



ACADEMIA ROMÂNĂ
ȘCOALA DE STUDII AVANSATE A ACADEMIEI ROMÂNE
INSTITUTUL DE BIOCHIMIE

REZUMATUL TEZEI DE DOCTORAT

Analiza bioinformatică a rețelelor farmaco-genomice ca metodă de a studia modularea duratei de viață

Doctorand:
Gabriela Bunu

Conducător de doctorat:
Dr. Andrei-José Petrescu

Bucharest

2023

Cuprins

1. Introducere generală.....	2
2. Rezultate și discuții.....	3
2.1. Definiții propuse pentru sinergism.....	3
2.2. Baza de date SynergyAge.....	8
2.3. Modele de rețele în longevitate.....	12
2.4. Meta-analiză a semnăturii de îmbătrânire.....	15
2.5. Boli și procese asociate vârstei.....	16
2.5.1. Profilul transcripțional comun în AD-Cluster și PD-Cluster.....	17
2.5.2. Fibroza pulmonară.....	19
2.6. Repoziționarea medicamentelor.....	21
2.6.1. Medicamente cu potențial de a inversa schimbarea profilului transcripțional asociat cu bolile neurodegenerative.....	22
2.6.2. Molecule mici pentru reprogramarea celulară.....	23
2.6.3. Medicamente care induc pluripotenta ca agenți potențiali împotriva patologiilor neurodegenerative.....	26
3. Concluzii finale.....	28
Contribuții personale.....	29
Bibliografie.....	31

1. Introducere generală

Îmbătrânirea este un proces complex și multifactorial care determină odată cu vârsta pierderea viabilității și creșterea fragilității. Este, de asemenea, factorul principal de risc pentru multe boli asociate vârstei, cum ar fi boala Alzheimer, boala Parkinson, cancerul și bolile cardiovasculare. Pe măsură ce distribuția populației se schimbă și proporția persoanelor în vârstă în societate crește, costul ridicat al asistenței medicale la bătrânețe poate deveni o povară pentru sistemul de sănătate.

Când ne referim la procesul de îmbătrânire, există două concepte importante: i) durata de viață a unei persoane sau animal (lifespan), și ii) durata de timp pe care un organism o petrece într-o stare sănătoasă (healthspan). În ultimele decenii, speranța de viață umană a crescut semnificativ, în timp ce durata de viață sănătoasă încă nu a avut aceeași amploare a schimbării (Ge et al., 2022). În general, cercetarea privind îmbătrânirea urmărește creșterea duratei de viață sănătoasă, pentru a reduce semnificativ apariția bolilor asociate vârstei, cu scopul ideal de a identifica tratamente sau terapii care să asigure o viață mai lungă, dar și mai sănătoasă la om (DeVito et al., 2022). Cu toate acestea, paradigma actuală de tratament atunci când se face referire la bolile asociate vârstei este încă de a lua în considerare fiecare boală separat, fără prea multă atenție pe una dintre cauzele principale, care ar fi îmbătrânirea în sine (Blagosklonny, 2012).

Una dintre provocările majore în înțelegerea îmbătrânirii este identificarea mecanismelor moleculare și celulare ce stau la baza ei, o direcție științifică în care s-au făcut mari progrese în ultimii ani. De exemplu, se știe că o multitudine de gene sunt implicate în determinarea duratei de viață, în procesele de îmbătrânire, precum și în bolile asociate vârstei. În mod similar, la nivel epigenetic, au fost descoperite tot mai multe mecanisme de reglare. Cu toate acestea, coroborarea tuturor acestor date pentru a modela pe deplin procesul de îmbătrânire sau relația sa cu diferite patologii este încă un proces în desfășurare. Totuși, progresele recente în bioinformatică au oferit noi oportunități de a studia îmbătrânirea și bolile asociate vârstei la nivel de sistem (Xue et al., 2007; Fernandes et al., 2016; Butler et al., 2008; Tacutu et al., 2011). Printre aceste metodologii noi, analizele bioinformatiche ale rețelelor mixte de gene și medicamente pot oferi câteva noi perspective unice asupra proceselor biologice care stau la baza îmbătrânirii și a bolilor legate de vârstă (Fuentelba et al., 2019; Xie et al., 2022). Poate mai important, aceste analize permit integrarea datelor multi-modale și identificarea de noi ținte și căi de semnalizare, prin relațiile funcționale dintre gene și medicamente, direcționând astfel cercetarea către tratamente și terapii noi.

Această lucrare își propune să exploreze astfel de analize bioinformatică, pentru a identifica: i) semnaturile genetice și transcripționale ale îmbătrânirii și bolilor asociate acesteia, ii) ținte potențiale pentru intervenții terapeutice care pot întârzia apariția sau progresia bolilor asociate vârstei și iii) potențiale medicamente aprobate care pot fi repositionate ca modulatori ai longevității. Această cercetare va contribui la înțelegerea proceselor biologice complexe care stau la baza îmbătrânirii și își propune să aibă implicații importante pentru repositionarea medicamentelor în bolile asociate vârstei.

2. Rezultate și discuții

2.1. Definiții propuse pentru sinergism

Multe studii intervenționale au arătat că îmbătrânirea este un proces care poate fi modulat genetic. În ciuda faptului că se cunosc până acum sute de intervenții asupra câte unei singure gene, analiza intervențiilor individuale ar putea să nu ofere o viziune suficient de largă asupra mecanismelor de reglare a procesului de îmbătrânire. Genele nu interacționează într-un mod simplu, liniar, fiind astfel dificil de prezis rezultatul combinării a două sau mai multe intervenții. Din cauza interacțiunilor complexe dintre două sau mai multe intervenții genetice, pentru început, am propus noi definiții cuprinzătoare pentru evaluarea rezultatului acestora.

Pentru a evalua efectul unei combinații de două intervenții, trebuie determinate patru valori ale duratei de viață, pentru următoarele tulpini: tipul sălbatic (wild type, WT), mutații simpli G1 și G2, precum și mutații dubli G1;G2. Efectul fiecărui mutant este definit ca procentul de modificare a duratei de viață (LS) în comparație cu durata de viață a tipului sălbatic, după cum urmează:

$$\Delta G1 = (LS(G1) - LS(WT)) * 100 / LS(WT)$$

$$\Delta G2 = (LS(G2) - LS(WT)) * 100 / LS(WT)$$

În continuare, efectul unei combinații de intervenții cu aceeași direcționalitate ($\Delta G1 * \Delta G2 > 0$) a fost definit după cum urmează (Fig. 1 și Fig. 2):

- complet sinergic, dacă $|\Delta(G1,G2)| > |\Delta G1| + |\Delta G2|$ (Fig. 2A, B)
- aditiv, dacă $|\Delta(G1,G2)| \cong |\Delta G1| + |\Delta G2|$ (Fig. 2C, K)
- aproape aditiv, dacă $|\Delta(G1,G2)| > \max(|\Delta G1|, |\Delta G2|)$ și $|\Delta G1,G2| < |\Delta G1| + |\Delta G2|$ (Fig. 2D, E)

- antagonist, dependent, dacă $|\Delta G1,G2| < \max(|\Delta G1|, |\Delta G2|)$ și $|\Delta G1,G2| > \min(|\Delta G1|, |\Delta G2|)$ (Fig. 2H)
- complet antagonist, dacă $|\Delta G1,G2| < \min(|\Delta G1|, |\Delta G2|)$ (Fig. 2F, G)

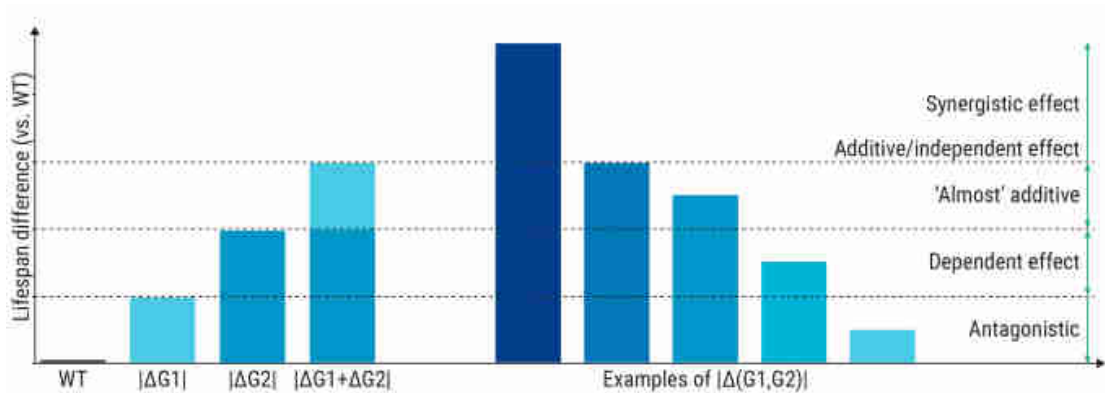


Figura 1. Tipuri de interacțiuni pentru un mutant dublu pentru care intervențiile asupra fiecărei gene sunt în același sens. Barele din stânga reprezintă valoarea absolută (modulul) a efectului asupra duratei de viață a fiecărei intervenții, în comparație cu tipul sălbatic. Suma efectelor este prezentată pentru referință. Barele din dreapta reprezintă exemple ale fiecărui tip de combinație: sinergică, aditivă, aproape aditivă, dependentă și antagonistă. Figură adaptată din (Bunu et al., 2020), utilizată sub licență CC BY 4.0.

Monotonie

Având în vedere că, în general, nu sunt evaluați toți mutații posibili într-un studiu asupra duratei de viață, în locul sinergie totale poate fi evaluată monotonia bazată pe datele existente. Acest tip de analiză implică identificarea tuturor căilor de la tipul sălbatic la n-mutant ($n \geq 2$), unde o cale reprezintă o secvență de $n+1$ tulpini, pornind de la tipul sălbatic și adăugând, la fiecare pas k , mutantul construit din mutantul $(k-1)$; de ex. o cale ar putea fi: tip sălbatic $\rightarrow G1 \rightarrow G1;G2 \rightarrow G1;G2;G3$. Dacă direcția schimbării este consecventă pentru fiecare mutant dintr-o cale, i.e. fiecare k -mutant crește durata de viață a mutantului $(k-1)$, sau fiecare k -mutant scade durata de viață a mutantului $(k-1)$, atunci calea este considerată monotonă. Dacă toate căile identificate sunt monotone (conform tuturor observațiilor experimentale incluse într-un studiu), atunci n-mutantul are o interacțiune epistatică monotonă (Fig. 2M,N). În caz contrar, dacă cel puțin o cale nu îndeplinește criteriile de monotonie, atunci n-mutantul conține dependențe între cele n componente ale sale (Figura 2L).

