

Primjena koncepta makrokariotipa u istraživanju evolucije stanica kvasca

Ina Mihalj

Mentor: prof. dr. sc. Nenad Pavin

Prirodoslovno-matematički fakultet, Fizički odsjek, Bijenička cesta 32, Zagreb

Datum: 2.1.2024.

Objašnjenja je potreba za proučavanjem mitoze i evolucije kariotipa populacije stanica. Uveden je koncept makrokariotipa i objašnjenja je njegova prednost zbog smanjenja dimenzionalnosti problema praćenja evolucije kariotipa populacije stanica. Razvijene su Monte Carlo simulacije za praćenje evolucije populacije stanica. Opisan je pristup srednjeg polja čija je prednost mogućnost praćenja evolucije velikog broja stanica. Vjerojatnosti raznih ishoda dijeljenja stanica u stohastičkom pristupu povezane su sa stopama u pristupu srednjeg polja. Uspoređene su rezultirajuće populacije dobivene u Monte Carlo simulacijama i pristupu srednjeg polja u početno homogenim populacijama stanica. Izložena je motivacija za praćenje evolucije stanica kvasaca u kontekstu adaptacije na okolinu.

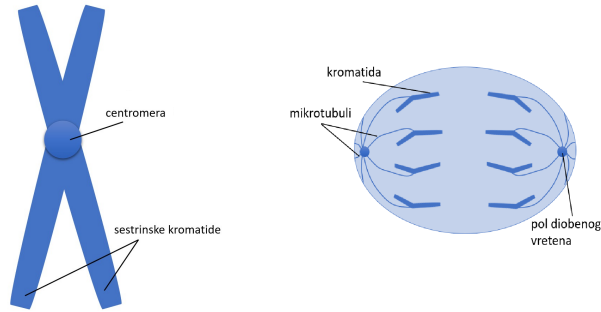
Teorijski uvod

Biološka motivacija

Stanična dioba je proces od temeljne važnosti za živa bića kojim iz postojećih roditeljskih stanica nastaju nove stanice kćeri.^[1] Tijekom diobe stanica formira se diobeno vreteno čija je zadaća simetrična podjela udvostručenog genetskog materijala u novonastale stanice kćeri.^[2] Za vrijeme trajanja diobe stanica genetski materijal tvori strukture koje oblikom podsjećaju na slovo 'X' i nazivaju se kromosomi. Ovakav kromosom sastoji se od dvije sestrinske kromatide, tj. od dvije kopije (slika 1). Tijekom faze životnog ciklusa stanice koja se naziva interfaza (točnije, S-faza interfaze) genetski se materijal udvorstruči - broj kromatida se udvostruči. Dioba završava odvajanjem sestrinskih kromatida, njihovom *segregacijom* u dvije stanice kćeri. Segregacija kromosoma kompleksan je proces jer stanica mora osigurati da točno jedan set kromosoma bude prenesen svakoj od dviju stanica kćeri.

Pojam kariotip odnosi se na potpuni set kromosoma stanice. U slučaju ispravnih dioba identičan kariotip prenosi se na sljedeće generacije stanica kćeri koje se nazivaju klonovi. Osim ispravne diobe, mogu se dogoditi greške u dijeljenju stanice. Primjer takve greške je missegregacija. Do missegregacije dolazi kad se u stanicu kćer prenese više ili manje kromosoma nego što je to slučaj u ispravnoj diobi. Missegregacija mijenja kariotip stanica u idućim generacijama. Također je moguća greška pri dijeljenju stanice zbog koje su veliki dijelovi kromosoma obrisani, udvostručeni ili premješteni s jednog kromosoma na drugi.

Svaka stanica nasljeđuje karakterističan set kromosoma. Stanice koje imaju dvije kopije svakog kromosoma (dva *seta* kromosoma) nazivaju se diploidne, one koje imaju jedan set kromosoma nazivaju se haploidne, one koje imaju tri seta nazivaju se triploidne, itd. Poliploidija je naziv za pojavu više od dva potpuna seta kromosoma u stanici. Ljudske somatske stanice diploidne su i sadrže ukupno 23 para kromosoma. Stanice kvasaca (lat. *Saccharomyces cerevisiae*) sadrže 16 kromosoma.^[3]



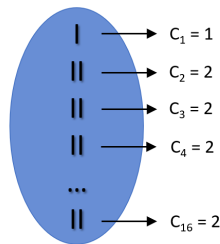
Slika 1: Kromosom i njegov izgled u stanici tijekom diobe. Na lijevom dijelu slike prikazan je kromosom u obliku slova 'X' koji se sastoji od dviju sestrinskih kromatida. Centromera je mjesto na kojem su spojene sestrinske kromatide. Na desnom dijelu slike prikazana je stanica tijekom faze diobe koja se naziva anafaza. Mogu se vidjeti odvojene sestrinske kromatide koje su privučene polovima diobenog vretena pomoću mikrotubula.

Koncept makrokariotipa i primjena na stanicama s maksimalno dvije kopije kromosoma

Kariotip stanice s n različitih kromosoma opisan je vektorom $\vec{K} \equiv (c_1, c_2, \dots, c_n)$, gdje je c_i broj kopija i -tog kromosoma. Postoje dva izazova u teorijskom istraživanju evolucije stanica: veliki broj različitih kariotipa i veliki broj stanica koje populacije stanica mogu sadržavati. To se može ilustrirati na primjeru stanica s maksimalnim brojem kopija kromosoma $L = 6$ i brojem kromosoma $n = 16$. Broj mogućih kariotipa jednak je L^n , što iznosi otprilike 2.8×10^{12} .

Osmišljen je novi pristup koji se temelji na konceptu makrokariotipa.^[5] Makrokariotip je vektor koji sadrži broj kromosoma s određenim brojem kopija, pri tome ne sadržava informaciju o kojim se točno kromosomima radi. Makrokariotip stanice s maksimalnim brojem kopija kromosoma L koje stanica može imati opisan je vektorom $\vec{M}(\vec{K}) \equiv (x_1, x_2, \dots, x_L)$. Ovdje je x_1 broj kromosoma koji imaju jednu kopiju, x_2 broj kromosoma koji imaju dvije kopije, itd. Broj različitih kromosoma po stanici očuvan je jer nulisomije nisu opisane u ovom modelu: $x_1 + x_2 + \dots + x_L = n$. Broj mogućih makrokariotipa jednak je $L(n-1)^{L-1}$. Za primjer naveden u prethodnom odjeljku, broj mogućih makrokariotipa iznosi otprilike 4.6×10^6 . Broj mogućih makrokariotipa manji je od broja mogućih kariotipa za 6 redova veličine. Prednost koncepta makrokariotipa je smanjenje dimenzionalnosti vektorskog prostora makrokariotipa u odnosu na vektorski prostor kariotipa u slučaju kad je $L \ll n$.

