



Det humane genom og ny teknologi fører til paradigmeskifte i biologien

Figur 1. Affymetrix-udstyr består af fluidics-station, scanner og computer. På billedet ses Chris Workman, en ph.d.-studerende fra CBS, der har stået for analysen af blærekræftdata, i færd med at sætte en DNA-chip i fluidics-stationen.

Offentliggørelsen af sekvensen af det humane genom i *Nature* (15. februar 2001, 860-921) og *Science* (16. februar 2001, 1304-1351) var kulminationen på en udvikling, der har taget fart i de sidste fem år: overgangen fra et reduktionistisk til et holistisk syn på levende cellers molekylære biologi

Af lektor Steen Knudsen, BioCentrum, DTU, steen.knudsen@biocentrum.dtu.dk

Udviklingen begyndte omkring 1995 med offentliggørelsen af det komplette genom af den simple archaeon *Mycoplasma genitalium* (*Science*, 20. oktober 1995, 397-403). Den ledsagende *Perspectives* artikel i *Science* hed »Life with 482 genes«. Det var første gang, man havde et komplet overblik over alle de gener, der var nødvendige for livet i en primitiv parasit. Efter i fire årtier at have studeret enkelte gener isoleret - deres sekvens, regulering og produkt - begyndte man nu at fokusere på, hvordan et samspil af gener kunne gå op i den højere enhed, vi kalder liv.

Omtrent samtidig med offentliggørelsen af *M. genitalium* genomet begyndte det amerikanske firma Affymetrix at tilbyde DNA-chips (figur 1) til samtidig måling af mRNA produceret af ca. 6000 humane gener (se boks, side 14). Før havde man fokuseret på målinger af mRNA fra enkelte gener.

I dag sælger firmaet chips, der samtidig kan undersøge mRNA fra de omtrent 30.000 gener, der blev fundet, da det humane genom blev sekventeret.

Det er især disse udviklinger, der har ført til, hvad jeg vil kalde et paradigmeskifte i den molekylære biologi. I dag overvejer flertallet af molekylærbiologer, hvordan de nye helcelle-værktøjer i form af DNA-chips og kendskabet til den komplette genomsekvens (som vi nu har for mere end 100 forskellige organismer) kan bruges til at se deres eget forskningsområde fra en ny synsvinkel.