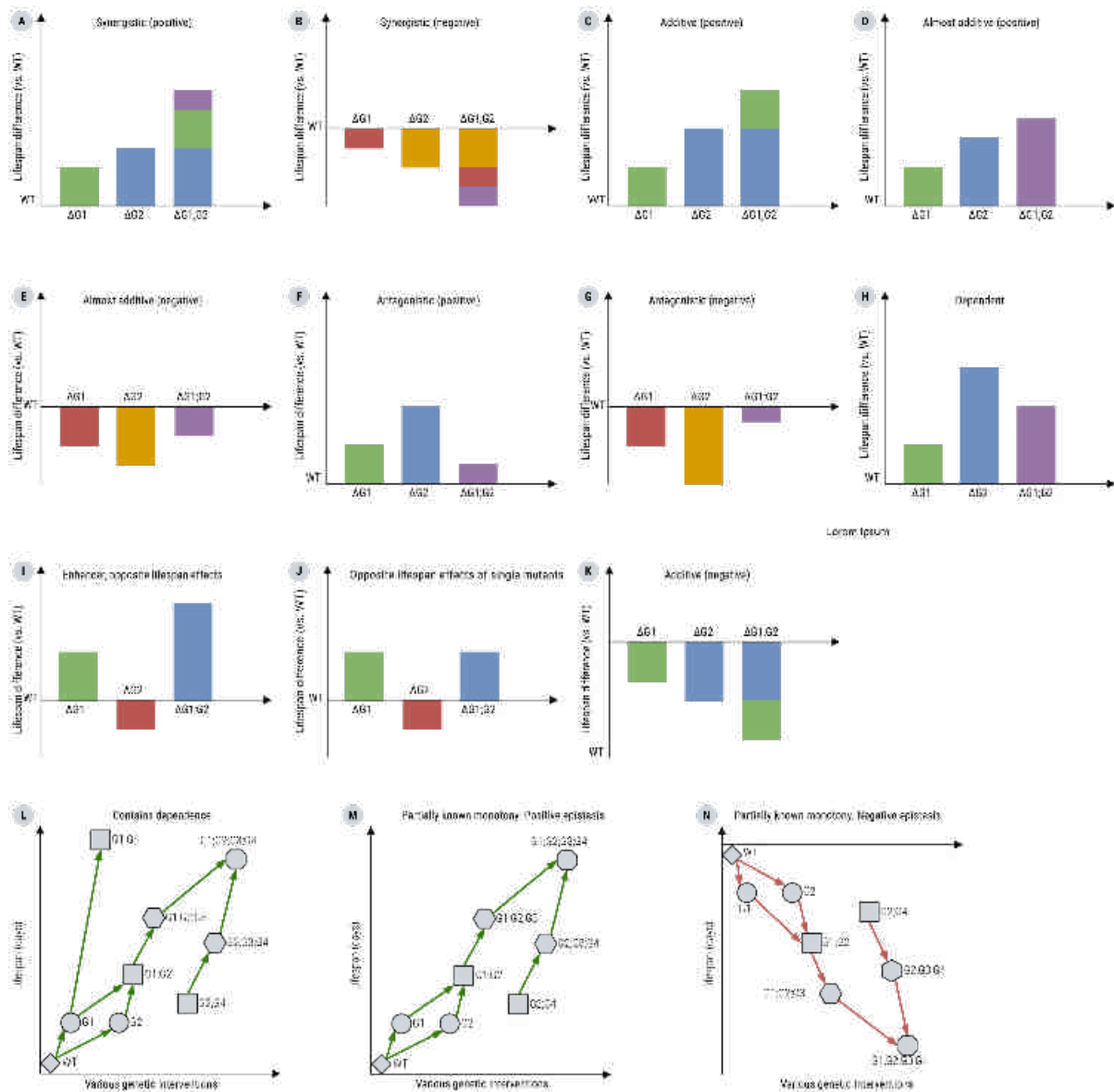


Figura 2. Reprezentarea grafică a tuturor categoriilor de efecte între două sau mai multe intervenții genetice evaluate în SynergyAge. **A-J.** Barele reprezintă efectele fiecărei intervenții ($\Delta G1$ și $\Delta G2$) și efectul combinației ($\Delta G1;G2$), ca procent în comparație cu tipul sălbatic. **A.** Exemplu de synergism pozitiv - efectul combinației este mai mare decât suma efectelor individuale. **B.** Exemplu de synergism negativ - valoarea absolută a efectului combinației este mai mare decât suma valorilor absolute ale efectelor individuale. **C.** Exemplu de efect aditiv pozitiv - efectul combinației este aproximativ egal cu suma efectelor individuale. **D.** Exemplu de efect pozitiv aproape aditiv - efectul combinației este mai mare decât oricare dintre cele două efecte individuale, dar mai mic decât suma lor. **E.** Exemplu de efect negativ aproape aditiv - valoarea absolută a efectului combinației este mai mare decât oricare dintre cele două valori absolute ale efectelor individuale, dar mai mică decât suma lor. **F.** Exemplu de efect antagonist pozitiv - efectul combinației este mai mic decât oricare dintre

cele două efecte individuale. **G.** Exemplu de efect antagonist negativ - valoarea absolută a efectului combinației este mai mică decât oricare dintre cele două valori absolute ale efectelor individuale. **H.** Exemplu de relație dependentă - efectul mutantului dublu (G1,G2) este mai mare decât efectul lui G1, dar mai mic decât efectul lui G2. **I.** Exemplu de amplificator (G2) cu efect opus asupra duratei de viață - efectul combinației este mai mare decât efectul unui singur mutant G1, chiar dacă G2 are un efect negativ. **J.** Exemplu de combinare cu efecte opuse pentru intervențiile individuale. **K.** Exemplu de efect aditiv negativ - valoarea absolută a efectului combinației este aproximativ egală cu suma valorilor absolute ale efectelor individuale. **L.** Exemplu de dependență care conține un n-mutant - G1;G4 are o viață mai lungă decât mutantul G1;G2;G3;G4. **M.** Exemplu de monotonie parțial cunoscută (pozitivă). **N.** Exemplu de monotonie parțial cunoscută (negativă).

Sinergism pentru n-mutant

Efectele cumulate ale interacțiunilor cu n componente sunt mai complexe și dificil de definit pentru $n \geq 3$. Cea mai comună definiție pentru synergism în acest caz este „mai mare decât suma părților sale”, adică $\Delta|G1,G2,...Gn| > |\Delta G1| + |\Delta G2| + \dots + |\Delta Gn|$ (Fig. 3B), totuși, această definiție ar putea să nu fie adecvată pentru a fi utilizată pentru interacțiuni genetice complexe. Fig. 3C arată un exemplu ipotetic despre cum ar putea fi înșelătoare: intervențiile individuale sunt în acord cu ipoteza menționată mai sus de „mai mare decât suma părților sale”, dar există doi mutați dubli (G1;G2 și G1;G3) cu o durată de viață mai lungă decât mutantul triplu G1;G2;G3. În acest caz, putem afirma că intervențiile G1 și G2 sunt synergice, dar această creștere a duratei de viață depinde de gena G3, care este necesar să fie neafectată (fără mutații, silențiere, supraexprimare etc.). De fapt, acesta este un caz de interacțiune dependentă, conform definițiilor propuse în această lucrare (vezi mai sus).

Pe de altă parte, pentru un n-mutant G1;G2;G3;...Gn, un synergism „complet” ar însemna că toate perechile (G1;G2...;Gk și Gk+1;Gk+2...;Gn) sunt ele însele synergice, i.e. $\Delta G1;G2;G3...Gn; > \Delta G1;G2...;Gk + \Delta Gk+1;Gk+2...;Gn$. (Fig. 3A). Pentru a evalua acest tip de synergism, trebuie determinate toate cele n! (n factorial) valori ale duratei de viață, ceea ce în practică ar reprezenta un efort intens de laborator și este adesea în afara domeniului de interes al studiilor privind durata de viață.

Din cauza lipsei datelor experimentale, am propus un alt tip de synergism: synergismul monoton parțial cunoscut (Fig. 2M,N, Fig. 3D). În acest caz, este posibil ca una sau mai multe căi să fie cunoscute, de exemplu: 1) WT -> G1 -> G1;G2 -> G1;G2;G3 și 2)

WT → G2 → G2;G3 → G1;G2;G3. O secvență de acest tip, în care la fiecare „pas” s-a observat o durată de viață mai mare decât pentru mutantul anterior, a fost definită ca fiind o cale monotonă (vezi Monotonie). Dacă toate căile investigate au această proprietate, atunci efectul celor n intervenții reprezintă un synergism monoton parțial cunoscut (Fig. 3D). În caz contrar, interacțiunea conține relații de dependență (Fig. 3E).

Antagonismul pentru un n-mutant a fost definit într-o manieră similară și implică „antagonism complet”, antagonism simplu sau antagonism monoton parțial cunoscut.

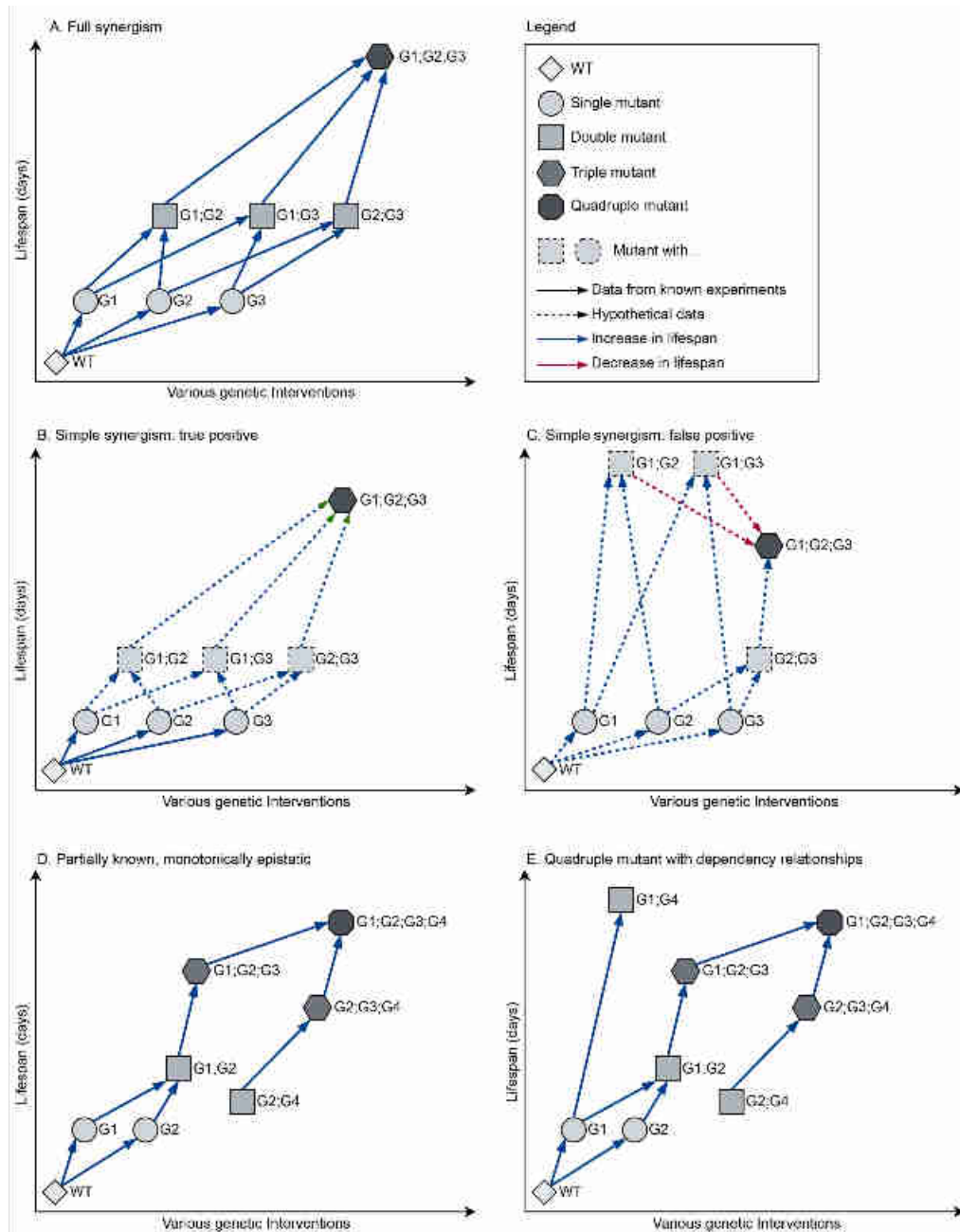


Figura 3. Reprezentarea diferitelor tipuri de synergism pentru n-mutanți cu cel puțin două intervenții. **A.** Exemplu de synergism complet între trei intervenții. **B.** Exemplu de synergism pozitiv în cazul evaluării synergismului simplu pentru trei intervenții. **C.** Exemplu de fals

pozitiv în cazul evaluării sinergismului simplu pentru trei intervenții: îndeplinește criteriile pentru sinergism simplu, dar conține o relație de dependență (i.e. creșterea duratei de viață obținută prin intervențiile G1;G2 sau G1;G3 depinde de faptul că G3 și respectiv G2 sunt nemodificate - în caz contrar efectul scade). **D.** Exemplu de sinergism monoton parțial cunoscut, între patru intervenții. **E.** Exemplu de mutant cvadruplu cu relații de dependență. Figură adaptată din (Bunu et al., 2020), utilizată sub licență CC BY 4.0.

2.2. Baza de date SynergyAge

Baza de date SynergyAge are scopul de a găzdui studii care evidențiază fie mutații longevivi, fie mutații cu viață scurtă, cu cel puțin două intervenții genetice concomitente și a fost construită cu date selectate manual din literatura științifică.

Baza de date SynergyAge conține aproximativ 7.000 de valori ale duratei de viață pentru trei organisme model: *C. elegans*, *D. melanogaster* și *M. musculus* (Tabelul 1). Cea mai mare parte a datelor disponibile au fost colectate din experimente cu nematode, *C. elegans* fiind modelul cel mai studiat pentru longevitate, datorită duratei de viață scurte și cerințelor privind condițiile de mediu.

Tabelul 1. Numărul de valori înregistrare în baza de date SynergyAge

	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>M. musculus</i>	Total
Durata de viață	6656	185	147	6988
Combinatii genetice	1770	27	3	1833
Gene cu cel puțin o înregistrare	714	36	44	794
Articole	144	20	29	193

Interfața web oferă o vizualizare a rețelei pentru fiecare organism model, precum și opțiuni multiple de căutare și filtrare (Fig. 4).