Za primjer se može uzeti stanica kvasca kariotipa $\vec{K} = (1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2)$. Ova stanica ima 1 kromosom s jednom kopijom i 15 kromosoma s dvije kopije (slika 2). Odgovarajući makrokariotip ove stanice, u slučaju kad je maksimalni broj kopija $L = 2$, jest $\vec{M}(\vec{K}) = (1, 15)$. Iz prethodnog primjera vidi se da ovako definiran makrokariotip uključuje više različitih kariotipa. To doprinosi smanjenju dimenzionalnosti problema.



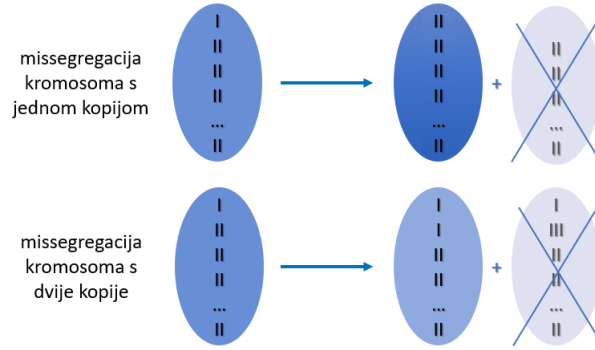
Slika 2: Primjer kariotipa stanice kvasca. Na slici je prikazan jedan kariotip stanice kvasca $\vec{K} = (1, 2, 2, \dots, 2)$. Odgovarajući vektor makrokariotipa za ovu stanicu jest $\vec{M}(\vec{K}) = (1, 15)$. Pogledom na makrokariotip vidi se da je stanica bliska diploidu.

Stohastički pristup evoluciji populacije stanica

Prvi pristup problemu evolucije stanica jest stohastički. Prednost ovog pristupa je to što je moguće pratiti dijeljenje pojedinih stanica i rezultirajuće kariotipe tih stanica - metoda nije aproksimativna.

Životni vijek stanice t_0 ovisi o kariotipu stanice i završava staničnom diobom. U slučaju ispravne stanične diobe koja rezultira dvama klonovima, nakon udvostručenja genetskog materijala slijedi simetrično dijeljenje kromosoma u dvije stanice kćeri. S druge strane, missegregacija znači da dolazi do asimetrije u podjeli udvostručenih kromosoma. U slučaju

kad je maksimalni broj kopija $L = 2$, može se dogoditi missegregacija nakon udvostručenja kromosoma s jednom kopijom ili missegregacija nakon udvostručenja kromosoma s dvije kopije (slika 3).



Slika 3: Primjer moguće missegregacije kromosoma za stanicu roditelja s maksimalno dvije kopije kromosoma. Roditeljska stanica s makrokariotipom $\vec{M} = (1, 15)$ može missegregirati kromosome na dva načina. U gornjem primjeru ishod su dvije stanice kćeri. Prva stanica kći ima makrokariotip $\vec{M} = (0, 16)$. Druga stanica kći završi svoj životni vijek apoptozom jer ima nulisomiju (0 kopija prvog kromosoma). U donjem primjeru ishod su dvije stanice kćeri. Prva stanica kći ima makrokariotip $\vec{M} = (2, 14)$. Druga stanica kći završi svoj životni vijek apoptozom jer ima trisomiju (3 kopije drugog kromosoma).

U Monte Carlo simulacijama ispravna dioba i missegregacija neovisni su događaji s vjerojatnostima prikazanim u tablici 1. Parametri navedeni u tablici zajedno s vremenom početka života stanice definiraju stohastički model za evoluciju makrokariotipa populacije stanica.

Parametar	Oznaka	Definicija
Missegregacija kromosoma s 1 kopijom	p_{m_1}	$1 \times x_1 \times p_0$
Missegregacija kromosoma s 2 kopije	p_{m_2}	$2 \times x_2 \times p_0$
Ispravna stanična dioba	p_d	$1 - p_m - p_v$
Životni vijek stanice	t_0	$1 + 0.005 \times (x_1)^{1.5}$

Tablica 1: Parametri korišteni u stohastičkim simulacijama. Parametri u Monte Carlo simulacijama definirani su za određeni makrokariotip $\vec{M}(\vec{K}) \equiv (x_1, x_2)$. Ovi aproksimativni izrazi vrijede u slučaju kad je missegregacija rijedak događaj, to jest kad je $p_0 \ll 1$. Izabran je $p_0 = 0.0025$.

Nedostatak Monte Carlo simulacija je nemogućnost praćenja velikih populacija stanica. Zbog toga je razvijen pristup srednjeg polja.

Pristup srednjeg polja

Razvijen je pristup srednjeg polja za praćenje evolucije kariotipa populacija stanica.^[4] Ovaj pristup povoljan je za praćenje evolucije populacija s velikim brojem stanica. Rješenja jednadžbe (1) za evoluciju broja stanica određenog makrokariotipa $\tilde{N}(\vec{M})$ realni su brojevi. Broj stanica u populaciji u eksperimentima može se mijenjati samo za diskretne vrijednosti - broj stanica je prirodan broj. S obzirom na to da se pristup srednjeg polja koristi za praćenje velikog broja stanica, ova razlika između teorijskih rješenja i eksperimenta zanemariva je.

