Povijesni i aktualni pregled koncepta gena

Babarović, Frane

Undergraduate thesis / Završni rad

2014

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:781902

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**



Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb







SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO – MATEMATIČKI FAKULTET
BIOLOŠKI ODSJEK

POVIJESNI I AKTUALNI PREGLED KONCEPTA GENA

SEMINARSKI RAD

Frane Babarović
Preddiplomski studij molekularne biologije
(Undergraduate Study of Biology)
Mentor: prof. dr. sc. Kristian Vlahoviček

Zagreb, 2014.

SADRŽAJ

1. POVIJESNI PREGLED KONCEPTA GENA

1.1. Pregled koncepta gena 1860 – 1900-e ……..4.

1.2. Pregled koncepta gena: 1910-ih godina ……..5.

1.3. Pregled koncepta gena 1940-ih godina ……..6.

1.4. Pregled koncepta gena 1950-ih godina ……..7.

1.5. Pregled koncepta gena 1960-ih godinav……..7.

1.6. Pregled koncepta gena 1970-ih – 1980-ih godina.……...8.

1.7. Pregled koncepta gena 1990-e – 2000-e ……...8.

2. POJMOVI KOJI KARAKTERIZIRAJU GENSKU AKTIVNOST

2.1. Regulacija aktivnosti gena ……..10.

2.2. Preklapanje ……..11.

2.3. Prikrajanje ……..11.

2.4. Trans – prikrajanje ……..12.

3. DOPRINOS ENCODE PROJEKTA DEFINIRANJU GENA

3.1. Transkripti koji prethodno nisu anotirani ……..13.

3.2. Neanotirani i alternativni TSS-ovi ……...14.

3.3. Raspodjela regulacijskih elemenata u genomu ……..14.

3.4. Pseudogeni ……...15.

3.5. Nekodirajuće RNA molekule ……..17.

3.6. Genski i intergenski: postoji li razlika? ……...18.

4.Zaključak

4.1. ENCODE definicija gena ……...19.

4.2. Zaključak i osobno viđenje problematike ……...19.

5. LITERATURA

6. SAŽETAK ……...22.

7. SUMMARY ……..22.

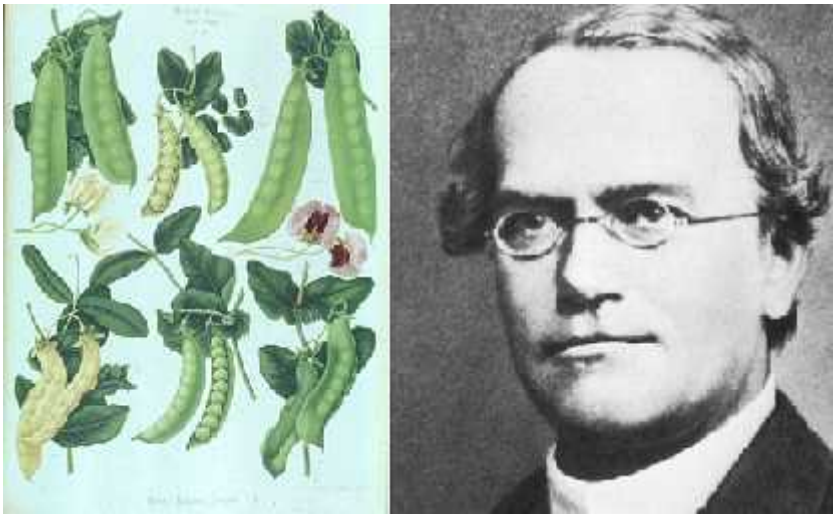
1. Povijesni pregled koncepta gena

1.1. Pregled koncepta gena 1860 – 1900-e

Koncept pojma gen se mijenjao i postajao sve kompleksniji od trenutka kada je prvi put predložen. Danas postoje mnoge definicije, iako se većina njih i dalje referira na pretpostavke postavljene u samom početku istraživanja istog, odnosno da definicija gena u sebi sadrži determinaciju određenog fenotipa organizma i mogućnost i karakteristike nasljeđivanja datog fenotipa (Gerstein, 2007).

Koncept gena prvi je pretpostavio Gregor Mendel 1860-ih godina. Gregor Mendel nikada nije upotrijebio riječ gen, ali je koristio pojmove poput *faktora* ili *elementa* kako bi obrazložio razloge nastanka i nasljeđivanja različitih fenotipskih karakteristika u potomcima. Mendel je eksperimentalno utvrdio da se tijekom križanja biljaka neka svojstva, poput visine biljke ili boje cvijeta (Sl. 1), u potomcima pojavljuju sa statističkom točnošću i da su nasljeđena poput određenih entiteta čiji fizički karakter, zbog eksperimentalne ograničenosti, nije mogao predvidjeti. Drugim riječima, Mendel je samo predvidio pravilnost nasljeđivanja određenih faktora i činjenicu da nasljeđivanje fenotipskih svojstava ima fizičku i mjerljivu pozadinu.

"Različite karakteristike dviju biljaka mogu jedino biti posljedica razlike postojanja i grupiranja elemenata koji intereagiraju u germinativnoj stanici." (Mendel, 1866).



Slika 1. Gregor Mendel

Iako Mendelov rad nije imao snažan utjecaj na tadašnju znanstvenu zajednicu, gotovo 40 godina nakon njegovih istraživanja, tri znanstvenika (Hugo de Vries, Carl Correns i Erich von Tschermak) su revidirala Mendelova istraživanja i od zaborava izbavili njihove rezultate.

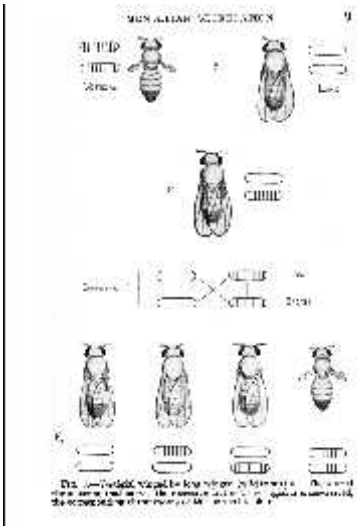
Pojam *gen* prvi je upotrijebio danski biolog Wilhelm Johannsen kako bi objasnio fizički entitet koji određuje nasljeđivanje određenih karakteristika nekog organizma. On je objasnio da je iskoristio riječ *gen* kao apstraktni pojam kako bi najlakše objasnio ekperimentalno predviđena svojstva fizičkog entiteta, odnosno,

pojave u germinativnim stanicama. Njime se najbolje može objasniti biološki fenomen da su mnoge karakteristike organizma određene u germinativnim stanicama pomoću specifičnih uvjeta, odnosno determinanti koje se u njima pojavljuju na jedinstven i odvojen način.

