

Den stora utmaningen blir att förstå proteinernas funktion

Kromosom 22 blev under det gångna året den första kromosomen i människans arvsmassa som kartlades, inom ramen för Human Genome Project, HUGO.

– Men den stora utmaningen återstår: att karakterisera alla aktiva gener från kromosomen, förstå proteinernas funktion och finna alla de gener som man vet orsakar sjukdomar, säger docent Jan Dumanski på enheten för klinisk genetik, Karolinska sjukhuset i Solna.

I en snabbt accelererande takt identifieras nu sekvenserna av de nukleotidbaser som utgör människans arvsmassa. Det tog omkring tio år efter beslutet om HUGO-projektet innan den första kromosomen, kromosom 22, blev kartlagd hösten 1999. Inom bara något år ska nukleotidsekvenserna i resterande 22 kromosompar vara bestämda.

Totalt består arvsmassan av cirka tre miljarder nukleotidbaspar. 33,4 miljoner av dessa finns i kromosom 22, som är tredje minsta av människans kromosomer. Endast den manliga könskromosomen Y och kromosom 21 är mindre. Kraftfullare datorer, nya program för bearbetning av sekvenser, bättre automatiserade maskiner för att generera sekvensdata och en rad förbättringar av kemi vid sekvensreaktioner har fått sekvenseringsarbetet att gå allt fortare.

Resultatet av arbetet med kromosom 22 presenteras i tidskriften *Nature* (1999: 402; 489-95). Den enda svenska grupp bland det 10-tal forskargrupper som deltagit i detta delprojekt av HUGO leds av docent Jan Dumanski vid Karolinska institutet.

– Sekvenseringsarbetet har gått betydligt snabbare än förväntat. Det beror bland annat på att kromosomen innehåller ett stort antal sjukdomsgener som har intresserat många forskare. Självt började jag arbeta med kromosom 22 redan 1987, med avseende på dess roll vid canceruppkomst.

– Vår grupp svarar uppskattningsvis för fem procent av hela kartläggningen, säger Jan Dumanski.

545 gener har lokaliserats till kromosom 22. Men den kan innehålla upp

till 1 000 gener, tror forskarna. Det grundar man bland annat på att det finns ett stort antal specifika nukleotidsekvenser som ofta är kännetecknande för promotorregioner, men man har ännu inte lyckats koppla någon aktiv gen till dem.

Ätminstone 27 olika sjukdomstillstånd beräknas vara associerade med mutationer i gener belägna på kromosom 22. Bland dessa nämns koloncancer, gliom, bröstcancer, äggstockscancer och schizofreni. Generna för dessa sjukdomar är dock svåra att lokalisera med befintlig teknik.

Men för många, mer ovanliga genetiskt betingade sjukdomstillstånd, har även den aktuella genen karakteriserats. Till dessa sjukdomar hör neurofibromatos typ 2.

I *Nature*-artikeln hösten 1999 rörande arbetet med kromosom 22 har författarna gjort en uppskattning av det totala antalet gener som borde finnas i människans arvsmassa, med utgångspunkt från antalet lokaliserade gener i kromosom 22. Det bör då, enligt författarna, åtminstone handla om 61 000 gener.

Men beräkningen är mycket osäker eftersom kromosom 22 endast utgör cirka 2 procent av människans arvsmassa. Dessutom tycks det vara så att kromosom 22 har en högre gentäthet än andra kromosomer.

– En indirekt antydning om att kromosom 22 har en högre täthet av aktiva gener än andra kromosomer är det höga antalet sjukdomsgener, säger Jan Dumanski.

– Jag är själv ganska försiktig med att försöka uppskatta antalet gener i hela arvsmassan enbart utifrån vad vi känner till om kromosom 22. Jag tror att vi kommer att hamna närmare 100 000 än 50 000 gener.

HUGO-projektet förändras

Den kompletta kartläggningen av kromosom 22 och publiceringen i tidskriften *Nature* har av vissa förvisso framställts som »den första fasen i en biologisk revolution». Men för forskarna själva innebär inte just detta något re-

Kromosom 22 är den tredje minsta av människans kromosomer. Trots att den bara innehåller cirka 2 procent av arvsmassan har hela 27 genetiska sjukdomar lokaliserats till kromosomen. »Det höga antalet sjukdomsgener antyder att kromosomen har högre täthet av aktiva gener än andra kromosomer», säger docent Jan Dumanski.

volutionerande. Resultaten från genomprojektet släpps löpande ut till databaser och informationen används kontinuerligt inom forskningen.

Den fullständiga sekvensbestämningen innebär dock att man nu har tillgång till en mycket väldetaljerad transkriptionskarta (en karta över vilka de aktiva generna är och var de sitter på kromosomen), och med hjälp av den underlättas sökandet efter de gener som man vet orsakar vissa sjukdomar men som har varit svåra att hitta. Till dessa hör flera tumörsuppressorgener.

Sekvensbestämningen gör det även lättare att karakterisera mutationer i specifika sjukdomsgener, och att identifiera faktorer som reglerar att sjukdomsgenen uttrycks i vissa vävnader och i vissa situationer.

Arbetet med kromosom 22 har också lett till insikter som gör att den fortsatta kartläggningen av människans arvsmassa inom HUGO ändrar form. Sekvensbestämningen av kromosom 22 har pågått sedan 1993 och det skulle ta alldeles för lång tid om resterande kromosomer kartlades på samma sätt, dvs avsnitt för avsnitt.

Tidsaspekten har blivit viktig efter-

Det är tänkbart att det kan handla om gener som uttrycks i mycket speciella situationer under exempelvis embryogenesen eller i specifika celler i vissa vävnader.

