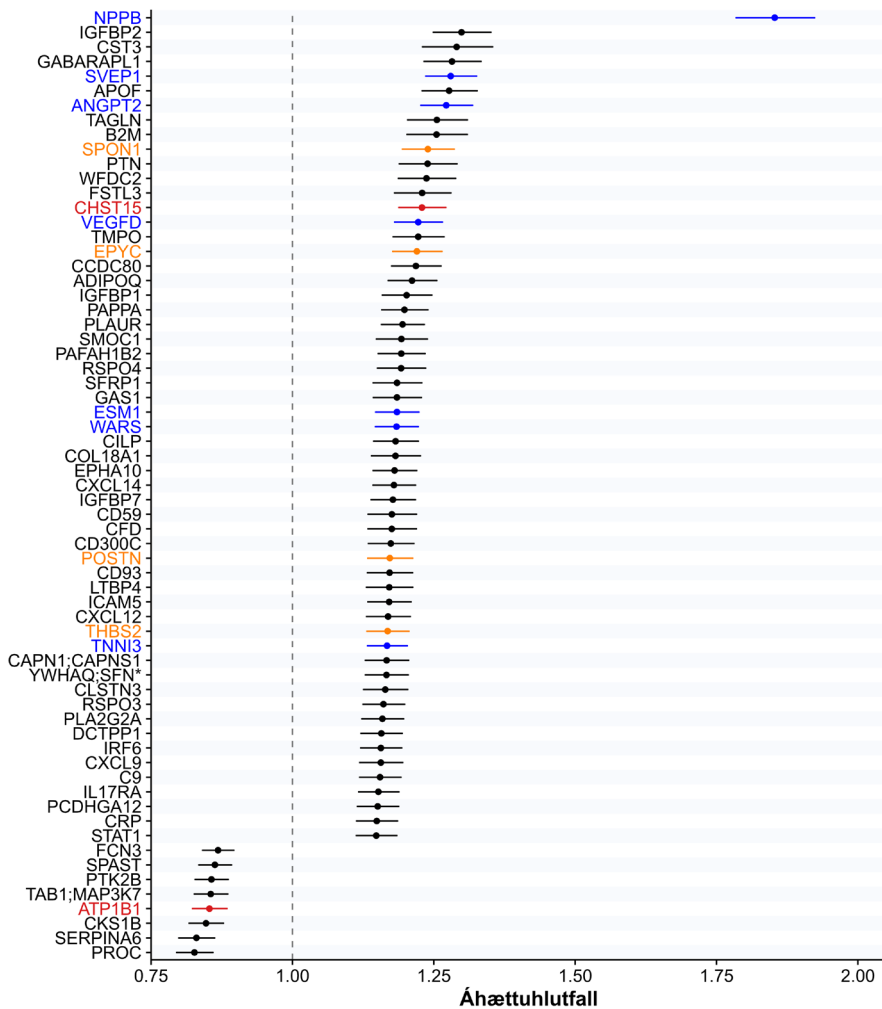
myndin er að þættir (ekki breytileika í DNA) erfist milli kynslóða og hafi áhrif á svipfar lífvera (líklega með því að tiltekin gen séu tjáð eða þöggud).

Í öðru lagi er hugtakið umframerfðir einnig notað um kerfi sem stýra aðgengi og eiginleikum litnis í frumum hvers einstaklings og hjálpa frumum að muna (eins og minnst var á að ofan). Í báðum tilfellum eru helstu sameindakerfin DNA metýlun (e. DNA methylation), histón breytingar (e. histone modifications) og táknalaus RNA (e. non-coding RNAs).²⁷ Dæmin hér að neðan tengjast aðallega áhrifum umframerfðakerfa í þroskun og lífeðlisfræði en ekki milli kynslóða.

Umframerfðir geta haft áhrif á líffræðilega ferla og breytileika í svipgerð.²⁸ Ein leið til að þetta geti gerst er með DNA metýlun. Meðal spendýra erfist mynstur DNA metýlunar almennt ekki til næstu kynslóðar líkt og DNA röðin sjálf en þó eru nokkar undantekningar. Kerfið virkar þannig að metýl hóp (CH3) er bætt við cytósín núkleótíð á DNA sameindinni og getur það í vissum tilfellum bælt genatjáníngu með því að koma í veg fyrir að umritunarþættir geti bundist DNA (t.d. í stýrlum). DNA metýlun á vissum setum í stjórnröðum eins og stýrlum eða efliröðum geta einnig bælt tjáningu nálgra gena.⁷ Rannsóknir hafa sýnt að breytingar á umframerfðamerkjum eins og DNA metýlun geta tengst ýmsum sjúkdómum þar sem þær hafa áhrif á tjáningu lykilgena en þetta gerist sérstaklega í krabbameini. Orsökina gæti þá verið stökkbreytingar sem leiða til mismunar í byggingu litnis í vefnum sem krabbameinið á upptök sín í²⁹ eða raskandi umhverfisþættir.

CpG eyjar (e. cytosine-phosphate-guanine islands) eru raðir með margar cytósín og gúanín tvenndir (basarnir tengdir í þessari röð). CpG eyjar eru oftast frá 100–1000 bp langar og innihalda hærra hlutfall af C og G miðað við aðrar raðir. Um það bil 70% af stýrlum manna hafa eina eða fleiri CpG eyjar og gegna þær hlutverki í genastjórn. Ef þær eru lítið eða ekki metýlraðar er genatjáníng stöðug. Mikil metýlun á þeim leiðir almennt til minni genatjáníngar (þöggunar). Stýrlar með CpG eyjar eru algengasta tegund stýrla í hryggdýrum en þeir tengjast stjórnun gena sem eru annað hvort vefjasértæk eða til búshalds (e. housekeeping genes) gena sem þurfa að vera tjáð í næstum öllum frumum fyrir eðlilega starfsemi.³⁰ Því er mikilvægt að búshaldsgenin verði ekki metýluð því það getur valdið þöggun í tjáningu þeirra og dregið úr lífslíkum.

Annað þrep umframerfða sem tengist líffræði fjölfruma heilkjörnunga er litni (e. chromatin) kjarnans. Kirnisögnin (e. nucleosome) í litni samanstendur af histón próteinum sem eru umvafin DNA þræði, oftast 146–147 bp um hvern klasa af histón prótínum. Kirnisagnir raðast síðan saman í stærri einingar sem leiðir til þéttar þökkunar á erfðafninu. Hvaða svæði erfðamengisins eru þéttþökkuð og hver minna eða óþökkuð ræðst af efnahópum sem hengdir eru á histón prótínum á ákveðnum stöðum og þeim ytri og innri boðum sem þessu stjórna. Þetta byggist því á sögu hvernar frumu og boðum sem hún hefur fengið. Efnahóparnir eru s.s. asetýling (e. acetylation), fosfórun (e. phosphorylation) og metýlun sem flokkast allt sem umbreyting á prótínum eftir



3. mynd. Niðurstöður úr rannsókn Jonmundsson et al 2023 á 4.765 einstaklingum úr AGES þýðinu þar sem 1.172 þátttakendur þróuðu með sér gáttatíf.⁵⁷ A) Hér má sjá áhættuhlutfall (e. hazard ratio) próteina sem eru marktækt tengd við gáttatíf, þ.e.a.s. magn þessara prótína í blóði er marktækt minna eða meira í þáttakendum sem þróuðu með sér gáttatíf miðað við aðra en búast má við vegna tilviljunar. Sterkasta tengingin var við NPPB (Natriuretic Peptid B) prótínið og sést á myndunum hvernig það hefur hæsta áhættuhlutfallið fyrir gáttatíf. – Results from the study by Jonmundsson et al. 2023⁵⁷ involving 4,765 individuals from the AGES cohort, of whom 1,172 developed atrial fibrillation. A) This panel shows the hazard ratios for proteins that were significantly associated with the development of atrial fibrillation. These proteins were either elevated or reduced in the blood of participants who later developed atrial fibrillation, beyond what would be expected by chance. The strongest association was observed for the protein NPPB (Natriuretic Peptide B), which had the highest hazard ratio for atrial fibrillation among all measured proteins.⁵⁷