Vidljivo je da je u ranim godinama svog postojanja gen egzistirao kao apstraktan koncept koji je u sebi sadržavao rezultate eksperimentalnih istraživanja, a kojim je ustvrđeno njegovo postojanje i važnost u nasljeđivanju fenotipa. Zaključak koji bi iz današnje perspektive gledanja sumirao sva tadašnja nagađanja i evaluaciju pojma gena jest da je fenotip determiniran genotipom. (The Gene: A Needed Revolution; Craig Holdrege; A publication of The Nature Institute)

1.2. Pregled koncepta gena: 1910-ih godina

Sljedeći korak u razvoju koncepta gena uslijedio je istraživanjima Thomasa Morgana i njegovog laboratorija na segregaciji mutacija u organizmu *Drosophila melanogaster*. Svojim istraživanjima potkrijepili su Mendelova opažanja, ali su također mogli, s obzirom na napredak znanosti, (ponajprije "explaining Mendelian phenomenon by means of manoeuvers of the chromosomes") ponuditi i mehanizme kojima se odvijaju predviđene jedinice nasljeđivanja. Koristeći mutante različitih fenotipa *Drosophile melanogaster* i istražujući principe njihove segregacije iz generacije u generaciju križanih jedinki, Morgan i suradnici su ne samo potvdili Mendelova nagađanja, već su ih i proširili. Istraživanja su se bazirala na međusobnoj povezanosti i određivanju međugenskih, ali I drugih utjecaja na segregaciju određenih gena, a samim time i svojstava do tada determiniranih gena unutar, kako su ih oni nazivali, Mendelovih regija (kod *Drosophile* su u vremenu pisanja citiranog rada postojale, primjerice, četiri Mendelove regije). Njihovi eksperimentalni podaci mogli su se objasniti modelom koji gene gleda kao entitete koji se nalaze na kromosomima, a koji se nasljeđuju zajedno u ovisnosti o njihovoj međusobnoj udaljenosti (Sl. 2; Morgan i suradnici, 1915).



Slika 2. Morganov prikaz segregacije promatranih fenotipskih svojstva, te vjerojatnost i kombinacije prilikom nasljeđivanja određenog svojstva.

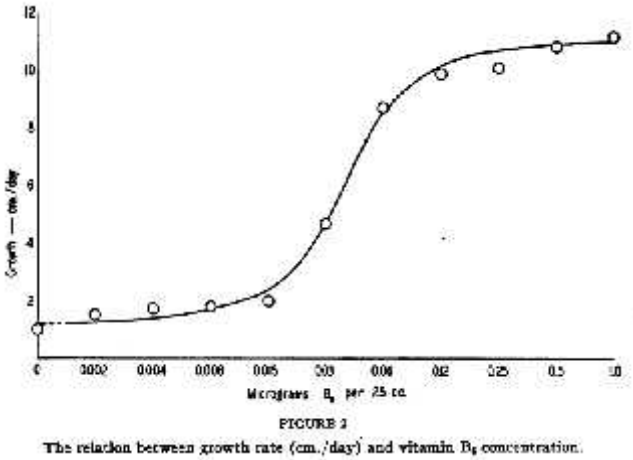
U stalnoj fluktuaciji poimanja koncepta gena, ovo desetljeće predstavlja odmak od kasnijih nastojanja što točnijeg utvrđivanja položaja gena, te ostavljanja njegovog postojanja na apstraktnoj razini. Pojam apstraktna razina najkraće pojašnjava niz eksperimenata u kojima je gen jednostavno pretpostavljen kao lokus na kromosomima, a sve ostalo što se moglo zaključiti o njemu proizlazi iz njegova ponašanja naspram drugih gena tokom rekombinacije i mutacije.

U ovom periodu s karakterističnom znanstvenom tendencijom ka definiranju mehanizama, tj. potvrđivanja "Mendelova nasljeđivanja" gena, konačnu potvrdu dala je Barbara McClintock koja donosi zaključak da spareni kromosomi koji su heteromorfni u dvima regijama izmjenjuju svoje dijelove u isto vrijeme kada izmjenjuju gene povezane s njima (McClintock, 1931).

1.3. Pregled koncepta gena 1940-ih godina

Definiciju pojma gen 1940-ih godina uvelike je okarakteriziralo istraživanje znanstvenika Georgea W. Bedlea i Edwarda L. Tatuma. Svoj su rad započeli na organizmu *Drosophila melanogaster*, a nakon što su uvidjeli veliku kompleksnost samog organizma, svoje istraživanje su nastavili na modelnom organizmu *Neurospora crassa* (http://www.genomenewsnetwork.org/resources/timeline/1941_Beadle_Tatum.php).

Bedle i Tatum su rengenskim zračenjem uzrokovali genetske mutacije, od kojih su neke onemogućavale sporama *Neurospora crassa* da sintetiziraju određene nutrijente, a time i da rastu. Potom su u medij dodavali jedan po jedan suplement za koji su pretpostavljali da postoji mutacija te su time determinirali koji enzim nedostaje. Npr., Mutaciju gena koji odgovara za sintezu enzima u putu biosinteze vitamin B6 ustvrdili su dodavajući vitamin B6 u podlogu i detekcijom rasta spora s dodatkom sve veće koncentracije vitamina B6 (Sl. 3). Sparivanjem mutirane jedinke s normalnom jedinkom pokazali su da je metabolički defekt recesivno svojstvo, a time su dokazali alternaciju određenog gena (Beadle i Tatum, 1941).



Slika 3. Ovisnost stupnja rasta jedinke o koncentraciji dostupnog vitamina B6

Misao vodilja njihova zajedničkog rada je ustvrđivanje povezanosti pojma gena i njihove funkcije, odnosno traženje poveznice između samog gena i "genskog produkta". Konačna tvrdnja koja je postavljena tokom njihova rada i koja je svojedobno definirala pojam gena bila je da jedan gen odgovaran za nastanak jednog enzima (Beadle i Tatum, 1941**).**

1.4. Pregled koncepta gena 1950-ih godina

Znanstvena težnja za identificiranjem kemijske i fizikalne osnove nasljeđivanja, odnosno determinacijom tipa molekule kojoj pripadaju geni, tekla je kroz nekoliko desetljeća uz nekolicinu bitnih otkrića. 1926. godine, istraživanje mehanizma genetičkog nasljeđivanja došlo je do molekularne razine. Prethodna otkrića uputila su znanstvenike na kromosome koji se nalaze u jezgri. Ali potraga za tipom molekule kojem pripada genetski materijal još je uvijek bilo pitanje bez odgovora.