Jednadžba za evoluciju broja stanica s određenim makrokariotipom $\tilde{N}(\vec{M})$ jest:

$$\frac{d\tilde{N}(\vec{M})}{dt} = \left[\tilde{k}_d(\vec{M}) - \tilde{k}_m(\vec{M}) - \tilde{k}_a(\vec{M}) \right] \tilde{N}(\vec{M}) + \sum_{l=2}^L \tilde{k}_{m_l}(\vec{M} - \vec{e}_{l-1} + \vec{e}_l) \tilde{N}(\vec{M} - \vec{e}_{l-1} + \vec{e}_l) + \sum_{l=1}^{L-1} \tilde{k}_{m_l}(\vec{M} + \vec{e}_l - \vec{e}_{l+1}) \tilde{N}(\vec{M} + \vec{e}_l - \vec{e}_{l+1}). \quad (1)$$

U prethodnoj jednadžbi $\tilde{k}_d(\vec{M})$ je stopa ispravne diobe, $\tilde{k}_m(\vec{M})$ stopa missegregacije, a $\tilde{k}_a(\vec{M})$ stopa apoptoze za određeni makrokariotip. Kao što je ranije navedeno u tekstu, L je maksimalni broj kopija kromosoma. Vektor \vec{e}_l jest jedinični vektor čija je l -ta koordinata jednaka 1, a ostale su 0. Stope za kariotipe i njima odgovarajuće makrokariotipe jednake su $k_{d,m,a}(\vec{K}) = \tilde{k}_{d,m,a}(\vec{M})$.

Stopa proliferacije stanice β i vrijeme života stanice t_0 odnose se kao $\beta = \ln 2/t_0$. Stopa missegregacije za dani kariotip glasi $k_m(\vec{K}) = (\ln 4)^{-1} p_0 \beta \sum_{i=1}^n c_i$. Lako se vidi da je stopa missegregacije za odgovarajući makrokariotip jednaka $\tilde{k}_m(\vec{M}) = (\ln 4)^{-1} p_0 \beta \sum_{i=1}^L l x_i$. Sume u ovim stopama označavaju ukupan broj kromosoma u stanici. Stopa missegregacije i -tog kromosoma jednaka je $k_{m_i} = \beta p_{m_i} / \ln 4$, gdje je p_{m_i} vjerojatnost missegregacije i -tog kromosoma definirana u stohastičkom pristupu. Stopa apoptoze za stanicu jednaka je $k_a(\vec{K}) = \beta p_a(\vec{K}) / \ln 4$, gdje je $p_a(\vec{K})$ vjerojatnost apoptoze za određenu stanicu definirana u stohastičkom pristupu. U ovom seminaru vjerojatnost apoptoze jednaka je 0 za sve stanice - ne razmatra se mogućnost stanične smrti. Stopa ispravne diobe implicitno je zadana izrazom $\beta = k_d + k_m + k_a$. Izvod jednadžbe i detaljniji opis stopa navedeni su u Dodatku (A).

S obzirom na to da je pristup srednjeg polja vrlo općenit, može se primijeniti na evolucije kariotipa u različitim organizmima, od kvasaca do ljudskih stanica - u tom slučaju može pomoći razmišljanju nastanka tumorskih tkiva kod ljudi.^[5]

Adaptacija stanica kvasca na okolinu

Koncept makrokariotipa i prethodno navedeni pristupi praćenja evolucije populacija stanica mogu se primijeniti za istraživanje prilagodbe kvasaca na okolinu.

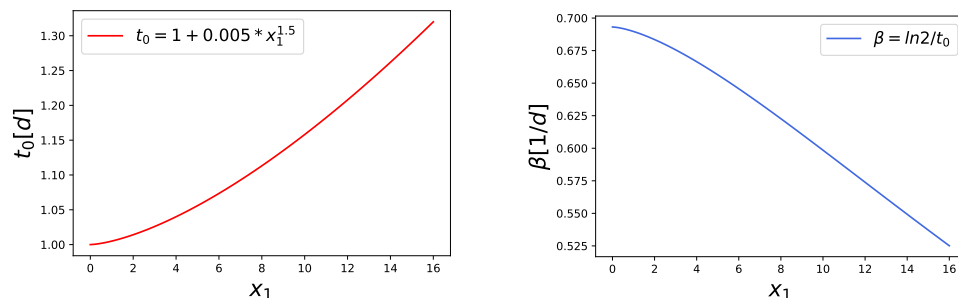
Utjecaj poliploidije na evoluciju ne razumije se potpuno. U in vitro evolucijskim eksperimentima izravno je ispitano može li poliploidija ubrzati evolucijsku prilagodbu kvasaca na okolinu. U usporedbi s haploidima i diploidima, tetraploidi se znatno brže prilagođavaju okolini.^[6] Za procjenu pogodnih prilagodbi okolini koristi se pojam *fitness*. Fitness stanice određen je brojnim parametrima poput duljine staničnog vijeka, metaboličke stope, aktivnosti signaliziranja između susjednih stanica...

Slijedi opis dijela eksperimentalnih rezultata objavljenih u časopisu Nature 2015. godine.^[6] Izolirano je 48 tetraploidnih klonova iz evolucijskih eksperimenata. Uspoređen je fitness izoliranih tetraploidnih klonova s diploidnim klonovima koji su evoluirali tijekom 250 generacija. Unatoč tome što su u početku imali sličan fitness, klonovi koji su nastali iz tetraploidne populacije brže su se prilagođavali okolini nego diploidi. Većina (71%) tetraploidnih klonova izgubilo je kromosome i postali su približno diploidi. Bilo bi zanimljivo znati koji faktori uzrokuju prelazak populacije stanica s tetraploidnog na gotovo diploidni kariotip.