þýðingu (e. post-translational modification). Þetta getur valdið breytingum á genatjáningu þar sem gen verða aðgengileg fyrir aðra stjórnþætti þegar losnar um litni og öfugt þegar þau bæld í þéttlitni.³¹ Nýlega hafa nokkrar tegundir RNA saneinda verið tengdar við stýringu á litnispökkun, eins og lncRNA.³²

UPPSPRETTUR BREYTILEIKANS

Öll ofangreind þrep genatjáningar geta verið breytileg vegna mismunar í erfðabáttum, ákveðnum umhverfisbreytum og sértækum tilviljunarkenndum atburðum og þannig stuðlað að svipfarsbreytileika. Erfðabreytileiki getur orsakast vegna margvíslegra stökkbreytinga í genum sem tilheyra ólíkum stjórnþrepum genatjáningar: stakar basabreytingar (SNP), innskot (e. insertions), úrfellingar (e. deletions)

o.fl. (flóknari breytingar verða ekki ræddar hér). Umhverfisþættir geta líka haft áhrif á kerfin sem stýra tjáningu gena og þar með svipfars. Ljóst er að svipfarsbreytileiki manna á milli er afurð víxlverkunnar milli margra erfðafraðilegra- og umhverfisþátta sem og hendingar. Svipfarseiginleikar og sjúkdómar eru einnig misjafnir hvað varðar mikilvægi erfðafamlags (arfgengi) og umhverfisþátta. Jafnvel þó tveir sjúkdómar hafi svipað arfgengi geta fá gen legið bak við einn en mörg annan. Sumir sjúkdómar myndast aðallega vegna galla í einum erfðabætti eins og t.d. stökkbreytingar sem tengjast cystic fibrosis sjúkdómnum (stökkbreyting í CFTR geninu). Aðrir sjúkdómar birtast vegna áhrifa margra þátta sem geta þá bæði verið erfðabreytileikar og ólíkir umhverfisþættir. Til

dæmis þróa einungis um 10–30% einstaklinga sem eru með stökkbreytingu í geni sem kallast complement component 2 með sér sjálfsofnæmissjúkdóminn helluroða (e. lupus erythematosus).³³ Hér fyrir neðan verða tekin dæmi um hvernig breytileiki sem verður vegna erfða eða umhverfis getur verið sértækur fyrir ákveðna vefi eða haft víðtækari áhrif.

Kímlínustökkbreytingar berast frá einni kynslóð til annarrar með egg- og sæðisfrumum þannig að allar frumur afkvæmisins bera þennan breytileika (í flestum tilfellum). Afleiðingar stökkbreytinga eru háðar hlutverki og tjáningu gensins sem þær verða í. Til dæmis geta stökkbreytingar í genum sem gegna mikilvægum hlutverkum við þroskunarferla eða nauðsynlega frumustarfsemi haft alvarlegar afleiðingar. Einnig skiptir tímasetning virkninar sem stökkbreytingin raskar máli. Stökkbreyting í efliröð sem er nauðsynleg snemma í þroskun getur leitt til fæðingargalla eða dauða fósturs á meðan stökkbreyting í efliröð sem virkar seinna í þroskun eða bara í fullorðnum hefur bara áhrif á ákveðinn vef og er ekki banvæn. Þetta geta verið efliraðir sama gens. Gen sem tengjast grundvallarferlum frumunnar eru nauðsynleg fyrir (nær) allar frumugerðir líkamans. Fóstur arfhrein um skaðlegar stökkbreytingar í slíkum genum komast ekki í gegnum fyrstu skiptingarnar eða verða ólífvænleg mjög snemma í þroskun.

Tvílitna lífverur hafa tvö eintök af hverju geni, eitt frá hvoru foreldri. Lífsnauðsynleg gen (e. essential genes) eru gen sem taka þátt í frumustarfsemi eins og t.d. DNA afritun, umritun, þýðingu, stjórnun frumuhringa og efnaskiptaferlum. Þessi gen eru tjáð í öllum frumum og ef bæði eintök af geninu missa virkni sína (arfhreint form) leiðir það til dauða.³⁴ Stakskortur (e. haploinsufficiency) er þegar annað eintakið missir virkni sem getur gerst vegna úrfellingar eða vanvirknibreytingar (e. loss-of-function mutation). Í mörgum tilvikum getur eitt eintak af geninu ekki staðið undir framleiðslu nægilegs magns af prótíni fyrir eðlilega frumustarfsemi – sem getur leitt til sjúkdómsástands t.d. ef genið spilar lykhlutverk í þroskun.³⁵ Dæmi um slíkan stakskort er Williams heilkennið (e. Williams syndrome) sem tengist 28 basapara úrfellingu á litningi 7q11.23. Einstaklingar með stakskort á þessu svæði hafa eitt virkt eintak af BAZ1B geninu sem gegnir hlutverki í tengslum við litnismótun (e. chromatin remodeling).³⁶ Vanvirknibreyting á báðum eintökum leiðir yfirleitt til dauða á meðan vanvirknibreyting á öðru eintakinu til skerðingar sem einstaklingurinn lifir yfirleitt með. Annað dæmi er mörg þeirra gena sem skrá fyrir prótínhluta ríbósómanna.³⁷

Öndverð dæmi eru líka til. Áhrif stökkbreytinga í mismunandi miRNA genum eru misalvarleg eftir því hvort miRNAið er vefjasértækt eða framleitt í öllum frumum líkamans. Tiltekið miRNA (miR-16) gegnir hlutverki í mörgum líffræðilegum ferlum, t.d. frumuhringnum og stýrðum frumudauða (e. apoptosis).

Úrfelling á miR-16 eða stökkbreyting veldur minni tjáningu þess sem getur ýtt undir myndun krabbameina eins og t.d. hægfara eitilfrumuhvítblæðis (e. chronic lymphocytic leukemia).³⁸ Almennt eru miRNA tjáð í mörgum líffræðilegum vefjum algengir krabbameinsvaldar. Hins vegar ef að miRNA sameindin sér um mikilvæga stjórnun í ákveðnum vef geta stökkbreytingar haft mikil áhrif og jafnvel verið banvænar. Í rannsókn á músum var hjarta- og vöðvafrumu-sértæka miRNA genið miR1-2 slegið út en það leiddi til galla í formþroskun (e. morphogenesis) hjartans eða sleglaskiptagalla (e. ventricular septal defect) sem veldur hárrí dánartíðni við fæðingu.²¹