Hermann J. Muller je 1926. godine započeo s nizom eksperimenata koji su u konačnici dali ideju o utjecaju radioaktivnog zračenja na gene, odnosno potpomogli su Mulleru da pojedine gene shvati kao odvojene molekularne jedinice. 1928. godine Frederick Griffith je među prvima postavio eksperiment i doveo do otkrića procesa transformacije kod bakterija, no kemijska priroda materijala prenešenog tokom transformacije je ostala nepoznata sve do 1944. godine. Članak *"Studies on the chemical nature of the substance inducing transformation of Pneumococcal types"* Oswalda Theodorea Averya Jr. iz 1944. Godine (Sl. 4) doveo je do zaključka da se radi o deoksiribonukleinskoj kiselini. (Avery i suradnici, 1944.)

3. The data obtained by chemical, enzymatic, and serological analyses together with the results of preliminary studies by electrophoresis, ultracentrifugation, and ultraviolet spectroscopy indicate that, within the limits of the methods, the active fraction contains no demonstrable protein, unbound lipid, or serologically reactive polysaccharide and consists principally, if not solely, of a highly polymerized, viscous form of desoxyribonucleic acid.

Slika 4. Dio zaključka iz originalnog rada Oswalda Theodorea Averya (Avery i suradnici, 1944.)

1.5. Pregled koncepta gena 1960-ih godina

Nekolicina je ekeperimenata i zaključaka pridonijela shvaćanju gena kao koda koji se nalazi na nukleinskoj kiselini, a u konačnici daje fukncionalan produkt; 1953. godine James D. Watson i Francis D. Crick pridonijeli su razjašnjenju načina na koji DNA funkcionira kao molekula nasljeđivanja rješavanjem trodimenzionalne strukture DNA (http://www.thehistoryblog.com/archives/25193). Važno otkriće bilo je i dešifriranje genetičkog koda. Iako se ova otkrića nisu direktno ticala pojma gena, svakako su pomogla u smještanju gena, kako mehanicistički, tako i konceptualno u priču života. Doprinos problematiziranju koncepta gena svakako ima i Francis Crick koji u svome radu *"On Protein Synthesis"* iznosi pregled dotadašnjeg znanja o genima, produktima gena te protoku informacija u ekspresiji gena iz nukleinske kiseline do proteina. To se smatra početkom sagledavanje pojma *centralne biološke*

dogme. Centralna biološka dogma sagledava protok genetičke informacije od molekule DNA do proteina (Sl. 5).



Slika 5. Centralna biološka dogma. (Koonin E.V., 2012.)

Međutim, bez obzira na brojna nova otkrića koja su osvijetlila pravu biološku ulogu i definiciju gena, postojali su određeni problemi, posebice s uklapanjem gena u koncept centralne biološke dogme. S obzirom na komplikacije, napravljene su i određene iznimke, npr. bilo je poznato da određeni geni ne kodiraju za proteinske produkte, već za funkcionalne RNA molekule, poput rRNA i tRNA, a isto tako bilo je poznato i RNA podrijetlo genetskog materijala nekih RNA virusa.

1.6. Pregled koncepta gena 1970-ih – 1980-ih godina.

1970-ih godina u molekularnoj biologiji se razvilo sekvenciranje molekule DNA. Prvi sekvencirani gen i genom prvog potpuno sekvenciranog organizma bio je na primjeru bakteriofaga MS2. Sekvenciranje u kombinaciji sa znanjem genetičkog koda dovelo je do novih spoznaja, kao npr. kako su geni organizirani i eksprimirani.

S obzirom na to da je sve veći broj sekvenciranih gena i genoma, a i paralelan razvoj računarne biologije i kompjuterskih metoda, došlo je do razvoja algoritama za identifikaciju gena prema njihovoj sekvenci. Sve veća količina sekvencirane DNA i sve bolja točnost algoritama za predikciju gena stvorila je novu podkategoriju u definiciji gena, a to je koncept *nominalnog gena*, odnosno gena koji se više definira predviđenom sekvencom u genomu, nego kao genetički lokus odgovoran za fenotip (Griffiths i Stotz, 2006.).

1.7. Pregled koncepta gena 1990-e – 2000-e

Današnje, najnovije definicije gena nastavljaju se kretati u smjeru gdje je glavna eksperimentalna pozadina za definiciju gena računalna biologija, odnosno računarne metode obrade sekvenciranih genoma. Stoga, postoji nekolicina postavljenih definicija.

Human Genome Nomenclature Organisation definira gen kao "segment DNA koji doprinosi fenotipu ili funkciji. U odsustvu određenog fenotipa ili funkcije, gen može biti karakteriziran skevencom, transkriptom ili homologijom." *Sequence Ontology Consortium* definira gen kao"regiju u genomskoj sekvenci koju je moguće locirati, koja odgovara definiciji gena kao jedinice nasljeđivanja, te je povezana s regulatornim regijama, transkribiranim regijama i/ili nekim drugim funkcionalnim regijama u genomu". *Esemble* u definiciju gena uvrštava i alternativno prikrajanje gena te sve transkripte nastale alternativnim prikrajanjem definira pripadnosti jednom genu.

Razvoj računalnih metoda omogućio nam je da raspoznamo brojnost I raznolikost elemenata molekule DNA, a koji su bitni za njenu funkcionalnost. U dotadašnjim istraživanjima nije bilo moguće sagledati genom u njegovoj cjelovitosti, te je pomak na sveobuhvatna istraživanja genoma istaknuo mnogobrojnost elemenata DNA koji su bili izostavljeni u dotadašnjem eksperimentalnom radu.

2. Pojmovi koji karakteriziraju gensku aktivnost

Tijekom istraživanja kojima se direktno i indirektno otkrivala biološku definiciju gena izronilo je na površinu nekoliko pojmova i bioloških entiteta koja još od tada pa sve do danas kompliciraju stvaranje jedinstvene definicije gena.

Tokom istraživanja koja su pridonjela poimanju gena provlači se ideja da je gen moguće definirati na jedinstven i jednoznačan način. Jedna od ideja koja je znanstvenicima dugi niz godina predstavljala sigurno tlo za početak rekonstrukcije pojma gen je ideja o njegovoj kontinuiranosti unutar genoma. Istraživanja regulacije genske aktivnosti, kao I otkriće preklapanja, prekrajanja i trans-prekrajanja uvjerili su nas u suprotno. S današnjeg stajališta gena, koje je objašnjeno u zadnjem dijelu seminara, povijesna viđenja I definiranja gena bila su previse robusna I s previše jednostavnom i naivnom upornošću u njegovo definiranje.

