

# Humán Genom Projekt: teljes emberi genom meghatározása

1953 áprilisában DNS szerkezete: Watson és Crick (*Nature*)

70-es évek közepe: DNS szekvenálási technikák

1990 október 1. Humán Genom Projekt kezdete.

2003 teljes emberi genom meghatározása

## Humán Genom Projekt céljai:

1. A humán DNS összes génjének (a korábbi becslés  $\sim 100.000$  volt, a mai ismert szám  $< 23.000$ ) azonosítása.
2. A humán DNS nagyjából 3 milliárd bázispárjának meghatározása.
3. A kinyert információ mennyiség adatbázisban történő tárolása.
4. Az adatfeldolgozási eljárások korszerűsítése.
5. A projekt során létrehozott illetve korszerűsített technológiák átadása a magánszektornak.
6. A felmerülő etikai, jogi és társadalmi kérdések tisztázása.

## **Előzmények:**

1985. május: University of California, Santa Cruz (UCSC): 12 szakértő bevonásával tartott tanácskozás a Humán Genom Projektről.

## **Konklúziók:**

A projekt technikailag kivitelezhető, de nagy kihívás.

6-an szavaztak végrehajtása mellett 6-an ellene.

## **Ellenzők indokai:**

Sok pénzt igényel, amit a többi kutatástól kell elvenni, és amit a nagyrészt „junk”, genom megszekvenálására használnánk.

A meglevő technológiák nem ideálisak ehhez a komplex projekthez.

Szekvenálás és térképezés monoton feladat, nem lesz vonzó élvonalbeli laborok számára.

## **A humán genom vázlatos (*draft*) szekvenciája**

2000 júniusában a Humán Genom Konzorcium és a Celera együtt mutatta be a humán genom első vázlatos szekvenciáját.

Az eukromatinnak kb 90%-át szekvenálták meg „résekkel” és sok hibával.

A Humán Genom Konzorcium 2001-ben a Nature-ben, publikálta a szekvenciát és teljes mértékben elérhetővé tette azt.

A Celera Genomics szintén 2001-ben a Science-ben, és nem tette teljes mértékben elérhetővé.

### Újdonságok:

Ismétlődő szekvenciákról,

a gének genomon belüli elhelyezkedéséről,

gének számáról: az előrejelzett 100 000 helyett csak alig 23000 gén található meg a humán genomban, körülbelül annyi, mint az *ecetmuslica* genomjában.

# Teljes humán genom szekvencia meghatározása

2003-ban 2 évvel a határidő előtt fejeződött be a projekt.

Az eredmény felhasználási lehetőségei:

evolúciós múlt felderítése: rokonságok, homológ szervek kérdése

eszköztár kialakítása, amely egyéb fajok genomjának meghatározásához is felhasználható

eszköztár a betegségek genetikai alapjainak tisztázásához

„Big Science” megközelítés alkalmazása: adatmegosztás, adatok nyilvánossá tétele, adatfeldolgozáshoz nyitott forráskódú szoftverek alkalmazása.

Inspiráció további genom projektekhez. HapMap Projekt (a haploid humán genom térkép megszerkesztése variációkkal), 1000 Genome projekt, The Cancer Genome Atlas, Human Brain Projekt, Humane Proteom Projekt.

## **Etikai kérdések avagy kinek a DNS-ét használták a szekvenáláshoz?**

Eredeti cél az volt, hogy a szekvenált DNS anonim legyen, ne lehessen visszakövetni, hogy kié volt.

A HGP során lehet tudni, hogy a minta 70%-a 1 donortól maradék 30% több donortól van. Válogatás véletlenszerűen történt.

A Celera is eredetileg több etnikumból gyűjtött mintát és véletlenszerűen sorsolta ki azokat, de aztán kiderült, hogy a Celera vezetője Craig Venter variált a mintákkal és a saját DNS-ét használta fel a szekvenáláshoz. Azt is bejelentette, hogy szívbetegségre és Alzheimer kórra van hajlama.

1000 dollár ma az exonok és 1500 a teljes eukromatinok szekvenálásának költsége.

