A GENOM

Venetianer Pál

A genom fogalma
A genom megismerése
Az emberi genom
Az emberi genom egyedi különbségei

A genom fogalma

A huszadik század első felében a klasszikus örökléstan (mendeli-morgani genetika) úgy tudta, hogy az egyes örökletes tulajdonságok meghatározói, a mintegy az átöröklés atomjainak tekinthető gének (amelyek természetéről, anyagáról akkor még semmit sem lehetett tudni), a minden sejt magjában megtalálható fonalszerű kromoszómákban egymás mellett gyöngysorszerűen helyezkednek el, és többé-kevésbé változatlanok. Az örökléstan, a genetika tudománya elsősorban arra törekedett, hogy az egyes gének egymáshoz viszonyított elhelyezkedését (térképezés), a következő generációkban történő együttes vagy külön-külön való megjelenését, hatásuk érvényesülését jellemezze. A genom fogalma (Winkler 1920) ekkor az adott élőlény génjeinek összességét jelentette. A gén természetének megismerésében jelentős lépés volt annak tisztázása, hogy a gének azáltal képesek tulajdonságokat determinálni, hogy egyes enzimek (azaz fehérjék, mert minden enzim fehérje) szerkezetét határozzák meg. A genetika döntő fordulatát az jelentette, hogy kiderült: az átöröklés egyetlen jól ismert anyaghoz a dezoxiribonukleinsavhoz (DNS) köthető (Avery 1945), és ennek az anyagnak a szerkezete (Watson–Crick 1953) magyarázatot kínál arra, hogyan kettőződnek meg a gének a sejtosztódás során, és hogyan határozhatják meg a fehérjék szerkezetét. A DNS hosszú láncmolekulái ugyanis négy különböző alapegység (nukleotidok) egymáshoz kapcsolódásával épülnek fel, és e négy nukleotidféleség sorrendje az az információ, amely az egész élővilágban egységes rejtjelkulcs, a genetikai kód segítségével meghatározza minden egyes fehérjében a fehérjeláncot alkotó aminosavak sorrendjét (20 különböző aminosavból épülnek fel az átlagosan 300-400 aminosav hosszúságú fehérjeláncok). Amióta ezt tudjuk, a genom fogalma az egyes élőlények, illetve egyes sejtek összes DNS-ét jelenti, a gének pedig a DNS-lánc meghatározott, hosszabb-rövidebb szakaszait, amelyek egy-egy fehérjelánc szerkezetét határozzák meg (kódolják).

A genom megismerése

Már régen tudjuk, hogy a genom mérete nagyjából arányos az élőlények fejlettségével, komplexitásával. A legkisebb vírusoknál néhány ezer nukleotid, a baktériumoknál néhány millió, élesztőnél tíz-, muslicánál százmilliós nagyságrendű, a gerinces állatok többségénél, így az embernél is 3 milliárd nukleotid. (A fenti adatok az ún. haploid genomra vonatkoznak, azaz a magasabbrendű élőlényeknél egy kromoszóma

szerelvényre. Mint tudjuk, ilyet csak az ivarsejtekben találunk, a testi sejtek két kromoszómaszerelvényt tartalmaznak, azaz diploid genomot. Mivel a magasabbrendű élőlények mitokondriumai és a növények kloroplasztiszai is tartalmaznak DNS-t, az eddig említett nukleáris genom mellett szokás beszélni külön mitokondriális, illetve kloroplasztisz genomról is.) Az is régen ismert, hogy az élővilágban találkozunk a sorból kilógó kivételesen nagy, illetve kivételesen kis genomokkal is *(1. táblázat).*

1. táblázat. A genomméret extremitásai (a méretadatok DNS-nukleotidpárt jelentenek, illetve az MS2-fágnál RNS-nukleotidot)

	Legkisebb		**Legnagyobb**	
	Név	**Méret**	**Név**	**Méret**
Vírus	MS2-fág	3 569	Mimivírus	1 181 404
Baktérium	*Carsonella ruddii*	159 662	*Solibacter usitatus*	9 970 000
Állat	*Protylenchus coffeae* (féreg)	20 000 000	*Polychaos dubium* (amőba)	670 000 000 000
Gerinces	*Tetraodon nigroviridis* (fuguhal)	385 000 000	*Pratipterus aethiopicus* (tüdőshal)	130 000 000 000