Rezultati

Početno homogena populacija diploidnih stanica kvasca

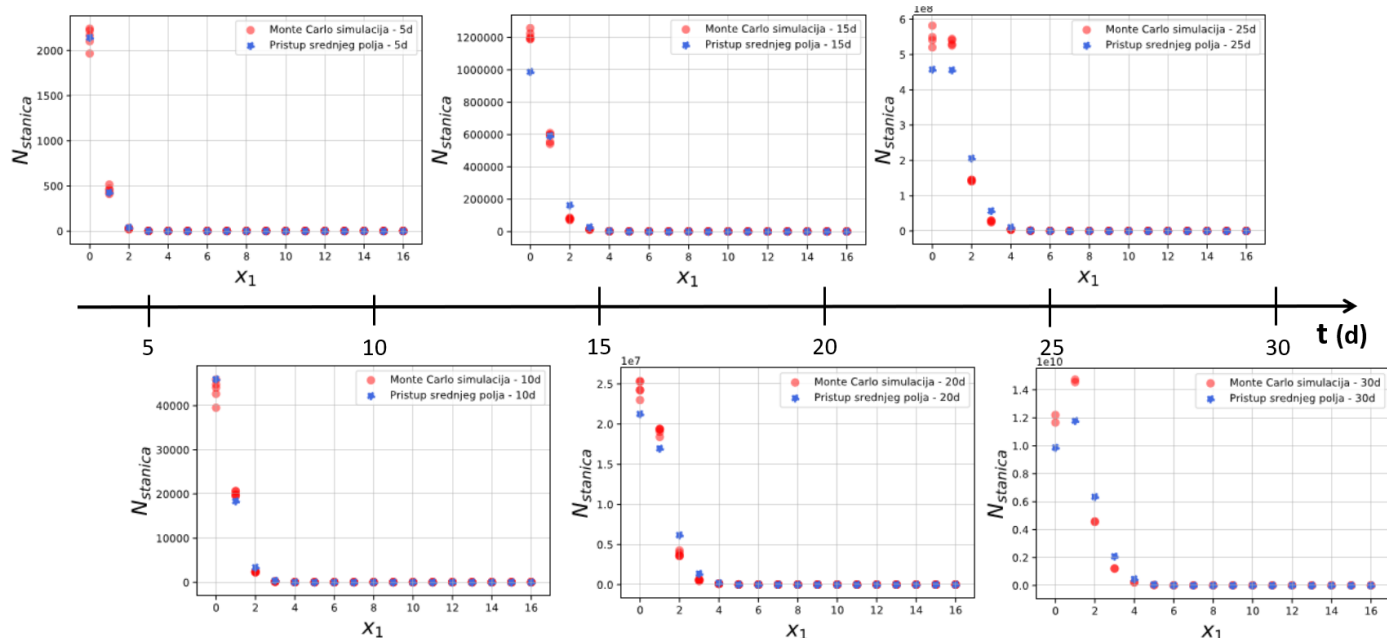
Proučena je evolucija 100 diploidnih stanica kvasca početnog makrokariotipa $\vec{M} = (0, 16)$. Izabrana funkcijska ovisnost životnog vijeka stanica u ovisnosti o kariotipu prikazana je na slici 4. Vjerojatnost ishoda dijeljenja pojedine stanice u ovisnosti o kariotipu stanice prikazana je u tablici 1.



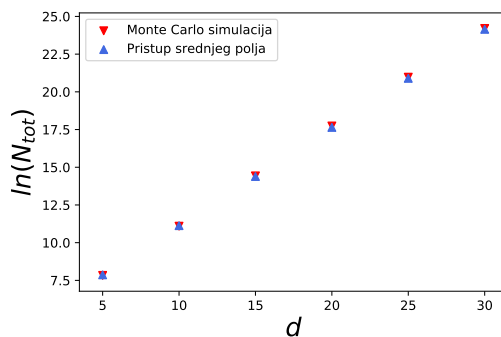
Slika 4: Izabrana funkcijska ovisnost životnog vijeka stanica i stope proliferacije o kariotipu. Na prvom prikazu vidi se da u ovom primjeru diploidne stanice imaju najkraći životni vijek. S druge strane, diploidne stanice imaju najveću stopu proliferacije.

Rezultati Monte Carlo simulacija prikazani su crvenom bojom na slici 5. Izgled populacije nakon određenog broja dana mijenja se od ishoda do ishoda zbog stohastičnosti simulacije. Na slici su prikazani neki od mogućih ishoda evolucije populacije stanica. Na slici 5 plavom je bojom prikazana evolucija iste diploidne populacije 100 stanica koja je predviđena

pristupom srednjeg polja. Trendovi promjene kariotipa u populaciji slažu se u oba pristupa. Nakon 5 dana u populaciji stanica prevladavaju diploidne stanice. Nakon 30 dana u populaciji prevladavaju aneuploidne stanice s makrokariotipom $M = (1, 15)$.



Slika 5: Vremenska evolucija početno homogene populacije diploidnih stanica kvasca. Početna populacija stanica sastoji se od 100 stanica makrokariotipa $\vec{M} = (0, 16)$. Koordinata x_1 označava broj kromosoma s jednom kopijom. Broj kromosoma s dvije kopije jednoznačno je određen jer je broj kromosoma očuvan: $x_1 + x_2 = 16$. Na prvom prikazu vidi se da većina stanica u populaciji nakon 5 dana ima 0 kromosoma s jednom kopijom, većina stanica je diploidna. Nakon 30 dana vidi se da u populaciji prevladavaju stanice koje imaju jedan kromosom s jednom kopijom. Na ovaj način se tijekom vremena mijenja kariotip populacije kvasaca. Trend evolucije kariotipa u pristupu srednjeg polja slaže se s Monte Carlo pristupom.



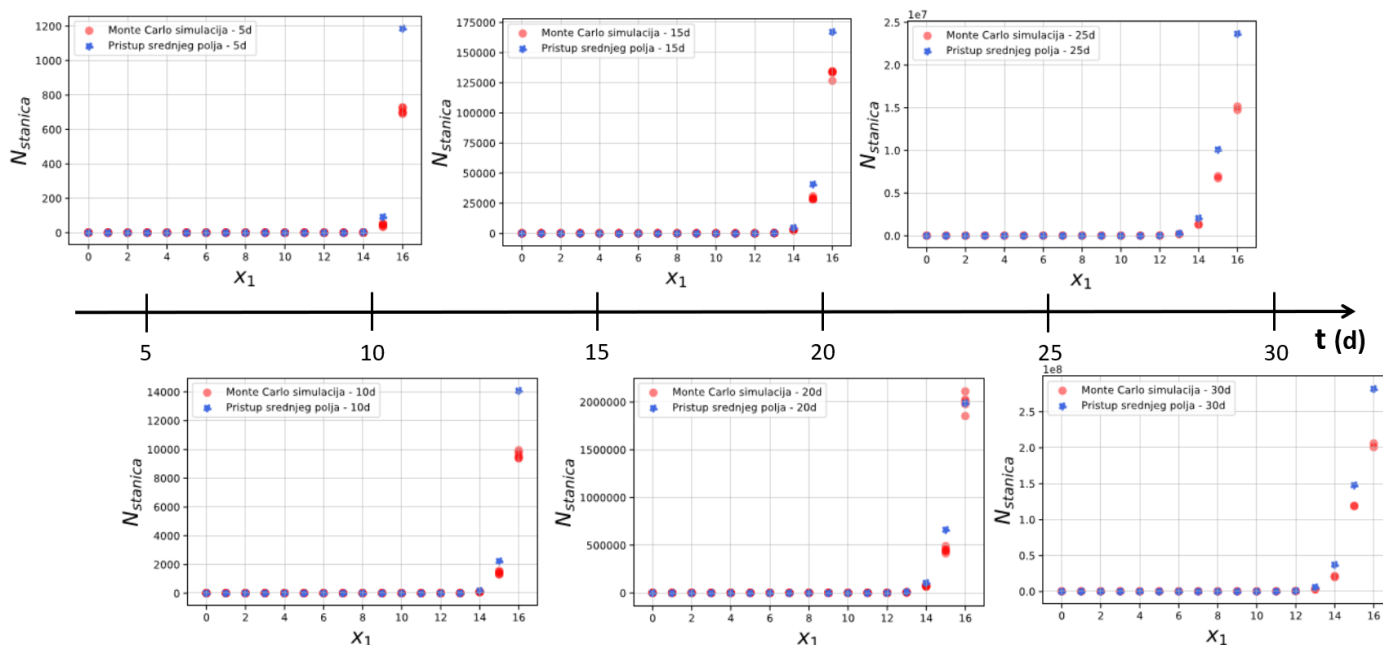
Slika 6: Logaritam ukupnog broja stanica N_{tot} u početno diploidnoj populaciji nakon d dana. Ovaj grafički prikaz služi za procjenu uspjeha stohastičkog pristupa i pristupa srednjeg polja. Ukupan broj stanica u populaciji slaže se u oba pristupa. Ukupan broj stanica u populaciji u Monte Carlo simulaciji ovisi o pojedinom ishodu simulacije. Za bolje procjene slaganja simulacije i pristupa srednjeg polja bilo bi potrebno uzeti prosjek mnogih ishoda Monte Carlo simulacija. Podatci na grafičkom prikazu odgovaraju podacima na slici 5.

