

NYILATKOZAT

Név: Berdefi Asztrid

ELTE Természettudományi Kar, szak: matematika BSc

NEPTUN azonosító: rsqtbx

Szakdolgozat címe: PageRang és alkalmazása biológiai hálózatok elemzésére

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló szellemi alkotásom, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2021. május 30.



a hallgató aláírása

PageRang és alkalmazása biológiai hálózatok elemzésére

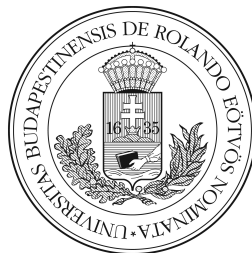
BSc szakdolgozat

Berdefi Asztrid

Témavezető: Grolmusz Vince, egyetemi tanár

Számítógéptudományi Tanszék, Matematikai Intézet

Eötvös Loránd Tudományegyetem, Természettudományi Kar



2021.

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Grolmusz Vincének, hogy megismertette velem a biológiai hálózatok témakörét, és ötleteivel, tanácsaival segítette a szakdolgozat elkészülését.

Tartalomjegyzék

Köszönetnyilvánítás	iii
Tartalomjegyzék	1
1. A PageRang algoritmus	3
2. Biológiai hálózatok	9
2.1. Fehérjék	9
2.2. Fehérjehálózatok	10
2.3. Agygráfok	11
3. A perszonalizált PageRang alkalmazása humán interakciós fehérjehálózaton	15
3.1. Az öregedés mechanizmusa: a mobilis genetikai elemek elsődleges szerepe a genom dezintegrációjában	15
3.2. Eredmények	18
4. Függelék	21
Hivatkozások	27

1. fejezet

A PageRang algoritmus

A gráfelmélet, mint tudományág az 1730-as években alakult ki. Leonhard Euler 1736-ban megjelent, a königsbergi hidak problémájával foglalkozó cikkét tekintik az első gráfelméleti tanulmánynak. A XX. század folyamán több fontos elméleti eredmény született, de a konkrét nagy gráfok vizsgálata a '90-es évek végén a WWW webgráfjának tanulmányozásával kezdődött. Ebben az időben a világháló 150 millió oldalt tartalmazott, ami kihívást jelentett a legtöbb gráfalgoritmus számára.

A PageRang algoritmust 1998-ban publikálta Larry Page és Sergey Brin. Céljuk egy hatékony keresőmotor előállítása volt, mely fontossági sorrendet állít fel a weboldalak között, így kiszűri az irreleváns találatokat. [1] Az algoritmus a következőképpen épül fel:

Tekintsünk egy irányított gráfot, melynek csúcsai a weboldalak, élei pedig a köztük vezető linkek.

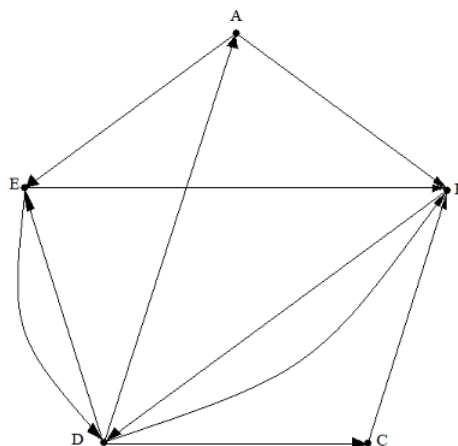
Egy oldal fontosságát először az oldal befokaként, vagyis a rámutató linkek számával határozták meg. A megközelítés hátránya, hogy nem veszi figyelembe a hivatkozások minőségét, így könnyen befolyásolható: létrehozhatunk üres oldalakat, amik növelik egy weboldal befokát, megváltoztatva ezzel a fontossági sorrendet. Tehát egy fontos oldalnak a magas befok mellett fontos szomszédokkal is

kell rendelkeznie. Ebből a meghatározásból iteratív algoritmust konstruálhatunk. Kezdetben legyen minden csúcs fontossága egységnyi. Ezt a fontosságot a csúcsok továbbadják azon szomszédaiknak, akikre mutatnak. Ezt a lépést ismételve legtöbbször stabilizálódott határeloszláshoz jutunk.

Legyen u egy weboldal. F_u jelölje az oldalakat, melyekre u -ról vezet link, B_u pedig azon oldalakat, melyekről u -ra vezet link. N_u az u -ra mutató linkek száma, vagyis u csúcs befoka. Legyen c normalizáló faktor ($c \in (0; 1)$), mely biztosítja, hogy a PageRang értékek összege konstans legyen. Ekkor a PageRang egy egyszerűsített formája:

$$R(u) = c \sum_{v \in B_u} \frac{R(v)}{N_v}$$

Példaként tekintsük az alábbi gráfot:



1.1. ábra

A fenti gráfra az egyszerűsített PageRang értékek a következőképpen számolhatók:

$$R(A) = c \left(\frac{R(D)}{4} \right)$$

$$R(B) = c \left(\frac{R(A)}{2} + \frac{R(C)}{1} + \frac{R(D)}{4} + \frac{R(E)}{2} \right)$$

$$R(C) = c \left(\frac{R(D)}{4} \right)$$

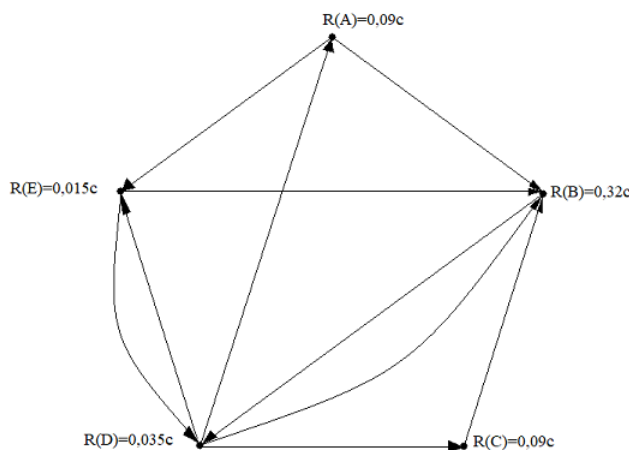
$$R(D) = c \left(\frac{R(B)}{1} + \frac{R(E)}{2} \right)$$

$$R(E) = c \left(\frac{R(A)}{2} + \frac{R(D)}{4} \right)$$

A pontos értékek az alábbi linken találhatóak.