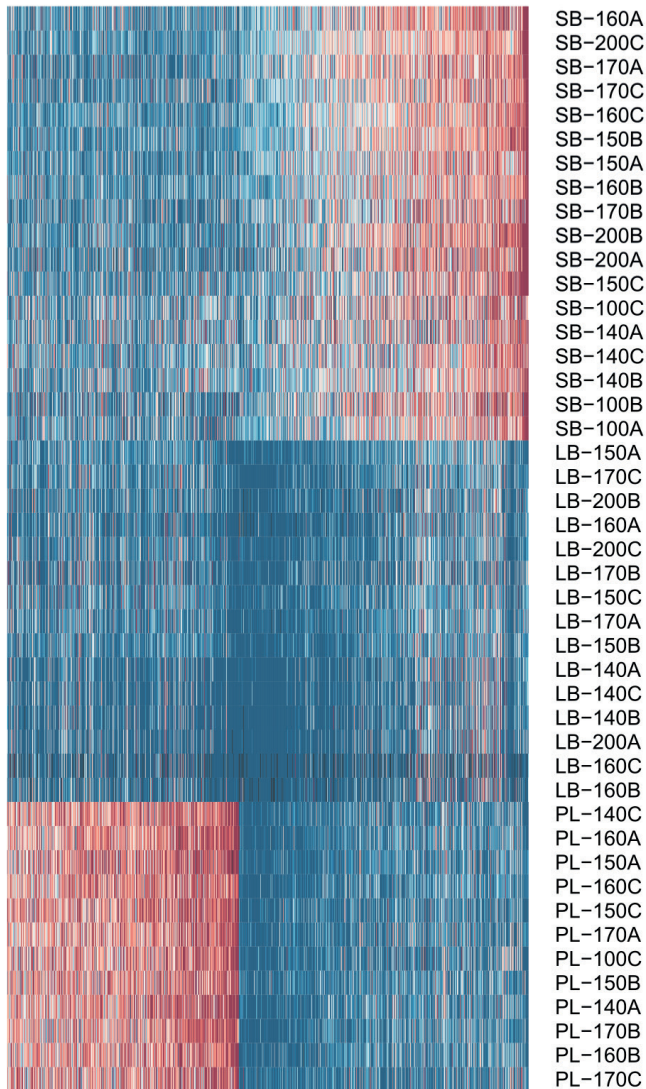
Mismunandi umhverfisþættir hafa einnig áhrif á tjáningu gena og svipfarsbreytileika. Líkt og stökkbreytingar geta þeir haft sértæk áhrif á ákveðna þætti í þroskun eða víðtækari afleiðingar. Fósturskemmdir vegna áfengis (e. fetal alcohol syndrome) er dæmi um umhverfisþátt sem getur haft víðtæk áhrif á fóstrið. Neysla móður á umtalsverðu áfengi á meðgöngu (þegar fósturþroska stendur yfir) getur valdið margs konar afleiðingum eins og greindarskerðingu og fæðingargöllum sem ekki koma fram við venjulegar aðstæður.³⁹ Aðrir umhverfisþættir geta haft sértækari áhrif. Til að mynda geta staðbundnar sýkingar eða mengun ýtt undir sjúkdóma sem eru sértækir við ákveðna vefi líkamans eins og t.d. öndunarferasjúkdóma líkt og astma.³³

Áhrif erfðabreytileika eða vissra umhverfisþátta á sjúkdóma og svipfar á millistig genatjáningar er hægt að skoða með ýmis konar aðferðum eins og t.d. mismunandi tilbrigðum RNA-raðgreiningar (e. RNA-sequencing) á ákveðnum vefjum og styrkmælingum á prótínum úr blóði.

LÍFMENGJAFRÆÐI

Á hverju stigi genastjórnunar starfa margir þættir og breytileiki í þeim getur orðið vegna erfða, umhverfis eða tilviljunar. Með tilkomu nýrra aðferða sem byggja á háhraðaraðgreiningu (e. high-throughput-sequencing) hafa orðið til margvísleg gögn sem nýtast til að rannsaka ólík þrep genastjórnunar og tengsl við t.d. þroskun, ónæmissvörun, sjúkdómsástand og þróun lífvera.⁴⁰ Nokkrar tegundir af lífmengjafraðilegum gögnum verða ræddar hér með sérstakri áherslu á hvernig þau geta gefið vísbendingar um hvernig breytileiki í genatjáningu mótar svipgerð.

Lífmengjafraði (e. omics) býr yfir aðferðum sem nýtast til að rannsaka ferli genastjórnunar með því að mæla magn mismunandi sameinda á ólíkum stigum frá DNA til prótína. Með þessum aðferðum er hægt að fá víðtæka sýn á genatjáningu með því að skoða styrk eða jafnvel virkni þúsunda sameinda af ákveðnum gerðum, t.d. mRNA, miRNA, bindingu umritunarþátta við DNA og styrk prótína, í frumum eða lífverum. Lífmengjafraði hefur einnig verið notuð til að greina annars konar fyrirbæri, til dæmis



Erfðabreytileiki í 3 gerðum bleikja úr þingvallavatni. Hver lína er eitt sýni, og dálkarnir tákna ólíka staði í umritunarmenginu. Litirnir tákna tíðni samsæta í hverju sýni. Úr grein Jóhannesar Guðbrandssonar og félaga frá 2019, í *Ecology and Evolution*. – Genetic variation in three morphs of Arctic charr from Þingvallavatn. Each row represents a sample, and columns different locations in the transcriptome. The colours indicate allele frequency per sample. From Guðbrandsson et al 2019, in *Ecology and Evolution*.

bakteríuflóru í mögum fólks eða örverur í vistkerfum.⁴¹ Í þessum svokölluðu umhverfiserfðamengjum (e. metagenomics) er rannsakað erfðaeftni úr umhverfinu, jarðvegi, vatni eða lofti.⁴² Þessum aðferðum fylgir mikið magn af gögnum og upplýsingum og nauðsynlegt að nota vandaðar tölulegar aðferðir.⁴¹

Gríðarleg fjölbreytni er í lífmengjafræðilegum aðferðum sem notaðar eru við greiningu mismunandi sameinda s.s. DNA, RNA, mRNA, miRNA, DNA metýlun, prótínnum, breytingum á prótínnum og histónnum, o.s.frv.⁴⁰ Hér fyrir neðan verður farið í nokkur dæmi um bæði gögn og aðferðir í lífmengjafræði og hvernig þær nýtast við rannsóknir á millistigum genatjáningar í tengslum við breytileika í svipgerð og sjúkdómum.

ERFÐAMENGJAFRÆÐI

Erfðamengjafræði (e. genomics) fjallar um rannsóknir á heilum erfðamengjum og virkni þeirra,⁴¹ þ.á.m. stökkbreytingum og tengingu þeirra við sjúkdóma. Flestir erfðabreytileikar eru einkennalausir (eða a.m.k. hlutlausir með tilliti til þróunarfræðilegrar hæfni) en sumir SNP-er geta tengst auknum líkum á að einstaklingar fái ákveðinn sjúkdóm, svari lyfjum mismunandi eða að einstaklingur sýni ákveðna hegðun.¹⁸ Með því að skima marga breytileika (SNP) í erfðamengi og spyrja um tengsl þeirra við sjúkdóm getum við öðlast meiri skilning á sameindalíffræði sjúkdómsins og mögulega þróað aðferðir til greiningar eða til persónubundnar eða fyrirbyggjandi meðferðar. Eins og áður sagði er sambandsgreiningum (GWAS) beitt til að leita gena sem tengjast svipgerðum. Sérstaklega hafa stærri þýði og fleiri endurtekningar leitt til mikils árangurs í kortlagningu eiginleika sem eru undir áhrifum margra gena (eins og hæð og flestir sjúkdómar) eða sjaldgæfra breytileika sem hafa mikil áhrif á sjúkdóma.⁴³ GWAS getur borið kennsl á svæði á litningum sem innihalda erfðabreytileika sem tengjast aukinni áhættu á ákveðnum sjúkdómum eða einkennum. Eitt fyrsta dæmið voru tengsl erfðabreytileika í FTO geninu við offitu.⁴⁴ Þó mikill árangur hafi náðst með þessari aðferð þá er vert að nefna að oftast er erfitt að vita nákvæmlega hvaða stökkbreyting á tilteknu svæði í erfðamenginu hefur sterkust áhrif (mesta áhættu fyrir sjúkdóm) og því er notast við tölfræðilega marktækni til að meta áreiðanleika tengslanna.² Einnig er stundum erfitt að segja til um hvaða gen eða ferli tiltekin stökkbreyting hefur áhrif á sem kallar á greiningar á t.d. umritun gena eða virkni prótína.