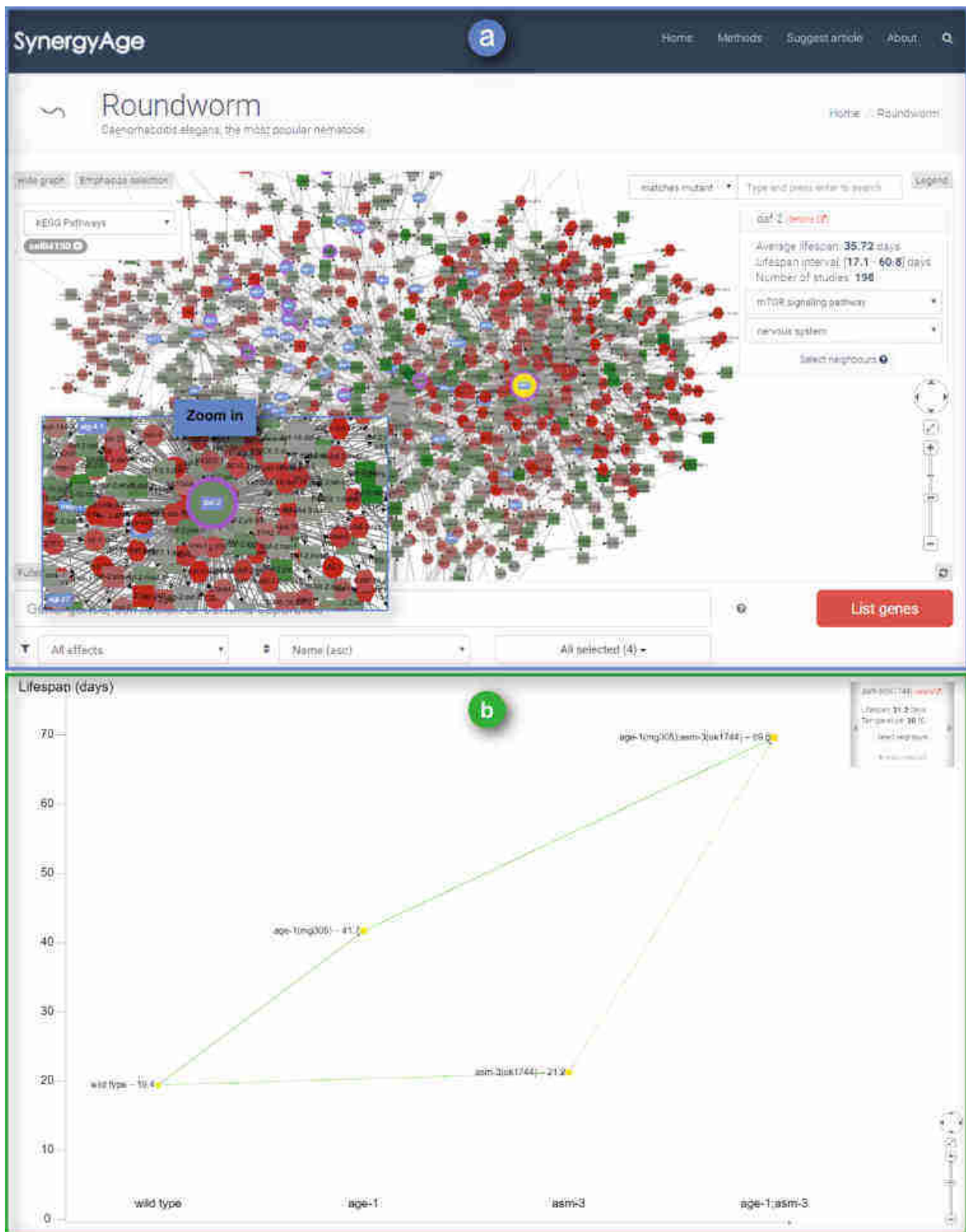


Figura 4. Vizualizări sub formă de rețea în site-ul web SynergyAge. (a) Vizualizare sub formă de rețea a mutațiilor *C. elegans*. Fiecare nod reprezintă un mutant (sau tipul sălbatic) și sumarizează toate înregistrările sale din baza de date (vezi panoul din dreapta sus). Gradientul de culoare codifică efectul său mediu în comparație cu tipul sălbatic - verde pentru efecte pozitive, roșu pentru efecte negative și gri pentru efecte de intensitate scăzută.

Dimensiunea nodului este proporțională cu numărul de înregistrări din baza de date. Muchiile sunt trasate de la un n-mutant la un (n+1)-mutant: liniile continue sunt folosite pentru comparații directe, care au fost studiate experimental, iar liniile întrerupte pentru comparațiile deduse din date din studii diferite. În acest exemplu, chenarul violet corespunde genelor care aparțin căii selectate (cel04150, panoul din stânga sus) (b) Vizualizarea rețelei a mutantului dublu *age-1;asm-3*. Fiecare nod reprezintă o cohortă; muchiile au fost trasate numai între cohortele studiate în același experiment. Muchiile verzi reprezintă o creștere a duratei de viață obținute prin intervenția suplimentară, în timp ce muchiile roșii descriu intervenții care determină un fenotip cu viață mai scurtă. Figură preluată din (Bunu et al., 2020), utilizată sub licență CC BY 4.0.

În continuare, am folosit datele din SynergyAge pentru a propune o combinație de intervenții pentru validarea experimentală. Au fost folosite datele colectate pentru a selecta gene care sunt cel puțin aproape aditive cu *daf-2*, pentru a crește probabilitatea de a identifica combinații de gene care sunt independente de calea insulinei/*daf-2*. Pentru toate perechile de gene din lista inițială, am: a) determinat numărul de căi de semnalizare KEGG comune dintre ele și b) identificat dacă aparțin aceluiași cluster în reprezentările furnizate de un model de învățare automată dezvoltat în laboratorul nostru.

Lista inițială de predicții a inclus 25 de gene și 9 combinații potențiale, fiecare având câte trei gene. După o evaluare manuală suplimentară și cercetare în literatură, am selectat 3 combinații care au fost testate într-un experiment pilot de către colegii mei din laborator. După experimentul pilot, combinația *odr-3;ife-2;cku-70* a fost selectată pentru testare ulterioară, în experimente mai riguroase.

Cele trei gene selectate pentru validare experimentală au fost testate în experimente de longevitate în nematodul *C. elegans*, în toate combinațiile posibile, iar efectul lor a fost comparat cu intervențiile asupra unei singure gene și cu tipul sălbatic. În conformitate cu datele anterioare, mutația *odr-3* și interferența ARN asupra *ife-2* au crescut durata medie de viață în comparație cu tipul sălbatic cu 26,2% și, respectiv, 18% (Fig. 5, Fig. 6). Creșterea duratei medii de viață a mutantilor *odr-3(n1605);ife-2(ARNi)* față de tipul sălbatic a fost de 40,3%, ceea ce reprezintă un efect aditiv, deoarece efectul intervenției duble este apropiat de suma celor două efecte individuale (44,2%). În cazul ARNi cu un amestec 1:1 de bacterii vector (EV) și clonă *ife-2* ARNi, combinația a fost sinergică: creșterea duratei medii de viață a fost de 46,1%, iar suma celor două efecte individuale a fost de doar 37,4%. Interferența ARN în *cku-70* nu a prelungit durata de viață a mutantului dublu.

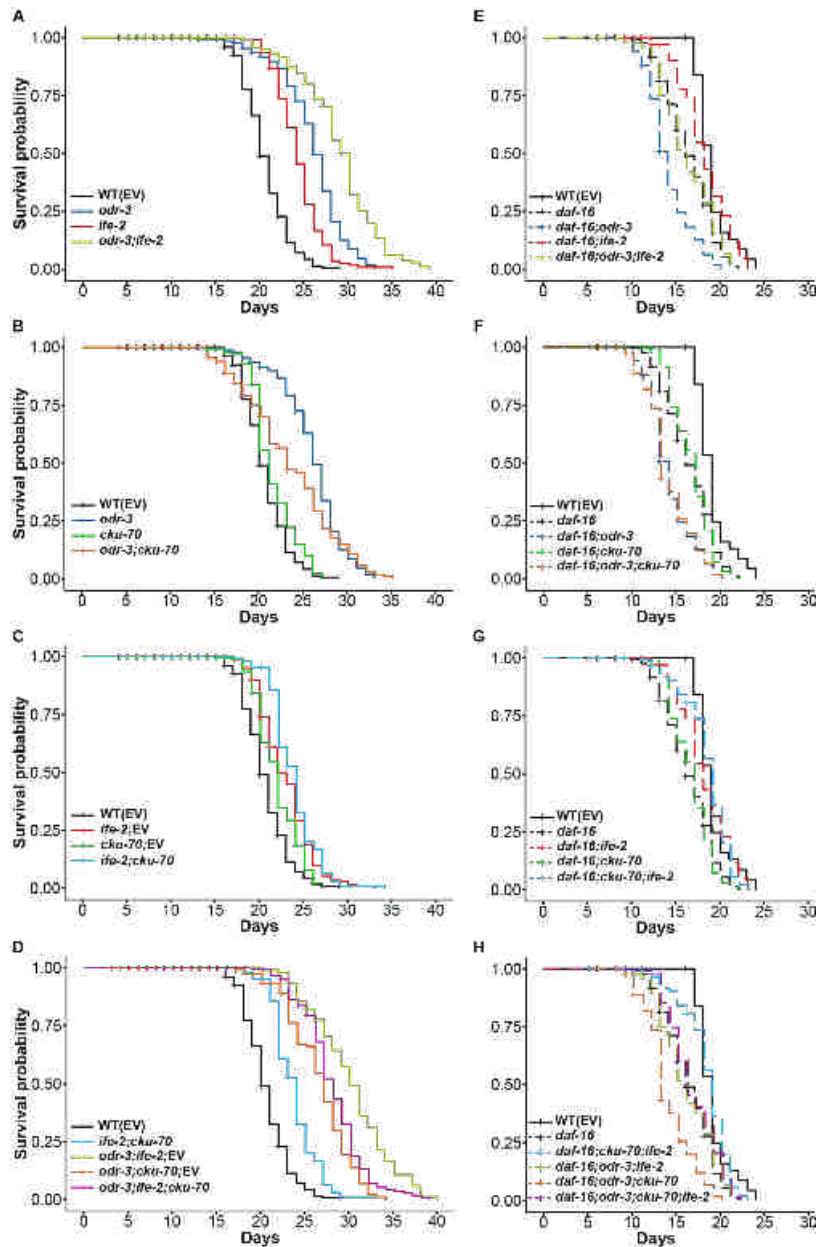


Figura 5. Curbele de supraviețuire Kaplan-Meier care descriu efectele intervențiilor genetice asupra *odr-3*, *ife-2* și *cku-70*. A-D. Efectele intervențiilor simple, duble și triple în comparație cu tipul sălbatic. E-H. Efectele intervențiilor simple, duble și triple în mutantul *daf-16(m26)*. În cazul interferenței ARN unice, controlul este tratat cu un amestec 1:1 de bacterii ARNi și EV. Au fost efectuate trei experimente independente în tipul sălbatic și două experimente independente în *daf-16*. Graficele prezintă populațiile reunite din aceste experimente. *odr-3* reprezintă mutantul *odr-3(n1605)*, iar *daf-16* reprezintă mutantul *daf-16(m26)*; ambii mutații au fost hrăniți cu EV. Figură preluată din (Matei et al., 2021), utilizată sub licență CC BY 3.0.

