

**Liss Anne Solberg Lavik**

Bioingeniør med master i molekylærmedisin. Molekylærbiolog ved Avdeling for medisinsk genetik (AMG), St. Olavs hospital.

**Marie Lundbæk**

Cand.scient. og ph.d. Molekylærbiolog ved AMG ved St. Olavs hospital, under spesialisering til klinisk laboratoriegenetiker (CLG).

**Kristine Misund**

Cand.scient. og ph.d. Molekylærbiolog ved AMG, St. Olavs hospital og forsker ved Institutt for klinisk og molekylær medisin (IKOM), Norges teknisk-naturvitenskapelige universitet (NTNU)..

**Wenche Sjørusen**

Bioingeniør, cand.scient. og ph.d. CLG ved AMG ved St. Olavs hospital og professor ved IKOM, NTNU. wenche.sjursen@stolav.no

Revolusjonen innen sekvenseringsteknologi:

# Framgang med sjumilssteg

Denne artikkelen gir innblikk i den rivende utviklingen innen sekvenseringsteknologien de siste tiårene og hvordan teknologien benyttes innen medisinsk genetik. Den beskriver også hvordan genetisk utredning foregår, fra første kontakt med helsetjenesten frem til endelig svar.

## Innledning

I fagfeltet medisinsk genetik diagnostiserer man sykdommer som har oppstått enten fordi en genfeil er arvet, eller fordi en genfeil har oppstått spontant. Genetiske feil kan være endringer som påvirker kromosomenes eller genenes struktur eller funksjon, noe som gjør oss mer sårbare for utvikling av sykdom. Genetiske faktorer kan virke alene eller i samspill med miljøfaktorer. Vi antas å ha cirka 21 000 proteinkodende gener, og disse er bare delvis kartlagt med tanke på assosiasjon til sykdom. I dag kjenner man til mer enn 6000 ulike genetiske tilstander

(1) og man finner stadig nye genetiske endringer som kan knyttes til utvikling av sykdommer, noe som ny sekvenseringsteknologi har bidratt vesentlig til. I tillegg til tradisjonelle PCR-baserte metoder som sangersekvensering av DNA og RNA, og fragmentanalyse av spesifikke genetiske markører, har ny sekvenseringsteknologi ført oss sjumilssteg fremover. Nestegenerasjonssekvensering (NGS) (også kalt andregenerasjonssekvensering) har gjort det mulig å undersøke mange gener eller hele det menneskelige genom i én enkelt analyse.

Omtrent åtte prosent av alle mennesker utvikler en genetisk sykdom i løpet av livet (2). Ved Avdeling for medisinsk genetik (AMG) ved St. Olavs hospital benytter vi NGS til diagnostisering av arvelig kreft, syndromer og utviklingsavvik i fosterlivet. Hensikten med denne artikkelen er å beskrive den rivende utviklingen som har vært de siste årene innen genteresting ved hjelp av sekvensering.

## Arvelig kreft

Kreft oppstår på grunn av ukontrollert cellevekst. Økt risiko for kreft kan være arvelig, og det er kartlagt mange risikogener innen arvelig kreft. Arvelige kreftsyndromer hvor årsaksgenet er

kjent, utgjør cirka fem til ti prosent av alle kreftformer. For eksempel har noen familier økt risiko for bryst- og eggstokkreft på grunn av en genfeil i BRCA1- eller BRCA2-genet, og andre familier har økt risiko for tarmkreft på grunn av feil i ett av fire gener involvert i DNA mismatch-reparasjon (Lynch syndrom) (3). Per i dag kjenner vi mer enn 50 ulike arvelige kreftsyndromer og mer enn 80 assosierte gener (4).