UMRITUNARMENGJAFRÆÐI

Umrítunarmengjafræði (e. transcriptomics) snýst um rannsóknir á gerðum og fjölda RNA umrita í frumum eða vefjum. Aðferðin hefur nýst til að skilja hvernig breytingar á genatjáningu geta stuðlað að flóknum svipgerðum sem þarfnast virkni margra gena.⁴⁵ Við svona greiningar voru fyrst notaðar svokallaðar DNA-örflögur (DNA microarray/chips) en nú er RNA-raðgreiningu mest beitt til að mæla fjölda RNA umrita og meta byggingu þeirra, t.d. úr frumum, vef eða beint úr umhverfinu.³ Umrítunarmengi reyndust mun flóknari en upphaflega var talið, sérstaklega meðal heilkjörnunga vegna mismunandi splæsingar sem getur gefið ólík mRNA af sama DNA móti sem gerir greiningar krefjandi.⁴⁶ Hér verða RNA-raðgreiningar kynntar stuttlega og hvernig hægt er að nýta þær til að rannsaka tjáningu gena.

Í RNA-raðgreiningu eru umrit (eða afrit af þeim) raðgreind og hægt að meta magn genatjáningar með því að mæla fjölda þeirra.³ Byrjað er á því að útbúa svokallað cDNA genasafn en cDNA stendur

fyrir mótlægt DNA (e. complementary DNA). cDNA umrit eru mynduð með sameindaerfðafræðilegum aðferðum⁴⁶ og síðan er safn þeirra úr ákveðnum vef eða frumgerð raðgreind, oftast með Illumina aðferð. Slík raðgreining myndar safn strengja sem eru yfirleitt 100–250 bp bútar (kallast „reads“ upp á ensku). Með svokallaðri stöflun (e. read mapping) eru þessir bútar bornir við viðmiðunarerfðamengi (e. reference genome) og þannig hægt að meta magn þeirra í mismunandi sýnum.³ Þar sem RNA-raðgreiningar ala af sér stór gagnasett þarf öflugar tölulegar greiningaraðferðir. Algengast er tilgátupróf fyrir hvert gen sem getur sagt til um hvaða þáttur (t.d. aldur, arfgerð eða umhverfisþáttur) hefur áhrif á það tiltekna gen. Einnig er beitt aðferðum eins og meginþáttgreiningu (e. principal component analysis, PCA) og klösum (e. clustering) sem draga fram mynstur í gagnasettinu bæði vegna tæknilegra og líffræðilegra þátta. Aðferðirnar greina undirliggjandi mynstur og samband milli breyta sem skýra mestan breytileikan í gögnunum (tjáningu genanna í sýnunum sem skoðuð eru). Til dæmis hafa meginþáttgreiningar á RNA-raðgreiningargögnum afhjúpað nýjar undirgerðir krabbameina (mynd 2).⁴⁷ RNA-raðgreiningar hafa opnað nýjar dyr í rannsóknum á genatjáningu og veita nákvæmari upplýsingar en DNA-örflögur.

PRÓTÍNMEGJAFRÆÐI

Prótínmenngjafræði (e. proteomics) gerir kleift að greina styrk hundruða eða þúsunda prótína og peptíða í hverju sýni. Prótínmengi (e. proteome) einstaklinga mótast bæði af genum og umhverfinu og geta verið breytileg eftir lífeðlisfræðilegu ástandi, kyni og aldri.⁴⁸ Með próteinmenngjafræði er hægt að bera kennsl og magngreina flest (ekki alveg öll) prótín sem eru til staðar í einni frumu, vef eða jafnvel heilli lífveru.⁴⁹ Með því að greina styrk próteina úr blóði er líka hægt að afla mikilvægra upplýsinga um heilsu og sjúkdómsástand einstaklinga.⁵⁰ Betri skilningur á orsakaþáttum breytileika magns prótína í blóði gæti nýst sérstaklega fyrir persónubundnar lækningar, ekki einungis til að meta heilsu- og sjúkdómsástand heldur einnig til að geta fyrr greint forstíga sjúkdóma ásamt því að geta fylgst með þróun þeirra á meðan að meðferð stendur.⁵¹ Prótínmengi geta einnig gefið mikilvægar upplýsingar um breytileika í genatjáningu líkt og umritunarmengi. Fylgni er á milli styrks mRNA og samsvarandi prótína í frumum en hún er ófullkomin.⁵² Kosturinn við greiningar prótínmenngja er að þau eru virku einingarnar í frumunni en mRNA t.d. er yfirleitt talið óvirkt millistig.⁴⁹ Prótínmengin eru því nær endanlegum svipgerðum einstaklinga en t.d. breytileiki í mRNA eða piwiRNA. Hins vegar hafa próteinmengi vissar takmarkanir: ekki er hægt að mæla öll peptíð sem mynda prótínmengið með hverri aðferð og eru ýmsar ástæður fyrir því

(sambærilegar hömlur eru ekki á mRNA mengjum). Til dæmis svokölluð himnuþrótein (e. membrane protein) (nauðsynleg fyrir fjölbreytta frumferla) sem mynda u.þ.b. 30% af próteinmenginu eiga það til að klumpast saman og falla út (e. precipitate) í lausn sem veldur ónákvæmni í greiningum.⁵³ Einnig eru margar aðferðir prótínmenngjafræði kostnaðarsamar og geta verið flóknaðar í framkvæmd.

Á síðustu öld voru notaðar aðferðir eins og ELISA, western blot og aðferðir byggðar á litskiljun (e. chromatography) sem nú teljast hefðbundnar greiningar á prótínum. Tæknilegar framfarir í greiningu þeirra hafa verið töluverðar, bæði í efnafræði, eðlisfræði og tækjabúnaði. Þróaðar hafa verið nokkrar aðferðir til að greina prótínmengi s.s. prótein örflögur, massagreining, Edman raðgreining og greiningar byggðar á hlaupi (e. gel based approaches). Sérstaklega forvitnilegar eru magnbundnar greiningar s.s. ICAT, SILAC og iTRAQ og afkastamiklar aðferðir s.s. röngtengislaggreining kristalla (e. X-ray crystallography), NMR-litrófsgreiningar og cryoEM. Að síðustu hafa verið framfarir í lífupplýsingafræðilegum greiningum á prótínmenngja gögnunum.⁴⁹ Val á aðferð fer eftir áherslum rannsóknar en hver þeirra hefur sína kosti og galla. Ekki verður farið nánar ofan í aðferðafræðina að þessu sinni.

Nýlegt dæmi þar sem notast var við próteinmenngja aðferðir er AGES (Age, Gene/Environment, Susceptibility) Reykjavík rannsóknin sem er framkvæmd af rannsóknarhópi Hjartaverndar. Rannsóknin er lýðgrunduð langsníðsrannsókn á 5.764 Íslendingum á aldrinum 66–93 ára þar sem hluti þátttakenda mætti í 5 ára eftirfylgnirannsókn. Rannsóknarhópurinn hefur mælt og greint rúmlega 7.000 prótín í sermi með aðferðinni SOMAScan sem byggir á prótein örflögum.⁵⁴ Síðustu ár hefur farið fram greining og vinnsla á göngunum og fundist tengsl ákveðinna prótín í sermi við svipgerðir og sjúkdóma. Rannsókn á 5.438 einstaklingum úr þýðinu sýndi fram á tengsl styrks hundruða prótína í sermi við sykursýki 2⁵⁵ á meðan önnur rannsókn á 5.294 einstaklingum (einnig AGES) fann tengingu u.þ.b. 303 prótína úr sermi við Alzheimer's sjúkdóminn.⁵⁶ Prótínmenngin úr AGES rannsókninni hafa að auki verið notuð til að kanna hvaða prótín séu marktækt tengd við gáttatif (e. atrial fibrillation) (hjartasjúkdómur sem veldur hröðum og óreglulegum hjartslætti).⁵⁷ Mynd 3 sýnir hluta af niðurstöðum rannsóknarinnar og er gott dæmi um hvernig tölfræðilegar greiningar nýtast við rannsóknir á tengingu sjúkdóma við styrk prótína. Að auki er vert að nefna að aðrar nýlegar rannsóknir (ekki AGES) hafa skoðað tengsl styrks próteina í blóði við lífaldur (e. chronological age) og öldrunarsjúkdóma og hvernig hægt væri að nýta þessi prótín sem lífmerki fyrir öldrun.⁵⁸

