

Postgenomisk epidemiologi i et FUGE-perspektiv

Rolv T. Lie¹ og Per Magnus²

1. Seksjon for medisinsk statistikk, Institutt for samfunnsmedisinske fag, Universitetet i Bergen

2. Divisjon for Epidemiologi, Nasjonalt folkehelseinstitutt, Oslo

FUGE står for "functional genomics", og er betegnelsen på forskningsrådets nye program for funksjonell genomforskning. Etter at det menneskelige genom er kartlagt begynner det store arbeidet med å beskrive funksjonen til de vel 30 000 genene. I tillegg til arbeidet med å identifisere gener og avsløre genetiske årsaker til sykdom, vil arbeidet rettes mot studier av sammenhengen mellom gener og sykdom, og samspillet mellom gener og miljøfaktorer. Her bør selvfølgelig norsk forskning spille en rolle. Forskningsrådet har derfor nylig fordelt en årlig pott på 100 millioner kroner på en rekke prosjekter under FUGE-programmet.

Oppbygging av teknologi og kompetanse er naturlig nok viktige elementer i FUGE-satsingen, som har gått relativt bredt ut og gitt støtte til en rekke prosjekter. For norsk genetisk epidemiologi generelt er nok behovet for oppbygging av nasjonale biobanker mest avgjørende. For de epidemiologene som sitter på biologisk materiale oppleves nok tilgangen til molekylærgenetisk analysekapasitet som den største hindringen for faglig utvikling. Men hvilken rolle bør egentlig genetisk epidemiologi spille i FUGE-satsingen, og hvordan kan vi best utnytte de fortrinnene som ligger i å basere studier på den norske befolkningen?

Det pågår en debatt i de internasjonale fagmiljøene om hvordan forskningen om genetiske årsaker til sykdom best skal komme videre i postgenom-fasen. Til nå er det uten tvil koblingsstudier (linkage) som har vært det viktigste virkemiddelet. En enkel forklaring av prinsippene for slike analyser finnes i en oversiktsartikkel i dette nummeret (1). Koblingsstudier prøver å identifisere segmenter av genomet som oftere enn forventet deles av to eller flere syke personer fra samme familie. Ved hjelp av koblingsstudier har man identifisert posisjonen til gener som er årsak til en rekke sjeldne mendelske sykdommer. De fleste genene som til nå er blitt knyttet til relativt vanlige sykdommer er også blitt funnet ved hjelp av koblingsstudier (2). Eksempler som nevnes er Alzheimers sykdom, tidlig onset type 2 diabetes, brystkreft og kolonkreft. Typisk for disse sykdommene er at variantene av genene som disponerer for sykdom er sjeldne og effekten av dem er kraftig, slik at arvelighetsmønsteret også her blir nesten mendelsk. Dette vil imidlertid ikke være tilfelle for en rekke av de relativt vanlige sykdommene som nå studeres innenfor genetisk epidemiologi. I forskning på genetiske årsaker til slike sykdommer må det antas at en rekke gener spiller inn både i samspill med hverandre og i samspill med miljøfaktorer. For slike

sykdommer er det ikke opplagt at genene kan finnes ved hjelp av koblingsstudier.

Den alternative metoden som skisseres for å finne de genene som har betydning for disse sykdommene er assosiasjonsstudier (linkage-disequilibrium) av den typen epidemiologer er vant til å gjennomføre (2,3). Det er antagelig en fordel at studiene gjennomføres i befolkninger som er relativt homogene, slik som den norske, og at studiene er representative for den befolkningen som studeres (4,5). Betydningen av dette er omdiskutert, men dersom det stemmer er det klart at Norge har spesielle forutsetninger for å drive genetisk forskning. I tillegg til å ha en relativt homogen befolkning har Norge også et godt og enhetlig organisert helsevesen, nasjonale helseregistre og en relativt samarbeidsvillig befolkning. Befolkningen er også stor nok til at man kan gjøre gode studier av ikke helt vanlige sykdommer. Dersom forskningen klarer å ta vare på disse fordelene, ikke minst en rimelig god oppslutning fra befolkningen, kan Norge oppnå en viktig posisjon i genetisk forskning i det neste tiåret.

Nå er kanskje ikke veien videre fullt så enkel som at man bare kan sette igang et stort apparat med forskning på den norske befolkningen og håpe på å kunne lete seg frem til de viktige genetiske årsakene til en rekke sykdommer. Risch (2) skisserer et scenario der man skal lete seg igjennom tusenvis av varianter av alleler for å finne de som har effekt. Statistisk sett får det hele preg av å lete etter en nål i en høystakk. Det begynner også å komme rapporter med eksempler på at de genetiske effektene man må vente å lete etter er adskillig mer komplekse enn det man har tenkt på til nå (6). I et nylig publisert eksempel måtte man lete seg frem til et komplisert samspill av såkalte haplotyper som strakk seg utenfor regionen for hvert enkelt aktuelt gen for å kunne finne de genetiske effektene (7). Haplotyper er definert ved en bestemt rekkefølge av alleler som ligger nær hverandre på samme kromosom.