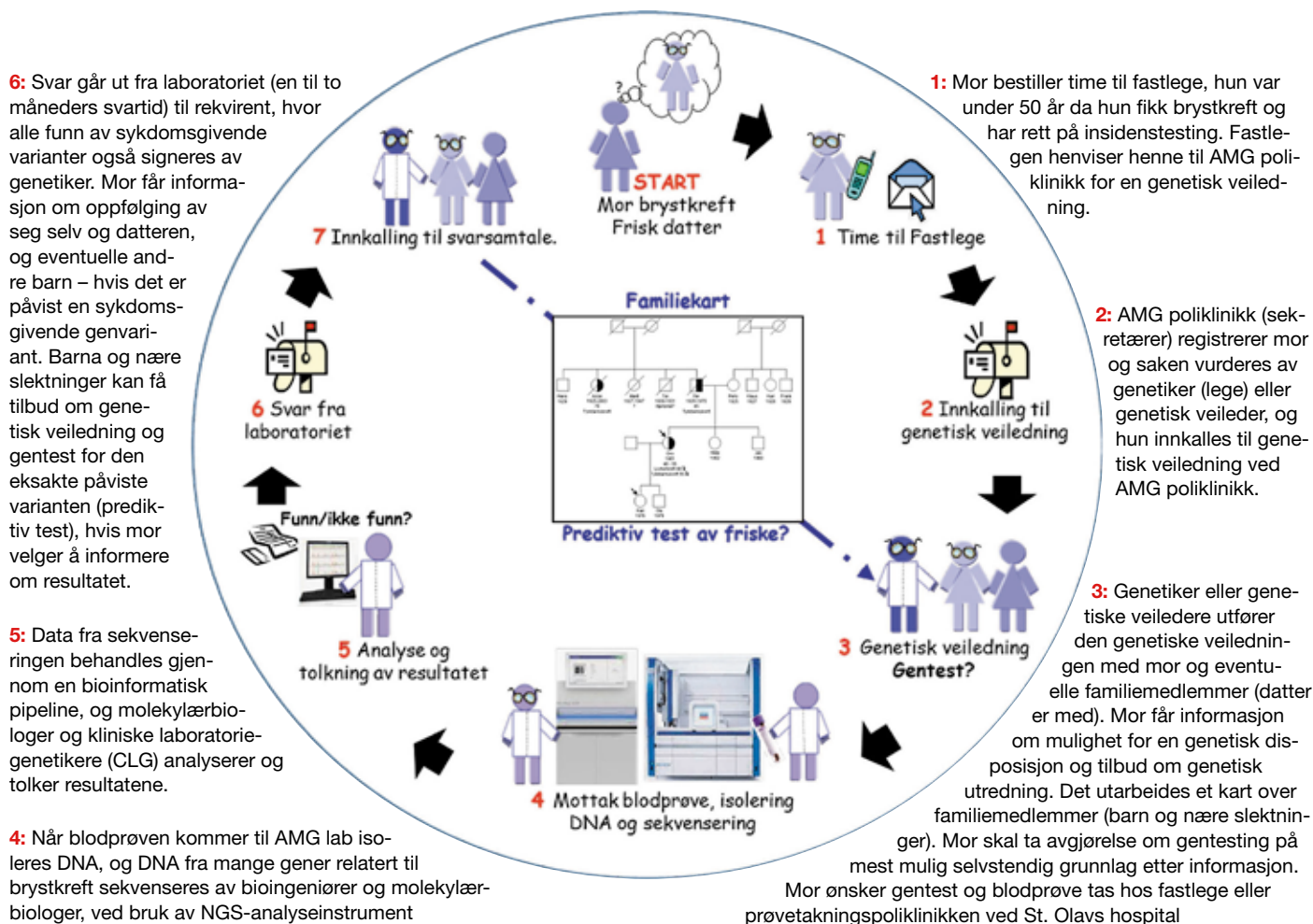
## Syndromer

Genetiske syndromer er definert som en sammenstilling av symptomer (fysiske tegn og mulige utviklingsavvik) som ofte opptrer sammen, og hvor genetisk årsak er kjent. Årsaken kan være nedarvet eller nyoppstått og omfatte en enkel genetisk variant (for eksempel Duchenne muskelatrofi, med feil i DMD-genet), eller endringer på kromosomnivå (for eksempel Downs syndrom – tre utgaver av kromosom 21). Flere tusen ulike gener involvert i syndromer er kjent, og de klassifiseres ut fra fenotyper.

## Fosterdiagnostikk

Det er utviklet NGS-baserte metoder for undersøkelse av kromosomavvik som kan oppstå i fosterstadiet. Ved indika-

**START:** Mor (index) er 50 år og har brystkreft, datter 30 år er bekymret, kan det være arvelig? Mor vil gjerne finne ut av dette.