Látható, hogy egy oldal (c -vel redukált) PageRang értéke egyenlően oszlik meg az általa mutatott oldalak között.

A PageRang számolás rekurzív, majdnem bármilyen értékről indulva iterálható, amíg nem konvergál. A fenti példán



1.2. ábra

A PageRang algoritmus egy intuitív megközelítése a Random Surfer modell. Az egyszerűsített képlet megfelel egy véletlen bolyongás valószínűségi eloszlásá-

nak a webgráfon. Tekintsünk egy random felhasználót, aki véletlenszerű oldalról indulva halad egyenletes valószínűséggel a linkeken. Egy oldal PageRang értéke megmutatja, hogy a felhasználó mekkora valószínűséggel jut el az adott oldalra.

Tekintsünk két weboldalt, amelyekből egymásra mutat link, de más oldalra nem, és egy harmadik weboldalt, amelyről a fenti kettő egyikére mutat link. Ekkor a gráfban képződött hurokban az oldalak súlya felhalmozódik, és nem adódik tovább más oldalnak. Többek között ezen probléma kiküszöbölésére vezették be a PageRang képletébe az $E(u)$ vektort, melyet a weboldalakon értelmezünk.

$$R'(u) = (1 - c) \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u)$$

Ebben az esetben a felhasználó c valószínűséggel teleportálhat a web egy véletlenszerű oldalára, ezáltal az algoritmus hibátűrése nő.

A hibátűrés azt mutatja meg, hogy mennyit változik a PageRang-vektor, ha néhány (nem olyan fontos) csúcs PageRangja változik. Az alábbi képlet szerint ha a rosszul kimért csúcsok PageRangjának összege nem túl nagy, akkor a PageRang vektorok közti különbség sem lesz nagy.

$$\|p - \hat{p}\|_1 \leq \frac{2(1 - c)}{c} \sum p_j$$

Bal oldalon a PageRang vektorok különbségének egyes normája látható (p az eredeti, \hat{p} a módosított), ami kisebb, mint a megváltozott PageRangok összege (p_j) szorozva egy c -től (a teleportálási valószínűségtől) függő konstanssal. A teleportálási valószínűség általában egy 0,2 körüli érték, vagyis a felhasználó 0,8 valószínűséggel az éleken halad, 0,2 valószínűséggel pedig teleportál. Ha c túlságosan nagy (közelít egyhez), akkor a sétánk szinte független lenne a gráftól ($c=0,99$ esetén $\|p - \hat{p}\|_1 \leq 0,02 \sum p_j$), ha pedig c nagyon kicsi (közelít nullához), akkor a teleportálás nem javítja az algoritmus hibátűrését ($c=0,01$ esetén $\|p - \hat{p}\|_1 \leq 198 \sum p_j$).

A PageRang algoritmus egy változata a perszonalizált PageRang. Ennek segítségével az egyes felhasználók személyes érdeklődését is figyelembe vehetjük a következőképpen: bevezetünk egy új konstanszt (d). A felhasználó $1 - (c + d)$ valószínűséggel a csúcsokon halad, c valószínűséggel egyenletes eloszlás szerint teleportál, d valószínűséggel pedig a kiválasztott (személyes érdeklődésének megfelelő) csúcsok valamelyikére ugrik.

A PageRang nagy hálózatokon gyorsan számolható, hibatűrő algoritmus. Olyan óriási hálózatokon, mint a WWW vagy az agygráfok, léteznek jó közelítő algoritmusok a számolására. Biológiai hálózatok vizsgálatánál különösen fontos hibatűrése, hiszen a biológiai mérések gyakran erősen laborfüggő értékeket szolgáltatnak.

2. fejezet

Biológiai hálózatok

2.1. Fehérjék

A fehérjék a sejt lineáris, aminosavakból felépülő polipeptidláncok által alkotott szerves makromolekulái. Sokféle funkciót tölthetnek be: lehetnek például transzport fehérjék, védőfehérjék, vagy szerkezeti fehérjék. A legtöbb sejtre jellemző hőmérsékleten az őket fenntartó kémiai reakciók rendkívül lassan mennének végbe. A fehérjék egyik legfontosabb feladata, hogy enzimként katalizálják a sejtet életben tartó metabolikus reakciókat.

Négy egymásra épülő szerkezeti szintjüket különböztetjük meg. A fehérjéket alkotó húszféle aminosav amino- és karboxilcsoportja peptidkötéssel kapcsolódik össze, ezek sorrendje határozza meg az elsődleges szerkezetet. A polipeptidlánc gerincét így három csoport ismétlődése alkotja, ennek térbeli elhelyezkedését nevezük másodlagos struktúrának. A bizonyos szénatomok kötési körüli rotáció lehetővé teszi a polipeptidlánc felcsavarodását, vagyis háromdimenziós fehérjeszerkezet kialakulását, ezt nevezük terciér struktúrának. A szerkezetet az egymáshoz közel kerülő aminosav-oldalláncok között létesülő kölcsönhatások stabilizálják. Negyedleges szerkezetről azon fehérjék esetén beszélhetünk, amelyek több polipeptidláncból épülnek fel.