2.1. Regulacija aktivnosti gena

Otkrivanje elemenata regulacije aktivnosti gena dovelo je u pitanje pojam gena kao jedinstvenog genetskog lokusa.

Već 1961. godine Francois Jacob i Jacques Monod u svom preglednom radu na temu lac operona (Jacob i Monod, 1961.) u uvodu spominju i uvode distinkciju između pojmova *structural genes* te *regulator and operator genes*. U zaključku rada, dvojac regulatorni aspekt gena karakterizira kao element koji ima utjecaj nad kontrolom stope protoka informacije strukturnih gena do proteina, a bez ikakva upliva u same proteine. Iako u tom radu tek spominju i razvijaju ideju regulacije gena i samih pojmova u definiciji koji ju prati, sličan problem na temu obuhvaćanja, imenovanja i distinkcije regulatornih elemenata u ideji gena ostaje do danas (Sl. 6).

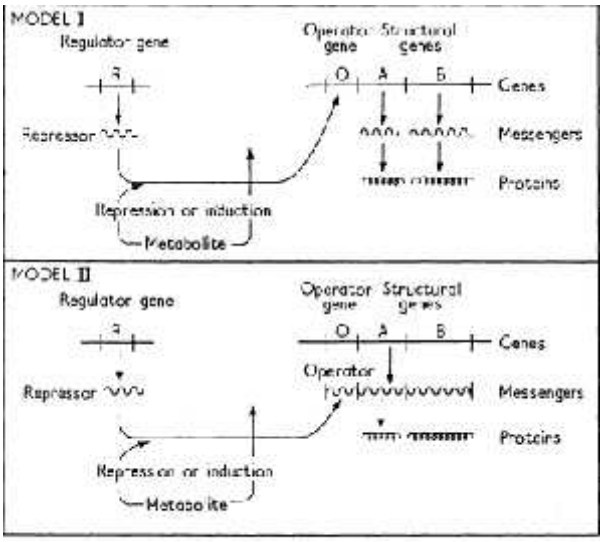


FIG. 1. Models of the regulation of protein synthesis.

Slika 6. Shema kojom su Jacob F. I Monod J. 1961. godine idejno predstavili mehanizme regulacije genske aktivnosti.

U nadolazećim godinama otkrivene su i druge sekvence koje utječu na mnoge aspekte regulacije gena, od transkripcije do degradacije mRNA, pa i post-transkripcijskih modifikacija. Ono što je zakompliciralo dotadašnje objašnjenje gena je otkriće da su određeni elementi koji reguliraju gensku ekspresiju ponekad čak vrlo udaljeni od kodirajuće sekvence za dotični gen. Naime, do tada je u znanstvenom svijetu postojala i bila prihvaćena hipoteza kako je gen jedinstven i kompaktan genetički lokus.

Jedan recentni udžbenik (*Molecular Cell Biology. Section 9.1 Molecular definition of a Gene*, Lodish i suradnici, 2000.) definira gen na idući način: "Sa stajališta molekularne biologije, gen je cijela DNA sekvenca potrebna za sintezu funkcionalnog proteina ili RNA molekule. Osim kodirajuće regije (eksona), gen uključuje i regije koje kontroliraju transkripciju, a ponekad i introne."

Ovakvom definicijom, koja bi uključila i regulacijske elemente, treba uzeti u obzir da se mnogi regulatorni elementi, kao i elementi pojačivači (*eng. enhancer*), ne nalaze neposredno uz regiju na koju djeluju, a koja će imati pre-mRNA produkt. Otkrivanjem šarolikosti regulacijskih elemenata unutar genoma i njihove varijabilnosti, te nužnim uvrštavanjem regulatornih elemenata u koncept gena, ideja da je gen jedinstvena sekvenca u genomu nije mogao opstati.

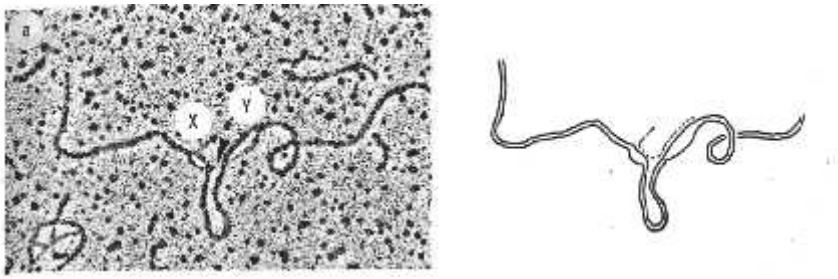
2.2. Preklapanje

Otkrivanjem i mapiranjem kodirajućih sekvenci pojedinih gena, te pregledavanjem dobivenih rezultata, model gena kao jedinica naslijeđa koje su kontinuirane duž DNA je nepreklapajuće "pao u vodu" otkrićem da određeni geni mogu dijeliti zajedničku sekvencu na molekuli DNA.

2.3. Prekrajanje

Iako je pojam kontinuiranosti genetičkog lokusa redefiniran otkrićem elemenata regulacije genske aktivnosti, otkriće prikrajanja je podcrtalo pojam gena kao diskontinuirane jedinice unutar genoma. Naime, nakon uvođenja pojmova intron i ekson, te razotkrivanja mehanizma prekrajanja transkripta i stapanja eksona u funkcionalni produkt, osim što elementi regulacije genske aktivnosti mogu biti odvojeni od same kodirajuće jedinice, pokazalo je da I transkript može nastati od isprekidane sekvence DNA.

Proces prikrajanja otkriven je 1977. godine na primjeru adenovirusnog *hexon* gena. Odnosno, u tri rada objavljena iste godine primjećeno je elektronskom mikroskopijom na primjeru istog gena pojavljivanje 5' terminalne domene RNA molekule koja nije bila povezana vodikovim vezama s pripadajućom DNA nakon hibridizacije (Berget i suradnici 1977).



Slika 6. Slika snimljena elektronskim mikroskopom na kojoj je vidljivo stvaranje omče na dijelu DNA koji se nije spario s komplementarnom mRNA (Chow i suradnici, 1977).