Venter a saját genomját elvileg nyilvánosságra hozhatja, de ezzel rokonairól is fontos információt árul el, amihez nincs joga az ő beleegyezésük nélkül.

Ha egy letális mutációt észlelünk az embrióba való beavatkozás az emberi evolúcióba történő mesterséges beavatkozás, előre nem látható géngyakoriság változásokat okozna.

Emberi genom szerkesztése és az ember klónozása nem engedélyezett.

## **Mi lesz a Humán Genom Projekt után?**

A DNS bázissorrendjének meghatározása utáni lépések: a gének jelentésének megfejtése, az egyes működések sorrendjének, kapcsolati hálójának megismerése.

A DNS bázissorrendjében ott a múlt:

A számbárszarvas hasonló a gímszarvashoz, de csak távoli rokona, de a milu evolúciós unokatestvére.



A számbárszarvas. Előfordulás: India, Sri Lanka, Banglades, Nepál, Bhután, Kína, Thaiföld, Laosz

Gímszarva: Európa, Kaukázus

Dávid-szarvas, Milu: eredetileg Kínában élt, vadon már kipusztult

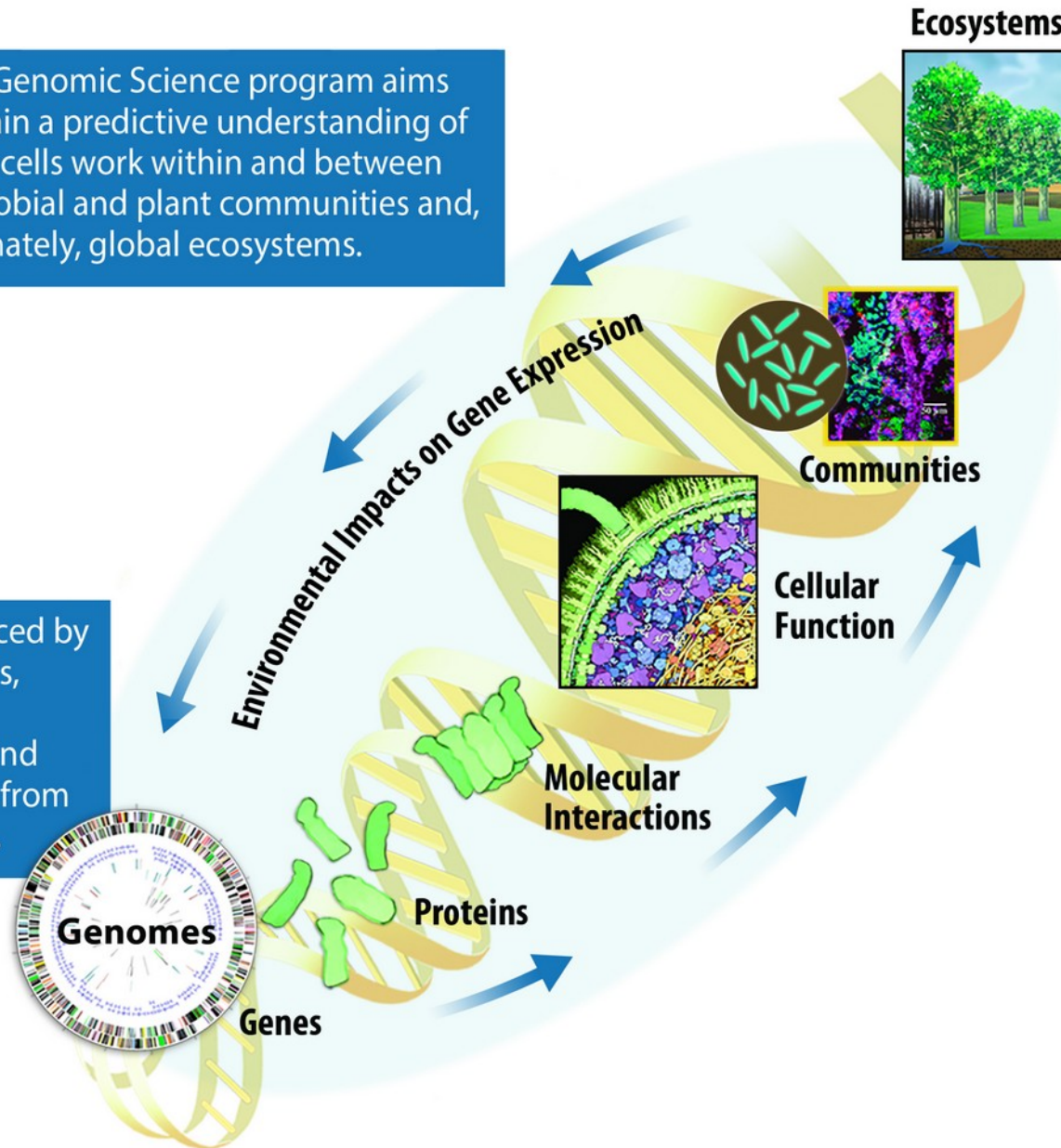


## **Genomic Science Program**

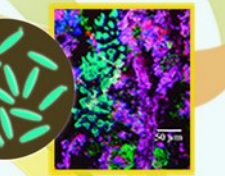
A világ energia és élelem ellátására, környezeti kihívásokra megoldást találhatunk a növényi és mikrobiális genomok kölcsönhatásának tanulmányozásával.

The Genomic Science program aims to gain a predictive understanding of how cells work within and between microbial and plant communities and, ultimately, global ecosystems.

The genome, influenced by environmental factors, determines dynamic biological structure and function at all scales, from genes to ecosystems.



**Ecosystems**



**Communities**

**Cellular Function**

**Molecular Interactions**

**Proteins**

**Genes**

**Genomes**

Department of Energy's Genomic Science program támogatja azokat a kutatásokat, amelyek növény- és mikrobaközösségek genomját vizsgálja, ezeknek a közösségeknek és környezetünknek a kölcsönhatását, továbbá azt, hogy hogyan változik a környezet hatásukra, és a környezet hogyan hat erre a metagenomra.

