

Små genetiske variationer mennesker imellem – hvad betyder de(t)?

Sekventeringen af det humane genom har forårsaget en voldsom stigning i forskning i genetiske variationer mellem mennesker – og især forskning i SNP'er (Single Nucleotide Polymorphisms). Hvad er disse polymorfier, og hvilken biologisk betydning har de?

Af Ulla Vogel, seniorforsker på Arbejdsmiljøinstituttet, ubv@ami.dk

Kortlægningen af det humane genom er ved at være tilendebragt. Der er kun enkelte områder tilbage, hvor generne stadig bytter plads med jævne mellemrum.

To interessante fund

I mine øjne gav sekventeringen to bemærkelsesværdige fund:
- Man fandt kun ca. 30-40.000 gener i det humane genom. Det lyder måske af mange, men til sammenligning har en simpel, encellet tarmbakterie som *Eschericia coli* 10.000 gener. Det kan være svært at forstå, at en kompliceret organisme som den menneskelige for at fungere kun kræver 3-4 gange så mange enzymer som en simpel, encellet organisme.
- Det andet interessante fund var, at DNA-sekvensen er forskellig fra person til person i et ud af hver 1000. nukleotid – og så sammenlignede man endda kun arvematerialet fra 24 forskellige personer. Det betyder, at vi genetisk set kun ligner hinanden 20 gange mere, end vi ligner chimpanser.

De fleste genetiske variationer mellem mennesker er enkelt-nukleotid-ændringer, dvs. at der på en given position i arvematerialet simpelthen er et nukleotid hos nogen og et andet hos andre. Nogen har f.eks. et A, hvor andre har et C. Disse enkelt-nukleotid-variationer kaldes SNP'er efter det engelske udtryk »Single Nucleotide Polymorphisms«.

Oftest finder man kun to forskellige nukleotider i en SNP. Det indebærer, at nukleotidændringen kun er sket i et enkelt menneske, og at den derefter er blevet nedarvet og spredt gennem generationer. Nogle SNP'er findes kun hos nogle få procent af befolkningen, mens andre bæres af næsten halvdelen af en befolkning. Faktisk er petitesse-grænsen, at en polymorfi skal forekomme hos 1% af befolkningen, før den kan kaldes en polymorfi.

Der er ofte også store raceforskelle. En SNP kan f. eks. være almindelig hos asiater men meget sjælden hos kaukasiere.

Fra SNP til årsag

Kortlægningen af alle disse SNP'er har givet anledning til eksplosiv vækst i forskning i sammenhængen mellem SNP'er og forekomsten af sygdomme – specielt cancer. Det er oftest undersøgelser af case-kontrol-typen:

Man undersøger forekomsten af en given SNP i en patient-gruppe og sammenligner med forekomsten hos en gruppe raske individer, der efter bedste overbevisning ligner patientgruppen mest muligt. Hvis SNP'en forekommer oftere hos patient-gruppen, kan man konkludere, at SNP'en er associeret med sygdommen (eller med beskyttelse mod sygdommen, hvis det modsatte er tilfældet). De fleste forskere går skridtet videre og diskuterer, hvordan SNP'en direkte kan være impliceret i den biologiske årsag til sygdommen. Det er imidlertid sjældent,

man er så heldig, at den SNP, man studerer, er den direkte årsag til den sygdom, den statistisk set er associeret til.

Arvematerialets opbygning

Alt arvemateriale er bygget op af fire byggesten, kaldet nukleotider, G, A, T og C, og rækkefølgen af dem indeholder al den information, der skal til for at danne hele organismen.

Menneskets arvemasse udgøres af 6 mia. nukleotider fordelt på 46 kromosomer. Disse kromosomer passer parvis sammen, og man arver det ene sæt kromosomer fra sin mor og det andet fra sin far. Nukleotiderne ligger som perler på en snor, som så udgør kromosomerne. En meget lang snor med gennemsnitligt 130 mio. »perler« pr. snor. De nedarves også i de fleste tilfælde som perler på en snor. Derfor nedarves en SNP noget forsimplet sagt sammen med op til 130.000 andre SNP'er. Man bør derfor være meget forsigtig, når man udtaler sig om den biologiske effekt af en enkelt SNP.