S tim su otkrićima u znanost uvedeni pojmovi introna i eksona, a također je I u potpunosti srušena ideja o neprekidnosti i kompaktnosti genetskog lokusa. Kasnije je Walter Gilbert predložio tzv. alternativno prikrajanje gdje različite kombinacije eksona mogu biti prekrojene zajedno i mogu dati različite izoforme određenog gena. Do sredine 1980-ih, alternativno prikrajanje je već bilo dokumentirano u nekolicini gena, a procjena gena viših eukariota koji imaju alternativno prikrajanje vrtila se oko 5%. Danas se kombinacijom mRNA-Seq I EST-cDNA podataka ustvrdilo da 95% transkripata multeksonskih gena pokazuje aletrnativno prekrajanje u ljudskom genomu. (Pan I suradnici; 2009.)

U radu *The Sequence of the Human Genome,* 2001. godine gen je definiran nedvosmislenom odrednicom kao lokus kotranskribiranih eksona. S daljnjim objašnjenjem da jedan gen može dati vise transkripata, te samim time i nekolicinu različitih proteina sa nekolicinom različitih funkcija, s naglaskom na nužnost uzimanja u obzir alternativnog prikrajanja. Pojava alternativnog prikrajanja polako, ali sigurno, uvodi nas u ideju o većoj bitnosti definiranja funkcionalnosti i elemenata postojanja RNA transkripta sa genoma, nego same sekvence na molekuli DNA.

2.4. Trans – prikrajanje

Trans – prikrajanje označava pojavu posebnog oblika procesiranja RNA molekule kod eukariota gdje se eksoni dvaju različitih primarnih transkripata spajaju i ligiraju. Eksperimentalno je utvrđeno postojanje različitih kombinacija transkripata koji se spoje prije nego budu prikrojeni, transkripti mogu biti s različitog DNA lanca, pa čak i s različitih kromosoma. Otkriće mehanizma trans – prikrajanja, a i naknadna i sveobuhvatna potkrijepa njegova postojanja te pojavljivanja u genomima mnogih organizama, uzburkala je koncept gena kao lokusa s obzirom na pojavu genskih produkata čije su DNA sekvence razdvojene preko cijelog genoma (Gerstein B. Mark, 2007).

Već spomenuto pravilo “jedan gen – jedan protein” kojeg su skovali Beadle i Tatum bilo je fundamentalno za molekularnu biologiju. To pravilo je podcrtavalo mnoga istraživanja, kao što je i usmjeravalo iščekivanja brojnih eksperimenata. Otkriće trans – prikrajanja, ali i *tandem Chimerism-a* potkopalo je tu, do tada bitnu, premisu molekularne biologije. Ovim otkrićima je spoznato da ograničen broj gena može potencijalno kodirati za mnogo veći broj proteina.

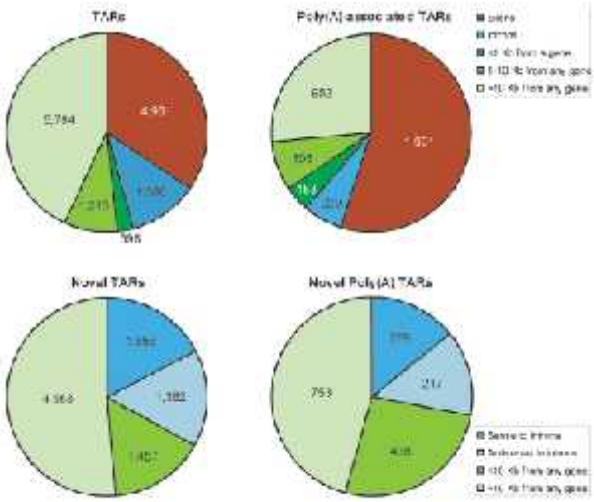
Naime, *tandem chimerism* je pojava kada su dva uzastopna gena transkribirana u jednu molekulu RNA. Translacija nakon spliceinga takve RNA molekule može proizvesti novi, fuzionirani protein, koji ima dijelove od oba originalna proteina.

3. Doprinos projekta ENCODE definiranju gena

The National Human Genome Research Institute (NHGRI) u rujnu 2003. je pokrenuo projekt ENCODE (Encyclopedia of DNA Elements) kojem je cilj bio ustvrditi sve funkcionalne elemente u sekvenci ljudskog genoma. Našu upućenost u bogatstvo I raznolikost elemenata genoma dodatno su zakomplicirali rezultati ovog projekta. Naime, sveobuhvatna analiza na razini cijelog genoma predstavila je neke zaključke koji pokazuju složenost genoma u većim detaljima, a koji u dotadašnjim eksperimentalnim rezultatima nisu bili vidljivi. Pregled zaključaka I rezultata iznesen je u ovom poglavlju.

3.1. Transkripti koji prethodno nisu anotirani

Među prvim otkrićima konzorcija ENCODE je neočekivano velika količina molekule DNA koja se u ljudskom genomu prepisuje, a u istraživanjima provedenim do objave istraživanja *Global Identification of Human Transcribed Sequences with Genome Tiling Arrays* 2004. godine nije bila poznata. Ove novootkrivene regije DNA koje se prepisuju nazivaju se TAR-ovi, odnosno transkripcijski aktivne regije. Istraživači u radu pohvaljuju metodu koja im je omogućila sveobuhvatno istražiti kompletni set regija u genomu koje kodiraju za RNA bilo koje vrste. Rezultati koje su dobili razotkrivaju 10585 novih transkribiranih sekvenci unutar genoma. Visok stupanj očuvanosti između novootkrivenih TAR-ova i dobro okarakteriziranih proteina miša daju naznaku da bi veliki dio TAR-ova mogao kodirati za funkcionalne transkripte (Sl. 7). Mnoge od očuvanih sekvenci su lokalizirane u regijama udaljenim od dosad poznatih gena, a spomena vrijedne frakcije ovakvih sekvenci su dovoljno dugačke da kodiraju proteine dugačke 300 ili više aminokiselina. Ostali TAR-ovi mogli bi kodirati za male proteine i RNA molekule još nepoznate funkcije (Berton i suradnici, 2004).



Slika 7. Distribucija TAR-ova u usporedbi s dosad anotiranim genima. (Bertone I suradnici; 2004.)

Ovakva i slična istraživanja provedena u godinama nakon svakako su nam ukazala na to koliko simplificirano poznajemo funkcionalnost, distribuciju I međuodnos pojedinih genomskih elemenata. Također, s dolaskom novih metoda i trendova, u molekularnoj biologiji još jednom je potvrđena kompleksnost genoma, njegovih elemenata, ali i vrlo živopisna upitnost prave naravi gena.