Néhány kutatási irány:

Költséghatékony, megújuló bioüzemanyag- és fűtőanyagforrások keresése.

Növényi gének és génfunkciók feltárása.

Génmódosított növények és mikróbák előállítás.

Biogeokémiai ciklusok háttérét biztosító ökoszisztémák egyes tagjai között fennálló kölcsönhatások felderítése.

## **Géndiagnosztika és génterápia létrejötte:**

HGP egyik legfontosabb eredménye, orvoslás forradalma, egyéni orvoslási eljárások kidolgozásának lehetősége.

### **Géndiagnosztika:**

Betegségekre hajlamosító gének feltárása, életmódbeli tanácsokkal, leghatékonyabb kezelés kiválasztásával segíti az egészség megőrzését

### **Génterápia:**

Testi sejtek génjeinek megváltoztatásával gyógyítja vagy mérsékli a betegséget.

## Betegségek genetikai háttere: fölös génkópiák

Kissejtes tüdőrák: epidermális növekedési faktor receptor (*EGFR*) génből fölös kópia

Mellrák: HER2 gén (human epidermális növekedési faktor receptor 2-es típusa) túlműködik.

## Fiatalkori Alzheimer kór kialakulása:

Amiloid prekursor protein génjének pontmutációihoz köthető típus: 21 kromoszómán (Down kórosakban is korai Alzheimer kór tapasztalható)

Preszenilin 1-es és 2-es enzimek génjeinek pontmutációi: 14-es és 1-es kromoszóma hosszú karján helyezkednek el, és ötvennél több pontmutációjuk ismert.

Az apoE4 allél, genetikai rizikótényező: A 19-es kromoszóma hosszú karján, az apoE molekula E4-es alléljának öröklése rizikótényező mind a korai, mind a késői AK formák esetében.

## **Egy nukleotidot érintő polimorfizmusok (SNPs: „snips”)**

Leggyakoribb genetikai variáció.

A HGP része volt a humán genom variációinak a vizsgálata is.