Početno homogena populacija haploidnih stanica kvasca

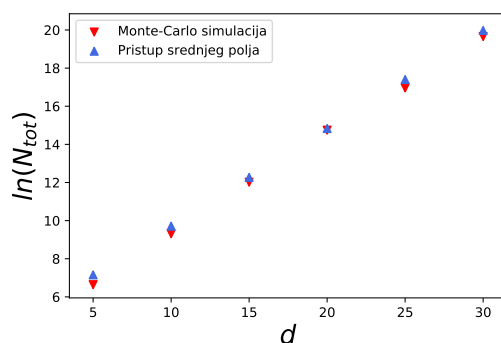
Svaka populacija stanica može se shvatiti kao zbroj homogenih populacija stanica određenih kariotipa. Evolucija bilo koje populacije stanica može se smatrati kao superpozicija evolucija homogenih populacija stanica. Zbog toga je pri razradi

ova dva pristupa dovoljno proučavati evolucije homogenih populacija stanica.

Kao drugi primjer može se pogledati suprotan primjer od prošloga, a to je početno homogena populacija haploidnih stanica. Proučena je evolucija 100 haploidnih stanica kvasca početnog makrokariotipa $\vec{M} = (16, 0)$. Izabrana funkcijska ovisnost životnog vijeka stanica u ovisnosti o kariotipu prikazana je na slici 4. Vjerojatnost ishoda dijeljenja pojedine stanice u ovisnosti o kariotipu stanice prikazana je u tablici 1.



Slika 7: Vremenska evolucija početno homogene populacije haploidnih stanica kvasca. Početna populacija stanica sastoji se od 100 stanica makrokariotipa $\vec{M} = (16, 0)$. Trendovi evolucije populacije stanica slažu se u rezultatima Monte Carlo populacije i u pristupu srednjeg polja. Za preciznu usporedbu ova dva pristupa u budućnosti je potrebno uključiti ishode mnogih Monte Carlo simulacija.



Slika 8: Logaritam ukupnog broja stanica N_{tot} u početno haploidnoj populaciji nakon d dana. Izabrani su nasumični ishodi Monte Carlo simulacija. Za bolje procjene slaganja simulacije i pristupa srednjeg polja bilo bi potrebno uzeti prosjek mnogih ishoda Monte Carlo simulacija. Podatci na grafičkom prikazu odgovaraju podacima na slici 7.

Zaključak

Koncept makrokariotipa pokazao se upješan u smanjivanju dimenzije vektorskog prostora za problem evolucije kariotipa populacije stanica.

Uspoređena su dva pristupa za praćenje evolucije kariotipa populacije stanica. Primijenjene su Monte Carlo simulacije i pristup srednjeg polja na primjeru evolucije populacije kvasaca. Na slikama 5 i 7 prikazana je evolucija početno homogene populacije stanica - početno diploidne i haploidne populacije. Trendovi evolucije populacija stanica slažu se za oba početna

uvjeta. Sa slika 7 i 8 može se zaključiti da konačna populacija dobivena u Monte Carlo simulacijama obično ima manje stanica od onog broja koji predviđa teorija srednjeg polja (slučaj početno haploidne populacije).

U budućnosti je potrebno usporediti pristup srednjeg polja s brojnim ishodima Monte Carlo simulacija zbog otklanjanja mogućih statističkih fluktuacija u rezultatima simulacija. Također je potrebno proširenje modela s maksimalno dvije kopije kromosoma $L = 2$ na maksimalno šest kopija kromosoma $L = 6$ zbog potreba objašnjavanja eksperimentalnih rezultata iznesenih u članku koji je naveden u referenci [6].

Reference

- [1] Martin E., Hine R. A dictionary of biology (6th ed.). Oxford: Oxford University Press (2020).
- [2] Prosser S., Pelletier L. Mitotic spindle assembly in animal cells: a fine balancing act. *Nat Rev Mol Cell Biol* 18, 187–201 (2017).
- [3] Steensels, J., Snoek T., Meersman E., Nicolino M., Voordeckers K., Verstrepen K. Improving industrial yeast strains: Exploiting natural and artificial diversity. *FEMS Microbiology Reviews* 38, (2014).
- [4] I. Ban. The study of microtubules and motor proteins self-organization in the formation of antiparallel bundles of mitotic spindle. PhD dissertation (2023).
- [5] I. Ban, L. Tomašić, M. Trakala, I.M. Tolić and N. Pavin. Proliferative advantage of specific aneuploid cells drives evolution of tumor karyotypes. *Biophys J* 122, 632-645 (2023).
- [6] A. M. Selmecki et al. Polyploidy can drive rapid adaptation in yeast. *Nature* 519, 349–352 (2015).

