

” *Leukemiutvecklingen styrs alltså huvudsakligen av tillskott av specifika kromosomer, men mekanismerna bakom detta är idag relativt okända.*

Den vanligaste cancerformen hos barn är akut lymfatisk leukemi, ALL. Nu har forskare vid Lunds universitet i samarbete med Karolinska Institutet, SciLifeLab och University of Cambridge genomfört den största analysen av ALL på proteinnivå hittills, genom att studera aktiviteten i mer än 8 000 gener och proteiner. Resultatet visar en avvikande veckning i DNA-strängen, något som i sin tur påverkar genernas aktivitet. Här sammanfattar docent **Kajsa Paulsson** vid Lunds universitet de intressanta resultaten från den första stora studien av proteiner hos patienter med ALL.

*Största studien
av proteiner i
barnleukemi*

– extra kromosomer

Akut lymfatisk leukemi (ALL) förekommer i alla åldrar men utgör den vanligaste cancerformen hos barn, med cirka 60–70 fall årligen i åldersgruppen 1–18 år i Sverige. Ungefär 85 procent av fallen är av så kallad B-cell precursor (BCP) fenotyp, där leukemicellerna liknar omogna förstadier till B-celler. Under de senaste årtiondena har överlevnaden i denna sjukdom förbättrats kraftigt och idag har barn-ALL en god prognos med långtidsöverlevnad på över 90 procent¹.

Redan i början av 1980-talet upptäckte man att BCP-ALL kan delas in i olika genetiska subtyper med stark koppling till kliniska parametrar och prognos. Subtyperna karakteriseras av specifika somatiska mutationer som initierat och driver leukemiutvecklingen. De två största genetiska subtyperna utgörs av fall med så kallad hyperdiploidi – 51–67 kromosomer – samt fall med *ETV6/RUNX1*-fusionsgen, vilka utgör ca 25 procent vardera av alla BCP-ALL. Båda dessa grupper har en stark koppling till åldersspannet 2–4 år och en mycket god prognos².