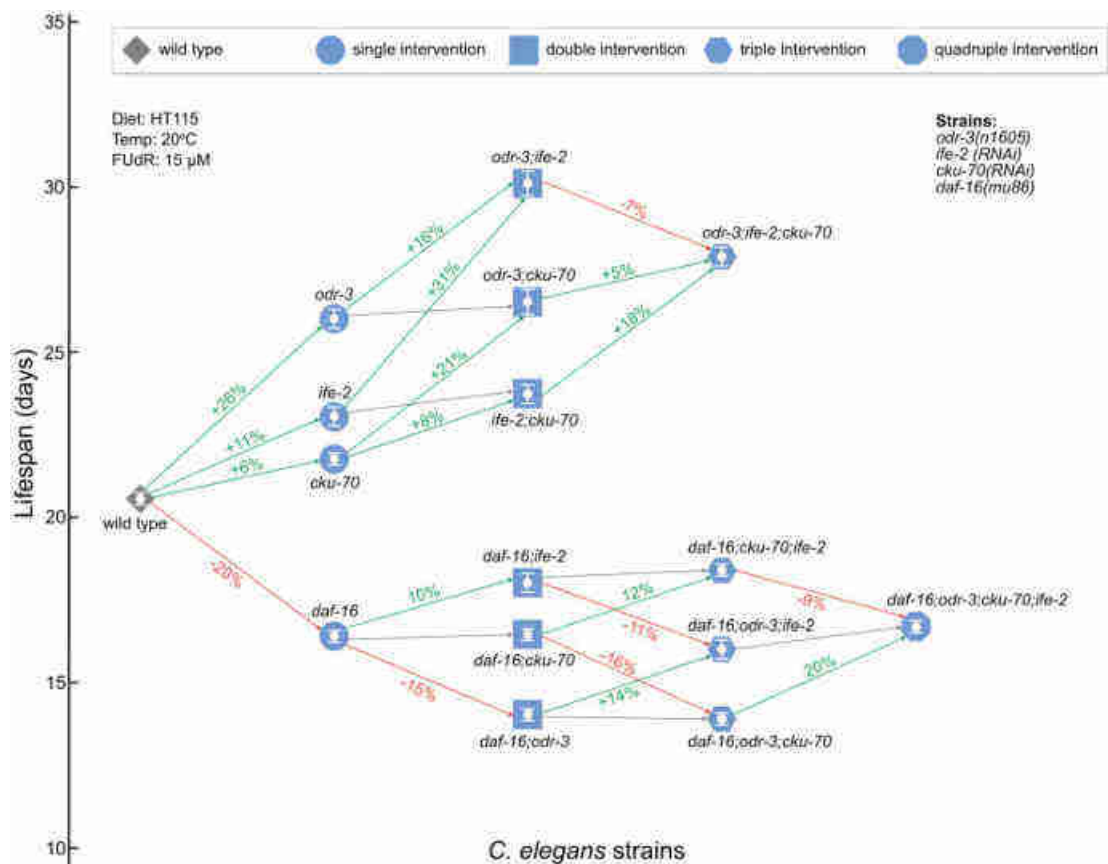


Figura 6. Reprezentarea sub formă de rețea a duratei medii de viață a fiecărei tulpini. Fiecare nod reprezintă o tulpină (romb - tip sălbatic, cerc - intervenții simple, pătrat - intervenții duble, hexagon - intervenții triple și octogon - intervenții cvadruple), iar poziția sa pe axa Oy este dată de durata medie de viață. Muchiile conectează n-mutanții la (n+1)-mutanți și cuantifică efectul intervenției (n+1) ca procent în comparație cu n-mutantul: verde pentru creștere, roșu pentru scădere și gri pentru efect nesemnificativ în durata medie de viață. *odr-3* reprezintă mutația *odr-3(n1605)*; *daf-16* reprezintă mutația *daf-16(mu86)*; *ife-2* definește *ife-2(ARNi)*, iar *cku-70* definește *cku-70(ARNi)*. Segmentele din interiorul nodurilor reprezintă media \pm SEM. Figură adaptată din (Bunu et al., 2020), utilizată sub licență CC BY 3.0.

2.3. Modele de rețele în longevitate

În ultimele decenii, cercetarea biologică a pivotat de la abordări reduționiste la cele holistice, pe măsură ce domeniul biologiei sistemelor s-a dezvoltat. Abordarea holistică constă în analiza sistemului ca întreg, spre deosebire de analiza componentelor individuale și implică adesea rețele de interacțiuni între componentele moleculare. În a doua secțiune a

acestei lucrări, prezentăm modelarea rețelelor mixte de proteine și medicamente asociate longevității pentru a explica mecanismele de îmbătrânire din perspectiva biologiei sistemelor.

Surse de date

Listele de gene și medicamente deja cunoscute pentru modularea procesului de îmbătrânire au fost descărcate din următoarele baze de date publice: i) GenAge (build 20) (Tacutu et al., 2018) - gene care au determinat creșterea/scăderea duratei de viață, atunci când sunt modificate printr-o intervenție (mutație, supraexprimare, interferență ARN etc) în organisme model, cunoscute și sub numele de gene asociate longevității (LAGs) și ii) DrugAge (Barardo et al., 2017) - molecule care au prezentat o creștere/scădere semnificativă a duratei de viață atunci când au fost administrate în organisme model.

Lista de molecule care sunt implicate în inducerea pluripotenței a fost selectată manual de colaboratorii noștri (Universitatea Ben-Gurion din Negev, Israel) din literatura științifică.

Datele pentru interacțiunile medicament-proteină au fost preluate din STITCH (Szklarczyk et al., 2016), folosind un script Python dezvoltat intern, pentru a interoga interacțiunile relevante din baza de date. Au fost utilizate doar interacțiunile cu scor de încredere cel puțin mediu, de tip „bază de date” și „experimentale”, excluzând predicțiile.

Interacțiunile proteină-proteină (PPI) au fost integrate din BioGRID 3.5 (Oughtred et al., 2019) și doar interacțiunile fizice au fost luate în considerare pentru analiza ulterioară. În Tabelul 2 este inclus un tabel rezumat al fiecărei surse de date.

Table 2. Statistici pentru datele selectate în această lucrare

Setul de date	Tipul de date	Organism(e)	Numărul de înregistrări	Numărul de ID-uri unice
GenAge	gene	<i>C. elegans</i>	1050	876
DrugAge	medicamente	<i>C. elegans</i>	659	283
Molecule implicate în inducerea pluripotenței	medicamente	<i>H. sapiens</i>	92	92
BioGRID	gene/proteine	<i>C. elegans</i>	6069	6069

Genele asociate longevității formează o rețea cu o componentă conexă mare, care conține 253 de noduri (44% din totalul de gene) - numită în continuare „rețeaua longevității” (Fig. 7A). Dacă sunt considerate împreună cu partenerii lor de ordinul întâi, procentul de noduri din componenta conexă crește la o valoare excepțională de 95% din rețea ("rețea de

longevitate extinsă") (Fig. 7B). Împreună, aceste rezultate arată natura cooperantă a LAG-urilor (definite mai sus ca gene asociate longevității). LAG-urile sunt suprareprezentate printre proteinele puternic conectate în interactom (GSEA, NES = 1,28, $p = 8,4e-15$), sugerând implicarea lor într-o mare varietate de procese. În mod interesant, partenerii LAG-urilor sunt, de asemenea, printre cele mai conectate proteine (GSEA, NES = 1,45, $p = 2e-50$), iar diferența de conectivitate globală în cadrul rețelei de interacții fizice proteine-proteine BioGRID între LAG-uri și parteneri nu este semnificativă statistic (testul t-Student, $p = 0,12$). Primele 372 de proteine cele mai conectate din BioGRID aparțin LAG-urilor sau partenerilor acestora, iar 489 din primele 500 de proteine cele mai conectate sunt LAG-urile sau partenerii lor.

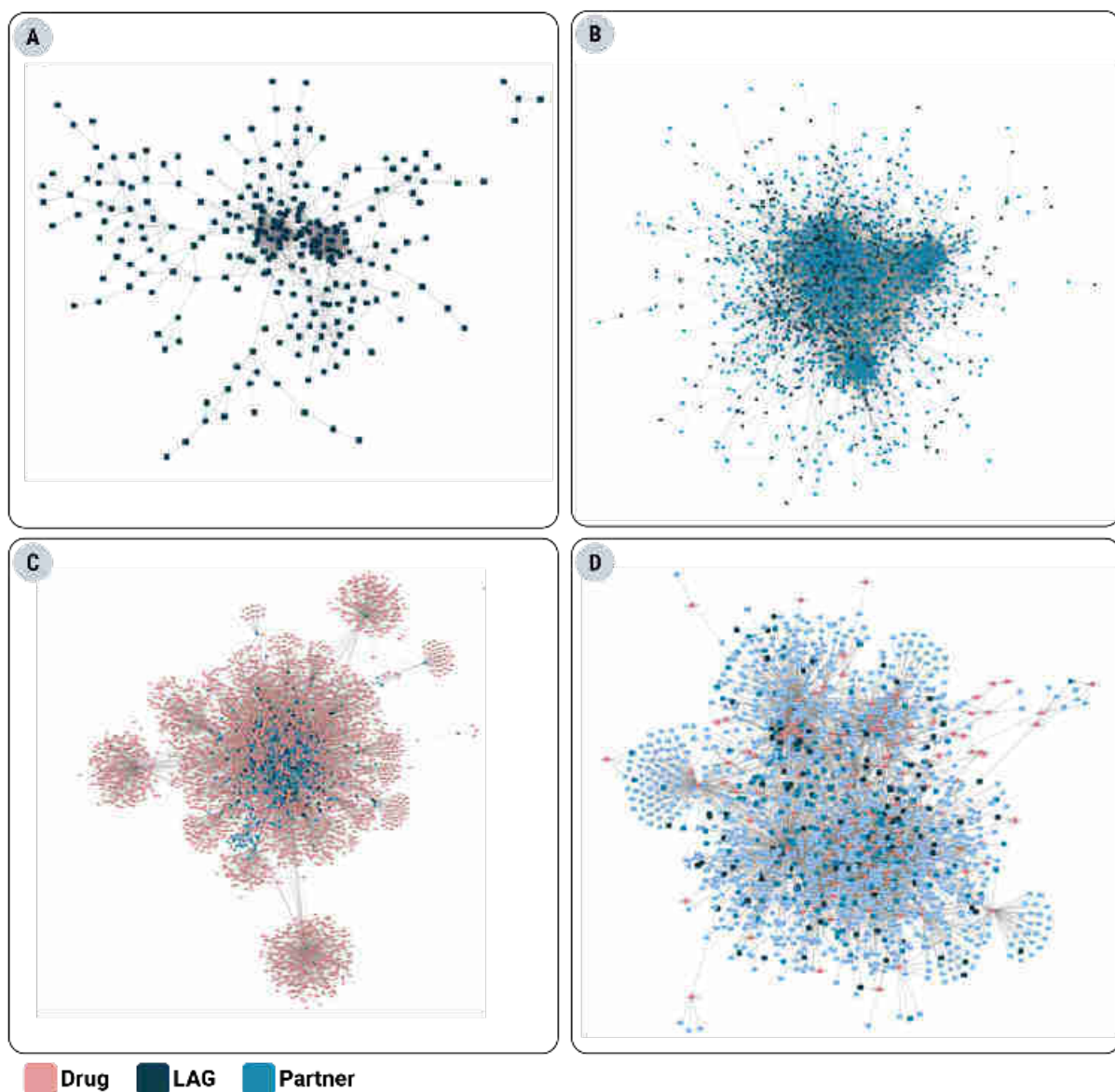


Figura 7. Modele de rețele de gene și medicamente asociate longevității A. Interacțiuni proteină-proteină între LAG-uri; B. Rețea de longevitate extinsă (LAG-uri și partenerii lor de

prim ordin); **C.** Rețea mixtă, inclusiv LAG-urile, partenerii lor de prim ordin și medicamentele care îi țintesc; **D.** Rețeaua de medicamente DrugAge și țintele lor genice, inclusiv interacțiunile proteină-proteină.

2.4. Meta-analiză a semnăturii de îmbătrânire

Pentru a înțelege mai bine procesul de îmbătrânire dintr-o perspectivă mai largă, sistemică, o abordare complementară este identificarea seturilor de markeri moleculari asociați cu vârsta sau o „semnătură de îmbătrânire”, astfel caracterizând întregul sistem pe baza măsurătorilor moleculare multiple. Mai multe studii au folosit date transcriptomice din diferite țesuturi pentru a identifica astfel de semnături și pentru a investiga asocierea lor cu bolile legate de vârstă și cu mortalitatea. Cu toate acestea, diferite tehnologii și metode de preprocesare/analiză pot genera rezultate variabile.

În acest studiu, am folosit 5 seturi de date asociate îmbătrânirii de la *C. elegans* și 24 de seturi de date asociate îmbătrânirii de la *H. sapiens*, descărcate din baza de date Gene Expression Omnibus (GEO) și reanalizate local folosind script-uri R dezvoltate intern, pentru a identifica o semnătură comună asociată îmbătrânirii. Meta-analiza seturilor de date a evidențiat că procesele metabolice, reglarea longevității și răspunsul imun au fost supraprezentate în semnătura genică caracteristică vârstei mai înaintate.