Különbség egy bázispárban van. 1000 nukleotidonként 1 különbség fordul elő, 4-5 millió a genomban. Tipikusan gének közötti szakaszokban. Általában SNP-ről beszélünk akkor, ha a variáció populációs gyakorisága meghaladja az 1%-ot, bár manapság minden különbség lehet SNP.

Egy részük biológiai markerként viselkedik, egyes betegségekre utalhatnak, környezeti faktorokra, gyógyszerekre adott választ határozhatják meg.

Az összes SNP-nek 0,1%-a változtat meg aminosavkódot, ennek 40%-a non-konzervatív.

A többi SNP első ránézésre semleges, fenotípusos megjelenés szempontjából (betegségekre való hajlam, környezeti tényezőkre, gyógyszerekre való reagálás is) nem meghatározó, hogy az SNP változtat-e meg aminosavkódot vagy nem. Az utóbbi időkben felfedezett betegségekhez kapcsolható SNP-k túlnyomó többsége nem változtat meg aminosavkódot.

## **Szekvenciavariációk (CNV: copy number variation)**

2006 végén, ahogy egyre javultak a genom vizsgálati módszerei, rájöttek, hogy a genomban rengeteg kisebb-nagyobb méretű kópiaszám-variáció fordul elő.

Vannak olyan, 1000-tól akár több megabázis nagyságú nukleotid szekvenciák, amelyek, ha a genomokat összehasonlítjuk, különböző kópiaszámban fordulnak elő.

Nincs olyan sok, mint az SNP-kből, azonban mivel nagyobb genomterületeket érintenek, összességében két ember között nagyobb variációért felelnek, mint az SNP-k.

A teljes genom kb 12%-át érintik ezek a variációk, és eddig 2900 gént találtak, amely érintett.

Egyes emberek különbözhetnek abban, hogy egy génből hány kópia található meg bennük. Az esetek többségében nem okoz látható tünetet.

DE: skizofrénia, a HIV/AIDS hajlam, a Crohn-betegség, vesebetegségek, Alzheimer-kór vagy az obezitás esetében a CNV-k szerepe bizonyított.

A betegségek mellett pl. a transzplantációban is szerepet játszhatnak a CNV-k. Például, vannak populációs szinten is gyakori, egész géneket érintő deléciók. Ha a beültetést kapó szervezetből hiányzik egy gén, (expresszáldó fehérje is), akkor a donorszervben megtalálható fehérje ellen immunválasz alakul ki.

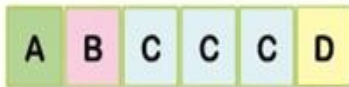
# A genomban előforduló szerkezeti variánsok



Referencia



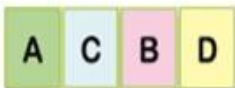
Szegmentális duplikáció - biallélikus CNV (C)2



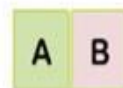
Multiallélikus CNV (C)3



Komplex CNV (D)4(CD)3



Inverzió (CB)



Deléció (CD)

Kromoszóma



## **A Humán Variom Projekt**

A különböző egyénekből származó változékony DNS szekvenciák összegyűjtése az orvostudomány számára, a személyre-szabott orvostudomány (personalized medicine) segítésére.

Ennek koncepciója, hogy az emberek közötti különbségek miatt hatékony gyógyítás egyéni gyógymódokkal lehet.

Ehhez ismerni kell a gyógyítás szempontjából fontos genetikai variációkat, gyógyszert kell rá fejleszteni. (Ehhez alapvetően át kell alakítani a gyógyszerkifejlesztés és engedélyezés jelenlegi gyakorlatát, klinikai tesztelési fázisát).

### **Az 1000 genom projekt**

nagyszámú egyén DNS-ének szekvenálása, deklaráltan az ember genetikai változékonyságának megismerése céljából. A projektet 2010-ben fejezték be.

Ma már a nagyobb volumenű genom projektek futnak: Nagy-Britannia százezer, az USA és Kína 1-1 millió emberi DNS megszekvenálásba kezdett bele.