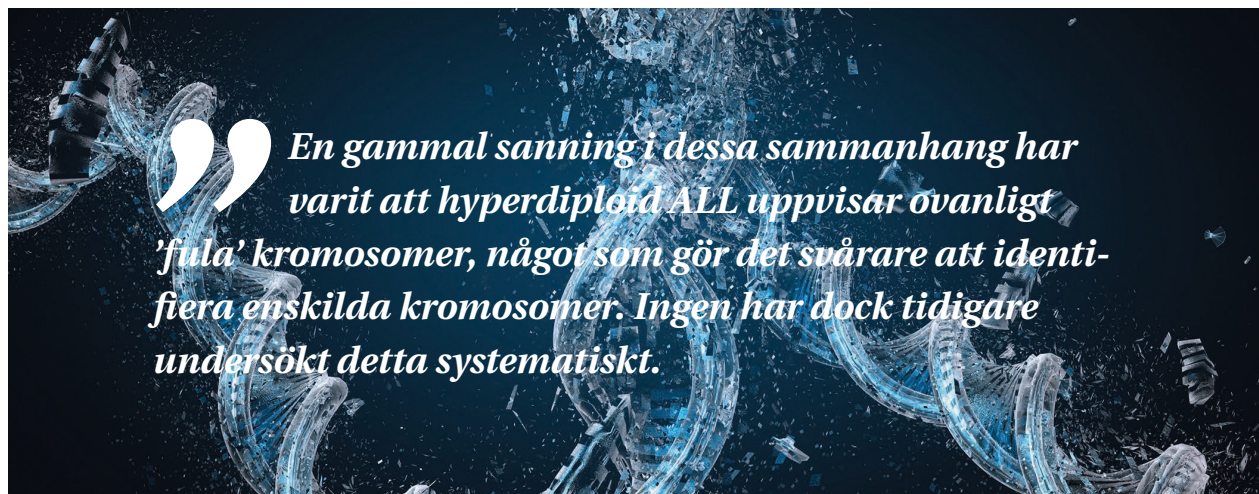


medför stor påverkan

FLER ÄN 46 KROMOSOMER

Hyperdiploidi innebär att cellen innehåller fler kromosomer än de normala 46. I hyperdiploid ALL har leukemicellerna 5–21 till synes normala extra kromosomer. De extra kromosomerna är inte slumpvis utvalda, utan mer än 75 procent av fallen har extra kopior av kromosomerna X, 4, 6, 10, 14, 17, 18 och 21. I övrigt är arvsmassan relativt ”tyst” och innehåller få andra mutationer³. Leukemiutvecklingen styrs alltså huvudsakligen av tillskott av specifika kromosomer, men mekanismerna bakom detta är idag relativt okända.

Man vet från studier av mRNA att uttrycket av gener på de extra kromosomerna är förhöjt i hyperdiploid ALL men exakt hur det påverkar leukemin har varit oklart. För att undersöka detta närmare utförde vi en så kallad proteogenomisk analys av hyperdiploid ALL⁴. Vi kunde studera nivåerna av mer än 8 000 proteiner i 18 fall av hyperdiploid ALL samt 9 fall av ETV6/RUNX1-positiva ALL, som vi använde som jämförelsegrupp. Dessutom hade vi data på mRNA-uttryck, mutationer och kopietal från samma fall samt en utökad serie av leukemier med mRNA-uttryck.



” *En gammal sanning i dessa sammanhang har varit att hyperdiploid ALL uppvisar ovanligt 'fula' kromosomer, något som gör det svårare att identifiera enskilda kromosomer. Ingen har dock tidigare undersökt detta systematiskt.*

Proteiner är betydligt svårare att studera än RNA och det är först under senare år som det har blivit möjligt att studera många proteiner samtidigt. Vi använde oss av en metod baserad på masspektrometri, där peptiderna (delar av proteiner) först separeras baserat på pH, något som leder till större datamängder och en högre representation av biologiskt relevanta proteiner⁵. Data från masspektrometri matchades sedan mot databaser innehållande genetisk information och RNA-data för att identifiera vilket protein peptiden kom från. På detta sätt kunde vi alltså få fram information om nivåerna av mer än 8 000 proteiner i varje enskilt leukemifall, vilket motsvarade cirka 2/3 av de RNA som vi hittade i samma fall. Analysen gav alltså en mycket god bild av proteininnehållet i leukemicellerna.

Vi undersökte först hur väl proteinnivåerna korrelerade med mRNA-nivåerna för motsvarande gen. Här fann vi en positiv korrelation i majoriteten (75 procent) av mRNA/protein-paren, medan en signifikant ($P < 0.05$ och justerat för multipla tester) korrelation bara kunde ses för 22 procent av paren. Liknande relativt låga korrelationer har tidigare observerats i andra cancerformer. För att se vad denna låga korrelation kunde bero på gjorde vi en rad analyser av olika processer involverade i regleringen av mRNA- och proteinnivåer. Vi fann att den relativt låga korrelationen orsakas av en kombination av flera olika processer, som till exempel en lägre korrelation för mRNA och proteiner som bryts ner snabbt i cellen och för gener som regleras av mikro-RNA. Vi kunde dra slutsatsen att mRNA-nivåer inte alltid går att direkt översätta till proteinnivåer och att proteinnivåerna alltså ger mer biologiskt relevant information än studier av mRNA-nivåer.