Vi må altså forberede oss på å studere både haplotyper (bestemte sekvenser av markører og ikke bare enkeltmarkører), sannsynlig interaksjon mellom flere gener (haplotyper) og i tillegg mulig interaksjon med miljøfaktorer. Dette arbeidet fremstår som mye mer komplisert enn bare et spørsmål om riktig epidemiologisk design, gode statistiske og bioinformatiske metoder og god kapasitet for laboratorieanalyser. Visst trenger vi å utvikle nye statistiske metoder for å studere haplotyper, mer kompliserte design og modeller med samspill. Men vel så viktig blir det, i følge de mer

pessimistiske genetikerne (8), å koble inn basalbiologisk kunnskap omkring virkningen av genene slik at man kan fremsette gode hypoteser om hvilke haplotyper som kan tenkes å øke risikoen for sykdom. Enkelte representanter for *functional genomics* går så langt som å antyde at hele løsningen på hvordan man skal finne de riktige genene ligger i å studere deres biokjemiske effekter (9).

Løsningen ligger nok ikke i å tenke for generelt om hvordan disse problemene skal løses. Midt i diskusjonene om design og statistikk mener vi at genetiske epidemiologer i større grad må vende seg mot miljøer som arbeider med den basale biologien rundt konkrete gener. Ved å målrette studier mot gener som har en rimelig sannsynlig virkning på den aktuelle sykdommen man studerer kan man muligens unngå å havne med hodet i høystakken. FUGE burde satse mye på genetisk epidemiologi fordi det er her Norge har spesielle fortrinn. Det er også klart at FUGE må bidra til å utvikle gode miljøer som arbeider med å studere funksjonen til bestemte gener eller grupper av gener på et ganske grunnleggende nivå. Men det viktigste må være at det skapes god kontakt og samarbeid mellom disse forskjellige miljøene.

Epidemiologi har lenge vært et tverrfaglig område, og epidemiologer er vant til å samarbeide. Den store utfordringen for genetisk epidemiologi ligger i å ut-

vikle tverrfagligheten videre til å omfatte kompetanse innen biokjemi, molekylærbiologi og bioinformatikk. Det er viktig at teknologiplattformer gjøres tilgjengelige for epidemiologiske prosjekter i Norge. Men det må også mer kompetanse på basal functional genomics inn i de epidemiologiske miljøene. Utnyttelse av de spesielle norske fortrinnene innenfor postgenomisk forskning vil ikke lykkes skikkelig uten at tradisjonelle barrierer mellom epidemiologi og basalforskning brytes ytterligere ned. Epidemiologer har i hvert fall begynt å få øynene opp for behovet for gjensidighet i dette arbeidet (10).

Forskningsmyndighetene bør derfor stimulere til at også basalmiljøene ser de spesielle fortrinnene som ligger i å studere populasjonsbaserte norske materialer. Samtidig må det stimuleres til at epidemiologiske prosjekter støtter seg på solid faglig kunnskap om den typen genetikk man ønsker å studere. Dette kan bidra til at tradisjonelt adskilte fagmiljøer kan se felles nytte i å utvikle samarbeidsmodeller. Fra vårt FUGE-perspektiv kan det derfor se ut til at de største utfordringene for norsk genetisk forskning og for genetisk epidemiologi ligger i at det er for stor oppsplitting av fagmiljøer og at det finnes et akutt behov for å utvikle mer samarbeid mellom fagområder som tradisjonelt ikke er vant til å samarbeide, ikke en gang utenfor Norge.

REFERANSER

1. Magnus P, Lie RT. En oversikt over design i klassisk genetisk epidemiologi. *Norsk Epidemiologi* 2002; **12** (2): xxx-yyy.
2. Risch N. Search for genetic determinants in the new millennium. *Nature* 2000; **405**: 847-56.
3. Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002; **11**: 505-12.
4. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev* 2002; **11**: 513-20.
5. Goldstein DB, Weale ME. Population genomics: linkage disequilibrium holds the key. *Curr Biol* 2001; **11**: R576-9.
6. Altshuler D, Daly M, Kruglyak L. Guilt by association. *Nat Genet* 2000; **26**: 135-7.
7. Horikawa Y, Oda N, Cox NJ, et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 2000; **26**: 163-75.
8. Weiss KM, Terwilliger JD. How many diseases does it take to map a gene with SNPs? *Nat Genet* 2000; **26**: 151-7.
9. Woychik RP, Klebig ML, Justice MJ, et al. Functional genomics in the post-genome era. *Mutat Res* 1998; **400**: 3-14.
10. Millikan R. The changing face of epidemiology in the genomics era. *Epidemiology*. 2002 **13**: 472-80.