Hb9

Hb8

```
CGTTTTTAAGATCCGTTTGGTTTGGTTTGGTTTGGTCCGCGATGGCAATTCACGTTTTTACGAGCTCGTTCCTTCCG-----GTCCAAAATTAAGCCAGTT
CGTTTTTAAGATCCGTTTGGTTTGGTTTGGTTTGGTCCGCGATGGCAATTCACGTTTTTACGAGCTCGTTCCTTCCGGGTTCCAACCATGTCCAAAATTAAGCCAGTT
CGTTTTTAAGATCCGTTTGGTTTGGTTTGGTTTGGTCCGCGATGGCAATTCACGTTTTTACGAGCTCGTTCCTTCCGGGTTCCAACCATGTCCAGAATTAAGCCAGTT
CGTTTTTAAGATCCGTTTGGTTT-----ATGCCAGTT
```

Hbs2

Hbs1

```
CTTTTTTAAATGCCCCACTGCTAAATGCAGCTAAATTCGTCGATTTGTCAGAAAAGCGTTCTACAAGCTGAAGCACCTTTTTTAAGATGTTTGGTAGAAAAATTA
CTTTTTTAAATGCCCCACTGCTAAATGCAGCTAAATTCGTCGATTTG-----TTAAAAATTA
CTTTTTTAAATGCCCCACTGCTAAATGCAGCTAAATTCGTCGATTTGTCAGAAAAGCGTTCTACAAGCTGAAGCACCTTTTTTAAGATGTTTGGTAGAAAAATTA
```

Setraðir sem spanna tvo hluta eve gensins í ávaxtaflugum. Í báðum hlutum gensins finnast stórar úrfellingar, sem raska bindisetum fyrir umritunarþáttinn Hb. Úr grein Arnars Pálssonar ofl frá 2014 í PLoS One. – Haplotypes span two parts of the eve gene in fruitflies. Both parts contain large deletions, that disrupt binding sites for the transcription factor Hb. From Pálsson et al 2014 PLoS One.

UMFRAMERFDAMENGJAFRÆÐI

Umframerfðamengjafiræði nýttist til að rannsaka hvernigbygginglitnisogmynsturDNA-metýlunarhefur áhrif á tjáningu gena og svipfar.⁵⁹ Umframerfðamengi getur verið ólíkt á milli fruma og innan vefja og haft áhrif á genatjáningu á nokkra mismunandi vegu (rætt að ofan). Til þess að öðlast betri skilning á stjórnun umritunar er hægt að skoða bindingu próteina og DNA en einnig umframerfðamerki. Aðferð sem kallast ChIP-raðgreining (e. chromatin immunoprecipitation sequencing) byggist á greiningum á bindingu prótína við DNA í erfðamenginu⁶⁰ en eins og áður sagði geta sumir umritunarþættir bundist mjög sértækt. ChIP-raðgreining hefur reynst gagnleg m.a. í að rannsaka hvar og hvenær ákveðinn umritunarþáttur bindst í erfðamenginu sem er vísbending um hvaða gen sá þáttur hefur áhrif á. Missterk binding slíks prótíns við stjórnroð ákveðins gens gæti leitt til breytileika í umritun þess og myndun sjúkdóma. Víkjum aftur að dæminu um stakskort á BAZ1B og Williams heilkennið. Ónógt BAZ1B leiðir til heilkennisins. ChIP-raðgreiningu var beitt til þess að finna stjórnsvæði þeirra gena sem BAZ1B binnt við. Gögnin sýna að BAZ1B próteinið tengist við stjórnsvæði margra gena sem eru mikilvæg fyrir þroskun og sérhæfingu tauga. Stakskortur á BAZ1B geninu leiðir til víðtækra breytinga á genatjáningu í forverafrumum tauga og er þetta vísbending um að brenglun á umritun margra gena í taugum tengist líffræði Williams heilkennisins.³⁶ Þetta er eitt dæmi um hvernig ChIP-raðgreining hefur verið notuð til að rannsaka áhrif umframerfða á genastjórnun.

LOKAORÐ

Hér hefur verið fjallað um mikilvægi breytileika í genastjórnun og tengsl þessara ferla við breytileika í svipgerð og sjúkdómum. Helstu stig genastjórnunar frá umritun yfir í þýðingu voru tekin fyrir en ekki fjallað um ferli á stigum þroskunar, vefja og lífeðlisfræði sem eru einnig mjög mikilvæg. Stjórnun genatjáningar er flókið ferli með mörgum þáttum. Margt getur haft áhrif á ferlið s.s. breytileiki í erfðum, umhverfi ásamt tilviljun sem allt getur valdið breytileika í svipgerð og jafnvel stuðlað að myndun sjúkdóma.

Notagildi aðferða lífmengjafiræðinnar hefur vaxið á síðustu áratugum enda veita þær innsýn, með því að skoða þúsundir eða tugþúsundir þátta á sama tíma, á ákveðin stig genastjórnunar (jafnvel tvö eða fleiri). Samþætting margra gerða lífmengjafiræðilegra gagna getur þannig aukið skilning á samspili sameinda, stjórnkerfum fruma, þroskunar og lífeðlisfræði og hvernig áhrif ákveðinna þátta (erfðabreytileika eða umhverfisþátta) getur hríslast um kerfi fruma og lífvera. Til dæmis er hægt að gera kortlagningu á erfðabáttum sem hafa áhrif á mRNA styrk, fyrir öll gen lífveru. Í ávaxtaflugum hafði breytileiki í cis-röðum áhrif á tjáningu um 2.000 gena (af ~13.600) en merkilegt var að innúr (e. indels) stökkbreytingar valda að meðaltali meiri breytingum á styrk viðkomandi gena.⁶¹ Einng er nú mögulegt að rannsaka stöðugleika kerfanna. Munur á hópum birtist í mun á meðaltölum en dreifni (e. variance) segir til um breytileika í kringum meðaltöl. Í skemmtilegum íslenskum dæmi fannst fyrst munur á tjáningu um 2000 gena milli þriggja afbrigða bleikju úr Þingvallavatni.⁶² Næsta

rannsókn skoðaði einnig mun á dreifni í tjáningu (auk meðaltals) og sýndi að munur var á dreifni í tjáningu nokkur hundruð gena milli afbrigða bleikju í Þingvallavatni.⁶³ Einnig hafa fundist erfðabreytileikar í ávaxtaflugum sem tengjast dreifni í tjáningu gena.⁶⁴

Auk erfðamengja og umritunarmengja fleygir fram rannsóknnum á umframerfða- og prótínmenngjum. Tekið var dæmi um hvernig ChIP-raðgreining varpaði ljósi á hver áhrif stakskorts á BAZ1B geninu væru á Williams heilkennið.³⁶ Með prótínmenngjafræði hefur starfsfólki Hjartaverndar tekist að tengja styrk próteina í blóði við ýmsa sjúkdóma s.s. Alzheimer's, sykursýki 2 og gáttatif.⁵⁵⁻⁵⁷ Með ofangreindum dæmum er hægt að sjá hvernig hvert stig lífmengjafræðinnar getur leitt í ljós mikilvægar upplýsingar um orsakir fyrir breytileika í svipgerðum og sjúkdómum.