A sejtekben lezajló komplex biológiai folyamatokért nem kisebb fehérjekomplexe-k, vagy különálló útvonalak, hanem összetett fehérjehálózatok felelősek. Együttes működésük megértéséhez olyan modellre van szükség, amely leírja a fehérjék közötti kapcsolatokat. A számítógépes kapacitás és a különböző fehérje adatbázisok fejlődése tette lehetővé a fehérjehálózatok kutatását. Egy-egy folyamat, például egy betegség, vagy egy baktérium működésének befolyásolásához ezen fehérjehálózatok fontos csúcsait kell megkeresnünk.

2.2. Fehérjehálózatok

Kétféle fehérjehálózatról beszélhetünk.

Az interakciós hálózatban a generált gráf csúcsai maguk a fehérjék. Két csúcs között akkor vezet él, ha a két fehérje egymással (bizonyos körülmények között) interakcióba lép. Az interakciókat különböző módszerekkel kimérhetik, vagy jósolhatják. Az interakciók jóslásának egy módszere a szövegbányászat. Ekkor a megjelent fehérjékkel foglalkozó cikkeken egy algoritmust futtatnak, amely élet hűz azon fehérjék közé, melyek neve egy cikkben egymáshoz közel szerepel. Ha az algoritmus eredményeit szakmailag nem bírálják felül, a jósolt interakciók jóval kevésbé megbízhatóak lehetnek a kimérteknél. Az interakciós hálózatok irányítatlan gráfok, mivel a mérések két fehérje kapcsolódását írják le, így nincs értelme irányról beszélni. A gráfpontok címkéi fehérjeazonosítók, melyek a használt adatbázistól függően lehetnek a génnevek, vagy a fehérjék UniProt azonosítója.

Az IntAct (<https://www.ebi.ac.uk/intact/>) egy nyílt, mindenki számára elérhető fehérje interakciós hálózat, melyben 22229 publikáció alapján megtalálható 121387 fehérje és 1156385 köztük vezető él. Megbízhatósága abból fakad, hogy szakértők egy csapata egyesével is gyűjti a szakirodalomból a kapcsolati adatokat. Az IntAct egy-egy interakcióról további adatokat is tartalmaz, mint a publikációs forrás, és a kimutatási módszer. Több más adatbázis adatait is tartalmazza. Az

adatbázis teljes egészében is letölthető, de a weblapon keresztül szűrhetjük is a találatokat fehérje vagy kimutatási módszer szerint.

A String (<http://string-db.org>) az egyik legnagyobb fehérje interakciós adatbázis. Több forrásból tartalmaz adatokat: kimért kapcsolatokat, valamint más adatbázisok adatait is tartalmazza, és jószolt kapcsolatokat is. Hátránya, hogy mivel sok jószolt élet is tartalmaz, használatakor oda kell figyelnünk az adatok megbízhatóságára.

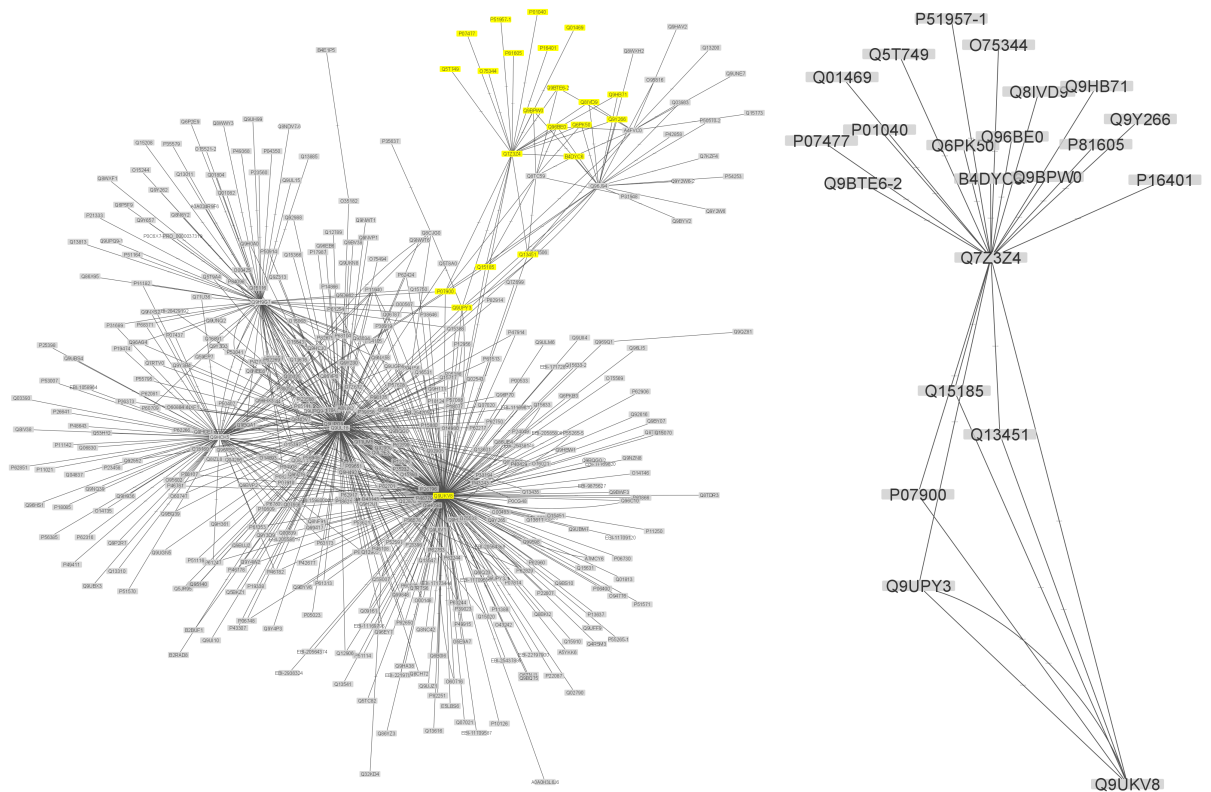
Vizsgálhatunk különböző algoritmusok által generált hálózatokat is, habár ezek általában kevésbé megbízhatóak, mint a mért adatokból felépülő hálózatok. Generált hálózat például a Nascent (<https://pitgroup.org/nascent/>), amely génneveken és szekvenciahasonlóságon alapul. Az IntAct adatbázis alapján egy forrásorganizmus megfelelő génnevű vagy hasonló szekvenciájú fehérjéit beazonosítja a célszervezetben, és a forrásnál jelenlévő interakciókat a célszervezet fehérjehálózatában is élként tünteti fel. Lehetővé teszi, hogy a gyakran vizsgált modellorganizmusokon (*C.elegans*, zebradánió) kívül más fajok fehérjehálózatáról is képet kapjunk.