## **A gének szerkezetének ismerete – a betegségek felismerése és gyógyítása**

### **Molekuláris medicina:**

Egyénre szabott terápia biztosítása

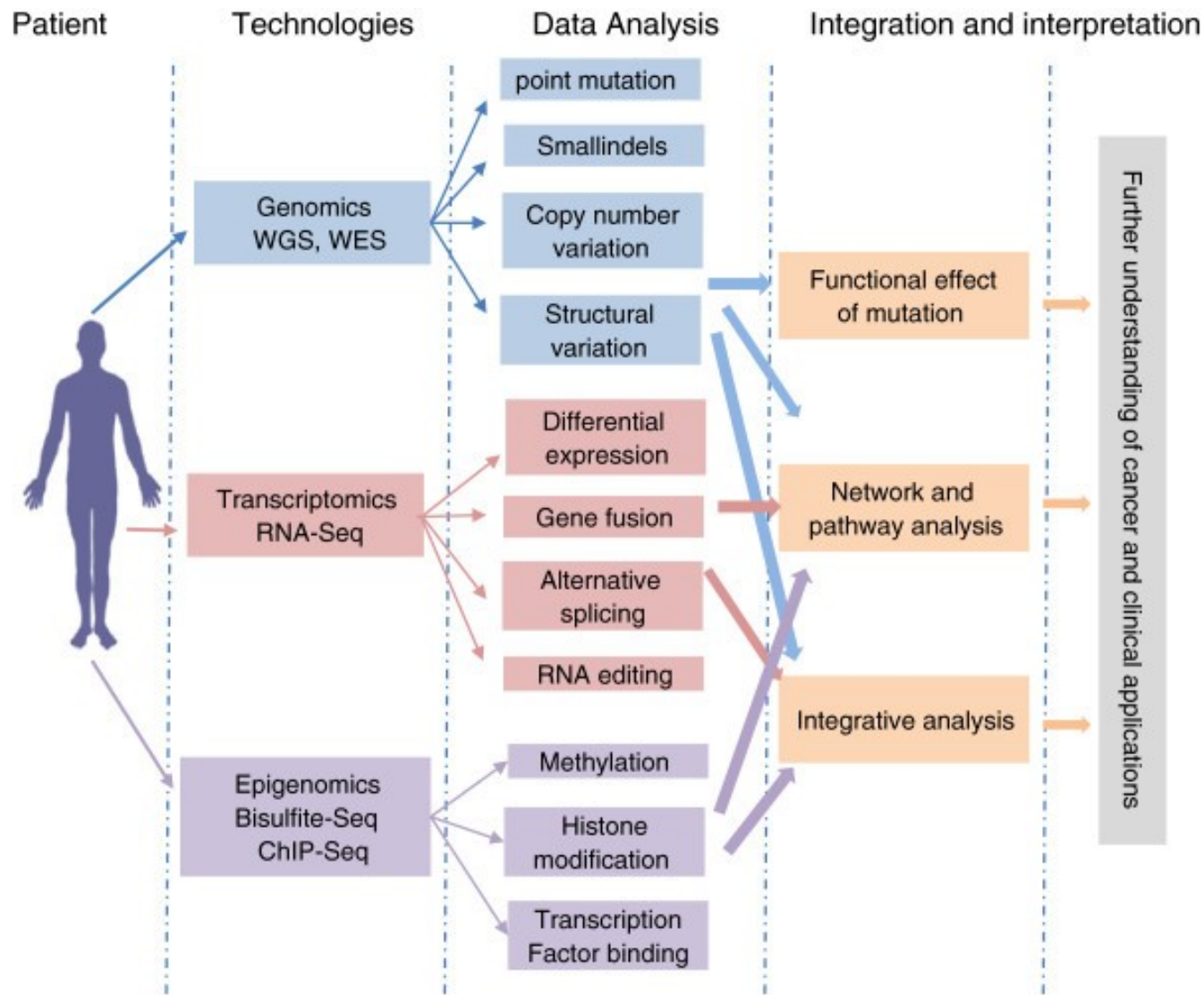
Megelőzés előtérbe kerülhet: hajlamok feltárása

Pontos diagnózis: betegségekben létrejött mutációk kimutatása

### **Korlátok:**

1. A gén szerkezetéből nem vezethető le a gén működése, csupán a kódolt fehérje elsődleges szerkezete. A gének szerkezetében alig van különbség az emlősfajok között. A fenotípusbeli különbségekért elsősorban gének működésbeli különbségei felelősek.
2. Nem minden betegség genetikai betegség.
3. Genetikai betegségeknel lehet néhány vagy nagyon sok mutációtípus is. Az utóbbi térképezése elég nehéz.

