

## Poglavlje 6

# Uporaba bioinformatike v prašičereji na primeru gena *leptin*

Daša Jevšinek Skok<sup>1,2</sup>, Tanja Kunej<sup>1</sup>, Milena Kovač<sup>1</sup>

### Izvleček

Prekomerno nalaganje mašcobe je iz ekonomskih razlogov nezaželeno tudi pri prireji prašičega mesa zaradi vse večjega povpraševanja potrošnikov po manj mastnih proizvodih. Gen *leptin* (*LEP*) je bil predhodno že povezan z nalaganjem mašcobe tudi pri prašičih, zato smo na primeru tega gena izdelali molekularno-genetski diagnostični test za analizo polimorfizma posameznih nukleotidov (SNP) z metodo, ki se imenuje polimorfizem dolžin restrikcijskih fragmentov (PCR-RFLP). Pred reakcijo je potrebno test sestaviti s pomočjo bioinformacijskih metod, ki jih imenujemo tudi metode *in silico*. Razviti diagnostični testi omogočajo nadaljnje laboratorijske študije povezav genetskih variabilnosti s fenotipskimi lastnostmi.

Ključne besede: prašiči, *leptin* (*LEP*), nalaganje mašcobe, bioinformatika, *in silico*

### Abstract

Title of the paper: **Application of bioinformatics in pig breeding: a case study on *leptin* gene.**

Growing consumer demand for products with lower fat content makes excessive fat deposition in the production of pig meat undesirable also for economic reasons. *Leptin* gene that was previously associated with fat deposition in pigs was used as an example to create a molecular-genetic diagnostic test for the analysis of single nucleotide polymorphism (SNP) in this gene. This test is called restriction fragment length polymorphism (PCR-RFLP). Before the reaction the test is designed using bioinformatics tools, also called *in silico* methods. Developed diagnostic tests enable laboratory analysis of associations between genetic variability and phenotypic traits.

Keywords: pigs, *leptin* (*LEP*), fat deposition, bioinformatics, *in silico*

---

<sup>1</sup>Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko, Groblje 3, 1230 Domžale

<sup>2</sup>E-pošta: dasa.jevsinek.skok@bf.uni-lj.si

## 6.1 Uvod

Molekularna biologija in genetika sta v zadnjih letih močno napredovali. Za odločitvijo o neki analizi stojijo številne predhodne raziskave in ure preživete pred računalnikom. Tako izberemo vse znane informacije o kandidatnem genu, ki ga želimo analizirati. Razvoj funkcionalne genomike je privadel do odkritja številnih genov, katerih izražanje je povezano s kopičenjem in presnova mašcobe. Precej omejen napredek pri raziskavah polimorfizmov, povezanih z nalaganjem mašcobe kaže na to, da je genetsko ozadje le-teh zelo zapleteno. V primerjavi s človekom in mišjo je bilo do sedaj pri prašičih izvedenih bistveno manj študij o povezavah med zamenjavami posameznih nukleotidov (SNP) (angl. *single nucleotide polymorphism*) in lastnostmi nalaganja mašcobe (Switonski in sod., 2010).

Bioinformatika je področje, ki nam omogoča obsežne raziskave bioloških procesov z računalniškimi orodji, s čimer se stroški in porabljen čas za raziskave bistveno zmanjšajo. Vključuje zbiranje, analizo in uporabo bioloških podatkov preko računalnika v namen raziskovanja. Z omenjenimi metodami, ki jih imenujemo metode *in silico*, lahko med seboj združimo ter obdelamo več podatkovnih zbirk in ostalih rezultatov raziskav s področja, ki ga proučujemo. Na področju genetike in molekularne biologije dnevno potekajo raziskave, katerih rezultate raziskovalci vnašajo v podatkovne zbirke. Uporabniki lahko do teh podatkov dostopajo preko osebnega računalnika. Takšne podatkovne zbirke, so v veliko pomoč za analizo pred delom v laboratoriju. Pred eksperimentalnim delom je potrebno predhodne rezultate in obstoječe podatke zbrati in preučiti. Številne bioinformacijske metode nam to tudi omogočajo. Na podlagi predhodnih bioinformacijskih raziskav izberemo metode, ki jih bomo v raziskavi uporabili. Rezultate raziskav *in silico* nato dokažemo s poskusom. Količina objavljenih podatkov se glede na lastnost in živalsko vrsto razlikuje. Največ raziskav je opravljenih na področju raka, najbolj zastopani preiskovani vrsti med sesalci pa sta človek in miš. Z analizo človeškega genoma (od leta 1990) so uspeli povezati vplive genov na različne fenotipske značilnosti. Ta odkritja so v veliko pomoč pri raznih kliničnih študijah. Za številne lastnosti lahko na podlagi predhodnih raziskav na modelnih organizmih analiziramo gene in njihovo variabilnost. Takšni geni so za druge vrste kandidatni geni za neko lastnost. Miš je zaradi hitrega reproducijskoga ciklusa in rasti dober modelni organizem za sesalce. Fiziološke podobnosti (npr. presnova) med človekom in prašičem kažejo na uporabnost raziskav pri obeh vrstah, saj lahko glede na rezultate pri eni vrsti sklepamo na rezultate pri drugi. Veliko pristopov *in silico* nam omogoča, da na osnovi obstoječih rezultatih kandidatne gene obdelamo, nato pa jih na podlagi teh bioinformacijskih analiz tudi eksperimentalno dokažemo.