A metabolikus hálózatok csúcsai biokémiai reakciók. Két csúcs között akkor vezet (irányított) él, ha az első folyamat termékét a második felhasználja. Ebben az esetben az él címkéje lesz a fehérje. Metabolikus hálózat például a KEGG (Kyoto Encyclopedia of Genes and Genomes) (<https://www.genome.jp/kegg/>)

2.3. Agygráfok

Agygráfnak, illetve konnektómnak nevezzük az agy kapcsolati hálózatát. A gráf csúcsai az egyes idegsejtek, élei pedig a köztük lévő szinapszisok.

Az emberi agyban 80 milliárd neuron található, az összes axon lefutásának feltérképezése jelenleg létező eszközökkel lehetetlen lenne, ezért az ember esetében nem vizsgálhatjuk ilyen értelemben a konnektómot. Helyette csúcsnak bizonyos



2.1. ábra. A Piwi-piRNS útvonal interakciós hálózata (Cytoscape) - a jobboldali ábrán a kiemelt részlet nagyítva látható

agyterületeket tekintünk. Diffúziós MRI (Magnetic Resonance Imaging felvételek segítségével követhető az axonok lefutása, így feltérképezhetővé válnak az agykéreg nagyobb egységei, és ezek közötti kapcsolati gráfja, vagyis konnektómja.

A Human Connectome Project keretein belül 1200 egészséges alany agyi dMRI felvételeit tartalmazó adathalmazból száz darab, egyenként ezercsúcsú agygráf készült. Ennél nagyobb élszám esetén csak lineáris algoritmusok futtathatók, ilyen paraméterek mellett viszont lehet mélyebb, lassúbb, érdekesebb paramétereket számoló algoritmusokat is futtatni.

Ilyen paraméter például a gráf expander tulajdonsága: minden (nem túl nagy) csúcshalmaz szomszédjainak elemszáma konstansszor nagyobb, mint az eredeti csúcshalmaz elemszáma. Ezt a tulajdonságot elemezve a kutatócsoport többek

között azt találta, hogy a nők agya jobb expander, mint a férfiaké és az idősebbek agya (35 év körül) jobb expander, mint a fiataloké (21 év körül).

3. fejezet

A perszonalizált PageRang alkalmazása humán interakciós fehérjehálózaton

3.1. Az öregedés mechanizmusa: a mobilis genetikai elemek elsődleges szerepe a genom dezintegrációjában

Az öregedés molekuláris mechanizmusának megértése alapvető probléma a biológiában. A szomatikus sejtekben a genom instabilitása az életkor előrehaladtával növekszik, és a kutatások azt mutatják, hogy ez a mobilis genetikai elemek, más néven transzpozabilis elemek vagy „ugráló gének” mobilizációjával jár.

Nonreplikatív transzpozíció esetén a DNS-szakasz kihal a genomból és egy másik helyen integrálódik. Replikatív transzpozíció során a DNS-szakasz megketőződik, az egyik kópia a helyén marad, a másik máshol épül be. A transzpozíció másik típusa RNS-intermedier közvetítésével zajlik. Retrotranszpozonok esetén a DNS-régió RNS-be íródik át, majd azon reverz transzkriptáz segítségével DNS szintetizálódik, ami beépül a DNS-be. A mobilis elemek integrációja többnyire véletlenszerűen történik, ezért kódoló vagy szabályozó régiókat is megszakíthatnak, inszerciós mutációkat hozhatnak létre.