4. A gének szabályozó régióiban történő változékonyság jelentőségét jelenleg alig ismerjük, azt sem tudjuk, hogy egy adott szekvencia részt vesz-e a génkifejeződés szabályozásában. Nem elegendő a bázissorrend ismerete, tudnunk kell, hogy egy-egy változás a DNS-ben milyen szerepet játszhat egy adott betegség kialakításában. Jelenleg génexpressziós vizsgálatok folynak (nem szekvenálás), az érintett szövetek (mRNS és/vagy fehérje tartalmát) vizsgálják.
5. Nem egyértelmű a gének és a környezet kapcsolata: a környezet a gének kifejeződésére hat, ez az információ azonban nincs benne a DNS bázis sorrendjében. A környezeti hatások sok esetben a DNS metilációra és a hisztonok kémiai mintázataira hatnak, ezzel a génkifejeződést szabályozzák.



**WGS:** whole genome sequencing

**WES:** whole exome sequencing

**Small indels:** small insertions and deletions

**ChIP Seq:** method for identification of regulatory elements in the genome. It combines chromatin immunoprecipitation (ChIP) with parallel DNA sequencing to identify the binding sites of DNA-associated proteins.

**Bisulfite-Seq:** determination of DNA methylation

Valójában betegségekben a fehérjék összekapcsolt hálózata változik egyszerre, a sejt áthangolódik, és másképp kezd működni – ezeket a funkcionális interakciós hálózatokat vizsgálja a proteomika.

Hamar kiderült, nem az az igazán érdekes, hogy milyen fehérjék vannak egy sejtben, hanem az, hogy ezek mennyisége, aránya hogyan változik egy adott behatásra, egy adott állapotváltozás, öregedés, betegség, gyógyszeres kezelés kapcsán.

Ma egy 2 mm<sup>3</sup>-es szövetmintából 4–5 ezer fehérjét tudunk kimutatni, és a '90-es évek második felében kifejlesztett differenciál-gélelektroforézis révén a változásokat is nyomon tudjuk követni.

A rendszerbiológia kialakulásával párhuzamosan a biokémia, ami száz évig azzal foglalkozott, hogy szétszedte a sejtet molekulákra, eljutott a serdülőkorba, azaz már bele tud kezdeni a mérhetetlen mennyiségű információ szintetizálásába.

Így az igény, hogy megértsük a bonyolultabb molekuláris szerkezeteket, találkozott azzal a lehetőséggel, hogy összerakjuk a nagyobb képet, mivel az egyes molekulákról már ehhez elégséges adat áll rendelkezésünkre.

Hamar kiderült, nem az az igazán érdekes, hogy milyen fehérjék vannak egy sejtben, hanem az, hogy ezek mennyisége, aránya hogyan változik egy adott behatásra, egy adott állapotváltozásra, vagy öregedés, betegség, gyógyszeres kezelés kapcsán.

Ma egy 2 köbmm-es szövetmintából 4–5 ezer fehérjét tudunk kimutatni, és a '90-es évek második felében kifejlesztett differenciál-gélelektroforézis révén a változásokat is nyomon tudjuk követni.

A neurodegeneratív és a pszichiátriai betegségek területén a rendszerbiológiai, (neuro- és pszicho) proteomikai megközelítéstől várnak áttörést.

A mai pszichiátria nagy problémája, hogy a beteget nem molekuláris szinten, hanem a viselkedéses jelek alapján diagnosztizálja, majd választ egy gyógyszert, amely a molekulák szintjén hat. Így rendkívül jól lehet például a skizofrénből parkinsonost csinálni, ami persze kellemesebb a környezetnek, de nem jelent gyógyulást.