**FIGUR 1:** Eksempel på genetisk utredning for arvelig kreft ved AMG, St. Olavs hospital.

sjoner på avvik hos fosteret ved ultralyd eller ved høy alder på mor, kan kromosomavvik undersøkes i blodprøve fra mor med NIPT (ikke-invasiv prenatal test, som utføres i samarbeid med Avdeling for medisinsk biokjemi). DNA isolert fra mors blod inneholder en liten andel av sirkulerende DNA fra fosteret, og det er tilstrekkelig til at man kan analysere for trisomi 13, 18 og 21.

### Hvorfor er det viktig med gentesting?

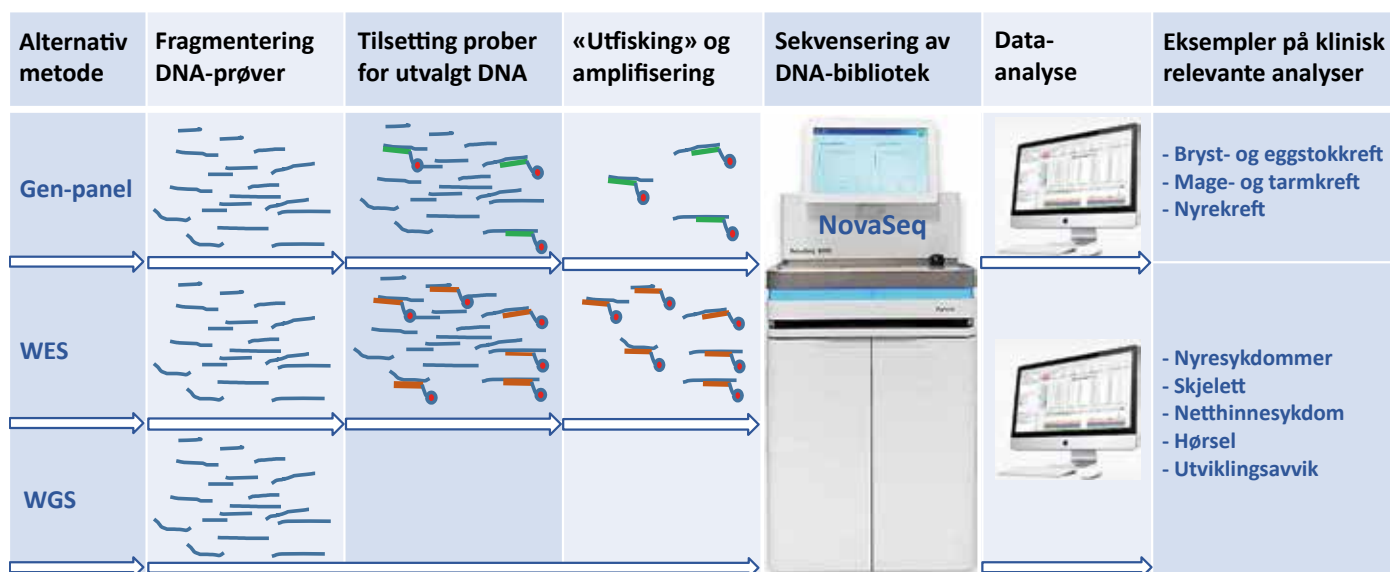
Gentesting kan være en målrettet undersøkelse av en kjent genetisk variant som er påvist i en familie, eller det kan være en undersøkelse av ett eller flere gener

for å prøve å finne en årsak til sykdom.

Når den sykdomsgivende genetiske varianten i en familie er identifisert, er det mulig å tilby gentesting av slektninger. Ved å identifisere bærere av denne varianten, kan disse tilbys oppfølging fra helsetjenesten i form av regelmessige undersøkelser og kontroller som for eksempel MR (magnetisk resonans) og CT (computertomografi). På denne måten kan man oppdage for eksempel kreftutvikling i et tidlig stadium og dermed forhindre videre utvikling og død. Figur 1 gir en skjematisk oversikt over genetisk utredning for arvelig kreft ved vår avdeling, og faktaboksen «Genetisk utred-

ning» gir en forklaring på aktuelle begreper ved gentesting.