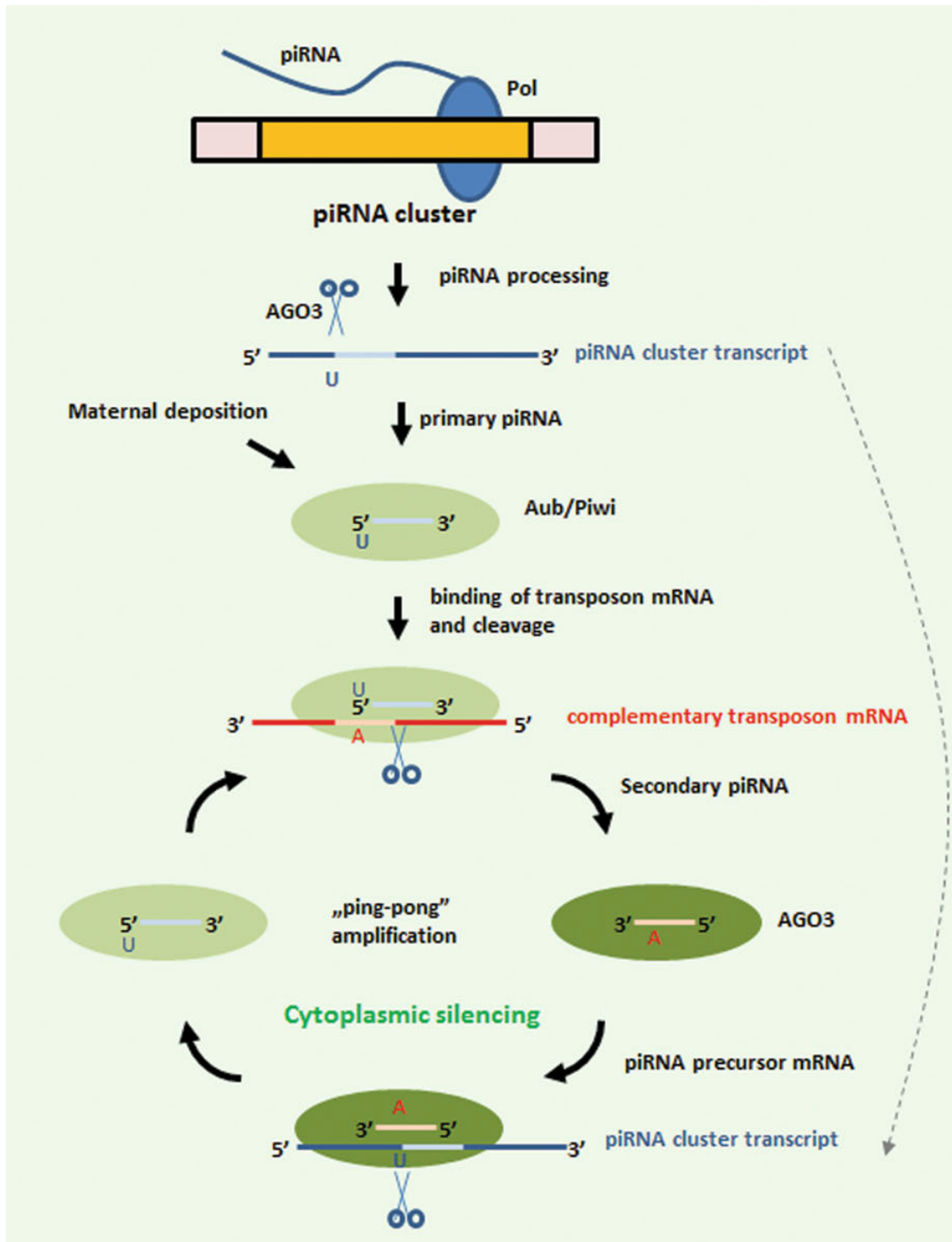
A Piwi-piRNS (P-element-induced wimpy testis-interacting RNAs) útvonal

RNS-alapú génszabályozó mechanizmus, megtalálható minden nem öregedő biológiai rendszerben, mint a csíravonalban és a rákos sejtekben. [2] Az útvonal funkciója elsősorban a transzpozabilis elemek aktivitásának repressziója. Azokban a szomatikus sejtekben, amelyekből hiányzik a Piwi-piRNS útvonal, a szervezet öregedése során a fokozatos kromatin dekondenzáció egyre inkább lehetővé teszi a transzpozabilis elemek mobilizációját.

A Piwi-piRNS útvonal egy primer piRNS-sel kezdődik, amely egy úgynevezett piRNS-klaszteren képződik. A primer piRNS-t egy endonukleáz 26-31 nukleotid hosszú darabokra hasítja, ezek már érett piRNS-ek lesznek, 5' végükön uracillal. A piRNS fragmentum kapcsolódik az Argonaute-család egy fehérjéjéhez (Aubergine/Piwi), és a komplex megköt egy a piRNS-sel komplementer transzpozon RNS-t. A transzpozon RNS-t egy endonukleáz hasítja 10 nukleotiddal egy adenin felett, melynek helye éppen megfelel az érett piRNS 5' végi uraciljának. Így szekunder sense transzpozon piRNS keletkezik, ami az AGO3 komplexhez kapcsolódva újabb piRNS klaszter transzkriptumot köt. A folyamat végére azok a transzpozon transzkriptumok, melyek bekerültek a piRNS klaszterbe, megsemmisülnek.

A piRNS útvonal túlnyomórészt a csíravonalban működik, csak egyes specifikus szomatikus sejtek (például rákos sejtek) képesek Piwi géneket és piRNS-eket expresszálni, ebből következtethetünk az útvonal szövetek regenerációjában és rákos megbetegedésekben betöltött szerepére.

A piRNS útvonalak transzkripció szinten is képesek szabályozni a transzpozabilis elemek aktivitását kromatin módosítások által, ezt RNS által közvetített epigenetikus csendesítésnek hívjuk.



3.1. ábra. A Piwi-piRNS útvonal [2]

3.2. Eredmények

A proteomikai mérések azokat a fehérjeinterakciókat képesek kimutatni, amelyek résztvevői mérhető koncentrációban vannak jelen. Ha egy fehérje koncentrációja túl kicsi, akkor változása proteomikai mérésekkel nem igazolható.