Genele exprimate diferențiat (DEG-uri) au o intersecție semnificativă statistic cu partenerii de ordinul întâi ai LAG-urilor, sugerând că mecanismele modulatorie ale îmbătrânirii (reprezentate de LAG-uri) și mecanismele afectate de procesul de îmbătrânire (reprezentate prin DEG-uri) nu sunt identice, dar sunt foarte bine interconectate.

Meta-analiza seturilor de date umane a arătat că diferențele de expresie variază dramatic între diferite țesuturi (Fig. 8). Cu toate acestea, pentru genele care își schimbă expresia în mai multe țesuturi, în majoritatea cazurilor direcția acestor modificări este aceeași (Fig. 9). Cu alte cuvinte, dacă o genă este exprimată diferențiat cu vârsta, expresia ei crește sau scade în toate țesuturile afectate.

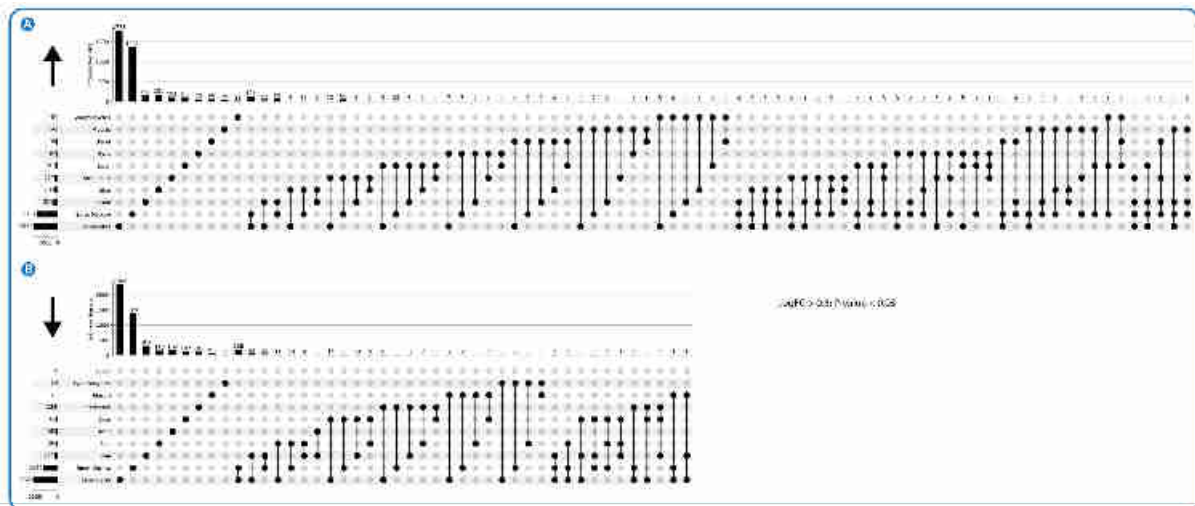


Figura 8. Modificări comune sau unice ale expresiei genelor în diferite țesuturi. A. Genele supraexprimate în procesul îmbătrânirii. **B.** Genele subexprimate în procesul îmbătrânirii. **A-B.** Numărul de gene supra- sau sub-exprimate din fiecare țesut este afișat în stânga jos; în partea de sus sunt prezentate într-o histogramă numărul de gene exprimate în același mod în unul sau mai multe țesuturi. Pentru această analiză, limita a fost stabilită la $\text{LogFC} > 0,5$ și $p < 0,05$.

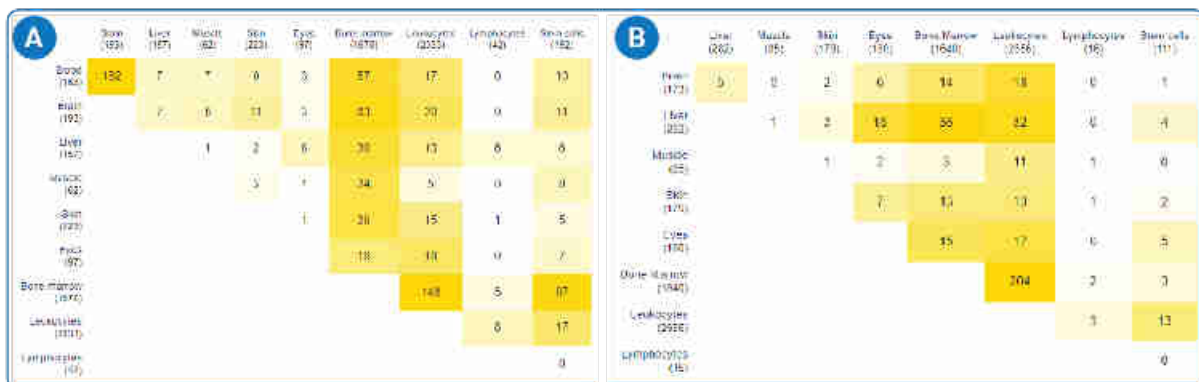


Figura 9. A. Intersecția genelor supra-exprimate cu vârsta în diferite țesuturi. **B.** Intersecția genelor sub-exprimate cu vârsta în diferite țesuturi.

2.5. Boli și procese asociate vârstei

Deoarece se știe că îmbătrânirea și bolile asociate vârstei (ARD) au mecanisme parțial comune (Fernandes et al. 2016), a fost efectuată o analiză bioinformatică pentru a identifica potențiala semnătură comună între profilele transcriptomice ale îmbătrânirii și ale bolilor asociate vârstei. Au fost selectate următoarele patologii: ateroscleroza, diabetul zaharat de tip II, boala Alzheimer (AD), boala Parkinson (PD) și osteoporoza.

Am calculat scorurile de similaritate dintre fiecare pereche de profiluri

transcripționale folosind GSEA, iar scorurile au fost vizualizate sub forma unei rețele formate din 94 de noduri (i.e. comparații de expresie genică) și 826 de muchii (i.e. scoruri de similaritate). Folosind algoritmul MCL, rețeaua din Fig. 10a a fost împărțită în 4 clustere, de dimensiuni diferite, plus o serie de noduri izolate (Fig. 10b).

Dintre acestea s-au remarcat două clustere (Fig. 10b1, b2), prezentând un număr mai mare de conexiuni între nodurile din componența lor. Primul cluster este axat pe comparații specifice bolilor neurodegenerative (predominant Alzheimer) și pe comparații de îmbătrânire (predominant în regiunile creierului, dar și în sânge) - care denotă anumite mecanisme specifice de îmbătrânire a creierului - foarte relevante pentru bolile neurodegenerative. Astfel, primul cluster a fost denumit AD-Cluster. Conține 21 de noduri, inclusiv 14 comparații de îmbătrânire și 6 comparații AD vs control, 94% dintre ele având scoruri de similaritate pozitive, ceea ce arată un consens ridicat.

Al doilea cluster este preponderent caracteristic bolii Parkinson (6 comparații), dar include, de asemenea, modificări legate de vârstă și, în mod interesant, profiluri ale osteoporozei (1 comparație) și ale aterosclerozei (2 comparații). Țesuturile reprezentate pentru comparațiile de îmbătrânire sunt diverse și includ creier, oase, mușchi, ficat și retină. În continuare facem referire la acest cluster ca PD-Cluster.

2.5.1. Profilul transcripțional comun în AD-Cluster și PD-Cluster

Pentru a identifica modificările comune ale profilurilor transcripționale pentru îmbătrânire și bolile asociate vârstei, am efectuat două meta-analize separate: pentru AD-Cluster și pentru PD-Cluster. Pentru fiecare dintre ele, deoarece s-a dorit prioritizarea direcției comune față de magnitudinea schimbării de expresie, am selectat doar genele semnificative din fiecare set de date ($p < 0,05$), care au aceeași direcție în toate seturile de date și o magnitudine a schimbării de cel puțin 2 ori față de control, în cel puțin o comparație.

Analizele funcționale au arătat că genele supraexprimate din AD-Cluster sunt caracteristice proceselor legate de vindecarea rănilor, producția de citokine, răspunsul la stres la ionii de cupru și dezvoltarea vasculară, în timp ce genele subexprimate sunt caracteristice pentru dezvoltarea axonilor, transmiterea sinaptică și transportul transmembranar. Genele supraexprimate PD-Cluster sunt caracteristice pentru sumoilarea proteinelor, diferențierea celulelor fotoreceptoare, răspunsul la stimulul factorului de creștere, iar genele subexprimate sunt caracteristice pentru transportul acizilor grași și menținerea structurii sinapselor.

comparație între două stări/condiții analizate în fiecare studiu. Interacțiunile dintre noduri reprezintă scoruri de similaritate între două comparații, care sunt mai mari decât un prag prestabilit. Scorurile de similaritate pozitive sunt reprezentate prin linie continuă; scorurile negative sunt reprezentate prin linie punctată.

2.5.2. Fibroza pulmonară

Una dintre bolile asociate vârstei, poate mai puțin evidențiată, dar cu un cost ridicat la nivel de societate și consecințe majore asupra organismului, este fibroza pulmonară. Fibroza tisulară este un proces a cărui incidență crește odată cu vârsta. Poate afecta mai multe organe, plămânii fiind unul dintre cele mai sensibile.

În această lucrare, au fost colectate 212 gene aparținând *Mus musculus* (numite PFRG-uri - pulmonary fibrosis-related genes), grupate în trei categorii: pro-fibrotice (50%), anti-fibrotice (43,5) și mixte (6,5%), i.e. prezentând atât dovezi pro- cât și anti-fibrotice. O genă pro-fibrotică reprezintă o genă care crește fibroza atunci când este supraexprimată și/sau reduce fibroza atunci când este subexprimată. În mod similar, o genă anti-fibrotică este o genă care crește fibroza atunci când este subexprimată și/sau scade fibroza atunci când este supraexprimată.

Analiza bioinformatică a fost efectuată pentru întregul set de PFRG-uri, precum și pentru genele pro- și anti-fibrotice separat. PFRG-urile au evidențiat procese precum proliferarea, inflamația, funcțiile imune și cancerul. Procesul de reglarea proliferării a arătat o corelație pentru genele pro- și anti-fibrotice: genele pro-fibrotice sunt reprezentative pentru reglarea pozitivă, iar genele anti-fibrotice sunt reprezentative pentru reglarea negativă a proliferării (Tabelul 3). Mai mult, mai multe căi de semnalizare legate de îmbătrânire au fost semnificative pentru genele pro-fibrotice, cum ar fi: rezistența la insulină, răspunsul la stres, căile de semnalizare PI3K-Akt, FoxO și apoptoza. Aceste rezultate evidențiază asocierea dintre fibroza pulmonară și procesul de îmbătrânire.

Tablelul 3. Rezultatele analizei funcționale pentru procesele biologice asociate cu PFRG-urile

Toate PFRG-urile	PFRG-uri pro-fibrotice	PFRG-uri anti-fibrotice
<i>Reglarea proliferării</i>	<i>Reglarea pozitivă a proliferării</i>	<i>Reglarea negativă a proliferării</i>
<i>Semnalizarea citokinelor</i>	<i>Semnalizarea citokinelor</i>	
<i>Inflamația</i>	<i>Inflamația</i>	
<i>Funcția imună</i>	<i>Funcția imună</i>	
<i>Cancer</i>	<i>Cancer</i>	<i>Cancer</i>
<i>Reacția la patogeni</i>	<i>Reacția la patogeni</i>	
<i>Homeostazia oxigenului</i>	<i>Homeostazia oxigenului</i>	
<i>Calea de semnalizare MAPK</i>		<i>Calea de semnalizare MAPK</i>
<i>Calea de semnalizare TNF</i>	<i>Calea de semnalizare TNF</i>	
<i>Calea de semnalizare Jak-STAT</i>	<i>Calea de semnalizare Jak-STAT</i>	
Astm		
Îmbătrânire		
	Rezidența la insulină	
	Răspunsul la stres	
		Calea de semnalizare a estrogenului
		Răspunsul la stimuli mecanici
	Calea de semnalizare PI3K-Akt	
	Calea de semnalizare VEGF	
	Calea de semnalizare FoxO	
	Apoptoza	

În continuare, am construit o rețea de interacții proteină-proteină (PPI) pentru proteinele umane, ortologi ai PFRG-urilor și am constatat că 56,3% dintre PFRG-uri (107 noduri) formează o componentă conexă. Acest procent este extrem de mare în comparație cu procentul așteptat de gene care ar forma o rețea conexă din eșantionarea aleatorie a 190 de noduri din interactom (Z-score de 5,36, Fig. 12D,E). Pe lângă faptul că sunt mai interconectate decât se aștepta întâmplător, PFRG-urile au și o conectivitate medie mai mare în întregul interactom (64,4), în comparație cu conectivitatea medie a tuturor genelor (45,5) (Fig. 12B), acestea situându-se printre genele cu cea mai mare conectivitate totală (NES = 1,18, $p < 0,05$) (Fig. 12C). Este de remarcat faptul că printre cele mai conectate noduri din

rețeaua PFRG-urilor se numără gene binecunoscute asociate îmbătrânirii, cum ar fi AKT1, HIF1A, SIRT1, HSPA5.