Pri prašiču je znanih le malo genov povezanih z nalaganjem mašcobe, medtem ko je pri človeku in miši ta lastnost dobro preučena in je zanjo znano veliko število genov. Ti geni so obenem močni kandidatni geni za omenjeno lastnost tudi pri prašiču. Analize genoma prašiča so pokazale, da je več kot 500 odsekov za kvantitativne lastnosti (QTL) (angl. *quantitative trait loci*) povezanih z nalaganjem mašcobe. Poleg QTL so analizirali številne SNP-je znotraj kandidatnih genov za nalaganje mašcobe, med katerimi je tudi *leptin* (Switonski in

sod., 2010). Gen *leptin* je pri prašiču lociran na kromosomu 18 znotraj QTL-a povezanega z debelostjo (Stachowiak in sod., 2007).

Gen *leptin* kodira istoimenski protein (hormon), sestavljen iz 146 aminokislin, ki se sintetizira v maščobnem tkivu in se po odcepitvi 21 aminokislin dolgega signalnega peptida izloča v krvni obtok. Hormon leptin vpliva na zauživanje krme, neuroendokrinološke in imunološke procese (Barb in sod., 1999; Houseknecht in Portocarrero, 1998; Ahima in sod., 2000). Sprva je bil leptin identificiran kot manjkajoči produkt pri debelih miškah (Zhang in sod., 1994). Ramsay in sod. (1998) so višje koncentracije leptina v serumu opazili pri debelih svinjah v primerjavi z nedebelimi križanci landrace x yorkshire. Tudi Cameron in sod. (2000) so predpostavili, da nižje koncentracije leptina v krvi vplivajo na nalaganje maščob pri prašičih. V raziskavi, kjer so prašiče pasme large white razdelili na dve skupini glede na zauživanje krme, so zaznali povisano koncentracijo leptina, nalaganje maščobe in boljšo ječnost pri skupini s povečanim zauživanjem krme. Glede na rezultate teh raziskav, bi bilo korelacijo med serumskim leptinom in vsebnostjo naložene maščobe smiselno vključiti v izračun plemenske vrednosti za zamaščenost prašičev.

Namen prispevka je bil prikazati uporabo računalniških metod, ki jih moramo izvesti pred poskusom v laboratoriju. V ta namen smo na primeru gena *leptin* prikazati izdelavo molekularno-genetskega diagnostičnega testa za analizo SNP-jev. Diagnostične teste uporabljamo v namen povezovanja genotipa (genov, SNP-jev) s fenotipom (npr. zamaščenost, mesnatost). Test za dokazovanje nukleotidne zamenjave se imenuje PCR-RFLP (angl. *polymerase chain reaction - restriction fragment length polymorphism*) in temelji na pomnoževanju odseka DNA (angl. *deoxyribonucleic acid*), ki vsebuje SNP in cepljenju z restrikcijskimi encimi.

## 6.2 Podatkovne zbirke in bioinformacijska orodja

Biološke informacije so zbrane v tako imenovanih podatkovnih bankah, ki vsebujejo podatkovne zbirke z informacijami ter orodja za dostop in obdelavo teh informacij. Do podatkovnih bank praviloma dostopamo preko interneta. Glede na namen raziskave se uporaba podatkovnih zbirk in bioinformacijskih orodij razlikuje. Najpogosteje uporabljene podatkovne zbirke vsebujejo informacije o genih, kot so: nukleotidno zaporedje DNA, genetska variabilnost, lokacija v genomu ...