A perszonalizált PageRang algoritmussal azokat a fehérjéket keressük, amelyek az adott biológiai folyamatban igazoltan résztvevőkhöz képest szintén fontosak. Az algoritmus olyan fehérjékre irányíthatja a figyelmet, amelyek fontos, eddig nem ismert szerepet töltenek be a folyamatban. Szakdolgozatomban az IntAct humán fehérje interakciós hálózatát vizsgáltam (29916 csúcs, 590710 él). Perszonalizáltam a Piwi-piRNS útvonal fehérjéire (145 csúcs), és megkerestem a legnagyobb perszonalizált PageRangú csúcsokat. Ezután megvizsgáltam, hogy a talált fehérjék közül melyeket kapcsolja a szakirodalom az öregedés mechanizmusához.

azonosító	fehérje	gén	PageRang
Q5S007			668,5862
Q8TB24	Ras and Rab interactor 3	RIN3	125,0458
P03372			122,4627
P42858	Huntingtin	HTT	98,06092
P01106	Myc proto-oncogene protein	Myc	89,29307
P05412	Transcription factor AP-1	JUN	81,93825
P60033	CD81 antigen	CD81	78,25815
Q6FHY5	MEOX2 protein	MEOX-2	72,70676
Q6IQ23			67,83961
P62993	Growth factor receptor-bound protein 2	GRB2	66,80334

3.2. ábra. A legmagasabb PageRanggal rendelkező fehérjék (szürkével kiemelve: azok a fehérjék, amikre perszonalizálva futott az algoritmus)

A Ras fehérjék regulációs szerepet töltenek be az intracelluláris jelátviteli hálózatban. [4] Az általuk szabályozott jelátviteli útvonalak olyan folyamatokat irányítanak, mint a sejtek szaporodása és differenciálódása, vagy az apoptózis. Működési ciklusuk zavarai gyakran vezetnek fokozott sejtproliferációhoz, daganatképződés-

hez. A Ras és Rab interakciós fehérjét kapcsolatba hozták a Paget-betegséggel, mely a csontátépülés későn kezdődő krónikus zavara. [5] A fehérjét mint a betegség megjelenésének késleltetőjét tanulmányozzák.

A mutáns Huntingtin fehérje okozza a Huntington-kór nevű neurodegeneratív betegséget, ami egy úgynevezett poliglutamin-betegség. [6] [7] A fehérje génjében a CAG ismétlődések expanziója toxikus fehérjeaggregátumok képződéséhez és az idegsejtek pusztulásához vezet. A Huntington-kór motoros, kognitív és pszichiátriai tünetei gyakran csak 40 éves kor után jelentkeznek.

A Myc proto-onkogén fehérjék olyan transzkripciós faktorok, amelyek számos gén expresszióját aktiválják. [8] Közvetlen szerepük van a DNS-replikáció szabályozásában, ezáltal hozzájárulhatnak a rákos sejtek DNS-amplifikációjához.

Az AP-1 transzkripciós faktor szintén a génexpresszió szabályozásában játszik szerepet. [9] Számos folyamatban vesz részt, mint a differenciáció, proliferáció és az apoptózis. Kimutatták, hogy az AP-1 transzkripciós faktor több mechanizmus révén is részt vesz az emlőrák sejteinek növekedésében. Az AP-1 modulációt a rák megelőzésének és terápiájának lehetséges stratégiájaként tanulmányozzák.

A CD81 transzmembrán jelátviteli fehérje. [10] Génje a tumor-szuppresszor gén régióban lokalizálódik, és ezért rosszindulatú daganatok jelölt génje.

A legmagasabb PageRank értékkel rendelkező fehérjék közül a szakirodalom majdnem mindegyiket kapcsolatba hozza az öregedés mechanizmusával akár daganatos, vagy késői kialakulású betegségeken keresztül. Matematikai módszerrel sikerült tehát a fehérjék olyan sorrendjét felállítani, amelyben az elsők közt az öregedés folyamatával ismertén kapcsolatban álló fehérjék szerepelnek. Ennélfogva a módszerrel megbízhatóan található olyan új fehérjecélpontok is, amelyeket a jövőben érdemes lehet kutatni.

4. fejezet

Függelék

azonosító	PageRang
Q9UBX2	602,4201
P20963	599,744
P05067	597,2451
P04637	589,2305
P51957-1	583,671
Q8IWL3	572,2966
Q9NRI5	558,3156
O75807	554,4232
Q9Y6K9	537,6284
Q96RS6	512,1741
P61981	509,6963
P08238	503,9014
Q13618	502,4424
P54253	502,2644
Q08379	481,1083
Q07955	475,1011
P67809	473,9586
P62136	458,9098
Q99459	442,3638
Q99750	431,8909
Q93079	429,424
Q13573	425,3277
P19320	415,6242
P13569	413,7084
Q7Z699	412,8692
P63104	408,7251
O76024	394,2713
Q92905	392,1089
A8MQ03	382,5332
Q9BYV6	381,9119
P00533	378,793
EBI-25847655	371,6733
Q12933	362,0951
EBI-15183729	360,2013
P68400	357,4266
P04156	355,9513
P27824	351,8006
P10809	343,9365
P60520	340,3767
Q96HA8	337,0335
P12004	336,7651
P11142	330,6096
Q8NHP6	328,6672
Q04917	327,931