Til brug i dansk kræftforskning

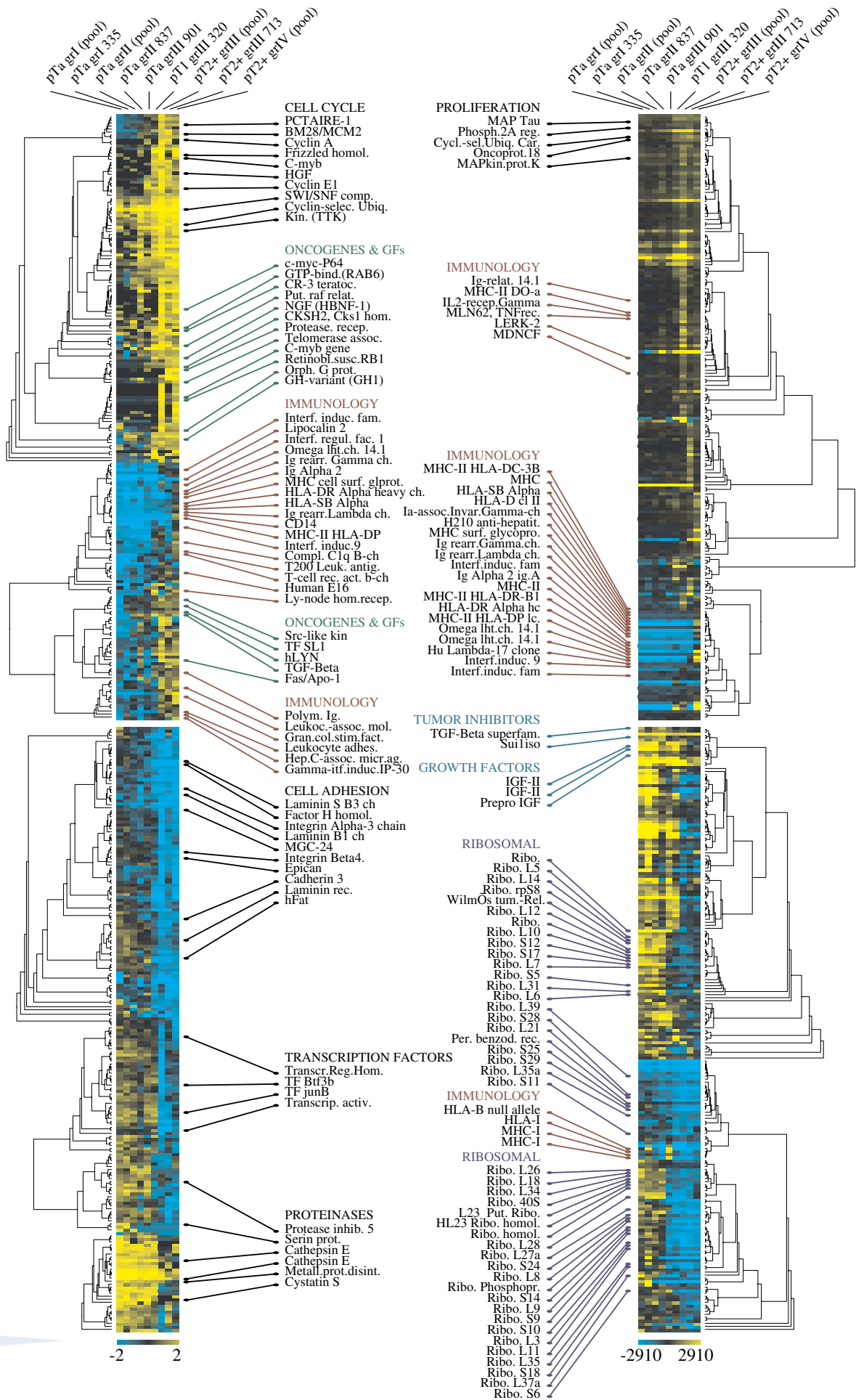
De humane DNA-chips blev hurtigt taget i brug til studiet af de store menneskelige sygdomme som cancer og AIDS. Herhjemme var overlæge Torben Ørntoft på Skejby Sygehus den