Podatkovna zbirka "Ensembl" (EMBL-EBI in Wellcome Trust Sanger Institute, 2010) ponuja obsežen nabor informacij za analizo bioloških podatkov znotraj vrste in med njimi. Vsi geni, dostopni v zbirki, temeljijo na eksperimentalnih dokazih zaporedij mRNA in proteinov, pridobljenih iz posameznih projektov sekvenciranja genoma različnih vrst. Prav tako kot "Ensembl" tudi "NCBI" (National Center for Biotechnology Information, 2010a) vključuje več podatkovnih zbirk in bioinformacijskih orodij, s katerimi lahko zberemo in obdelamo številne biološke informacije. Računalniški program "Primer3" (Rozen in Skaletsky, 2010) načrtuje začetne oligonukleotide za reakcijo PCR na osnovi več dejavnikov. Na vnešenem nukleotidnem zaporedju računalniški orodji "NEBCutter" (New England Biolabs, 2003) ali

“Webcutter” (Heiman, 1997), poiščeta možne restrikcijske encime za izvedbo metode PCR-RFLP.

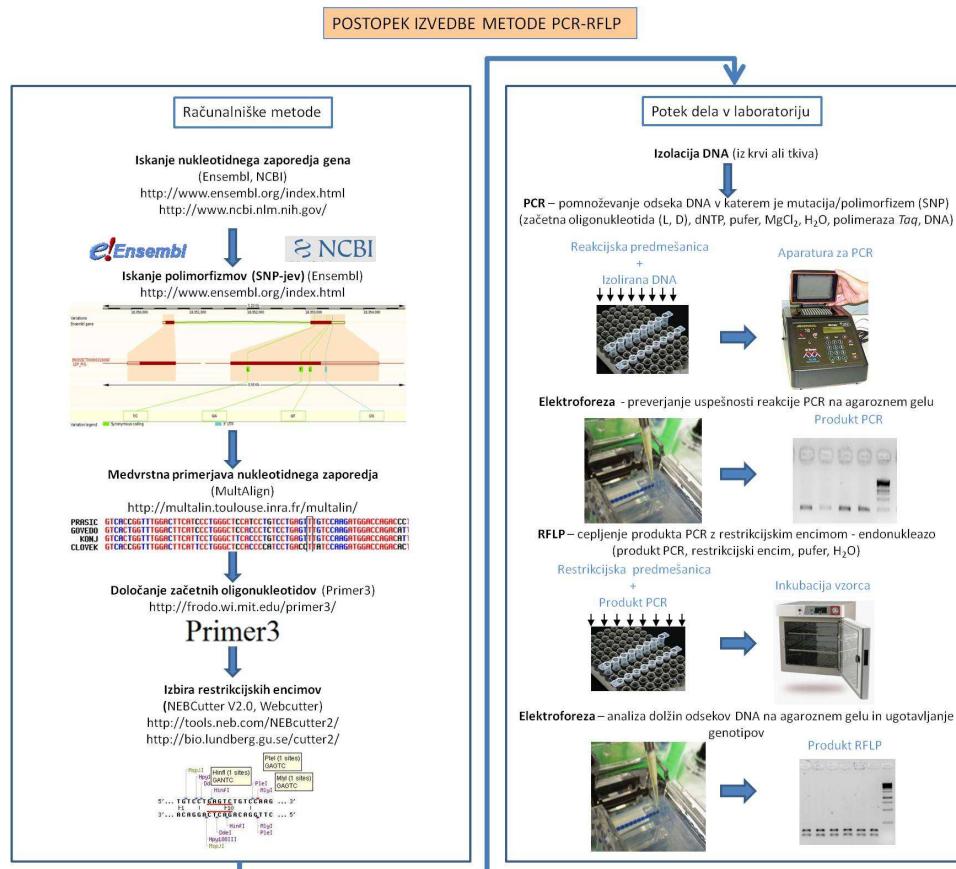
V podatkovni zbirki “Ensembl” (release 60) in “NCBI” smo našli nukleotidno zaporedje gena *leptin* pri prašiču in njegovo genetsko variabilnost. Ohranjenost gena pri desetih vrstah smo preverili z bioinformacijskim orodjem “MultAlin” (Corpet, 2010). Z bioinformacijskim orodjem “Primer3” (v. 0.4.0) smo določili začetne oligonukleotide za gen *leptin* pri prašiču. Za iskanje in izbiro restrikcijskih encimov smo uporabili bioinformacijsko orodje “NEBCutter” (v 2.0).

### 6.3 Dokazovanje nukleotidne zamenjave z metodo PCR-RFLP

Dokazovanje nukleotidne zamenjave z metodo PCR-RFLP lahko razdelimo v dva sklopa. V bioinformacijskem delu z računalniškimi metodami *in silico* poiščemo in obdelamo potrebne biološke informacije o genu, v laboratoriju pa SNP po točno določenem protokolu tudi eksperimentalno dokažemo (slika 1).