Vonir standa til að aukin þekking af þessu tagi gæti nýst fyrir persónubundnar lækningar þar sem margskonar gögn úr sama einstakling geta afhjúpað undirgerðir sjúkdóma sem má e.t.v. meðhöndla á ólíka vegu. Einstaklingar með sama sjúkdóm geta sýnt mismunandi svörum við lyfjum eða meðferðum sem getur stafað af þáttum sem valda breytileika í gena-tjáningu. Sumar meðferðir geta því gagnast ákveðnum hluta sjúklinga en að sama skapi verið hættulegar öðrum.⁶⁵ Öll ofangreind dæmi endurspeglar mikilvægi lífmengjarannsóknna og hvernig þróun á því sviði mun veita betri innsýn í hvernig líffræðileg kerfi vinna saman í heild sinni að mótun lífvera. Næsta skref er að samþætta núverandi þekkingu úr hverju skrefi lífmengjafræðinnar til að skoða áhrif gena, próteina og umframerfðabátta í heild sinni á ólíkar svipgerðir og undirgerðir sjúkdóma.⁶⁶

Fyrirséð er að rannsóknir á lífmengjafræði muni verða algengari í framtíðinni og mun vitneskja okkar um breytileika í genastjórnun vera enn víðtækari. Eitt er víst, að breytileiki í genatjáningu hefur áhrif á tilurð breytileika í svipgerð og sjúkdómum á mjög marga ólíka vegu.

ABSTRACT

Phenotype formation is a complex and multifaceted process governed by gene expression, where genetic variation, environmental influences, and random effects interact to shape biological outcomes. These interactions contribute not only to phenotypic diversity but also to disease risk. The emergence of omics technologies has enabled large-scale analysis of gene regulation, providing insights into key molecular processes from DNA to proteins. This review focuses on how genotype influences phenotype through variations in gene expression levels, highlighting how such variations may contribute to phenotypic variation and disease. Recent advances in molecular biology and integrative data analysis have enabled researchers to map genetic influences on mRNA expression across the genome, identify causal variants

and risk loci for diseases, and uncover protein-level associations with conditions such as Alzheimer's disease, type 2 diabetes, and cardiovascular disease. By integrating data from different omics techniques, we gain a deeper understanding of cellular regulation, development, and physiology, as well as the interplay between genetic and environmental factors. This knowledge is essential for developing personalized medicine, where molecular profiling can help classify disease subtypes and guide targeted therapies. Further research into gene expression variability and its regulatory mechanisms holds great promise for enhancing our understanding of complex biological systems and advancing medical treatments.

HEIMILDIR

- Zhang J., Yang J., Zhang L., Luo, J., Zhao, H., Zhang, J. & Wen, C. 2020. A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties. *Sci Rep* 10 (1). DOI: <https://doi.org/10.1038/s41598-020-62518-6>
- Gallagher, M. D. & Chen-Plotkin, A. S. 2018. The Post-GWAS Era: From Association to Function. *Am J Hum Genet* 102 (5). 717–730. DOI: 10.1016/j.ajhg.2018.04.002
- Lowe, R., Shirley, N., Bleackley, M., Dolan, S. & Shafee T. 2017. Transcriptomics technologies. *PLoS Comput Biol* 13 (5). DOI: 10.1371/journal.pcbi.1005457
- Buccitelli, C. & Selbach, M. 2020. mRNAs, proteins and the emerging principles of gene expression control. *Nat Rev Genet* 21 (10). 630–644. DOI: <https://doi.org/10.1038/s41576-020-0258-4>
- Wittkopp, P. J. & Kalay, G. 2012. Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet*. 13 (1). 59–69. DOI: <https://doi.org/10.1038/nrg3095>
- Clancy, S. & Brown, W. 2008. Translation : DNA to mRNA to Protein. *Nature Education* 1 (1). 101. Sótt af <https://www.nature.com/scitable/topicpage/translation-dna-to-mrna-to-protein-393/> (skoðað 20. júní 2024).
- Dor, Y. & Cedar, H. 2018. Principles of DNA methylation and their implications for biology and medicine. *The Lancet* 392 (10149). 777–786. DOI: 10.1016/S0140-6736(18)31268-6
- Eun, H-M. 1996. RNA Polymerases. Bls. 491–565 í: *Enzymology Primer for Recombinant DNA Technology* (ritstj. Eun, H-M.). Academic Press, Cambridge.
- Hernandez-Garcia, C.M. & Finer, J.J. 2014. Identification and validation of promoters and cis-acting regulatory elements. *Plant Science* 217–218. 109–119. DOI: 10.1016/j.plantsci.2013.12.007
- Mattioli, K., Oliveros, W., Gerhardinger, C., Andergassen, D., Maass, P. G., Rinn, J. L. & Melé, M. 2020. Cis and trans effects differentially contribute to the evolution of promoters and enhancers. *Genome Biol* 21 (1). DOI: <https://doi.org/10.1186/s13059-020-02110-3>
- Arnar Pálsson 2014. Stefnumót skilvirkni og breytileika - snertiflötur þroskunar og þróunar. *Náttúrufræðingurinn* 84 (1–2). 53–60.
- Huang, Y., Shang, R., Lu, G. A., Zeng, W., Huang, C., Zou, C. & Tang, T. 2022. Spatiotemporal Regulation of a Single Adaptively Evolving Trans-Regulatory Element Contributes to Spermatogenic Expression Divergence in *Drosophila*. *Mol Biol Evol* 39 (7). DOI: 10.1093/molbev/msac127
- Emerson, J.J. & Li, W.H. 2010. The genetic basis of evolutionary change in gene expression levels. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1552). 2581–2590. DOI: 10.1098/rstb.2010.0005