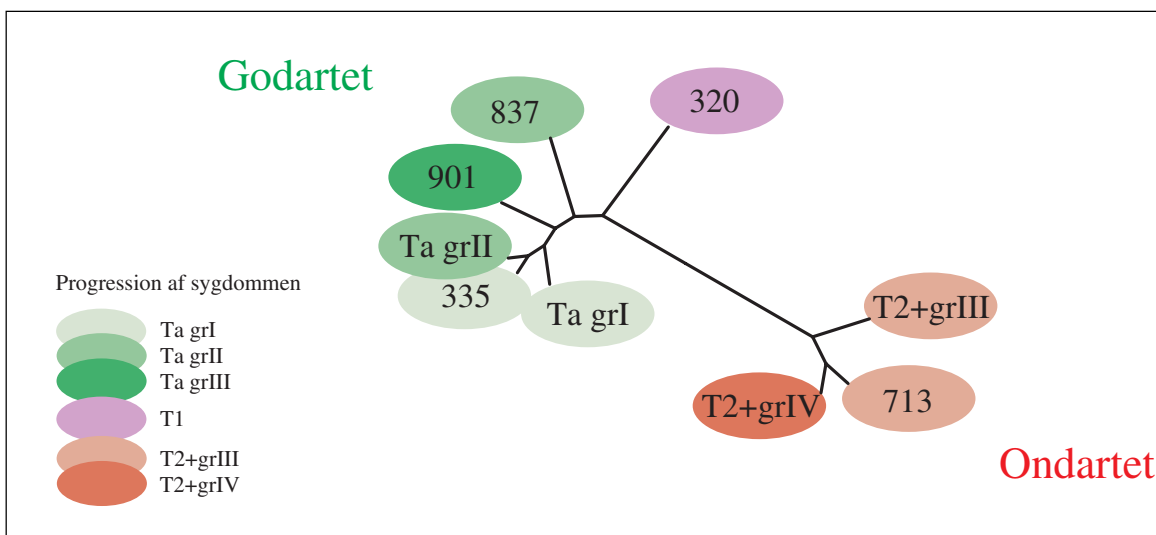
første til at se potentialet i den ny teknologi. Han indkøbte Affymetrix-udstyret og anvendte det til at studere blærekræft og tyktarmskræft. Resultatet af undersøgelsen af blærekræft blev i marts i år offentliggjort i det ansete tidsskrift *Cancer Research*. Konklusionen var, at man med DNA-chips kunne se forskelle på biopsier taget fra patienter med forskellige varianter af blærekræft - en godartet, overfladisk type, og en ondartet invasiv type. Forskellen var udmøntet i enkelte gener, hvis udtryk var ændret i de forskellige kræftformer. Studiet havde altså bragt vores viden om kræftudvikling på molekylært plan et stykke videre.

Et væld af data

Den bioinformatiske analyse af den store mængde data genereret af Ørntofts team blev foretaget på Center for Biologisk Sekvensanalyse (CBS) på BioCentrum, DTU. For hver patient havde Ørntofts team ca. 6000 målinger af mRNA, inklusive forskellige kontroller. For at få overblik over hvordan disse

Figur 2. Cluster af gener involveret i blærekræft. Hver række i farvediagrammet svarer til et gen. Farvekoden viser udtrykket af genet i de enkelte patienter (kolonner). Venstre side af figuren viser relative udtryk i forhold til normalt væv. Højre side viser absolutte forskelle i genudtryk. Baseret på målingerne på chippen indikerer et dendrogram med streger de indbyrdes forhold mellem generne. Gener med lignende målinger er grupperet sammen, og teksten indikerer funktionskategorier for kendte gener.





Figur 3. Cluster af blærekræftpatienter. Hver patient er vist med en oval og et nummer. Deres indbyrdes afstand er målt med chip og er indikeret med streger. Det kliniske billede af sygdommen er indikeret med farvekode.

data så ud i 9 patienter, beregnedes den euklidiske afstand mellem patienterne, hver repræsenteret ved et punkt med 6000 koordinater. I figur 3 illustreres de indbyrdes afstandsforhold mellem patienterne i et såkaldt dendrogram. Med farvekode er det vist, hvilken type blærekræft patienten lider af. Det viser sig, at de patienter, der ligger nærmest hinanden - på basis af målingerne af de 6000 geners udtryk - også ligner hinanden mest i kræftform.

Der ligger altså information på molekylært niveau om sygdommen, men i hvilke gener ligger den information? Der er brugt forskellige bioinformatiske metoder til at udvælge de 200 gener, der har størst betydning for sygdommen, og i figur 2 er de vist sammen med deres aktiviteter i de enkelte patienter. Ved at anvende en metode der kaldes clustering, kan vi gruppere generne efter, hvordan de opfører sig i sygdommen. Det viser sig, at der ved denne gruppering dukker grupper af gener frem, der spiller roller for forskellige processer i cellen - processer der alle er kendt for at være involveret i kræft: vækstkontrol, celleyklus, immunrespons.

Dvs. at chip-forsøgene, ud over at have bekræftet hvad vi allerede ved om sygdommen, også har kategoriseret kendte og ukendte gener i processer, vi kan forholde os til.