I de tilfeller man ikke kjenner den genetiske årsaken til sykdommen fra før, er det viktig for familie og helsetjenesten å finne en forklaring på sykdomsutviklingen – for eksempel ved medfødt utviklingsavvik. Genetiske undersøkelser kan være med å støtte opp under diagnosen ved syndromutredning, fordi symptomene ofte kan være overlappende og de kan variere i omfang og alvorlighetsgrad. Også for syndromer er det viktig med korrekt diagnose, med tanke på riktig behandling og oppfølging. Det kan også være gjentakelsesrisiko, noe som er ►



**FIGUR 2:** Prinsippskisse for NGS-sekvensering ved målrettet genskvensering (genpanel), eksomsekvensering (WES) og helgenomsekvensering (WGS). DNA fragmenteres, og for målrettet genskvensering og eksomsekvensering blir de aktuelle områdene tilsatt prober som «fisker» ut DNA. Det utvalgte DNAet blir amplifisert og sekvensert. Ved helgenomsekvensering blir DNA tilsatt direkte på sekvensatoren uten «utfisking». Etter sekvensering blir rådataene behandlet i en bioinformatisk pipeline, sekvensene for de aktuelle genpanelene blir filtrert ut (selektert) og deretter analysert for mulige genfeil.

prober som kun binder til utvalgte gener. prober som binder seg til alle ekson (proteinkodende områder i DNA).

viktig å være klar over ved eventuelle nye svangerskap.

Det er i dag flere behandlingsmetoder som differensieres ut fra om årsaken til sykdom skyldes spesifikke genetiske varianter eller ikke. Persontilpasset medisin er noe helsetjenesten vil satse på fremover, og her er både medfødte og ervervede genetiske faktorer av betydning.

For alle som skal rekvirere medisinsk genetiske analyser i Norge, er det tilgjengelig en nettbasert norsk portal (5). Portalen inneholder oversikt over alle tilgjengelige genetiske analyser innen medisinsk genetik i Norge. Man kan søke på indikasjon, underindikasjon, gen og analysenavn, og få informasjon om hvem som setter opp de aktuelle analysene. Portalen har tilgjengelig rekvisisjonsskjema fra alle de medisinsk genetiske avdelingene i landet.

### Utvikling av sekvenseringsmetoder – fra Sanger til NGS

Sekvenseringsteknologien hadde sin begynnelse på 1970-tallet med intro-

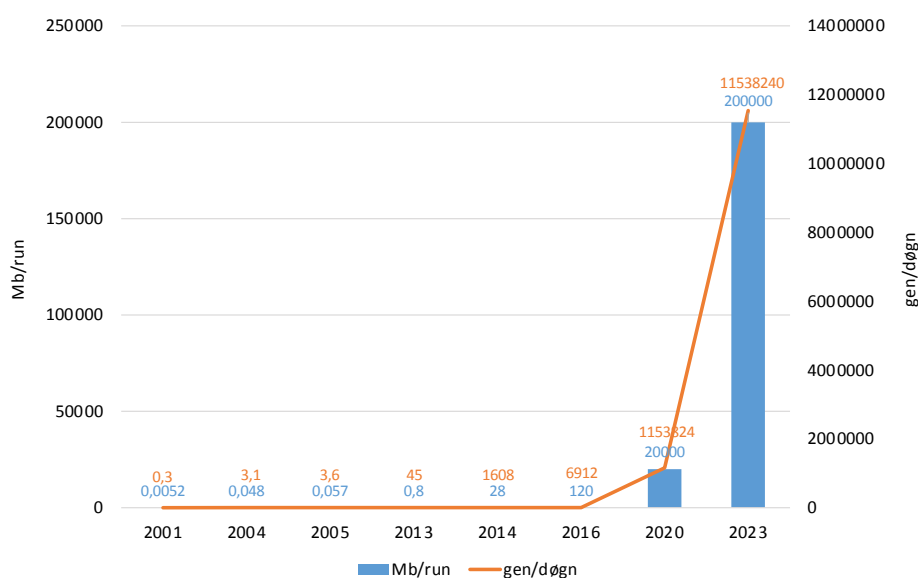
duksjonen av Sanger-sekvensering, som fortsatt er i bruk. Det første humane genomet ble sekvensert med Sanger-sekvensering. Det tok 13 år og ble publisert i 2001 (6). Dette førte til en utvikling av Sanger-sekvensering med kapillærelektroforese, og etter hvert sekvensatorer med flere parallelle kapillærer. I år 2000 ble den første NGS-teknologien lansert (7), men den ble benyttet kun til forskning det første tiåret. De siste ti til femten årene har det vært en rivende utvikling av NGS, og den har fått sin anvendelse i diagnostikken. Nå er det mulig å undersøke større deler av genomet til lavere kostnad enn tidligere.