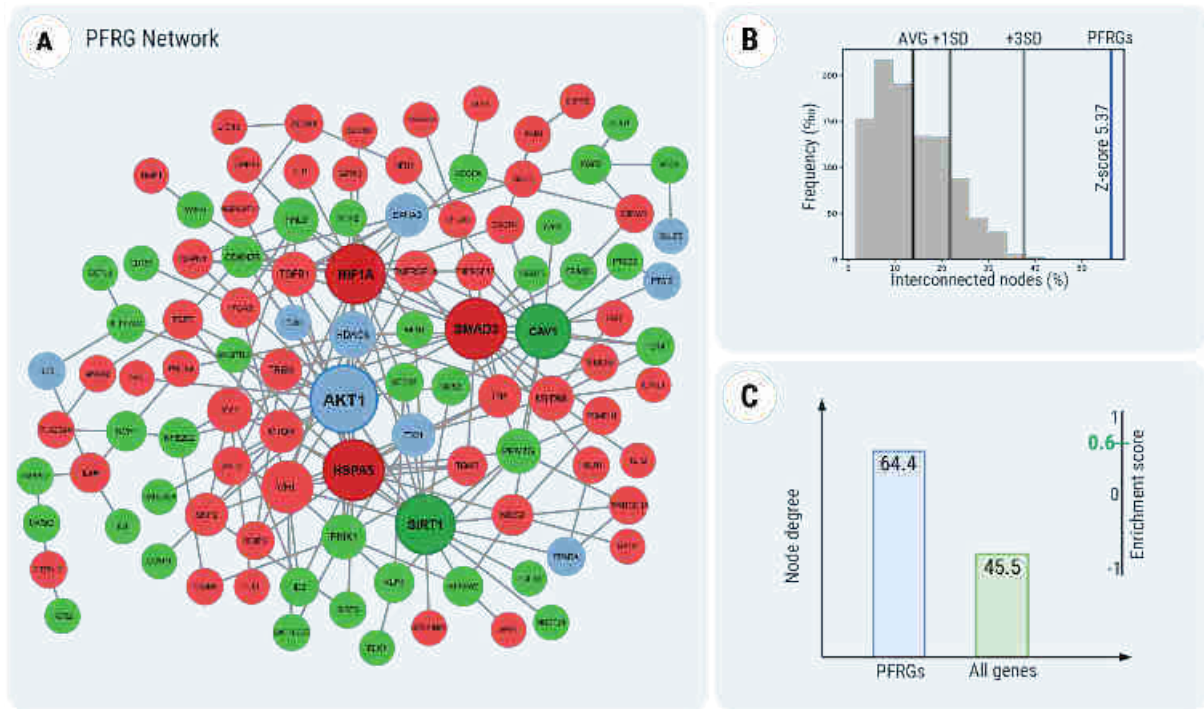


Figura 11. Vizualizarea în rețea a interacțiunilor proteină-proteină dintre PFRG-uri. (A) Cea mai mare componentă conexă a rețelelor de interacțiuni proteină-proteină pentru toate PFRG-urile ($n = 107$ noduri). Primele 5% noduri cu cel mai înalt grad sunt evidențiate cu o etichetă îngroșată. (B) Procentul de noduri care formează cea mai mare componentă conexă (procent de noduri interconectate) pentru un număr de noduri egal cu dimensiunea setului PFRG ($n = 190$), eșantionate aleatoriu din interactomul uman. Analiza a fost repetată de 100 de ori, iar histograma arată distribuția acestor 1.000 de valori care reprezintă procentul de noduri interconectate (media 13,9%, SD: 7,9%). Linile verticale descriu valorile pentru media distribuției (AVG), precum și pentru o deviație standard și pentru trei deviații standard de la medie. Dreapta albastră corespunde scorului Z al PFRG-urilor pe această distribuție. (C) Compararea gradului mediu în interactomul uman (BioGRID v 3.5.188) pentru PFRG-uri (bara albastră) și toate genele corespunzătoare întregului proteom (bara verde). Figură adaptată după (Toren et al., 2021), utilizată sub licență CC BY 4.0.

2.6. Repoziționarea medicamentelor

Cu privire la identificarea medicamentelor care modulează longevitatea, în această lucrare, accentul a fost pus pe două metode. Prima abordare este de a prioritiza moleculele care induc un profil transcripțional care este opus celui asociat cu procesul de îmbătrânire în

sine, adică medicamente care ar putea inversa schimbările transcripționale care apar odată cu îmbătrânirea. Cea de-a doua abordare este căutarea printre medicamentele care vizează căile legate de îmbătrânire, cele care ar putea avea un rol în reprogramarea celulară, și prin urmare și în procese anti-îmbătrânire.

2.6.1. Medicamente cu potențial de a inversa schimbarea profilului transcripțional asociat cu bolile neurodegenerative

Asemănările observate între semnătura de îmbătrânire și semnăturile AD/PD sugerează că profilele transcripționale comune indică o stare mai vulnerabilă și ar putea avea un rol în procesul degenerativ. Pentru a identifica substanțele chimice cu potențială utilizare terapeutică anti-neurodegenerativă, am căutat medicamente care induc un profil transcripțional opus celui asociat modificărilor patologice. În mod ideal, o abordare terapeutică ar fi inversarea completă a transcriptomului către o stare tânără, dar este improbabil să fie realizată în practică. În schimb, propunem să căutăm compuși care ar putea conduce sistemul către o stare mai tânără și mai sănătoasă.

Pentru această abordare, am folosit Connectivity Map (Subramanian et al., 2017), un instrument online care conține peste 1 milion de profile transcripționale pentru aproximativ 20.000 de compuși, obținute prin studii de perturbare în 70 de linii celulare umane. Acest instrument folosește metoda GSEA (Subramanian et al., 2005) pentru a calcula un scor pentru fiecare substanță, utilizând ca date de intrare semnătura alcătuită din genele supraexprimate și genele subexprimate. Un scor pozitiv arată că medicamentul are un profil similar cu profilul analizat, în timp ce un scor negativ sugerează un profil opus.

Analiza AD-cluster a evidențiat 15.076 de medicamente semnificative ($FDR < 0,05$) cu profiluri opuse modificărilor semnăturii care apar în AD sau stări similare de îmbătrânire. Am priorizat cele mai semnificative medicamente cu cele mai mari scoruri negative și am găsit, printre primele 200, 9 medicamente ale căror mecanisme au fost deja asociate cu procese legate de îmbătrânire, cum ar fi: modularea căilor de inflamație, prin antagonismul receptorului de chemokine CC (AMG-487) sau inhibarea lipoxigenazei (dietilcarbamazina); stimularea apoptozei (BRD-K68144790) sau inhibarea XIAP (GDC-0152); inhibarea MEK (selumetinib) - care poate preveni sau suprima îmbătrânirea celulară (Steelman et al., 2011); inducerea proteinelor de șoc termic (HSP) (homosalat) - care au un rol pro-longevitate și reglează senescența celulară și apoptoza (Tower, 2009); inhibarea mTOR (AZD-8055); modularea proteinelor PPAR (Pioglitazona, Rupacarib) - care sunt implicate în stresul

oxidativ și metabolismul energetic (Erol, 2007).

Analiza PD-cluster a evidențiat doar 194 de medicamente semnificative (FDR<0,05) cu profiluri opuse semnăturii PD. Printre acestea am regăsit și 8 medicamente aprobate pentru patologii reprezentate în cluster, precum și alte medicamente aflate în studii clinice sau compuși cu mecanisme similare cu cele deja aprobate, care ar putea fi agenți terapeutici promițători. Această analiză a evidențiat ca mecanisme potențiale: modularea receptorului de estrogen (Raloxifen - aprobat pentru tratamentul osteoporozei și Bazedoxifen - o nouă generație de modulatori, aprobat pentru prevenirea osteoporozei postmenopauză); medicamente anticolinergice (Metixen - aprobat ca antiparkinsonian, Solifenacin - aprobat pentru incontinență urinară, Mebeverină - aprobată ca antispastică și Deltalină); agonismul receptorilor dopaminergici (Metergolină) și inhibarea fosfolipazei (Darapladib - un medicament experimental pentru stabilizarea plăcii aterosclerotice).

Aceste rezultate arată potențialul metodei, sugerând ca și alte ținte identificate astfel cu scoruri foarte ridicate, ar trebui studiate în laborator, pe modele *in vitro* de AD și PD.

2.6.2. Molecule mici pentru reprogramarea celulară

Au fost selectate din literatura științifică 92 de moleculele mici (SM) pentru reprogramarea celulară, adnotate apoi cu proprietăți chimice din PubChem (Kim et al., 2021) și Human Metabolome Database (HMDB) (Wishart et al., 2018). Cei 92 de compuși au fost clasificați, pe baza rolului lor în reprogramare ca având rol inductor, amplificator sau, în unele cazuri, ambele. În plus, am colectat o listă de 10 cocktail-uri care includ doar combinații de medicamente care împreună induc reprogramarea parțială sau totală, fără a necesita prezența factorilor de transcripție.

Acești compuși pot fi împărțiți în trei categorii funcționale principale (modificatori de semnalizare - în principal inhibitori ai TGF β , modificatori epigenetici - metiltransferaze și histondeacetilaze, modificatori metabolici - inhibitori ai GSK3), precum și o mică fracțiune care nu aparține nici uneia dintre aceste categorii (Fig. 12).

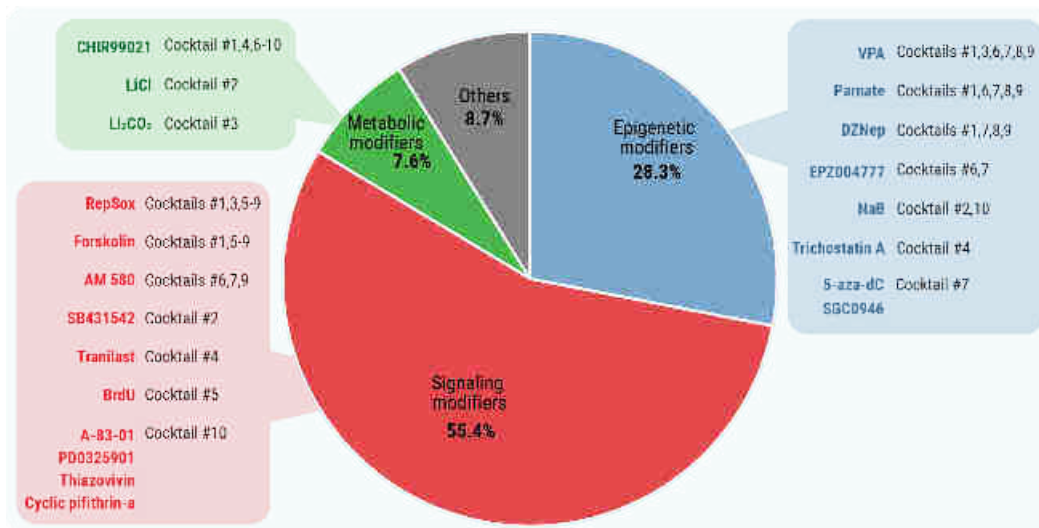


Figura 12. Distribuția celor 92 de amplificatori/inductori ai pluripotenței pe baza categoriei funcționale. Graficul prezintă procentul de medicamente care se încadrează în fiecare categorie (modificatori de semnalizare, modificatori epigenetici, modificatori metabolici și altele). Panourile laterale arată doar medicamentele din cocktail-urile din aceste categorii. Figură adaptată din (Knyazer et al., 2021), utilizată sub licență CC BY 3.0.

Pentru a oferi o perspectivă a biologiei sistemelor țintelor proteice ale SM-urilor, am construit o rețea PPI pentru proteinele țintă ale celor 92 de molecule: 991 de proteine și 6.072 de interacțiuni. Rețeaua aproximează o structură scale-free, caracteristică rețelelor biologice (Figura 13A), cu gradul nodului urmând o distribuție caracterizată de funcția $P(k) = 221 * k^{-1,16}$. Mai mult, rețeaua de 991 de proteine, precum și fiecare subrețea a celor 10 cocktail-uri sunt mai interconectate decât se aștepta întâmplător (Figura 13B). O astfel de interconectivitate ridicată pentru toate cocktail-urile susține ideea că acestea vizează o serie de proteine care în mod normal lucrează împreună pentru a-și îndeplini funcția.