A másik megoldandó kérdés a neurodegeneratív elváltozások korai diagnosztikája és a betegség kifejlődésének megakadályozása. A kórképek mögött komplex anyagcsere-proteomikai változást sejtene, vizsgálják például a mitokondriumok mozgását, szinapszis körüli elhelyezkedését – biomarkereket keresnek.

Gyakorlati sikerhez nem invazív molekuláris képalkotó diagnosztikai vizsgálatok kellene: PET-CT-vel feltárni, hogy az agyi hálózatokban változik-e valami úgy, hogy abból neurodegeneratív probléma lehet. Majd ezt az áthangolódást kellene visszafordítani.

A kissejtes tüdőrák és az emlőrák diagnosztikájában épp proteomikai módszerekkel értek el áttörést).

Ez lesz a neuro- és a pszichoproteomika területén is.

A gyógyszergyárak már megfogalmazták, hogy a több hatóanyagú gyógyszereké a jövő, nem egyet vagy kettőt, hanem 30–40 hatóanyagot fogunk egyszerre alkalmazni.

Gyógynövények: egyszerre sokféle hatóanyag van bennük, amelyek kombinációjának hatása teljesen más, mint az egyes hatóanyagoké külön-külön.

A pszichiátriában visszajön majd a gyógyszeres és a pszichoterápia kombinációja: gyógyszerkombinációval fellazítják az agyi fehérjehálózatokat, majd verbális terápiával visszaterelik a normális kerékvágásba. A rendszerszemléletű orvoslás fog elterjedni: sok ponton, egyszerre, koordináltan fogunk beavatkozni.



## **Jövő: a biológia átformálódása:**

Az új adatok jelentős részét az emberi genom által kódolt információ adhatja, ugyanis az előrejelzések azt mutatják, hogy tíz éven belül az emberek nagyobb fele (több mint 4 milliárd ember!) rendelkezhet a saját genetikai kódjával.

Emellett persze ott lesznek a kísérleti, diagnosztikai és más fajokból származó szekvencia- és egyéb információk is.

Rengeteg új állás lesz a „nagy adat”-nak köszönhetően. Jelenlegi kérdés, hogy ezekből mennyit tudnak végül betölteni. A bioinformatikusok többsége biológusból képezte át magát, így nem váltak programozóvá, de együtt tudnak dolgozni az informatikusokkal és akár önálló informatikai fejlesztésre is képesek lehetnek.

Olyan „komplementer” szakemberekre lesz szükség akik különböző területeken jártasak, de a másik terület kérdéseit is értik, a másik terület számára érthető megoldásokhoz tud jutni.

Orvoslás forradalma:

Adattermelés felgyorsult. 2003-ig a világ 1 billio gigabyte adatot termelt összesen.

Ma 1 trillió gigabyte adatot termelünk évente.

Egyes emberek genomjának, fiziológiájának digitalizálása (DNS szekvenáló gépeknek és bioszenzoroknak és képalkotó technikáknak köszönhetően) új lehetőségeket teremt az orvoslásban.

Mobiltelefonokra töltött egészségügyi adatok lehetőséget biztosítanak a gyors diagnózisra, állapot folyamatos követésére stb.

A beteg nem egy vérnyomás, vércukor stb labor eredménnyel fordul orvoshoz, hanem a genomjával és 3 havi vérnyomás, vércukor pulzus stb értékekkel.

A genom ismeretében hatástalan vagy súlyos mellékhatással járó kezelések elkerülhetőek.

Rák genetikai kompozíciója és az eredeti genom eltéréséből rákot okozó elváltozás felderíthető, jobban kezelhető.

Idiopátiás betegségek eliminációja: Genetikai hibák feltérképezésével jobban láthatóvá válhat az eredeti probléma ami a tüneteket okozza.

Felhasznált irodalom:

<https://genomicscience.energy.gov/>

Hood, L., Rowen, L. The Human Genome Project: big science transforms biology and medicine. *Genome Med* 5, 79 (2013).

<https://doi.org/10.1186/gm483> <https://genomemedicine.biomedcentral.com/articles/10.1186/gm483>