### NGS – sekvensering av korte fragmenter (short-read)

Short-read-sekvensering baserer seg på avlesning av korte sekvenser (vanligvis 150 basepar) og er per i dag den ledende NGS-teknologien på markedet, fullstendig dominert av det amerikanske bioteknologifirmaet Illumina. Det er denne teknologien som benyttes i rutinediagnostikken, og i laboratoriet kan man

velge å gjøre enten målrettet sekvensering (kun utvalgte områder av genomet sekvenseres), eksomsekvensering (de protein-kodende delene av genomet sekvenseres) eller helgenomsekvensering (hele genomet sekvenseres, inklusiv ikke-proteinkodende deler og mitokondriegenomet) (8) (Figur 2).

Ved målrettet sekvensering og eksomsekvensering må de delene av genomet som ønskes sekvensert «fiskes ut» på forhånd. En teknisk fordel med helgenomsekvensering er at det krever mindre arbeid å klargjøre prøvene for sekvensering (såkalt biblioteksprep, figur 2), siden det ikke skal gjøres en «utfisking» av gener på forhånd, som ved eksomsekvensering. Ved helgenomsekvensering vil antall ganger hver nukleotide (DNA-byggestein) blir sekvensert være mer likt fordelt gjennom genomet (mer jevn dekning, eller «coverage»). Dette gjør helgenomsekvensering bedre egnet til å analysere større endringer som delesjoner og duplikasjoner av ett eller flere gener, hvilket er en klinisk fordel, fordi mange nyoppståtte utvik-



**FIGUR 3:** Grafisk illustrasjon og tallgrunnlag for utviklingen i sekvenseringskapasitet ved AMG St. Olavs hospital fra 2001 og frem til i dag, angitt i antall megabaser (Mb) per analyseoppsett (venstre Y-akse (Y-1)) og antall gen per døgn (høyre Y-akse (Y-2)).

lingsavvik-syndromer er forårsaket av slike kopitallsvarianter (CNV). Ved å ta i bruk helgenomsekvensering vil man bedre kunne detektere både større og mindre DNA-endringer i samme analyse, og i motsetning til eksomsekvensering vil man også kunne analysere ikke-kodende regioner, som intron og regulatoriske regioner, samt gener i mitokondrielt DNA. Totalt sett vil innføring av helgenomsekvensering kunne føre til at flere får en forklaring på sin genetiske sykdom (9). En av ulempene ved helgenomsekvensering er at det blir mange flere varianter å tolke med hensyn til om de er sykdomsgivende eller ikke. Mange varianter vil ligge i ikke-kodende regioner, og de har vi mindre erfaring med og færre verktøy (software-programmer) tilgjengelige til å hjelpe oss i tolkningen. De vil derfor ofte ende som en variant av usikker klinisk betydning (VUS). Dette problemet vil bli mindre etter hvert som vi får erfaring med tolkning av ikke-kodende varianter, og tolkningsprogrammene vil bli bedre etter hvert som det benyttes kunstig intelligens (KI) for å pre-

dikere effekt av variantene. Et eksempel på KI som allerede anvendes til tolkning av varianter er prediksjon av proteinstruktur ved hjelp av AlphaFold(10).

### Sekvensering av lange fragmenter (long-read)

Long-read-sekvensering er en teknologi som har vakt stor interesse de siste par årene. Metoden representerer tredjegenereasjonssekvensering. Per i dag domineres markedet av to bioteknologifirmaer som begge tilbyr sekvensatorer og reagenser for long-read-sekvensering, men med forskjellig metodikk: Oxford Nanopore og PacBio. Med long-read-teknologi kan man lese av lengre sekvenser, typisk rundt 10 000 basepar i snitt (11). Dette gjør at man kan lese av og analysere regioner som er krevende med den gjeldende NGS-teknologien, og man kan blant annet avdekke strukturelle varianter, varierende antall tripplett-repetisjoner (for eksempel assosiert med Huntingtons sykdom) og om en variant befinner seg på den maternelle eller den paternelle genstrengen. Denne