azonosító	PageRang
Q13547	324,1304
Q6A162	312,2222
Q86X19	311,6719
P11021	311,2917
P40763	310,3777
Q9UHD9	309,9762
Q00005	305,7114
P04792	305,5413
Q9H0R8	300,6117
Q9Y4K3	300,0936
Q13616	298,6514
Q9UMX0	298,5518
Q15323	297,4759
Q9NRD5	295,1893
Q96CV9	295,0229
P68431	294,8411
Q9NZ94-2	294,8153
Q5U5Q3	293,6357
P04626	292,3645
P60410	289,9947
EBI-15182334	288,5369
P49639	286,0435
Q92731	285,8232
O00560	285,575
P84022	284,169
O95166	283,8078
Q16543	283,194
Q9BYV2	283,0175
Q8TBB1	281,1609
P40337	280,4876
EBI-16419249	279,4517
P35579	279,2696
Q8IUH5	277,938
Q86VP6	277,5486
Q9H1C4	277,4259
P06241	276,3083
Q04206	275,51
P28799	275,3569
P78362	274,4972
Q15077	272,6788
P06748	271,6393
O43765	271,1371
Q96SB4	270,5221
Q9C0B5	268,9208

azonosító	PageRang
Q06787	268,4978
P08670	261,971
EBI-4399559	261,2066
P46108	260,8571
P14373	259,6876
P05067- PRO_0000000092	258,1867
Q5JR59-3	257,6502
Q16659	257,5983
Q71U36	256,5821
P16403	256,4647
O43493	255,9893
Q969Q1	254,8815
P30153	251,4711
P22607	249,0636
P01889	248,6546
Q96JN8-2	247,3786
Q92624	246,1683
P46013	245,4088
Q16778	244,2635
P62805	243,2071
P01100	242,9518
Q15796	240,2199
P49407	239,9881
P32121	237,917
Q00839	237,3005
Q9BY14	237,2158
Q13643	236,8808
P02768	235,1378
O75031	234,5147
P60409	234,2016
Q8IVT5	233,2639
P10636-8	232,6813
EBI-10171679	232,2688
P10275	232,0454
O14964	231,5128
P50222	231,0157
P27348	230,4858
P16401	230,239
P23508	230,0615
P02751	229,4289
Q14197	228,6974
Q13485	227,2464
Q6NUS6	226,2248

azonosító	PageRang
P31946	221,6298
Q93034	219,8611
Q8NHQ1	218,2884
P27635	218,1596
P11279	217,5072
Q16695	216,0612
P40692	215,2226
Q9UKV8	215,1909
P48431	214,7516
O43639	214,1879
P16333	213,5133
Q13526	212,8432
Q13077	212,5366
P49841	211,8213
Q14164	211,3051
Q04864-2	210,9616
P13473-2	210,8114
Q93062	210,5667
Q13617	210,2212
Q02539	209,1346
Q9UBB9	208,4228
Q9NWT6	208,0874
Q96EQ0	207,7083
Q13520	207,3368
P35222	207,259
P15884	206,9075
P00441	206,2774
P47755	206,14
P06396	205,0742
P41235	204,5243
P16104	204,3461
O14880	204,3012
P15884-3	203,6364
Q99523	202,8055
Q13620	202,618
Q9BRK4	202,5133
Q96FW1	202,2265
Q96RS6-3	200,518
Q12888	200,131
Q9BZS1	199,802
Q92876	199,7384
Q96GX1	197,684
P19838	196,9525
P02545	195,8963

azonosító	PageRang
Q7Z3S9	195,1976
Q0VD86	193,0926
Q9P0L0	192,6166
Q92993	192,2228
P00519	192,2131
Q9H6Z9	192,1584
Q08117-2	192,0264
P60709	191,9396
Q8WXH2	190,8645
Q09472	189,8775
Q14684	189,0495
Q15654	189,0389
P25788	188,146
Q15910	188,0447
Q9UJV3-2	187,7367
P12931	187,59
P07900	186,829
P27986	186,803
P36957	186,3732
Q04864	186,028
P36873-2	185,7375
P63165	185,2827
P98066	185,1111
P37840	184,9746
P45973	184,6239
Q9UGL1	184,4787
Q9GZQ8	183,7342
P24941	183,1533
P53350	182,7956
Q15645	182,1373
Q8N6Y0	182,0457
O95758	181,3384
Q13263-2	180,577
EBI-12690028	179,9004
Q08AM6	178,2123
O60341	178,1512
P36406	178,0823
O76011	177,6968
Q7L5N1	177,6168
P63167	176,5608
Q9NUH8	176,0397
Q9H492	175,4669
Q9GZM8	175,2181
P62826	174,0788

azonosító	PageRang
O95273	173,9294
O14901	173,8469
Q9NPJ6	173,6905
P26371	173,3047
Q9NYB9-2	173,2167
Q14192	173,059
P60411	172,7943
Q9Y3E0	172,6987
Q13148	172,6143
P62140	172,511
P10412	171,8594
Q15149	171,705
Q6RFH8	171,5833
Q12959	171,4176
Q15051	171,3732
F5H1C8	170,6507
Q9H8W4	170,54
Q6UY14-3	170,1137
P62807	169,7165
O00213	169,5402
Q01844	169,3639
Q86VM9	168,8129
Q15118	168,5261
Q8N9N5-2	168,3682
P48443	168,2147
P08047	167,9502
Q8IZU0	167,5879
P17813	167,09
A6NEM1	167,0283
Q13509	166,4846
Q9P0N5	166,0007
Q96BA8	165,8341
P31946-2	165,5359
P57078	164,7351
O15155	164,2639
Q9Y371	164,1752
P63208	163,5713
P19784	162,1486
P60953	161,9563
P46940	161,5996
O75208	161,2108
P62258	160,9349
Q5JX71	160,1527
Q8TAP4-4	159,9896