3.2. Neanotirani i alternativni TSS-ovi

Projekt ENCODE je također u svojim rezultatima utvrdio velik broj dotad neotkrivenih TSS-ova (*Transcription Start Site*), tj. mjesta u genomu koja označavaju početak transkripcije. Od novootkrivenih TSS-ova mnoga imaju svojstva već poznatih TSS-ova. Ta svojstva uključuju strukturu kromatina I sekvence na kojima se vežu DNA vezni proteini, a koje su nam već poznate iz prakse.

3.3. Raspodjela regulacijskih elemenata u genomu

U rezultatima ENCODE projekta obraćena je pozornost i na regulacijske elemente gena. U tom istraživanju, a i u svim kasnijim istraživanjima koja uključuju bioinformatičku obradu podataka, donesen je opći zaključak u kojem se izražava briga oko slabog poznavanja regulatornih elemenata i njihovom potencijalnom pojednostavljivanju sa pokušajem da ih se ugura u jednostavni model gena.

Regulatorne regije jednog gena nisu nužno locirane neposredno uzvodno do dotičnog gena, već mogu biti podosta udaljene na genomu, pa čak mogu biti i bliže nekom drugom genu. Također, pojedini geni mogu koristiti više promotora i pojavu grupiranja određenih promotora (Kim i suradnici, 2005.).

Vezna mjesta transkripcijskih faktora protežu se čitavim genomom, ali, suprotno očekivanju, nisu nasumično posložena unutar genoma već su organizirana u više i manje bogate regije. Istraživanje rasprostranjenosti transkripcijskih regulatornih elemenata unutar ENCODE regija ljudskog genoma donosi ideju o transkripcijski bogatim *otocima* i transkripcijski siromašnim *pustinjama*. Pustinje se podudaraju s lokacijama u genomu s malom gustoćom gena, dok se otoci podudaraju s lokacijama u genomu sa velikom gustoćom gena (Sl. 8; Zhang i suradnici, 2007.).



Slika 8. Vizualni prikaz regulatorno bogatih otoka i regulatorno siromašnih pustinja

Otkriveno je da poneki regulatorni elementi mogu i sami biti prepisani, što se kosi s konvencionalnim modelom gena, u kojem se DNA elementi koji reguliraju ekspresiju gena ne prepisuju i sami po sebi nisu dio transkriptoma. Projekt ENCODE i druga istraživanja osigurala su nam dokaze na razini genoma da je konvencionalni model gena prejednostavan, te da regulatorne regije pojednih gena mogu postojati unutar prvog eksona, introna ili negdje duž cijele sekvence određenog gena.

3.4. Pseudogeni

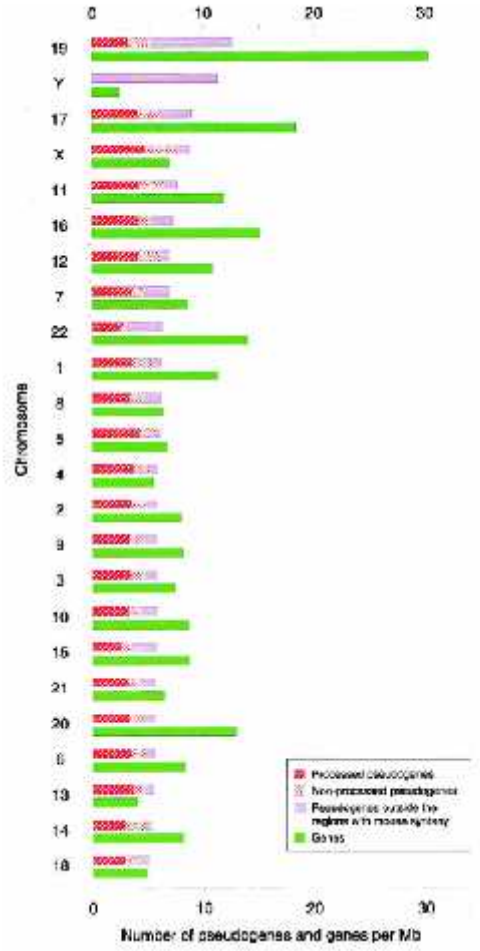
Prilikom definiranja gena u obzir treba uzeti i pseudogene. Ponajviše zbog toga što postojanje pseudogena dodatno komplicira preciziranje strukture i položaja gena.

Zastarjela definicija pseudogena glasi da su oni cjelovite i djelomične kopije gena koje nisu u mogućnosti kodirati za funkcionalni polipeptid. Postoje dva tipa pseudogena koja nastaju odvojenim mehanizmima. “Neprocesirani” peseudogeni nastaju nakon djelomične ili cjelovite duplikacije gena te naknadnog gubitka funkcije dupliciranog dijela (ili čitavog gena). “Procesirani” pseudogeni nastaju retrotranspozicijom zrele mRNA. Oni se ugrađuju nasumično u genom i zbog toga im nedostaje uzvodni promotor. Zbog preuranjena zaustavljanja reverzne transkripcije,

mnogi procesirani pseudogeni sadrže djelomičnu kodirajuću regiju ili ju uopće ne sadrže.

Kad su u pitanju pseudogeni i naša potreba da ih smjestimo ili bar izdvojimo negdje unutar definicije gena, potrebno je prvo prilagoditi definiciju pseudogena s obzirom na otkrića vezana uz njih. Otkriće i detaljna analiza NOS pseudogena (eng. *nitric oxide synthase*) u centralnom živčanom sustavu *Lymnaea stagnalis* pokrenula je i daljnje utvrđivanje i nagađanje funkcije transkripta sa zaključkom o potencijalnoj novoj klasi regulatornih gena živčanog sustava (Korneev, 1999). Osim na ovom, aktivna transkripcija pseudogena otkrivena je na brojnim drugim primjerima. Također, otkrivena je RNA pseudogena koja je sa susjednim genom tvorila gen-pseudogen kimerni transkript. Osim aktivnosti, postoje dokazi da pseudogeni mogu biti uključeni u regulaciju ekspresije gena (Balkariev, 2003).

U ljudskom genomu otkriveno je približno 20, 000 pseudogena. 72% identificiranih pseudogena je nastalo retrotranspozicijom, a 28% segmentalnom duplikacijom. (Torrents I suradnici, 2003.) Slika 9. predstavlja vizualni prikaz raspodjeljenosti pseudogena po ljudskim kromosomima.



Slika 9. Pseudogeni u humanom genomu

Kao što je već spomenuto u poglavlju *Transkripti koji prethodno nisu anotirani* bioinformatičkim metodama, u genomu je otkriven veliki broj novih TAR-ova. Zbog slabe distinkcije između gena i pseudogena, kao i zbog velike varijabilnosti u funkcionalnosti pseudogena, nužna je njihova preciznija definicija koja će olakšati distinkciju njih i gena. Također, potrebna je identifikacija novih (novootkrivenih) TAR-ova genetičkim i biokemijskim eksperimentima.