Dodatak

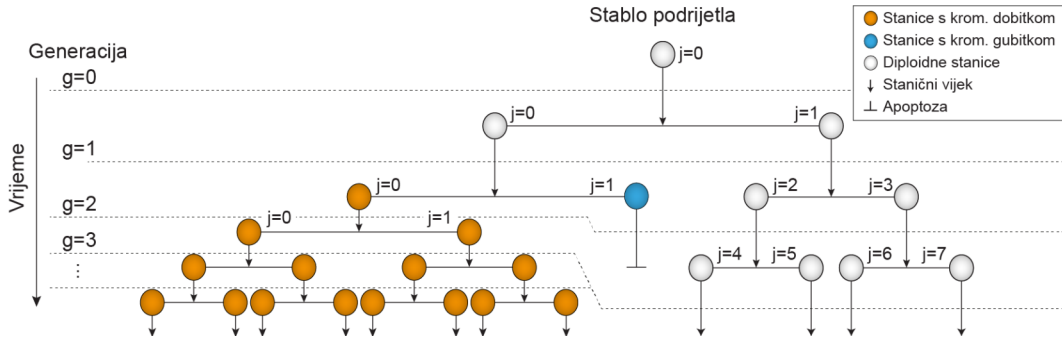
(A) Izvod jednadžbe za evoluciju broja stanica s određenim makrokariotipom

Vremenska evolucija vjerojatnosti da se stanica s kariotipom \vec{K} nalazi na poziciji j u generaciji g u stablu podrijetla (slika 9): [5]

$$\begin{aligned} \frac{dP_{g,j}(\vec{K})}{dt} = & k_d(\vec{K})P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K}) + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} + \vec{e}_i)P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K} + \vec{e}_i) + \\ & + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} - \vec{e}_i)P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K} - \vec{e}_i) - \left[k_d(\vec{K}) + k_m(\vec{K}) + k_a(\vec{K}) \right] P_{g,j}(\vec{K}). \quad (2) \end{aligned}$$

Stanice u generaciji g pojave se kao rezultat dijeljenja stanica u prethodnoj generaciji $g-1$, ali i nestaju prilikom njihove diobe ili apoptoze. Prvi član u jednadžbi (2) opisuje ispravnu diobu stanice u prethodnoj generaciji $g-1$ s kariotipom \vec{K} . Drugi i treći članovi u jednadžbi opisuju slučajeve kad se roditeljske stanice s kariotipom $\vec{K} \pm \vec{e}_i$ missegregiraju na način da izgube ili dobiju jednu kopiju i -tog kromosoma. Pri tome nastanu *dvi* stanice, *jedna* od kojih ima kariotip \vec{K} . Zadnji član opisuje podjelu ili apoptozu stanica u generaciji g na poziciji j . Pojavljuje se s negativnim predznakom jer smanjuje vjerojatnost da se u generaciji g na poziciji j nađu stanice - ti procesi dovode do *nestajanja* stanica iz generacije.

U jednadžbi (2) $k_d(\vec{K})$ je stopa ispravne diobe, $k_a(\vec{K})$ stopa apoptoze, $k_m(\vec{K}) = \sum_{i=1}^n k_{m_i}(\vec{K})$ stopa missegregacije, gdje je $k_{m_i}(\vec{K})$ stopa missegregacije i -tog kromosoma. Veza između stopa missegregacije i apoptoze te pripadnih vjerojatnosti jest: $k_{a,m_i} = \beta p_{a,m_i} / \ln 4$. Stopa proliferacije stanice β povezana je s životnim vijekom stanice t_0 : $\beta = \ln 2 / t_0$.



Slika 9: Stablo podrijetla s nekoliko generacija stanica. Stanica roditelj je diploidna stanica u generaciji $g = 0$. U slučaju ispravne diobe, stanice kćeri su klonovi stanice roditelja. U slučaju missegregacije stanice kćeri mogu imati višak ili manjak kromosoma.[4]

Broj stanica s određenim kariotipom iznosi:

$$N(\vec{K}) \equiv \sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g,j}(\vec{K}), \quad (3)$$

gdje je $P_{g,j}(\vec{K})$ vjerojatnost da se u generaciji g na mjestu j pojavi stanica s kariotipom \vec{K} (slika 9). Vremenska evolucija tog broja stanica računa se neovisno o poziciji u stablu podrijetla:

$$\begin{aligned} \frac{d}{dt} \left(\sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g,j}(\vec{K}) \right) = & k_d(\vec{K}) \sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K}) + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} + \vec{e}_i) \sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K} + \vec{e}_i) + \\ & + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} - \vec{e}_i) \sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K} - \vec{e}_i) - \left[k_d(\vec{K}) + k_m(\vec{K}) + k_a(\vec{K}) \right] \sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g,j}(\vec{K}). \quad (4) \end{aligned}$$

U jednadžbi (4) lako se prepozna izraz za broj stanica (3). Za računanje drugog i trećeg člana u jednadžbi (4) supstituiru se $f = g - 1$:

$$\sum_{g=0}^{\infty} \sum_{j=0}^{2^g-1} P_{g-1, \lfloor \frac{j}{2} \rfloor}(\vec{K}) = \sum_{f=0}^{\infty} \sum_{j=0}^{2 \cdot 2^f-1} P_{f, \lfloor \frac{j}{2} \rfloor}(\vec{K}). \quad (5)$$

Suma po indeksu j dijeli se na dvije sume, jednu za $j = 2a$, a drugu za $j = 2b + 1$. Sada vrijedi:

$$\sum_{f=0}^{\infty} \sum_{a=0}^{2^f-1} P_{f, \lfloor \frac{j}{2} \rfloor}(\vec{K}) + \sum_{f=0}^{\infty} \sum_{b=0}^{2^f-1} P_{f, \lfloor \frac{j}{2} \rfloor}(\vec{K}) = 2N(\vec{K}). \quad (6)$$

Uvrštavanjem posljednjeg izraza u jednadžbu (4) nakon preraspodjele članova u jednadžbi dobije se izraz za vremensku evoluciju broja stanica određenog *kariotipa*:

$$\begin{aligned} \frac{dN(\vec{K})}{dt} = k_d \cdot 2N(\vec{K}) + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} + \vec{e}_i) \cdot 2N(\vec{K} + \vec{e}_i) + \frac{1}{2} \sum_{i=1}^n k_{m_i}(\vec{K} - \vec{e}_i) \cdot 2N(\vec{K} - \vec{e}_i) + \\ - \left[k_d(\vec{K}) + k_m(\vec{K}) + k_a(\vec{K}) \right] N(\vec{K}). \quad (7) \end{aligned}$$