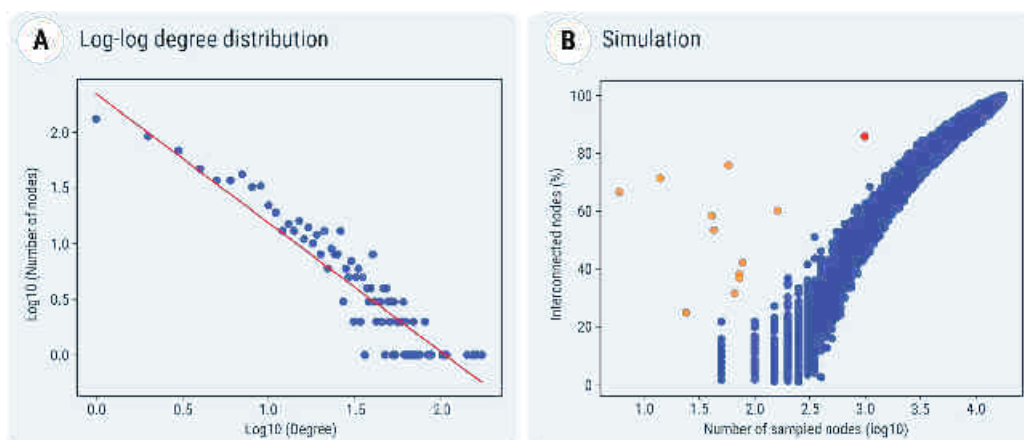


Figura 13. Rețeaua de interacțiuni proteină-proteină pentru ținte proteice ale

moleculelor mici care induc sau amplifică pluripotența. (A) Procentul de noduri care formează cea mai mare componentă conexă în subrețele de diferite dimensiuni eșantionate aleatoriu din interatom. Dimensiunea setului de noduri variază de la 50 la 17.600, cu un pas de 50, reprezentat în scară logaritmică pe axa Ox. Punctele portocalii reprezintă interconectivitatea observată pentru rețelele PPI ale celor 10 cocktail-uri, iar punctul roșu reprezintă interconectivitatea observată pentru rețeaua de 991 ținte proteice; **(B)** Proprietatea scale-free pentru rețeaua de 991 de proteine țintă. Graficul reprezintă distribuția gradului într-o scară logaritmică. Dreapta roșie descrie funcția liniară care aproximează distribuția (în scara log-log). Figură adaptată din (Knyazer et al., 2021), utilizată sub licență CC BY 3.0.

Am efectuat o comparație între țintele celor 10 cocktail-uri și țintele factorilor Yamanaka, atât la nivel de proteine, cât și la nivel de căi de semnalizare. În mod interesant, intersecția dintre țintele de proteine a fost limitată și nesemnificativă pentru toate cocktail-urile, cu excepția cocktail-ului 7 ($n=15$, $p=0,0033$, Fig. 14A). Cu toate acestea, la nivelul căilor de semnalizare, suprapunerea este semnificativă pentru patru cocktail-uri (Fig. 14B) și include mai multe căi de semnalizare legate de cancer.

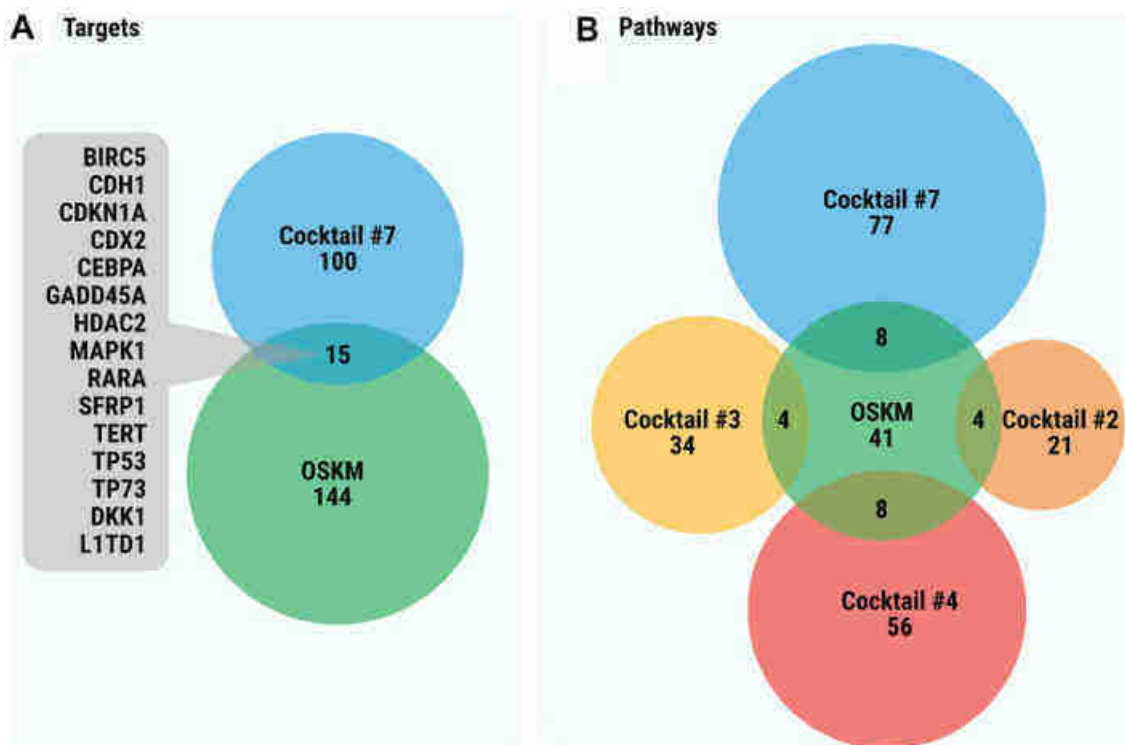


Figura 14. Diagrama Venn a intersecției dintre (A) țintele proteice ale cocktail-urilor și țintele proteice ale factorilor de transcripție OSKM; (B) căile de semnalizare suprareprezentate în țintele proteice ale cocktail-urilor și căile de semnalizare

suprareprezentate în țintele proteice ale OSKM. Pentru simplificare, intersecțiile dintre perechile de cocktail-uri nu au fost afișate în figură. Figură adaptată din (Knyazer et al., 2021), utilizată sub licență CC BY 3.0.

2.6.3. Medicamente care induc pluripotenta ca agenți potențiali împotriva patologiilor neurodegenerative

Folosind lista de molecule mici pentru reprogramarea celulară descrisă anterior, denumită PASM - „Molecule mici asociate cu pluripotența”, am prioritizat țintele genetice pentru bolile studiate asociate vârstei și am investigat dacă medicamentele care induc reprogramare celulară ar putea îmbunătăți fenotipul neurodegenerativ, pe baza unei analize bioinformatică a țintelor genice ale acestora.

Cele 10 cocktail-uri PASM au un anumit grad de suprapunere a medicamentelor constitutive. În plus, utilizarea unui număr mai mare de molecule mici și, în consecință, a unui număr mare de ținte ca set de intrare pentru o analiză bioinformatică ar crește probabilitatea de rezultate fals pozitive. Prin urmare, analiza s-a concentrat pe cele mai „compacte” cocktail-uri, adică pe cel mai mic număr de molecule mici care pot realiza reprogramarea la stadiul de celule progenitoare neuronale. Au fost selectate trei cocktailuri: 1) Acid valproic, CHIR99021, RepSOX (VCR); 2) NaB, LiCl, SB431542 (NLS); 3) Trichostatină A, Li₂CO₃ / LiCl, Tranilast (TLT).

Câteva dintre aceste medicamente sunt deja cunoscute ca modulatori ai longevității, mai precis ele măresc durata de viață a organismelor model. DrugAge (Barardo et al., 2017) raportează că acidul valproic și clorura de litiu cresc durata medie de viață în *C. elegans* cu 35%, respectiv 36%. Butiratul de sodiu crește durata medie de viață în *D. melanogaster* cu 14%, iar tricostatina A a arătat o creștere cu 22% a duratei medii de viață a nematodelor.

Este important de menționat că nici unul dintre medicamentele din lista PASM nu este aprobat de FDA pentru boala Alzheimer. Două dintre ele sunt aprobate pentru alte patologii (acidul valproic ca anticonvulsivant și clorura de litiu ca antipsihotic); tricostatina A și tranilast au potențial terapeutic la om, dar nu sunt aprobate; celelalte substanțe din listă nu au nicio utilizare terapeutică raportată. Intersecția dintre țintele proteice ale fiecărui cocktail și țintele proteice ale medicamentelor aprobate pentru boala Alzheimer este mică, dar semnificativă statistic pentru toate (Fig. 15A,B,C), iar existența unor ținte comune arată potențialul de identificare a unui nou medicament. Fig. 15D evidențiază intersecțiile dintre țintele acestor trei cocktail-uri.

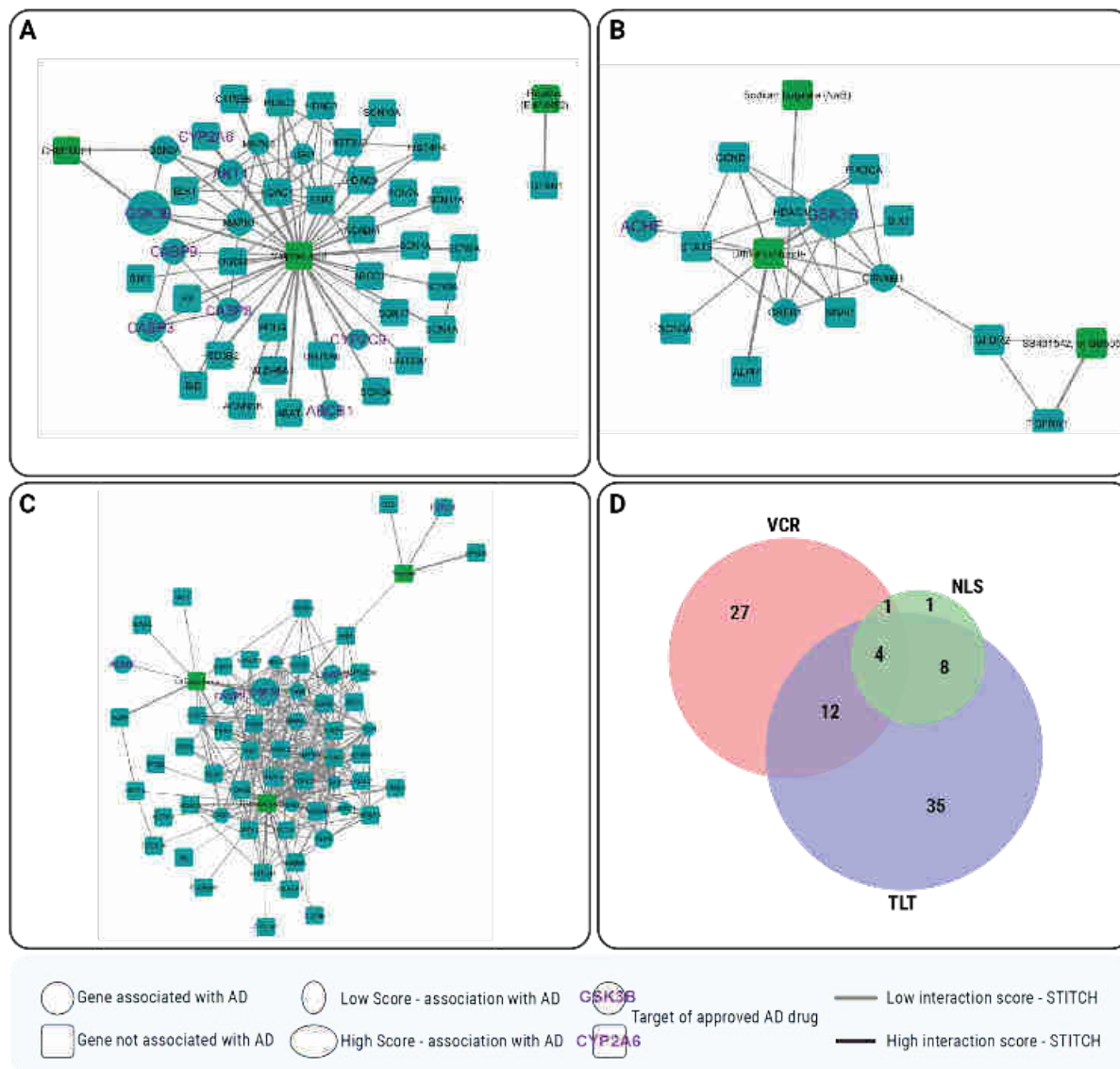


Figura 15. PASM-uri în cele trei cocktail-uri selectate și țintele acestora: A. Cocktail VCR: Acid valproic, CHIR99021, RepSOX; **B.** Cocktail NLS: NaB, LiCl, SB431542; **C.** Cocktail TLT: Trichostatin A, LiCl, Tranilast. **D.** Diagrama Venn a țintelor comune între cele trei cocktail-uri

Analiza rețelelor de medicamente și proteine sugerează un efect potențial al PASM-urilor ca modulatori AD (Fig. 15A,B,C). PASM-urile țintesc unele gene despre care se cunoaște deja că sunt implicate în patologia legată de boala Alzheimer sau gene care aparțin unui potențial mecanism modulator al bolii. Printre aceste ținte, GSK3B, ACHE și AKT1 sunt toate ținte ale medicamentelor deja aprobate. Pe baza direcției de interacțiune dintre medicamentele aprobate de FDA și proteinele codificate de aceste trei gene, o potențială abordare terapeutică ar putea implica inactivarea GSK3B și ACHE și supraexprimarea AKT1. În acest moment, aceasta ipoteză este investigată în laboratorul nostru.