3.5. Nekodirajuće molekule RNA

Centralna uloga RNA molekule u transkripciji bila je poznata još 1950-ih godina. Tada je eksperimentalno bilo poznato da se, iako je molekula DNA smještena u jezgri, sinteza proteina odvija u citoplazmi u prisustvu velikog broja RNA molekula. Kasnije je ustvrđeno da su mjesta sinteze proteina ustvari ribosomi. U to je vrijeme rođena hipoteza "jedan gen, jedan protein, jedan ribosom" i hipoteza "centralne dogme". Ono u čemu su obje hipoteze bile u pravu je uloga mRNA u prenošenju informacije s molekule DNA do ribosoma, ali te su se hipoteze u kasnijem istraživačkom radu pokazale krivima s otkrićem sve većeg broja specifičnih RNA molekula s posebnim ulogama u stanici.

Drugu klasu funkcionalnih RNA molekula predvidio je Francis Crick svojom "hipotezom adaptora". Crick je predvidio postojanje molekule koja je medijator između tripleta genetičkog koda i aminokiseline. Intuitivnost Crickove predikcije seže u toliku dubinu da je predvidio kako je molekula koja je adaptor - molekula RNA i da je molekula RNA za tu funkciju preferentnija od proteina čak i zbog toga što bi njena struktura i kemizam odgovarali njenoj, tada predviđenoj funkciji. To je zato jer bi ona sparivanjem baza mogla biti pogodna kao mala, molekula RNA specifična za prepoznavanje određenog kodona. Mahlon Hoagland i njegov tim potvrdili su Crickovu hipotezu te otkrili i definirali molekule tRNA.

Kasnija istraživanja molekula RNA dala su nam pregled njihove raznolikosti funkcije, a također i ostavila upitnik na koncept gena kao jedinice koja će kodirati proteine koji doprinose funkcionalnosti stanice i organizma.

Mnoge male molekule RNA su uridin (U) – bogate RNA molekule koje s proteinima formiraju RNP-ove (ribonukelinske komplekse). Mnogi otkriveni RNP-ovi su kasnije nazvani U1, U2, U4, U5 I U6, te je otkriveno da su komponente *spliceosoma*, koji je uključen u prikrajanje mRNA.

Otkrivena je i uloga RNA molekula u strukturi kromatina. Fundamentalni primjer je ljudska XIST RNA koja ima ulogu u *dosage compensation*-u i inaktivaciji X kromosoma (*dosage compensation* – genetički regulatorni mehanizam koji djeluje na način da izjednačuje ekspresiju gena koji se nalaze na X kromosomu tako da u konačnici budu isti i kod žena XX i kod muškaraca XY).

RNA molekule imaju funkciju i u procesiranju RNA molekula na ribosomima.

Eksperimentalne potvrde su pronađene da bi kod eukariota male jezgrene RNA molekule (snoRNA) mogle biti katalizatori 2'-O-metilacije i pseudouridilacije. Eksperimentalna potvrda tog fenomena dala je naznaku o količini još neotkrivenih

malih jezgrenih RNA molekula. Otkrivanje novih malih jezgrenih RNA molekula vrši se bioinformatičkim metodama i imunoprecipitacijom s antitijelima specifičnima za enzim fibrilarin (metiltransferaza). Pretraga ljudskog genoma za sačuvanim sekundarnim strukturama koje su karakteristične za RNA molekule dala je pretpostavku od oko 30000 strukturnih elemenata u ljudskom genomu koji su pretpostavili postojanje transkripcije RNA molekula, od čega ih se približno 40% podudaralo s eksperimentalnim podacima, tj. s već predviđenim mjestima transkripcije u ljudskom genomu (Washietl i suradnici 2005).

Otkrivene su i RNA molekule (RNAi, eng. *RNA interference*) koje dovode do degradacije komplementarnih mRNA molekula i gubitka specifičnog fenotipa. Smatra se da ove RNA molekule predstavljaju primitivni imunološki sustav protiv RNA virusa i retrotranspozona.

Sve nabrojene RNA molekule i njihova fukncionalnost dočaravaju bogatstvo svoje uloge i važnosti kako u stanici, tako i u organizmu, tokom svih životnih stadija. Otkriće i rasvjetljenje njihove uloge stavilo je naglasak na veću problematiku funkcionalnosti gena nego njegovu nužnu karakterizaciju u kvantitativnom i kvalitativnom pogledu (Eddy, S.R. 2001.).

3.6. Genski i intergenski: postoji li razlika?

U klasičnom konceptu gena jedna od temeljnih zamisli bila je da su geni odvojeni poput čvorova na konopu te da se između njih nalazi dio koji ih odvaja, odnosno intergenski dio.

ENCODE je pokazao da je ono što smo do sada smatrali intragenskim dijelom, odnosno dijelom izmedu dva anotirana gena, ustvari transkripcijski aktivno. Transkribirane RNA molekule koje ne kodiraju za proteine i pseudogeni su elementi koji se često nalaze u intragenskom dijelu genoma, a velik broj njih je pozicioniran u intronima protein-kodirajućih gena. Ove komponente unutar introna mogle bi utjecati na ekspresiju njihova gena domaćina, izravno ili neizravno.

4.Zaključak

4.1. ENCODE definicija gena.

S obzirom da je najveći pomak u rješavanju problema definicije gena i svih njegovih svojstava i raznolikosti napravljen tokom ENCODE projekta, a i naknadno na njegovim rezultatima prilažem njihovu definiciju gena kao onu koja najobrazloženije I s najvećom eksperimentalnom pozadinom razmatra gen.

Consortium ENCODE predlaže definiciju gena kao skupa genomskih sekvenci koje kodiraju koherentnu skupinu potencijalno preklapajućih funkcionalnih produkata. Tijekom postavljanja definicije odredili su nekoliko glavnih stavki koje će pratiti. Gen je genomska sekvenca, DNA i RNA, koja kodira funkcionalne produkte, a to uključuje ili proteine ili RNA molekule. Ako više funkcionalnih produkata dijeli regije koje se preklapaju, u obzir se uzimaju sve preklapajuće genomske sekvence koje za njih kodiraju. Skup sekvenci koje čine gen mora biti koherentan, što znači da se konačni produkti transkripcije trebaju gledati zasebno, a oni uz to moraju imati preklapajuće dijelove u genomu, s time da svaki od njih ne mora imati sa svakim od njih preklapajući dio.