Veza između broja stanica s određenim kariotipom i makrokariotipom jest:

$$\tilde{N}(\vec{M}) = \frac{1}{x_1! x_2! \dots x_L!} \times \sum_{perm.} N(\vec{K}). \quad (8)$$

Za dobivanje jednadžbe koja opisuje vremensku evoluciju broja stanica određenog *makrokariotipa* potrebno je sumirati jednadžbu (7) po svim kromosomskim permutacijama množenjem jednadžbe s faktorom $\frac{1}{x_1! x_2! \dots x_L!} \times \sum_{perm.}$. Stope koje su funkcije makrokariotipa imaju iste vrijednosti kao stope koje su funkcije kariotipa koji pripadaju tom makrokariotipu: $\tilde{k}_{d,m,a}(\vec{M}) = k_{d,m,a}(\vec{K})$, gdje je $\vec{M} = \vec{M}(\vec{K})$.

Stopa missegregacije za određeni kariotip iznosi $k_m(\vec{K}) = (\ln 4)^{-1} p_0 \beta \sum_{i=1}^n c_i$. Izraz stope za makrokariotip glasi: $k_m(\vec{K}) = (\ln 4)^{-1} p_0 \beta \sum_{i=1}^n l x_i$. Sume u oba izraza označavaju ukupan broj kromosoma.

U slučaju drugog s desne strane jednadžbe (7) vrijedi:

$$\tilde{N}(x_1, \dots, x_{c_i} - 1, x_{c_i} + 1, \dots, x_L) = \frac{1}{x_1! \dots (x_{c_i} - 1)! (x_{c_i} + 1)! \dots x_L!} \times \sum_{perm.} N(\vec{K} + \vec{e}_i). \quad (9)$$

Koristi se identitet: $\sum_{\nu=1}^L \delta_{\nu, c_i} = 1$ i definicija $x_{\nu} \equiv \sum_{i=1}^n \delta_{\nu, c_i}$. Jednadžba (7) sada poprima oblik:

$$\begin{aligned} \frac{d\tilde{N}(\vec{M})}{dt} = \left[\tilde{k}_d(\vec{M}) - \tilde{k}_m(\vec{M}) - \tilde{k}_a(\vec{M}) \right] \tilde{N}(\vec{M}) + \frac{p_0}{\ln 4} \sum_{\nu=1}^{L-1} (\nu + 1)(x_{\nu+1} + 1) \beta (\vec{M} - \vec{e}_{\nu} + \vec{e}_{\nu+1}) \tilde{N}(\vec{M} - \vec{e}_{\nu} + \vec{e}_{\nu+1}) + \\ + \frac{p_0}{\ln 4} \sum_{\nu=2}^L (\nu - 1)(x_{\nu-1} + 1) \beta (\vec{M} - \vec{e}_{\nu-1} - \vec{e}_{\nu}) \tilde{N}(\vec{M} - \vec{e}_{\nu-1} - \vec{e}_{\nu}). \quad (10) \end{aligned}$$

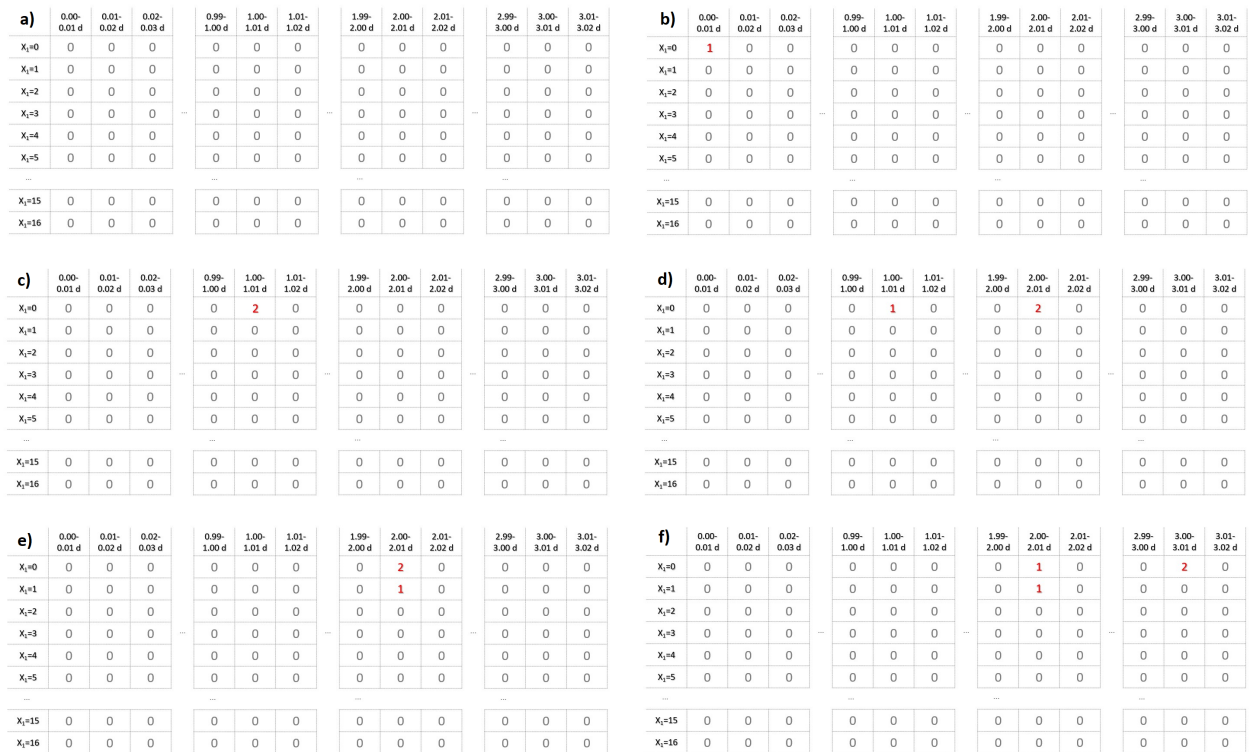
Nakon uvođenja supstitucija u sumama u jednadžbi (10) i prepoznavanja definicija za stope missegregacije dobije se konačni oblik jednadžbe (1).