Det næste skridt er at udvikle diagnostik og prognose af kræft på basis af DNA-chips. Dette projekt arbejdes der på i øjeblikket.

I dag findes der også Affymetrix-udstyr på Rigshospitalet og på DTU. Desuden er udstyr til hjemmefabrikation af såkaldte cDNA-arrays, der kan konkurrere med Affymetrix' DNA-chips, ved at være udbredt i Danmark.

Integrering af data

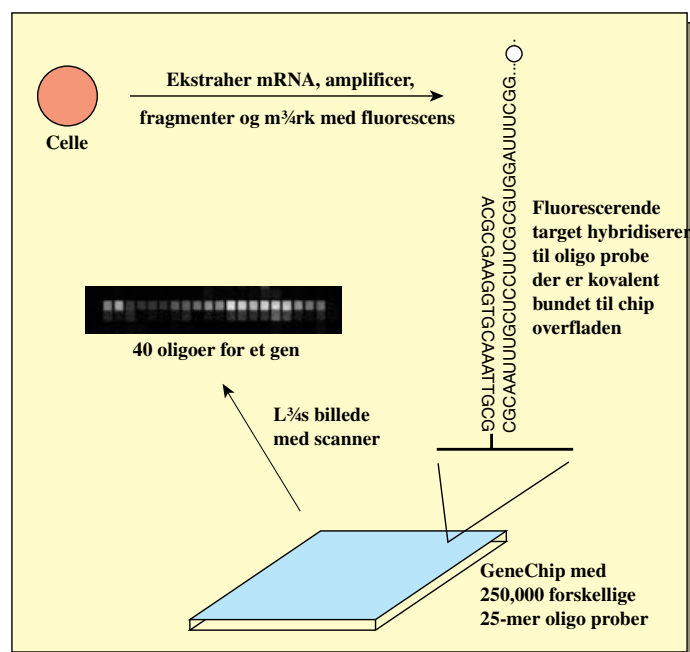
Når først man begynder at samkøre data fra DNA-chips-eksperimenter med den information, der ligger i det humane genom, er der næsten ingen grænser for, hvad man kan opdage. Det gen-udtryk man observerer på chippen må være styret af signaler, der er kodet i DNA-sekvensen, og ved snedige korrelationsanalyser mellem chip-data og sekvens kan man finde de signaler. Derved kan man opbygge modeller, der på helcelle-niveau søger at forklare de data, man observerer i DNA-chips-eksperimenterne.

Når man derudover sammenholder disse modeller med kendskab til funktionen af de proteiner, der syntetiseres fra mRNA, og kendskab til de biologiske processer proteinerne deltager i, begynder vi at nærme os en cellemodel, der retfærdiggør brugen af ordet holistisk i sin indgangsvinkel til molekylærbiologien.

Der forskes i øjeblikket intenst i en sådan integrering af data. Et vigtigt arbejde, der forestår, er at klarlægge funktionen af de ca. 30.000 gener, der blev fundet i det humane genom. Vi kender funktionen af mindre end halvdelen. Bioinformatik kan hjælpe os et godt stykke på vej ved at forudsige en funktion, der dog skal eftervises i laboratoriet.

Fremtiden

Vi har kun haft kendskab til hele sekvensen af det humane genom i få måneder og er kun lige begyndt at udforske de muligheder, der ligger i at sammenholde sekvensinformation med eksperimentelle data, både når det gælder om at forstå menneskelige sygdomme, grundvidenskabelige problemer, og hvordan mikroorganismer fungerer. De avancerede analysemetoder, der vil være i brug om fem år, er sikkert ikke engang udviklet endnu. En ting er dog sikker. Computeren er blevet et af de vigtigste værktøjer for molekylærbiologen. Mængden af data kan ikke længere overskues og fortolkes uden.



Sådan virker Affymetrix DNA-chips.