14. Singhm, R.K. & Cooper, T. A. 2012. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med* 18 (8). 472–482. DOI: 10.1016/j.molmed.2012.06.006
15. Majewski, J. & Pastinen, T. 2011. The study of eQTL variations by RNA-seq: From SNPs to phenotypes. *Trends in Genetics* 27 (2). 72–79. DOI: 10.1016/j.tig.2010.10.006
16. Li, Y. 2021. Modern epigenetics methods in biological research. *Methods* 187. 104–113. DOI: 10.1016/j.ymeth.2020.06.022
17. Li, J. & Liu, C. 2019. Coding or noncoding, the converging concepts of RNAs. *Front Genet* 10. DOI: <https://doi.org/10.3389/fgene.2019.00496>
18. Castellanos-Rubio, A. & Ghosh, S. 2019. Disease-associated SNPs in inflammation-related lncRNAs. *Front Immunol* 10. 420. DOI: 10.3389/fimmu.2019.00420
19. Frankish, A., Diekhans, M., Ferreira, A. M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J. M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Carbonell Sala, S., Chrast, J., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I. T., García Girón, C., Gonzalez, J. M., Grego, T., Hardy, M., Hourlier, T., Hunt, T., Izuogu, O. G., Lagarde, J., Martin, F. J., Martínez, L., Mohanan, S., Muir, P., Navarro, F. C. P., Parker, A., Pei, B., Pozo, F., Ruffier, M., Schmitt, B. M., Stapleton, E., Suner, M. M., Sycheva, I., Uszczynska-Ratajczak, B., Xu, J., Yates, A., Zerbino, D., Zhang, Y., Aken, B., Choudhary, J. S., Gerstein, M., Guigó, R., Hubbard, T. J. P., Kellis, M., Paten, B., Reymond, A., Tress, M. L. & Flicek, P. 2019. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47 (D1). D766–773. DOI: 10.1093/nar/gky955
20. Karolinska Institutet 2024. The Nobel Prize in Physiology or Medicine 2024 to Victor Ambros and Gary Ruvkun. Sótt af <https://news.ki.se/the-nobel-prize-in-physiology-or-medicine-2024-to-victor-ambros-and-gary-ruvkun> (skoðað 10. janúar 2025).
21. Kawahara, Y. 2014. Human diseases caused by germline and somatic abnormalities in microRNA and microRNA-related genes. *Congenit Anom (Kyoto)* 54 (1). 12–21. DOI: 10.1111/cga.12043
22. Syeda, Z. A., Langden, S. S., Munkhzul, C., Lee, M. & Song, S. J. 2020. Regulatory mechanism of microRNA expression in cancer. *Int J Mol Sci* 21 (5). DOI: 10.3390/ijms21051723
23. Bartsch, R. P., Liu, K. K. L., Bashan, A. & Ivanov, P. C. 2015. Network physiology: How organ systems dynamically interact. *PLoS One* 10 (11). DOI: <https://doi.org/10.1371/journal.pone.0142143>
24. Cholewa-Waclaw, J., Bird, A., von Schimmelmann, M., Schaefer, A., Yu, H., Song, H., Madabhushi, R. & Tsai, L-H. 2016. The role of epigenetic mechanisms in the regulation of gene expression in the nervous system. *Journal of Neuroscience. Society for Neuroscience* 36 (45). 11427–11434. DOI: 10.1523/JNEUROSCI.2492-16.2016
25. Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascensão, K., Rummel, C., Ovchinnikova, S., Mazin, P. V., Xenarios, I., Harshman, K., Mort, M., Cooper, D. N., Sandi, C., Soares, M. J., Ferreira, P. G., Afonso, S., Carneiro, M., Turner, J. M. A., VandeBerg, J. L., Fallahshahroudi, A., Jensen, P., Behr, R., Lisgo, S., Lindsay, S., Khaitovich, P., Huber, W., Baker, J., Anders, S., Zhang, Y. E. & Kaessmann, H. 2019. Gene expression across mammalian organ development. *Nature* 571 (7766). 505–509. DOI: <https://doi.org/10.1038/s41586-019-1338-5>
26. Heard, E. & Martienssen, R. A. 2014. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* 157 (1). 95–109. DOI: 10.1016/j.cell.2014.02.045
27. Glastad, K. M., Hunt, B. G. & Goodisman, M. A. D. 2018. Epigenetics in Insects: Genome Regulation and the Generation of Phenotypic Diversity. *Annual Review of Entomology* 64. 185–203. DOI: <https://doi.org/10.1146/annurev-ento-011118-111914>
28. Lappalainen, T. & Grealay, J. M. 2017. Associating cellular epigenetic models with human phenotypes. *Nat Rev Genet* 18 (7). 441–451. DOI: <https://doi.org/10.1038/nrg.2017.32>
29. Kagohara, L. T., Stein-O'Brien, G. L., Kelley, D., Flam, E., Wick, H. C., Danilova, L. V., Easwaran, H., Favorov, A. V., Qian, J., Gaykalova, D. A. & Fertig, E. J. 2018. Epigenetic regulation of gene expression in cancer: Techniques, resources and analysis. *Brief Funct Genomics* 17 (1). 49–63. DOI: 10.1093/bfpg/elx018
30. Deaton, A. M. & Bird, A. 2011. CpG islands and the regulation of transcription. *Genes Dev* 25 (10). 1010–1022. DOI: 10.1101/gad.2037511
31. Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., Tang, Y., Fange, Y. & Fang, D. 2021. Overview of Histone Modification. *Adv Exp Med Biol* 1283. 1–16. DOI: 0.1007/978-981-15-8104-5_1
32. Ibeagha-Awemu, E. M. & Zhao, X. 2015. Epigenetic marks: Regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. *Front Genet* 6. 302. DOI: 10.3389/fgene.2015.00302
33. Virolainen, S. J., VonHandorf, A., Viel, K. C. M. F., Weirauch, M. T. & Kottyan, L. C. 2023. Gene–environment interactions and their impact on human health. *Genes Immun* 24 (1). 1–11. DOI: <https://doi.org/10.1038/s41435-022-00192-6>
34. Xu, Y. C. & Guo, Y. L. 2020. Less Is More, Natural Loss-of-Function Mutation Is a Strategy for Adaptation. *Plant Commun* 1 (6). DOI: 10.1016/j.xplc.2020.100103
35. Johnson, A. F., Nguyen, H. T. & Veitia, R. A. 2019. Causes and effects of haploinsufficiency. *Biological Reviews* 94 (5). 1774–1785. DOI: 10.1111/brv.12527
36. Lalli, M. A., Jang, J., Park, J. H. C., Wang, Y., Guzman, E., Zhou, H., Audouard, M., Bridges, D., Tovar, K. R., Papuc, S. M., Tutulan-Cunita, A. C., Huang, Y., Budisteanu, M., Arghir, A. & Kosik, K. S. 2016. Haploinsufficiency of BAZ1B contributes to Williams syndrome through transcriptional dysregulation of neurodevelopmental pathways. *Hum Mol Genet* (7). 1294–1306. DOI: 10.1093/hmg/ddw010
37. Mills, E. W. & Green, R. 2017. Ribosomopathies: There's strength in numbers. *Science* 358 (6363). DOI: 10.1126/science.aan2755
38. Aqeilan, R. I., Calin, G. A. & Croce, C. M. 2010. miR-15a and miR-16-1 in cancer: Discovery, function and future perspectives. *Cell Death Differ* 17 (2). 215–220. DOI: <https://doi.org/10.1038/cdd.2009.69>
39. Kaminen-Ahola, N. 2020. Fetal alcohol spectrum disorders: Genetic and epigenetic mechanisms. *Prenat Diagn* 40 (9). 1185–1192. DOI: 10.1002/pd.5731
40. Graw, S., Chappell, K., Washam, C.L., Gies, A., Bird, J., Robeson, M. S. & Byrum, S. D. 2021. Multi-omics data integration considerations and study design for biological systems and disease. *Mol Omics* 17 (2). 170–185. DOI: <https://doi.org/10.1039/D0MO00041H>
41. Yamada, R., Okada, D., Wang, J., Basak, T. & Koyama, S. 2021. Interpretation of omics data analyses. *J Hum Genet*. 66 (1). 93–102. DOI: 10.1038/s10038-020-0763-5
42. Wooley, J. C., Godzik, A. & Friedberg, I. 2010. A primer on metagenomics. *PLoS Comput Biol* 6 (2). DOI: <https://doi.org/10.1371/journal.pcbi.1000667>
43. Barbeira, A. N., Dickinson, S. P., Bonazzola, R., Zheng, J., Wheeler, H. E., Torres, J. M., Shah, K. P., Garcia, T., Edwards, T. L., Stahl, E. A., Huckins, L. M., GTEx Consortium, Nicolae, D. L., Cox, N. L. & Im, H. K. 2018. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* 9 (1). DOI: <https://doi.org/10.1038/s41467-018-03621-1>
44. Huang, Y. H., Grasis, J. A., Miller, A. T., Xu, R., Soonthornvacharin, S., Andreotti, A. H., Tsoukas, C. D., Cooke, M. P. & Sauer, M. 2007. Positive regulation of Itk PH domain function by soluble IP4. *Science* 316 (5826). 886–889. DOI: 10.1126/science.1138684
45. Harrison, P. W., Wright, A. E. & Mank, J. E. 2012. The evolution of gene expression and the transcriptome-phenotype relationship. *Semin Cell Dev Biol* 23 (2). 222–229. DOI: 10.1016/j.semcdb.2011.12.004
46. Ozsolak, F. & Milos, P. M. 2011. RNA sequencing: Advances, challenges and opportunities. *Nat Rev Genet* 12 (2). 87–98. DOI: 10.1038/nrg2934
47. Vidman, L., Källberg, D. & Rydén, P. 2019. Cluster analysis on high dimensional RNA-seq data with applications to cancer research - An evaluation study. *PLoS One* 14 (12). DOI: <https://doi.org/10.1371/journal.pone.0219102>
48. Bodaghi, A., Fattahi, N. & Ramazani, A. 2023. Biomarkers: Promising and valuable tools towards diagnosis, prognosis and treatment of Covid-19 and other diseases. *Heliyon* 9 (2). DOI: 10.1016/j.heliyon.2023.e13323