Den biologiske effekt af SNP'er

Kun en lille del af det humane genom indeholder den information, der bliver omsat til proteiner og enzymer, eller har betydning for opretholdelse af arvematerialet. Resten er fyldt – enten mellem generne eller inde i generne (introns). SNP'er i fyldet har næsten aldrig biologisk betydning.

SNP inde i de kodende og de regulatoriske regioner kan både være neutrale eller have biologisk betydning som illustreret i figur 1.

Intet protein

Nukleotiderne i DNA koder for proteiner ved at angive rækkefølgen af aminosyrer i proteinet. Hvis en aminosyrekode ændres til et stopsignal, vil der kun dannes den første del af det oprindelige protein. Stumpen er oftest ustabil, og nedbrydes hurtigt. Men da man har to ens kromosomer, vil det andet kromosom i de fleste tilfælde også udtrykke et tilsvarende protein, således at effekten af det manglende protein mindskes. I disse tilfælde ville man forvente en kraftig biologisk effekt, hvis begge kromosomer bærer SNP'en, men en langt svagere effekt hvis man kun har SNP'en på det ene kromosom. SNP'er af denne type er ofte årsag til arvelige sygdomme, som f.eks. cystisk fibrose, arvelig brystkræft og arvelig tyktarmskræft.

Ændret protein

En SNP kan også medføre, at der på en position sættes en anden aminosyre ind i et protein. Det kan give anledning til dannelse af et protein med nedsat eller ændret funktion. Men igen vil det andet kromosom udtrykke et tilsvarende protein, således at effekten af det ændrede protein mindskes. I disse

tilfælde ville man (måske) forvente en biologisk effekt, hvis begge kromosomer bærer SNP'en, men en langt svagere effekt, hvis man kun har SNP'en på det ene kromosom. Det kan også være, at aminosyreudskiftningen ikke betyder noget.

I sjældne tilfælde giver en SNP anledning til dannelse af et dominant protein, hvor det ændrede protein forhindrer den korrekte version af proteinet i at fungere rigtigt. Det kan f.eks. være, hvis proteinet skal være en dimer eller en hexamer for at virke. Så kan det ændrede protein være årsag til, at hele hexameren ikke virker, selvom den indeholder både ændrede og korrekte versioner af proteinet. I disse tilfælde er det nok at bære en kopi af SNP'en for at få en biologisk effekt, og at bære SNP'en på begge kromosomer øger ikke effekten meget.

Ingen forskel

Endelig kan der være SNP'er, der ikke giver anledning til aminosyreændringer, selvom de er inde i den kodende region. Det er, fordi koden for aminosyrerne er inddelt i tripletter, således, at der skal tre nukleotider til at definere, hvilken af de 23 aminosyrer, der er tale om. Da der er 64 mulige kombinationer af de 4 nukleotider, vil en del tripletter kode for den samme aminosyre. Tripletterne er organiseret sådan, at det altid er 3. position i tripletten, der kan varieres. Dvs. at på ca. hver 3. plads i nukleotidrækkefølgen er en SNP neutral og vil ingen biologisk effekt have. Neutrale SNP'er, der ikke påvirker aminosyre-rækkefølgen, kan dog godt have biologisk effekt i enkelte tilfælde. Det sker ved at SNP'en påvirker stabiliteten af mRNA'et, mellemlæddet mellem DNA og proteiner. Hvis stabiliteten af et mRNA nedsættes, når cellen ikke at syntetisere lige så store mængder af proteinet som før.