ASSOCIERADE MED DOSEFFEKTER

Som nämnts ovan förväntar vi oss att när vi får extra kopior av en kromosom kommer mRNA-uttrycket av de gener som finns på den kromosomen att gå upp, vilket i sin tur bör leda till ökade proteinnivåer, något som kallas *doseffekter*. Vi undersökte därför hur mRNA- och proteinnivåerna påverkades av de extra kromosomkopiorna. Mycket riktigt kunde vi se doseffekter både på mRNA- och proteinnivå

även om inte alla gener och proteiner var påverkade på detta sätt. Vidare kunde vi se att de kromosomer som oftast förekommer i extra kopior – kromosom X, 14 och 21 – också var associerade med kraftigast doseffekter. Doseffekterna kan i sin tur påverka uttrycket av andra gener, vilket kallas *transeffekter*. Därför undersökte vi också hur extra kopior av enskilda kromosomer påverkade uttrycket av gener och proteiner på andra kromosomer. Vi kunde se tydliga transeffekter för samtliga undersökta kromosomer men mängden transeffekter skilde sig åt mellan olika kromosomer. Tillskott av kromosomer leder alltså både till direkt uppreglering av de gener/proteiner som kodas av de berörda kromosomerna samt i ett andra led till effekter på hela arvsmassan.

Ett sätt att undersöka vad som händer i en cell är att se vilka signalvägar som är aktiva och inaktiva. Vi gjorde detta genom att jämföra mRNA- och proteinnivåerna i hyperdiploida och *ETV6/RUNX1*-positiva leukemier. Vi såg bland annat att signalvägar involverade i translation och proteinmetabolism var relativt sett uppreglerade i hyperdiploid ALL, något som kan vara associerat med det högre uttrycket av gener som de extra kromosomkopiorna medför. I tidigare experimentella studier har man visat att förändringar i antalet kromosomer kan leda till så kallad proteotoxisk stress⁶, där cellen inte hinner bryta ner felaktiga proteiner, och våra data skulle kunna tyda på att detta förekommer i hyperdiploida leukemier.

Våra data kunde också användas för att titta på uttrycket av enskilda proteiner av speciellt intresse för leukemikutveckling. Vi kunde därvid bland annat se att genen *FLT3*, som tidigare visats ha ett högt mRNA-uttryck i hyperdiploid ALL, också uppvisade ett högt uttryck på proteinnivå i dessa leukemier. *FLT3* är en känd leukemi-associerad gen som ofta är muterad i ALL och våra resultat tyder på att den kan vara viktig för utvecklingen av hyperdiploid ALL också när den inte är muterad. Vidare noterade vi låga nivåer av proteinet CTCF samt medlemmar av proteinkomplexet cohesin i hyperdiploid ALL. CTCF och cohesin binder till kromatin (DNA-strängen samt strukturproteiner som binder till den) i interfasedelen av cellcykeln och styr hur kro-

matinet veckas i cellkärnan. Kromatinet veckas normalt i öglor innehållande <1 Mb av DNA, så kallade *topologically associating domains* (TADs), med CTCF och cohesin i basen av öglan. TADs är i stort sett konstanta mellan celltyper och spelar stor roll för regleringen av vilka gener som ska transkriberas eftersom promotorer och enhancers i samma TAD interagerar i större utsträckning⁷. Låga nivåer av CTCF och cohesin kan därför påverka det globala genuttrycket.

AVVIKANDE TAD-STRUKTUR

Vi undersökte hur mRNA-uttrycket av närliggande gener korrelerade utifrån om generna förväntades finnas i samma TAD eller i angränsande TADs. I *ETV6/RUNX1*-positiva ALL samt i flera andra kontrollvävnader som vi undersökte kunde vi se det förväntade mönstret med en högre korrelation för gener i samma TAD jämfört med gener i olika TADs. I de hyperdiploida leukemierna såg vi däremot ingen sådan skillnad. Det tyder på att de hyperdiploida leukemierna har en avvikande TAD-struktur, med en global dereglering av genuttrycket som följd. För att undersöka om så var fallet använde vi oss av en teknik kallad high resolution chromatin conformation capture (Hi-C). I denna metod behandlar man först cellerna med formaldehyd, vilket skapar bindningar mellan DNA-molekyler som fysiskt befinner sig nära varandra i cellkärnan. Efter flera fragmenterings- och ligeringssteg gör man en storskalig sekvensering⁸. Av de resulterande sekvenserna kan man med bioinformatisk behandling göra en karta över vilka delar av kromatinet som interagerar med varandra, vilket visar bland annat TAD-strukturen. Hi-C är en ganska avancerad och dyr metod och för att kunna göra detta samarbetade vi med forskare vid University of Cambridge som utvecklat en metod för att göra Hi-C på primärt patientmaterial. Vi undersökte fyra fall av hyperdiploid ALL samt två fall av *ETV6/RUNX1*-positiv ALL och gjorde två olika sorters analyser för att studera TAD-strukturen. I den ena analysen uppvisade två av fyra hyperdiploida fall en avvikande TAD-struktur och i den andra tre av fyra hyperdiploida fall, medan övriga prover visade normal TAD-struktur. Resultaten tydde alltså på att åtminstone en del hyperdiploida ALL har förändringar i hur kromatinet veckas i cellkärnan, men fler fall skulle behöva undersökas för att vi helt säkert ska kunna säga att så är fallet.