49. Aslam, B., Basit, M., Nisar, M. A., Khurshid, M. & Rasool, M. H. 2017. Proteomics: Technologies and their applications. *J Chromatogr Sci* 55 (2). 182–196. DOI: 10.1093/chromsci/bmw167
50. Dodig-Crnković, T., Hong, M. G., Thomas, C. E., Häussler, R. S., Bendes, A., Dale, M., Edfors, F., Forsström, B., Magnusson, P. K., Schuppe-Koistinen, I., Odeberg, J., Fagerberg, L., Gummesson, A., Bergström, G., Uhlén, M. & Schwenk, J. M. 2020. Facets of individual-specific health signatures determined from longitudinal plasma proteome profiling. *EBioMedicine* 57. BIO: 10.1016/j.ebiom.2020.102854
51. Zhong, W., Edfors, F., Gummesson, A., Bergström, G., Fagerberg, L. & Uhlén, M. 2021. Next generation plasma proteome profiling to monitor health and disease. *Nat Commun* 12 (1). DOI: <https://doi.org/10.1038/s41467-021-22767-z>
52. Liu, Y., Beyer, A. & Aebersold, R. 2016. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* 165 (3). 535–550. DOI: 10.1016/j.cell.2016.03.014
53. Chandramouli, K. & Qian, P.-Y. 2009. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Human Genomics and Proteomics* 1 (1). DOI: 10.4061/2009/239204
54. Harris, T. B., Launer, L. J., Gudny Eiríksdóttir, Olafur Kjartansson, Palmi V Jonsson, Gunnar Sigurdsson, Guðmundur Thorgerisson, Thor Aspelund, Garcia, M. E., Cotch, M. F., Hoffman, H. J. & Vilmundur Gudnason. 2007. Age, gene/environment susceptibility-reykjavik study: Multidisciplinary applied phenomics. *Am J Epidemiol* 165 (9). 1076–1087. DOI: 10.1093/aje/kwk115
55. Valborg Guðmundsdóttir, Zaghlool SB, Valur Emilsson, Thor Aspelund, Ilkov, M., Elias F Guðmundsson, Stefan M Jonsson, Zilhao, N. R., Lamb, J. R., Suhre, K., Jennings, L. L. & Vilmundur Gudnason 2020. Circulating protein signatures and causal candidates for type 2 diabetes. *Diabetes* 69 (8). 1843–1853. DOI: 10.2337/db19-1070
56. Elisabet A. Frick, Valur Emilsson, Thorarinn Jonmundsson, Anna E. Steindorsdóttir, Johnson, E. C. B., Puerta, R., Dammer, E. B., Shantaraman, A., Cano, A., Boada, M., Valero, S., García-González, P., Elias F. Guðmundsson, Alexander Guðjonsson, Pitts, R., Qiu, X., Finkel, N., Loureiro, J. J., Orth, A. P., Seyfried, N. T., Levey, A. I., Ruiz, A., Thor Aspelund, Jennings, L. L., Launer, L. J., Valborg Guðmundsdóttir & Vilmundur Gudnason 2024. Serum proteomics reveal APOE-ε4-dependent and APOE-ε4-independent protein signatures in Alzheimer's disease. *Nat Aging* 4. 1446–1464. DOI: <https://doi.org/10.1038/s43587-024-00693-1>
57. Thorarinn Jonmundsson, Anna E Steindorsdóttir, Austin, T. R., Elisabet A Frick, Gisli T Axelsson, Launer, L., Psaty, B. M., Loureiro, J., Orth, A. P., Thor Aspelund, Valur Emilsson, Floyd, J. S., Jennings, L., Vilmundur Gudnason & Valborg Guðmundsdóttir 2023. A proteomic analysis of atrial fibrillation in a prospective longitudinal cohort (AGES-Reykjavik study). *Europace* 25 (11). DOI: 10.1093/europace/euad320
58. Argentieri, M. A., Xiao, S., Bennett, D., Winchester, L., Nevado-Holgado, A. J., Ghose, U., Albukhari, A., Yao, P., Mazidi, M., Lv, J., Millwood, I., Fry, H., Rodosthenous, R. S., Partanen, J., Zheng, Z., Kurki, M., Daly, M. J., Palotie, A., Adams, C. J., Li, L., Clarke, R., Amin, N., Chen, Z. & van Duijn, C. M. 2023. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat Med* 30. 2450–2460. DOI: <https://doi.org/10.1038/s41591-024-03164-7>
59. Wang, K. C. & Chang, H. Y. 2018. Epigenomics technologies and applications. *Circ Res* 122 (9). 1191–1199. DOI: 10.1161/CIRCRESA-HA.118.310998
60. Park, P. J. 2009. ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet* 10 (10). 669–680. DOI: 10.1038/nrg2641
61. Massouras, A., Waszak, S. M., Albarca-Aguilera, M., Hens, K., Holcombe, W., Ayroles, J. F., Dermitzakis, E. T., Stone, E. A., Jensen, J. D., Mackay, T. F. C. & Deplancke, B. 2012. Genomic Variation and Its Impact on Gene Expression in *Drosophila melanogaster*. *PLoS Genet* 8 (11). DOI: 10.1371/journal.pgen.1003055
62. Jóhannes Guðbrandsson, Sigríður Rut Franzdóttir, Bjarni Kristófer Kristjánsson, Ahi, E. P., Maier, V. H., Kapralova, K. H., Sigurður Sveinn Snorrason, Zophonías Oddur Jónsson & Arnar Pálsson 2018. Differential gene expression during early development in recently evolved and sympatric Arctic charr morphs. *PeerJ* 6. DOI: <https://doi.org/10.7717/peerj.4345>
63. Horta-Lacueva, Q. J. B., Zophonías Oddur Jónsson, Dagny A. V. Thorholludóttir, Benedikt Hallgrímsson, Kapralova, K. H. 2023. Rapid and biased evolution of canalization during adaptive divergence revealed by dominance in gene expression variability during Arctic charr early development. *Commun Biol* 6 (1). DOI: <https://doi.org/10.1038/s42003-023-05264-5>
64. Huang, W., Carbonea, M. A., Magwire, M. M., Pieffer, J. A., Lyman, R. F., Stone, E. A., Anholt, R. R. H. & Mackay, T. F. C. 2015. Genetic basis of transcriptome diversity in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 112 (44). E6010–E6019. DOI: <https://doi.org/10.1073/pnas.1519159112>
65. Simonovsky, E., Schuster, R. & Yeger-Lotem, E. 2019. Large-scale analysis of human gene expression variability associates highly variable drug targets with lower drug effectiveness and safety. *Bioinformatics* 35 (17). 3028–3037. DOI: <https://doi.org/10.1093/bioinformatics/btz023>
66. Subramanian, I., Verma, S., Kumar, S., Jere, A. & Anamika, K. 2020. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform Biol Insights* 14. DOI: 10.1177/1177932219899051

UM HÖFUNDA

Hulda Karen Ingvarsdóttir (f. 2001) er með meistaraþróf í Lífupplýsingafræði og nú doktorsnemi í Líf- og læknávisindum við Háskóla Íslands. Í doktorsverkefninu vinnur Hulda með rannsóknarhópi Hjartaverndar en markmiðið er að greina og skilgreina peptíð og próteóform í blóðrás sem tengjast snemmbúinni þróun á sykursýki 2. Einnig verða erfðamengjafræðileg gögn notuð til að meta möguleg orsakasambönd milli próteoforma og sykursýki 2.

hki1@hi.is



Arnar Pálsson (f. 1970) er með doktorsþróf í erfðafræði. Hann starfar sem prófessor í lífupplýsingafræði við Líf- og umhverfisvísindadeild Háskóla Íslands, og rannsakar þróun, þroskun og erfðir í bleikjum, urriðum, ávaxtaflugum, maurum og lúsmý.

apalsson@hi.is