### FAKTA |

#### Genetisk utredning – begreper

- **Prediktiv gentest** betegner testing av friske personer som enten har en kjent sykdomsgivende variant i familien eller har en familiehistorie som kan være forenlig med arvelig sykdom.
- **Diagnostisk gentest** betegner testing av syke personer med indikasjon på eller mistanke om arvelig betinget sykdom.
- **Proband** betegner den person i en familie som er henvist til genetisk utredning.
- **Autosomt recessiv arv** betegner tilstander der genfeil må forekomme på begge genstrengene for at sykdommen kommer til uttrykk.
- **Autosomt dominant arv** betegner tilstander der genfeil på kun den ene genstrengen er sykdomsgivende.
- **X-bunden arv** betegner tilstander der genfeilen forekommer på X-kromosomet. Affiserer primært gutter. Jenter kan være affisert, men har da gjerne mildere symptomer.
- **Mitokondriell arv** betegner tilstander der genfeilen forekommer i det mitokondrielle genom. Affiserer både gutter og jenter, men nedarves alltid maternelt.
- **Bærer** betegner en person som bærer en genfeil. Ved autosomt recessive sykdommer betegner bærertilstand at man er frisk bærer av en enkelt sykdomsgivende variant.
- **VUS** betegner en variant av usikker klinisk betydning.

teknologien har derfor et stort potensial for klinisk bruk. Long-read-sekvensering brukes ennå ikke i rutinediagnostikken, men mange laboratorier har tatt teknologien i bruk i forskningsprosjekter (12-14). Også ved AMG har vi pågående forskningsprosjekter der vi benytter long-read-teknologi.

### Bioinformatisk kompetanse

Moderne sekvenseringsmetoder genererer store datamengder. Dette krever en infrastruktur som kan håndtere analyser av stadig større datasett og som har høy lagringskapasitet. Den stadige økningen i NGS-analyser krever også mer ►

bioinformatisk kompetanse i avdelingen. I tillegg til bioinformatikere, som etablerer og vedlikeholder dataflyt av store datamengder, må også laboratoriepersonell (som bioingeniører) ha kunnskap om den bioinformatiske behandlingen av sekvenseringsdata som foregår før resultatet av gentesten analyseres (for eksempel kvalitetskontroll og diverse filtreringstrinn).

### Sekvensering ved AMG

Ved AMG på St Olavs hospital startet vi med vår første sekvensator i 2001, en ékapiillær maskin. Det var da mulig å analysere ett genfragment fra en pasient av gangen, og det tok flere døgn å analysere ett gen. For pasienter som ble utredet for eksempel for arvelig tarmkreft, lå svartiden de første årene på opptil ett år. Etter hvert fikk vi flerkapillær-sekvensatorer; 16 kapillærer i 2005 og 48 kapillærer i 2014. Vår første NGS-maskin fikk vi i 2012, en Roche GS Junior (pyrosekvensering) med kapasitet til å analysere fire gener fra 48 pasienter. Analysen tok totalt en liten uke, og svartiden ble redusert til noen få måneder. Denne teknologien ble snart utdatert, og vi kjøpte vår første Illumina-maskin i 2016, en MiSeq. Siden utviklingen i feltet skjer så raskt, og det blir dyrt å stadig bytte ut sekvensatorene, har vi siden 2020 samarbeidet med kjernefasiliteten for sekvensering ved universitetet (NTNU). Fra 2020 fikk vi tilgang til Illumina NextSeq, og fra 2023 NovaSeq. Vi startet med eksomsekvensering i 2021, og er nå i ferd med å innføre helgenomsekvensering. Vi har som mål å erstatte majoriteten av eksomsekvenseringen med helgenomsekvensering i diagnostikken i løpet av 2024. Utviklingen ved vår avdeling med fokus på kapasitet er illustrert i Figur 3.

### Anvendelse av NGS i diagnostikken

Ved AMG benytter vi målrettet sekvensering ved kreftutredning og eksomsekvensering ved syndromutredning. Det betyr at rent laboratorieteknisk behandles alle pasientprøver for utredning av arvelig kreft likt, og vi sekvenserer et utvalg på cirka 80 gener assosiert med arvelig kreft. Når sekvenseringsdataene skal analyseres, analyserer vi der-

imot kun de genene som har relevans for den aktuelle kliniske problemstillingen. Disse filtreres ut bioinformatisk. Hvis indikasjonen er bryst- og/eller eggstokkreft, analyserer vi for eksempel et panel som består av 21 gener. I noen familier er det mange ulike krefttyper, og dette fører ofte til at vi må re-analysere for å se på et annet panel (for eksempel Mage-tarmkreft på 24 gener). Derfor har vi nylig innført et panel på 50 gener, som dekker de mest vanlige formene for arvelig kreft. Oversikt over alle genpanel kan rekvirentene finne i genetikkportalen (5).