Ursprungligen gjordes genetiska analyser främst genom att man tittade på infärgade kromosomer i mikroskop, så kallad G-bandsanalys eller cytogenetisk analys; något som också fortfarande används i klinisk verksamhet. En gammal sanning i dessa sammanhang har varit att hyperdiploid ALL uppvisar ovanligt ”fula” kromosomer, något som gör det svårare att identifiera enskilda kromosomer. Ingen har dock tidigare undersökt detta systematiskt. Med tanke på att ”fula” kromosomer skulle kunna tyda på att hyperdiploid ALL också har avvikande metafaskromosomer gjorde vi en

uttömmande analys där vi klassificerade kromosomer från cirka 40 hyperdiploida och 40 *ETV6/RUNX1*-positiva leukemier utifrån hur kontraherade kromosomerna var, hur distinkta gränser de hade och hur tydligt bandningsmönstret framträdde. Resultaten visade tydligt att kromosomerna i hyperdiploid ALL faktiskt signifikant oftare har en avvikande morfologi. Vi kan dock idag inte säga att detta orsaksmässigt hänger samman med de låga nivåerna av CTCF och cohesin, men vi planerar att undersöka detta vidare i kommande studier.

Sammanfattningsvis har vi genomfört den hittills största studien av proteiner i barnleukemi. Våra resultat visar att extra kromosomer medför stor påverkan på uttrycket av gener och proteiner, både i form av doseffekter och genom dessas påverkan på övriga gener. Vidare har hyperdiploid barn-ALL ett lågt uttryck av CTCF och cohesin, vilket verkar leda till en dereglering av genuttrycket som möjligtvis är kopplat till en förändring i kromatinets veckning. Varför CTCF och cohesin finns i så låga nivåer vet vi inte idag men det verkar inte vara associerat med mutationer i dessa gener. Vår studie visar att mätningar av proteinnivåerna är mer biologiskt relevanta än mätning av mRNA-nivåer. All proteindata vi har genererat är fritt tillgängliga för andra forskare och vi hoppas att de ska vara till nytta i många framtida studier.

REFERENSER

1. Toft N, et al. (2018) Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. *Leukemia* 32(3):606-615.
2. Harrison CJ (2000) The management of patients with leukemia: the role of cytogenetics in this molecular era. *Br J Haematol* 108(1):19-30.
3. Paulsson K, et al. (2015) The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Genet* 47:672-676.
4. Yang M, et al. (2019) Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Commun* 10(1):1519.
5. Branca RMM, et al. (2014) HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nat Methods* 11(1):59-62.
6. Santaguida S & Amon A (2015) Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nature reviews. Molecular cell biology* 16(8):473-485.
7. Merkenschlager M & Nora EP (2016) CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu Rev Genomics Hum Genet* 17:17-43.
8. Nagano T, et al. (2015) Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol* 16:175.

KAJSA PAULSSON, DOCENT, UNIVERSITETSLEKTOR I MEDICINSK GENETIK, AVDELNINGEN FÖR KLINISK GENETIK, LUNDS UNIVERSITET, KAJSA.PAULSSON@MED.LU.SE