(B) Opis Monte Carlo simulacije

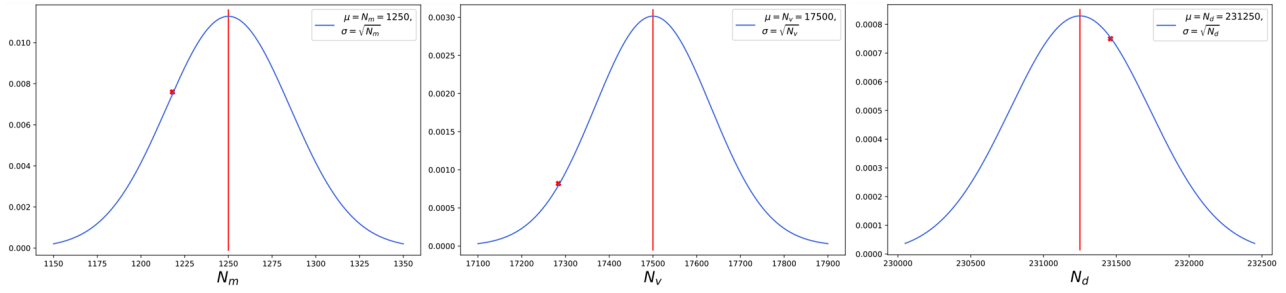
Zbog nepraktičnosti označavanja kariotipa i praćenja diobe pojedinih stanica stohastičkoj simulaciji evolucije velikih populacija stanica pristupilo se na sljedeći način.

U prvom dijelu simulacije definirane su vjerojatnosti triju mogućih ishoda podjele *pojedine* stanice (Tablica 1). Na slici 10 prikazan je princip rada Monte Carlo simulacije. U općenitom slučaju životni vijek stanice ovisi o kariotipu, što znači da će različite stanice početi svoj životni vijek u različitim trenucima. S obzirom na to da je životni vijek prosječne stanice otprilike 1 dan (1d), u simulaciji se koristi preciznost od 0.01d. Greška koja se može propagirati tijekom 50 dana u ovom je slučaju 0.25d. Preciznost se vrlo lako može modificirati na drugu vrijednost, ali se u provjerama pokazalo da ne utječe značajno na točnost rezultata. Nakon dugo vremena, u populaciji se pojavi velik broj stanica određenog makrokariotipa. Određena je granica dovoljno *velikog* broja stanica (npr. 100000).

U drugom dijelu simulacije stanice određenog makrokariotipa ne dijele se jedna po jedna, nego se diobi *velikog* broja stanica pristupa na drugi način. Za primjer se može uzeti populacija od $N_0 = 250000$ stanica makrokariotipa ($x_1 = 2, x_2 = 14$). Odrede se vjerojatnosti ishoda dijeljenja stanice s ovakvim makrokariotipom (Tablica 1). Zatim se odredi očekivana vrijednost broja stanica koje su missegregirale kromosom s 1 kopijom $N_m = N_0 \times p_{m_1} = 1250$, očekivana vrijednost broja stanica koje su missegregirale kromosom s 2 kopije $N_v = N_0 \times p_{m_2} = 17500$ i koje su rezultirale ispravnom staničnom diobom $N_d = N_0 \times p_d = 231250$. Nasumičnim izabiranjem vrijednosti varijable koja je normalno raspodijeljena oko očekivane vrijednosti N_m odredi se koliko stanica missegregira kromosom s 1 kopijom. Slično se napravi i za preostala dva ishoda (slika 11). Na taj način zadržava se stohastičnost procesa i proces se ubrzava.



Slika 10: Prvi dio simulacije. Na slici je prikazan početak jednostavnog ($t_0 = 1d$) procesa diobe stanice u Monte Carlo simulaciji. Algoritam radi tako da se generira matrica $M_{17 \times 100d}$, gdje d označava broj dana. Redaka ima 17 i obuhvaćaju sve moguće makrokariotipe. U verziji s maksimalno dvije kopije kromosoma $L = 2$ za vektor makrokariotipa vrijedi $\vec{M} = (x_1, x_2)$ i $x_1 + x_2 = 16$. Zbog toga je cijeli makrokariotip stanice zadan jednom koordinatom (x_1). U primjeru na ovoj slici sve stanice imaju jednako vrijeme života, neovisno o makrokariotipu. Na slici a) vidi se dio generirane matrice. Na slici b) vidi se kako matrica izgleda kad se zada početni uvjet jedne diploidne stanice kvasca. Zatim se - pomoću vjerojatnosti definiranih u Tablici 1 - odredi ishod diobe te stanice. Na slici c) vidi se ishod ispravne diobe. Dvije nastale stanice dijele se pojedinačno. Na slici d) vidi se da je ishod prve diobe ispravno dijeljenje, a na slici e) vidi se da je ishod druge diobe missegregacija. Pogledom na matricu na slici e) vidi se da nakon 2 dana u populaciji postoje 3 stanice, dvije od kojih imaju makrokariotip ($x_1 = 0, x_2 = 16$) i jedna od kojih ima makrokariotip ($x_1 = 1, x_2 = 15$). Na slici f) vidi se nastavak evolucije populacije stanica. Za nedovoljno *velik* broj stanica u polju matrice stanice se dijele *pojedinačno*.



Slika 11: Drugi dio simulacije. Na slici su prikazani ishodi dioba 250000 stanica. Na prvom prikazu nasumično je odabran broj stanica koje missegregiraju kromosom s jednom kopijom (1218). Na drugom prikazu nasumično je odabran broj stanica koje missegregiraju kromosom s dvije kopije (17284). Na trećem prikazu nasumično je odabran broj stanica koje svoj životni vijek završe ispravnom staničnom diobom (231460). Naravno, potrebno je osigurati da vrijedi $N_0 \geq N_{m,rand} + N_{v,rand} + N_{d,rand}$.

U ovdje opisanom primjeru (slika 4) vrijedi $N_0 = 250000$, $N_{m,rand} + N_{v,rand} + N_{d,rand} = 249962$. Preostalih 38 stanica dijeli se *pojedinačno* jer broj preostalih stanica u polju matrice ne premašuje granicu dovoljno *velikog* broja stanica.