azonosító	PageRang
Q562R1	159,8013
Q15287	159,7793
P0DPK4	159,7683
Q9NR12	159,3273
Q00987	158,704
P55212	158,4253
P38398	158,2271
Q16665	158,1118
P21333	157,9907
P67775	157,8261
Q8IV20	157,5503
Q15038	157,301
P10636	157,1871
P56945	157,0319
Q9NRR5	156,6987
Q15714	156,1133
Q07627	156,0385
P51114	155,2421
P42224	155,2304
P07910	154,8194
Q9Y282	154,6551
P49368	154,6498
Q9UKE5	154,6269
P78527	154,474
P55072	154,4
Q92997	153,5653
P14136	153,5552
Q8NA72	153,191
Q13137	152,9874
Q12800	152,7372
Q9UI95	152,6102
Q14103	152,4032
P46379	152,2328
P15311	152,1125
O95292	152,0653
P52298	151,9542
Q96EB6	151,8367
Q16082	151,7797
Q14160	151,5997
Q93009	151,5491
Q9HCE5	151,4645
Q969G2	151,3185
Q5T7N2	151,0185
Q3LI66	150,2346

azonosító	PageRang
Q9HC52	150,074
Q8TAP6	149,6981
P09874	149,6128
P07550	149,5162
Q9Y3C5	149,2121
P05787	149,1009
Q93062-3	148,8247
P31749	148,6402
P09622	148,379
P15336	148,3761
P63279	148,2862
O60333-2	148,2466
Q9H0I2	148,2062
P31947	147,9178
Q86U44	147,6375
P33993	147,5982
P62714	147,0409
Q16643	146,8661
P22087	146,7826
Q96D03	146,5978
P11308	146,5184
Q9H8Y8	146,0017
Q53F19	145,7707
P25054	145,5515
P30101	145,1038
Q5JR59	145,0839
P04406	145,0351
Q14457	144,6316
O95684	144,5412
O43741	144,3537
Q92734	144,1428
Q8IXL6	144,1174
Q8TDT2	144,0398
P14625	143,9568
P11912	143,4904
Q13418	143,371
Q15125	142,4329
Q8ND90	142,3849
P61964	141,9954
Q09161	141,8654
Q96RS6-2	141,8277
Q63ZY3	141,777
Q96GG9	141,3161
Q96N21	140,9372

azonosító	PageRang
O95831	140,3699
Q15428	140,3471
O60260-5	140,3434
P01112	139,989
Q9UKT9	139,5152
Q15843	139,4987
O14965	139,3451
P25963	139,1773
P78371	139,1758
P19793	139,158
Q8IYT4	139,0753
Q9Y5V3	138,9935
P83105	138,6708
Q14108	138,5662
P43360	138,1745
O75376	138,0419
P62158	137,6353
P20339	137,1214
Q9BTE6-2	137,0862
Q86YD7	136,485
P49336	136,1358
Q92569	135,8393
P21860	135,7356
P38936	135,704
P62753	135,5575
P04183	135,4609
P63000	135,4334
Q8WV44	135,1764
P19338	135,131
Q99832	134,9063
EBI-14348029	134,8304
Q9UMX0-2	134,7811
Q9NY72	134,3784
P40227	134,3459
O00214	133,8194
POCG47	133,383
P61088	133,0765
P32242	132,7791
O60506	132,6973
Q14145	132,4533
P11802	132,3203
P25786	132,2756
O43889-2	132,2349
Q07021	132,2197

azonosító	PageRang
Q12824	132,149
Q96BE0	132,0764
P26641	132,0722
Q8TED1	132,0508
G5E9A7	131,9964
EBI-25865403	131,5186
Q9UJX2	131,1824
Q9BVG8-5	130,8821
Q8N720	130,818
EBI-22310722	130,522
P21709	130,4731
Q96H55	130,2953
Q9BYF1	130,1218
Q13557-8	130,0178
Q99558	129,8964
Q9NZD8	129,8432
P13010	129,6844
P41182	129,2333
EBI-22311199	129,1402
Q14203	128,9319
Q00610	128,7348
O76003	128,6406
P54252	128,6079
Q92673	128,5983
P40855	128,5682
Q7Z3Y8	128,3121
Q8IUQ4	127,9979
Q9BUZ4	127,7599
Q9Y478	127,5507
P43243	127,3408
P08473	127,3017
Q6ZU80	127,1572
Q13011	127,1461
P27797	127,0489
Q96GM5	126,9316
Q9UL18	126,0779
P09429	126,042
P25791-3	126,0039
O95429	125,8314
Q9BWG6	125,8035
O14744	125,7261
Q96LK0	125,705
Q13895	125,5689

Irodalomjegyzék

- [1] The anatomy of a large-scale hypertextual Web search engine, Brin, Page, 1998.
- [2] The mechanism of ageing: primary role of transposable elements in genome disintegration, Sturm, Ivics, Vellai, 2015.
- [3] Molekuláris sejtbiológia, Szeberényi, 2014.
- [4] Ras GTPase https://en.wikipedia.org/wiki/Ras_GTPase
- [5] Genetic Variation in RIN3 in the Belgian Population Supports Its Involvement in the Pathogenesis of Paget's Disease of Bone and Modifies the Age of Onset, Ridder et al., 2019.
- [6] Huntingtin <https://en.wikipedia.org/wiki/Huntingtin>
- [7] A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group, 1993.
- [8] Myc <https://en.wikipedia.org/wiki/Myc>
- [9] AP-1 transcription factor https://en.wikipedia.org/wiki/AP-1_transcription_factor
- [10] CD-81 <https://en.wikipedia.org/wiki/CD81>