Stora utmaningen återstår

Inom något eller några få år kommer man att ha samma detaljkunskap om människans samtliga kromosomer som man idag har om kromosom 22. Då inträder nästa forskningsfas, vilken innebär en ännu större utmaning än HUGO-projektet: att karakterisera de proteiner som generna kodar för och förstå proteinernas funktion och interaktion med varandra.

Den forskningen pågår sedan länge i takt med att nya gener blivit kända. Men det kommer att ta ytterligare minst 50 år innan det arbetet – det som kallas functional genomics – närmar sig sitt slut, tror Jan Dumanski.

– Functional genomics är ett betydligt mer diffust arbete än HUGO-projektet, det används fler tekniker och det blir nog svårt att samla den forskningen under ett paraply så som man har gjort med HUGO, säger Jan Dumanski.

I Sverige finns det flera grupper som har höga ambitioner ifråga om HUGO-projektet och functional genomics, menar Jan Dumanski, men på grund av bland annat forskningsfinansierarnas rädsla för att satsa på den här typen av forskning finns det ingen samordning. Finansierarnas inställning har ofta varit att den här sortens forskningsinriktning tar för mycket resurser från andra grenar.

– Om vi tidigt fått resurser för att skapa ett nationellt centrum för genforskning hade vi efter HUGO-projektet snabbt kunnat övergå till nästa steg; functional genomics. Vi har överhuvudtaget varit ganska dåliga i Sverige på att dra nytta av genomprojektet, framför allt på sekvenseringssidan.

– Men det är en mycket kortsiktig forskningspolitik. En möjlighet att utföra ett antal större sekvenseringsprojekt innebär också en möjlighet för en högklassig utbildning av bioinformatiker. De forskare som idag arbetar med sekvenseringsprojekt kommer att vara bäst lämpade för morgondagens datorbaserade analyser av sekvenser från hela genom, exempelvis för att hitta kopplingar till sjukdomar. Bioinformatiker är idag en yrkesgrupp som vi har svårt att utbilda och ännu svårare att rekrytera utifrån, säger Jan Dumanski.

För den som vill få en uppfattning om omfattningen av sekvenseringsarbetet med kromosom 22 finns hela nukleotidsekvensen tillgänglig på Internet (<http://www.genome.ou.edu/human.html>).

Peter Örn

FOTO: HANS PETERSSON

Docent Jan Dumanski på enheten för klinisk genetik vid Karolinska sjukhuset i Solna har lett den enda svenska forskargrupp som deltagit i kartläggningen av kromosom 22. »Man kan se vårt arbete som ett pilotprojekt. För att kartläggningen av återstående kromosomer ska gå snabbare förändras nu modellen för hur sekvensbestämningen ska genomföras inom HUGO-projektet», säger Jan Dumanski.

som HUGO konkurrerar med privata intressen, framför allt med företaget Celera Genomics vars grundare är den tidigare HUGO-medarbetaren Craig Venter.

– Arbetet med kromosom 22 kan ses som ett pilotprojekt. Nu ändrar man inriktning; ett konsortium har bildats med fem stora laboratorier, och i stället för att ta avsnitt för avsnitt av specifika kromosomer i mindre databaser planerar man att sekvensera hela det humana genomet i en enda stor databas, säger Jan Dumanski.

Genom att lägga ihop den information som idag finns rörande gener och andra kända markörer på kromosomerna, så ska all ny information som kommer ur sekvenseringen i den stora databasen kunna separeras till specifika kromosomer.

– Det är oklart hur snabbt det här arbetet kan gå, eller hur stor del av geno-

met som kan sekvenseras på det sättet. Någon gång i slutet av år 2000 kanske 70–80 procent är sekvensbestämt. Sekvenseringen kommer att lämna efter sig ett stort antal luckor som måste fyllas igen med hjälp av samma tillväggångssätt som användes för kromosom 22.

Ett av problemen med att sekvensbestämma hela genomet på en gång i en stor databas blir att sortera de s k low copy repeats. Det är repeterande sekvenser, ibland ända upp till 100 000 baspar, som förekommer i ett fåtal kopior på kromosomen och som till 90–95 procent är identiska med andra segment av kromosomen.

Men inte heller kromosom 22 är helt sekvensbestämd, även om det oftast har framställts så. Det återstår elva luckor som inte har kunnat bestämmas med hjälp av befintlig sekvenseringsteknik. Ingen av luckorna utgör över 150 000 baspar.

– Av någon anledning innehåller dessa kromosomsegment antingen något som är »toxiskt» för bakterier eller jästceller, eller består av segment som är svåra att replikera med bakteriella eller jästsystem. Det kan exempelvis handla om omfattande partier, upp till 2 000 baspar, som består av en specifik svåravläst nukleotidkombination.

– Men tekniken går snabbt framåt och inom ett år kommer dessa problem förmodligen inte att finnas kvar. Vi arbetar själva med att fylla igen en av dessa elva luckor, och andra forskargrupper försöker kartlägga de övriga, säger Jan Dumanski.

Idag finns det inga belägg för att luckorna består av några viktiga gener.

90 procent kartlagt

Företaget Celera Genomics har nu kartlagt 81 procent av arvsmassan. Det meddelade företags VD Craig Venter i början av januari. Tillsammans med de cirka 10 procent av arvsmassan som karakteriserats inom ramen för HUGO-projektet är nu drygt 90 procent av människans arvsmassa känd. Fortsätter sekvenseringen i nuvarande takt kommer hela människans arvsmassa att vara sekvensbestämd inom år 2000, tror Craig Venter.