Det samme prinsippet gjelder for syndromutredning: Vi utfører eksomsekvensering for alle pasienter som er henvist for syndromutredning, men analyserer kun de genene som har relevans for den aktuelle kliniske problemstillingen. Her benytter vi genpaneler som er satt sammen av Genomics England PanelApp (15) som utgangspunkt for hvilke gener vi har med i våre genpaneler. Eksempler på paneler som tilbys er Nyresykdommer (259 gener), Netthinnesykdom (201 gener), Hørsel (113 gener), Skjelett (380 gener) og Utviklingsavvik (1973 gener).

### Oppsummering

Revolusjonen som har vært innen NGS har medført en dramatisk økning i kapasitet og mye kortere svartid. Dette har kommet pasientene til gode, både i form av kortere ventetid for svar på en genetisk utredning og i form av et bredere analysetilbud.

Da vi begynte med Sanger-sekvensering i 2001, brukte vi en uke på å sekvensere ett gen i en pasientprøve. Nå i 2023 sekvenserer vi på en uke 48 pasientprøver (80 kreftassosierte gener) med NGS til målrettet sekvensering, eller 12 pasientprøver (hundre til tusenvis av gener avhengig av fenotypen) med eksomsekvensering. Fra i år er målet å benytte helgenomsekvensering til syndromutredning. Da vil også kopitallsanalyse for alle gener kunne gjøres.

Ny teknologi har skapt behov for en ny type fagfolk og en ny type kompetanse. Tolkning av varianter er blitt et stort fagfelt, og bioinformatikere er blitt en uunnværlig del av staben ved AMG. ■

### Referanser:

1. Online Mendelian Inheritance in Man: <https://www.omim.org/statistics/entry> (2.11.23).
2. Paus B. Det store norske leksikon, Medisinsk genetik: [https://sml.sn.no/medisinsk\\_genetik](https://sml.sn.no/medisinsk_genetik) (7.6.23).
3. Hall MJ, Obeid EI, Schwartz SC, Mantia-Smaldone G, Forman AD, Daly MB. Genetic testing for hereditary cancer predisposition: BRCA1/2, Lynch syndrome, and beyond. *Gynecol Oncol*. 2016;140(3):565-74.
4. Cheng DT, Prasad M, Chekaluk Y, Benayed R, Sadowska J, Zehir A, et al. Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. *BMC Med Genomics*. 2017;10(1):33.
5. Norsk portal for medisinsk-genetiske analyser: <https://www.genetikkportalen.no/> (16.1.24).
6. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
7. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, et al. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol*. 2000;18(6):630-4.
8. Gorcenco S, Ilinca A, Almasoudi W, Kafantari E, Lindgren AG, Puschmann A. New generation genetic testing entering the clinic. *Parkinsonism Relat Disord*. 2020;73:72-84.
9. Stranneheim H, Lagerstedt-Robinson K, Magnusson M, Kvarnung M, Nilsson D, Lesko N, et al. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. *Genome Med*. 2021;13(1):40.
10. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-9.
11. Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. *Front Genet*. 2019;10:426.
12. Gupta P, Nakamichi K, Bonnell AC, Yanagihara R, Radulovich N, Hisama FM, et al. Familial co-segregation and the emerging role of long-read sequencing to re-classify variants of uncertain significance in inherited retinal diseases. *NPJ Genom Med*. 2023;8(1):20.
13. Reiner J, Pisani L, Qiao W, Singh R, Yang Y, Shi L, et al. Cytogenomic identification and long-read single molecule real-time (SMRT) sequencing of a Bardet-Biedl Syndrome 9 (BBS9) deletion. *NPJ Genom Med*. 2018;3:3.
14. Höijer I, Tsai YC, Clark TA, Kotturi P, Dahl N, Stattin EL, et al. Detailed analysis of HTT repeat elements in human blood using targeted amplification-free long-read sequencing. *Hum Mutat*. 2018;39(9):1262-72.
15. Genomics England PanelApp: <https://panelapp.genomicsengland.co.uk/> (2.11.23).