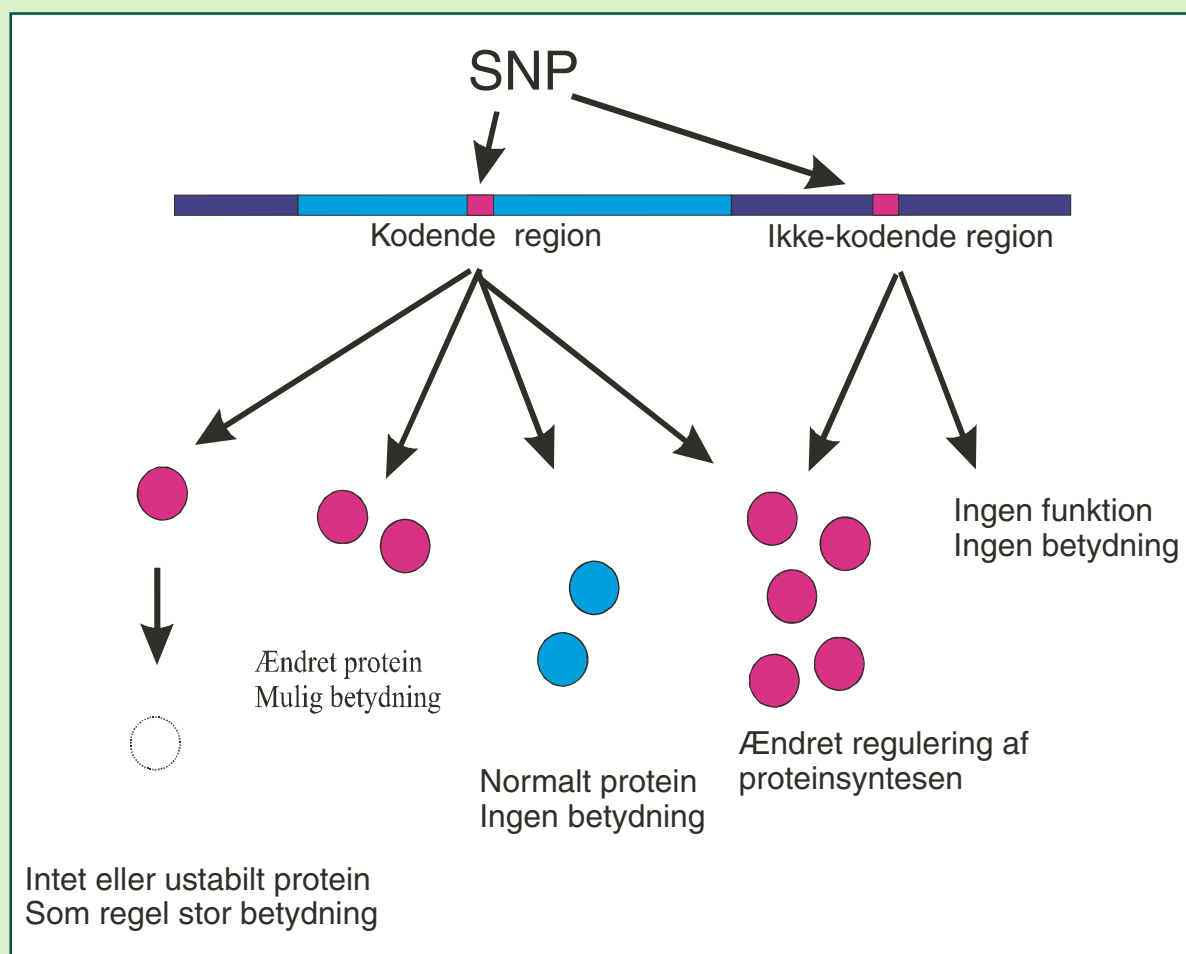
Ændret regulering

De SNP'er, der potentielt har størst biologisk betydning, er i de områder, der regulerer gen-ekspressionen, dvs. hvor meget protein der dannes, hvornår og i hvilke celletyper. SNP'er af denne type kan, hvis de ændrer koden i et bindingssted for et regulatorisk protein, påvirke genekspressionen af flere gener samtidig.

Som det måske fremgår af ovenstående, gør en SNP ofte ingen forskel. Enten ændres proteinet ikke, eller også virker det på trods af ændringen alligevel fint. Eller måske *har* proteinet nedsat funktionsevne, uden at det gør noget, fordi proteinet katalyserer en del-proces, der ikke er det begrænsende led i den samlede proces.

Vigtig information

Det, at en SNP ikke har direkte biologisk betydning, betyder ikke, at den er uinteressant. Overhovedet ikke. Den er interessant på en anden måde. En sammenhæng mellem en SNP og forekomsten af en sygdom indikerer, at der i dette kromosom-område findes en mutation, der har betydning for sygdommen – enten lige her, et andet sted i samme gen eller i et af de nærliggende gener. Det gør jagten på de processer og proteiner, der er involveret i f.eks. kræft, lettere, fordi SNP-nettet er både følsomt og finmasket. Der vil altid nedarves mange SNP'er sammen med den SNP, der virkelig gør forskellen. Det er altså endnu en metode til at finde ud af, hvilke funktioner hvert enkelt af de 30.000 proteiner varetager. Og det er vigtige brikker i løsningen af det store puslespil, som f.eks. kræft udgør.



Mulige biologiske effekter af en SNP.