Deleted: *ENCODE consortium*

4.2. Zaključak i osobno viđenje problematike

U zadnjih 100 godina izravnog i neizravnog istraživanja značaja i funkcije gena jednoznačan odgovor na pitanje “Što je gen?” ostaje nepoznat. O širini i količini različitih upliva biološkog pojmova gena govori činjenica da je sa svakim novim korakom znanosti i razvoju novih metoda i tehnika u znanstvenim istraživanjima pojam gena mijenjao svoj smisao.

Koncept gena je tijekom povijesti svojeg definiranja imao konstantnu fluktuaciju između prebacivanja naglaska s njegovih fizičkih na funkcionalne karakteristike. Ono što je definitivno koordiniralo trendovima u znanosti, a time i definicija gena je tehnološki napredak, ali i dostupnost metoda korištenih prilikom istraživanja. Računarnom biologijom uvidjeli smo koliko je genom kompleksan i koliko raznolikih segmenata ima. ENCODE projekt je progresivno pomogao u rješavanju prozvane tematike.

Najmoderniju definiciju gena nudi ENCODE projekt. Naime, u toj definiciji je naglasak stavljen na funkcionalni produkt gena, a ne na sekvencu RNA ili DNA. S ozbirom na raznolikost funkcija sekvence DNA (ili RNA, ako je RNA genom) pojam gen kao fizički čvrste jednice bi mnogo više odgovarala transkriptu s genoma, nego jedinici unutar samog genoma.

U konačnici, iako i preko 100 godina udaljeni od početne definicije gena, znanstveno se vraćamo na sami početak. Konačna definicija gena ponekad se čini dostupnom i na dohvat ruke, ali ona nikada ne uključi sve eksperimentalne rezultate, te bi ovom konstantacijom gen konceptualno i dalje ostao na objašnjenju Wilhellma Johannsena, a ta je da je to pojam kojim najlakše možemo objasniti eksperimentalne podatke.

5. Literatura

Avery, O.T., MacLeod, C.M., and McCarty, M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J. Exp. Med.

Balakirev, E.S. and Ayala, F.J. 2003. Pseudogenes: Are they "junk" or functional DNA? Annu. Rev. Genet. 37: 123–151

Beadle, G.W. and Tatum, E.L. 1941. Genetic control of biochemical reactions in Neurospora. Proc. Natl. Acad. Sci. 27: 499–506.

Benzer, S. 1955. Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci. 41: 344–354.

Berget, S.M., Moore, C., and Sharp, P.A. 1977. Spliced segments at the 5 terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. 74: 3171–3175

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al. 2004. Global identification of human transcribed sequences with genome tiling arrays. Science 306: 2242–2246

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. 1977. An amazing sequence arrangement at the 5_ ends of adenovirus 2 messenger RNA. Cell 12: 1–8.

Eddy, S.R. 2001. Non-coding RNA genes and the modern RNA world. Nat. Rev. Genet. 2: 919–929.

The ENCODE Project Consortium. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE

Gersten M. I suradnici; What is a gene, post-ENCODE? History and updated definition

Griffiths, P.E. and Stotz, K. 2006. Genes in the postgenomic era. Theor. Med. Bioeth. 27: 499–521.

Jacob, F. and Monod, J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3: 318–356

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. 2005. A high-resolution map of active promoters in the human genome. Nature 436: 876–880.

Koonin, E.V., Does the central dogma still stand?; 2011.

Korneev, S.A., Park, J.H., and O'Shea, M. 1999. Neuronal expression of

neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene.

McClintock B. I suradnici, A Correlation of Cytological and genetical Crossing-over in Zea Mays, 1931.

Mendel, J.G. 1866. Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn 4 Abhandlungen, 3–47

Molecular Cell Biology. Section 9.1 Molecular definition of a Gene, Lodish i suradnici, 2000

Morgan, T.H., Sturtevant, A.H., Muller, H.J., and Bridges, C.B. 1915. The mechanism of Mendelian heredity. Holt Rinehart & Winston, New York

Muller, H.J. 1927. Artificial transmutation of the gene. Science

Pan I suradnici; Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing; *Nature Genetics* 40, 1413 - 1415 (2008)

Rogic, S., Mackworth, A.K., and Ouellette, F.B. 2001. Evaluation of gene-finding programs on mammalian sequences. Genome Res. 11: 817–832

Torrents, D., Suyama, M., Zdobnov, E., and Bork, P. 2003. A genome-wide survey of human pseudogenes. Genome Res. 13: 2559–2567

Washietl, S., Pedersen, J.S., Korbel, J.O., Stocsits, C., Gruber, A.R., Hackermüller, J., Hertel, J., Lindemeyer, M., Reiche, K., Tanzer, A., et al. 2007. Structured RNAs in the ENCODE selected regions of the human genome. Genome Res. (this issue) doi: 10.1101/gr.5650707

Washietl, S., Hofacker, I.L., Lukasser, M., Huttenhofer, A., and Stadler, P.F. 2005. Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome. Nat. Biotechnol. 23: 1383–1390.

Zhang, Z.D., Paccanaro, A., Fu, Y., Weissman, S., Weng, Z., Chang, J., Snyder, M., and Gerstein, M.B. 2007. Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. Genome Res. (this issue) doi: 10.1101/gr.5573107.

6. *Sažetak*

Koncept gena predložen je prije vise od 100 godina. Od trenutka nastanka mijenjao je svoje značenje i implikacije na znanost i društvo. Razvojem metoda u molekularne biologije mijenjali su se pogledi na gen i njegovu funkcionalost unutar živog sustava.

U ovom radu izložen je pregled najvažnijih otkrića koja su definirala gen. Obuhvaćeno je viđenje gena u različitim kontekstima u kojima se spominje tijekom znanstvene povijesti. Krećući od nasljedne jedinice, jedinice koja daje funkcionalnost organizmu preko produkata transkripcije, do suvremenog bioinformatičkog pogleda na gen, izložen je uvid u složenost genetskih elemenata molekule DNA.

7. Summary

The concept of gene was proposed more than a hundred years ago. From the beginning, it changed its meaning and implications for science and society. The development of molecular biology methods changed the perception of gene and its functionality inside living systems.

In this paper the most important discoveries that defined gene are reviewed. The understanding of genes in different contexts, throughout the history of science, is analyzed. Starting with gene as a unit of heredity, a unit that gives functionality through transcription products, to contemporary bioinformatics gene definition, an insight into the complexity of DNA’s genetic elements are presented.