3. Concluzii finale

Această lucrare oferă o abordare a biologiei sistemelor pentru studiul îmbătrânirii și a bolilor asociate vârstei. Folosind metode bioinformatică legate de biologia rețelelor și meta-analize, am explorat mecanismele moleculare ale îmbătrânirii și bolilor asociate vârstei, precum și semnaturile comune și am priorizat medicamente care pot fi repositionate ca modulatori ai îmbătrânirii.

O realizare importantă este construirea bazei de date SynergyAge, prima de acest gen, care conține intervenții validate experimental asupra genelor asociate longevității. Mai mult, am propus definiții atât pentru interacțiunile dintre două gene, cât și pentru interacțiunile mai complexe dintre cel puțin trei gene. Baza de date SynergyAge facilitează predicții computaționale ale noilor combinații genetice, așa cum este demonstrat de dublul mutant propus *odr-3;ife-2*, reprezentativ pentru o combinație aditivă și validat experimental de colegii mei din laborator.

Meta-analizele procesului de îmbătrânire și ale bolilor selectate asociate vârstei au evidențiat mecanisme comune care au fost separate în două grupuri majore de schimbări transcripționale. Acest rezultat a fost susținut de o analiză de tip Connectivity Map, folosind aceste semnaturi comune; în urma analizei au fost identificate potențiale medicamente implicate în procesele de îmbătrânire, deja aprobate pentru boli asociate vârstei sau similare farmacologic cu medicamentele aprobate, dovedind potențialul metodei.

O altă abordare pentru repositionarea medicamentelor a fost alegerea combinațiilor de medicamente din lista de cocktail-uri ce induc reprogramarea celulară parțială. Au fost construite rețele mixte de medicamente și proteine pentru trei cocktail-uri, care au fost selectate pentru a maximiza țintele asociate longevității și pentru a minimiza țintele ce nu sunt asociate proceselor îmbătrânirii. Aceste cocktail-uri prezintă un potențial ridicat ca agenți anti-neurodegenerativi, având în vedere direcția interacțiunilor medicament-proteină și asocierile țintelor proteice cu AD, iar aceste intervenții sunt testate în acest moment în laborator.

Contribuții personale

1. *Lucrări publicate*

Bunu G*, Toren D*, Ion CF, Barardo D, Sârghie L, Grigore LG, de Magalhães JP, Fraifeld V, Tacutu R. “SynergyAge, a curated database for synergistic and antagonistic interactions of longevity-associated genes”, *Scientific Data*, 7:366 (2020), <https://doi.org/10.1038/s41597-020-00710-z>. **IF 5.54, Q1**

Knyazer A*, **Bunu G***, Toren D, Bucaciuc M, Racica T, Segev Y, Wolfson M, Muradian K, Tacutu R, Fraifeld V. “Small molecules for cell reprogramming: a systems biology analysis”, *Aging (Albany NY)*, 13:25739-25762 (2021), <https://doi.org/10.18632/aging.203791>. **IF 5.6, Q2**

Toren D, Yanai H, Abu Taha R, **Bunu G**, Ursu E, Ziesche R, Tacutu R, Fraifeld V. “Systems biology analysis of lung fibrosis-related genes in the bleomycin mouse model”, *Scientific Reports*, 11:19269 (2021), <https://doi.org/10.1038/s41598-021-98674-6>. **IF 4.38, Q2**

Matei IV, Samukange VNC, **Bunu G**, Toren D, Ghenea S, Tacutu R. “Knock-down of odr-3 and ife-2 additively extends lifespan and healthspan in *C. elegans*”, *Aging (Albany NY)*, 13:21040-21065 (2021), <https://doi.org/10.18632/aging.203518>. **IF 5.6, Q2**

* prim co-autor

2. *Capitole de carte*

Tacutu, R., Toren, D., Ursu, E., **Bunu, G.**, Mracica, T.B. (2020). Healthy Biological Systems. In: Sholl, J., Rattan, S.I. (eds) Explaining Health Across the Sciences. *Healthy Ageing and Longevity*, vol 12. Springer, Cham. https://doi.org/10.1007/978-3-030-52663-4_5

3. *Comunicări la conferințe naționale sau internaționale*

Prezentări orale

Bunu G, „SynergyAge: A Curated Database for Interactions of Longevity Genes”, Ending Age-Related Diseases 2020 - August 2020, Online

Bunu G, ”Synergism and antagonism in aging”, The Annual International Conference of the Romanian Society for Biochemistry & Molecular Biology, September 2019, Iași, Romania

Prezentări poster

Bunu G, Ion C, Tăcutu R, “Analysis of synergism and antagonism in aging”, International Perspectives on Geroscience” Israel, September 2019

Bunu G, Ion C, Tăcutu R, “Synergistic and antagonistic interventions between longevity genes - the SynergyAge database”, International Conference on Aging and Disease, October 2018, Nisa, Franța (**Oustanding Poster Award**)

Bunu G, Ion C, Tăcutu R, “SynergyAge: a database of synergistic and antagonistic genes in aging research”, The Annual International Conference of the Romanian Society for Biochemistry & Molecular Biology, September 2018, București, România

Bunu G, Voinea A, Kulaga A, Constantinescu V, Craciun I, Ion C, Tăcutu R, “Gerontomics: a multi-omics prediction system for prioritization of gerontological interventions”, The Annual International Conference of the Romanian Society for Biochemistry & Molecular Biology, June 2017, Timișoara, România

Bibliografie

- Barardo, D., Thornton, D., Thoppil, H., Walsh, M., Sharifi, S., Ferreira, S., Anžič, A., Fernandes, M., Monteiro, P., Grum, T., et al. (2017). The DrugAge database of aging-related drugs. *Aging Cell* 16, 594–597. doi:10.1111/accel.12585.
- Blagosklonny, M. V. (2012). Prospective treatment of age-related diseases by slowing down aging. *Am. J. Pathol.* 181, 1142–1146. doi:10.1016/j.ajpath.2012.06.024.
- Bunu, G., Toren, D., Ion, C.-F., Barardo, D., Sârghie, L., Grigore, L. G., de Magalhães, J. P., Fraifeld, V. E., and Tacutu, R. (2020). SynergyAge, a curated database for synergistic and antagonistic interactions of longevity-associated genes. *Sci. Data* 7, 366. doi:10.1038/s41597-020-00710-z.
- Butler, R. N., Miller, R. A., Perry, D., Carnes, B. A., Williams, T. F., Cassel, C., Brody, J., Bernard, M. A., Partridge, L., Kirkwood, T., et al. (2008). New model of health promotion and disease prevention for the 21st century. *BMJ* 337, a399. doi:10.1136/bmj.a399.
- DeVito, L. M., Barzilai, N., Cuervo, A. M., Niedernhofer, L. J., Milman, S., Levine, M., Promislow, D., Ferrucci, L., Kuchel, G. A., Mannick, J., et al. (2022). Extending human healthspan and longevity: a symposium report. *Ann. N. Y. Acad. Sci.* 1507, 70–83. doi:10.1111/nyas.14681.
- Erol, A. (2007). The functions of ppars in aging and longevity. *PPAR Res.* 2007, 39654. doi:10.1155/2007/39654.
- Fernandes, M., Wan, C., Tacutu, R., Barardo, D., Rajput, A., Wang, J., Thoppil, H., Thornton, D., Yang, C., Freitas, A., et al. (2016). Systematic analysis of the gerontome reveals links between aging and age-related diseases. *Hum. Mol. Genet.* 25, 4804–4818. doi:10.1093/hmg/ddw307.
- Fuentealba, M., Dönertaş, H. M., Williams, R., Labbadia, J., Thornton, J. M., and Partridge, L. (2019). Using the drug-protein interactome to identify anti-ageing compounds for humans. *PLoS Comput. Biol.* 15, e1006639. doi:10.1371/journal.pcbi.1006639.
- Ge, Y., Zhou, M., Chen, C., Wu, X., and Wang, X. (2022). Role of AMPK mediated pathways in autophagy and aging. *Biochimie* 195, 100–113. doi:10.1016/j.biochi.2021.11.008.
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., et al. (2021). PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.* 49, D1388–D1395. doi:10.1093/nar/gkaa971.
- Knyazer, A., Bunu, G., Toren, D., Mracica, T. B., Segev, Y., Wolfson, M., Muradian, K. K., Tacutu, R., and Fraifeld, V. E. (2021). Small molecules for cell reprogramming: a systems biology analysis. *Aging (Albany NY)* 13, 25739–25762. doi:10.18632/aging.203791.
- Matei, I. V., Samukange, V. N. C., Bunu, G., Toren, D., Ghenea, S., and Tacutu, R. (2021). Knock-down of odr-3 and ife-2 additively extends lifespan and healthspan in *C. elegans*.

- Aging (Albany NY)* 13, 21040–21065. doi:10.18632/aging.203518.
- Oughtred, R., Stark, C., Breitkreutz, B.-J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* 47, D529–D541. doi:10.1093/nar/gky1079.
- Steelman, L. S., Chappell, W. H., Abrams, S. L., Kempf, R. C., Long, J., Laidler, P., Mijatovic, S., Maksimovic-Ivanic, D., Stivala, F., Mazzarino, M. C., et al. (2011). Roles of the Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR pathways in controlling growth and sensitivity to therapy-implications for cancer and aging. *Aging (Albany NY)* 3, 192–222.
- Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., et al. (2017). A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* 171, 1437-1452.e17. doi:10.1016/j.cell.2017.10.049.
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102, 15545–15550. doi:10.1073/pnas.0506580102.
- Szkarczyk, D., Santos, A., von Mering, C., Jensen, L. J., Bork, P., and Kuhn, M. (2016). STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* 44, D380-4. doi:10.1093/nar/gkv1277.
- Tacutu, R., Budovsky, A., Yanai, H., and Fraifeld, V. E. (2011). Molecular links between cellular senescence, longevity and age-related diseases - a systems biology perspective. *Aging (Albany NY)* 3, 1178–1191.
- Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., Diana, E., Lehmann, G., Toren, D., Wang, J., et al. (2018). Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Res.* 46, D1083–D1090. doi:10.1093/nar/gkx1042.
- Toren, D., Yanai, H., Abu Taha, R., Bunu, G., Ursu, E., Ziesche, R., Tacutu, R., and Fraifeld, V. E. (2021). Systems biology analysis of lung fibrosis-related genes in the bleomycin mouse model. *Sci. Rep.* 11, 19269. doi:10.1038/s41598-021-98674-6.
- Tower, J. (2009). Hsps and aging. *Trends Endocrinol. Metab.* 20, 216–222. doi:10.1016/j.tem.2008.12.005.
- Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617. doi:10.1093/nar/gkx1089.
- Xie, E., Nadeem, U., Xie, B., D'Souza, M., Sulakhe, D., and Skondra, D. (2022). Using Computational Drug-Gene Analysis to Identify Novel Therapeutic Candidates for Retinal Neuroprotection. *Int. J. Mol. Sci.* 23. doi:10.3390/ijms232012648.
- Xue, H., Xian, B., Dong, D., Xia, K., Zhu, S., Zhang, Z., Hou, L., Zhang, Q., Zhang, Y., and

Han, J.-D. J. (2007). A modular network model of aging. *Mol. Syst. Biol.* 3, 147.
doi:10.1038/msb4100189.