A 20. század hetvenes éveiben dolgozták ki azokat a módszereket, amelyekkel a DNS nukleotid sorrendje meghatározható (Sanger 1977), így 1990-ben lehetővé vált a humán genom program megindítása, amely célul tűzte ki az ember 3 milliárd nukleotidnyi teljes haploid genomjának megismerését. A program 2004-ben fejeződött be, azóta viszont a DNS-szerkezet megismerése (szekvenálás) sokkal gyorsabbá és olcsóbbá vált, így ma már közel négyezer különböző faj teljes genomszerkezetét ismerjük, és ez a szám napról napra nő. Elmondhatjuk, hogy a teljes genomok vizsgálatával, összehasonlításával foglalkozó tudomány a genomika lassan elfoglalja a genetika helyét.

Az emberi genom

A klasszikus értelemben vett gének (fehérjét kódoló DNS-szakaszok) a genomnak kevesebb mint 2%-át teszik ki, és számuk kb. 20–25 000 (régebben sokkal nagyobb génszámot feltételeztek). A genom körülbelül felét olyan ismétlődő DNS-szakaszok adják, amelyek eredetileg mintegy parazitaként szállták meg a genomot, helyüket változtatni képes („ugráló”) gének voltak *(2. táblázat).*

2. táblázat. Az emberi genom szerkezeti elemeinek közelítő megoszlása

Fehérjét kódoló szekvenciák	2%
Intronok	26%
Egyéb egyedi szekvenciák	12%
Egyszerű szekvencia ismétlődések	3%
Nagyobb régiók megkettőződései	5%
Heterokromatin	8%
Mobilis ismétlődő elemek	44%

Ezeknek döntő többsége ma már stabil, és evolúciós múltunk maradványainak tekinthető. Vannak olyan rövid, ismétlődő DNS-szakaszok, amelyeknek nincs semmiféle ismert funkciójuk, de számuk nagy egyedi változatosságot mutat –ezek szolgálnak az egyedi DNS-azonosítás (DNS-ujjlenyomat-készítés) alapjául. A genom jelentős része (egyes becslések szerint 70%-a) átíródik (azaz RNS-másolat készülhet róla), noha ezek az RNS-ek nem kódolnak fehérjét. Egy, a közelmúltban felfedezett részük igen fontos, jelenleg a kutatások középpontjában álló szabályozó szerepet játszik a génműködésben, ezek az ún. kisRNS-ek, más részüknek egyelőre nem ismerjük a szerepét. Az mindenesetre biztos, hogy a fehérjét nem kódoló DNS, amelyet korábban (helytelenül) „szemét-DNS”-nek neveztek, igen fontos és részben még nem ismert szerepet játszik a sejtek működésében.

Az emberi genom egyedi különbségei

Mivel ma már tucatnyi emberi egyed teljes genomját ismerjük, megállapítható, hogy genomi szinten háromféle különbség adja az ember egyediségét. A nukleotidsorrendben (szekvenciában) mutatkozó különbség 0,1%, azaz átlagosan 1000 nukleotidonként egy különbözik minden egyénben, e különbségeket a szakmai zsargon SNP-nek (single nucleotide polymorphism ’egyes nukleotidok többalakúsága’) nevezi. Ennél nagyobb (1-3%) különbség mutatkozhat abban, hogy hosszabb DNS-szakaszok ismétlődhetnek, és az ismétlődések száma és helye egyénenként különböző. Ezek neve CNV (copy number variant ’példányszám különbség’). Végül a harmadik különbség nem a nukleotid sorrendben mutatkozik, hanem a DNS másodlagos kémiai módosulásaiban (pl. egy-egy metilcsoport jelen- vagy távolléte valamelyik nukleotidon), illetve a DNS-hez kapcsolódó fehérjék, a kromatin bizonyos módosulásaiban. Ezek neve: epigenetikus módosulások. Régebben ezeket kizárólag az egyedi élet során szerzett megváltozásnak tekintették, de a mai felfogás szerint ezeknek egy része is átörökíthető. A jelenlegi orvosi-genetikai kutatások központi kérdése, hogy az itt felsorolt tényezőknek milyen szerepük van a legfontosabb emberi betegségek kialakulásában.