Poskus, kjer dokazujemo SNP-je, se začne že pred delom v laboratoriju. Najprej z računalnikom v podatkovni zbirki “Ensembl” ali “NCBI” izberemo nukleotidno zaporedje gena ter poiščemo nukleotidne zamenjave v njem. Nukleotidne zamenjave lahko motijo prevajanje gena v protein in s tem njegovo funkcijo. Poznamo več vrst zamenjav, ki na izražanje fenotipskih lastnosti vplivajo različno ali pa sploh ne (tihe mutacije). Večji kot je delež ohranjenosti gena v območju s SNP-jem med vrstami, večja je verjetnost, da le-ta vpliva na fenotip (Conde in sod., 2005). Na podlagi pridobljenih informacij o ohranjenosti gena in nukleotidnih zamenjavah določimo odsek, ki ga želimo preučiti. Z računalniškimi orodji zanj določimo in izberemo začetne oligonukleotide in restrikcijske encime, ki jih potrebujemo za metodo PCR-RFLP ter jih naročimo pri podjetju, ki ji proizvaja. Po opravljenem delu z računalnikom lahko začnemo z eksperimentalnim dokazom v laboratoriju.

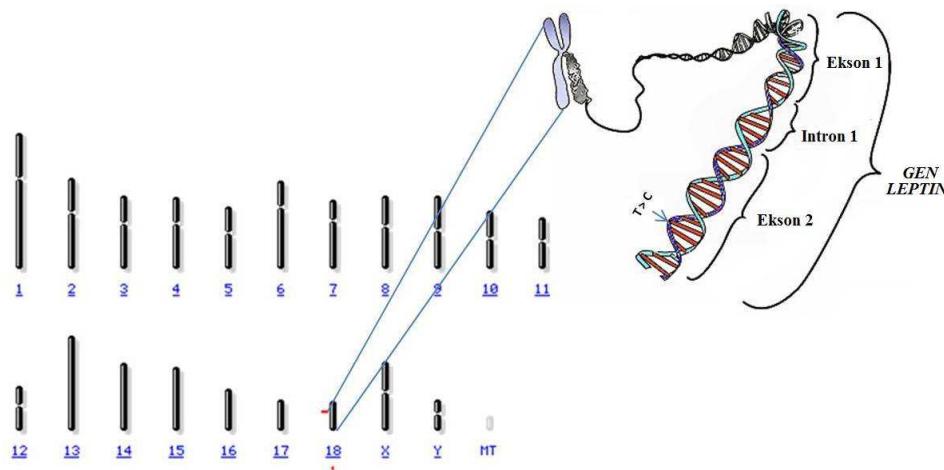
Pri metodi PCR potrebujemo izoliran genomski material (DNA), saj gre pri tej reakciji za pomnoževanje odsekov molekule DNA. Običajno DNA izoliramo iz tkiva (meso, uhlji ...) ali telesnih tekočin (kri, seme ...), nato si v mikrocentrifugirkah pripravimo reakcijsko predmešanico (angl. *master mix*), kateri dodamo izolirano DNA. S tem je vzorec pripravljen za reakcijo PCR. Reakcijo izvedemo tako, da mikrocentrifugirke vstavimo v aparaturo za PCR, ki uravnava temperaturni in časovni program. Kot rezultat PCR dobimo številne kopije odsekov DNA. Dolžino odsekov predhodno določimo z izbiro začetnih oligonukleotidov, ki se prilegajo zaporedju DNA. Uspešnost reakcije PCR preverimo na agaroznem gelu z elektroforezo. Restrikcijsko analizo (RFLP) naredimo po uspešni reakciji PCR tako, da produktu dodamo restrikcijski encim, ki dobljeni odsek DNA razreže na krajše odseke. Kje bo encim DNA rezal, je odvisno od nukleotidnega zaporedja in prepoznavnega mesta tega encima. Produkti PCR in RFLP so vidni na agaroznem gelu kot odseki. DNA je nabita negativno, zato pri elektoforezi potuje proti pozitivno nabiti elektrodi. V agaroznem gelu odseki potujejo glede na velikost različno hitro (manjši oz. krajevi potujejo hitreje, daljši počasneje), zato jih na sliki opazimo v določenem zaporedju.



Slika 1: Shematski prikaz izdelave molekularno-genetskega diagnostičnega testa PCR-RFLP

### 6.3.1 Zbiranje bioloških informacij o genu

V jedru celice je DNA organizirana v kromosome. Prašič ima 18 parov avtosomalnih in dva spolna kromosoma (X ali Y). Z bioinformacijskim orodjem "NCBI-MapViewer" (National Center for Biotechnology Information, 2010b) smo pri prašiču gen *leptin* našli na kromosому 18 (slika 2).



Slika 2: Genomska lokacija gena *leptin* na kromosomu 18 pri prašiču. MT: mitohondrijska DNA

V podatkovnih zbirkah "Ensembl" in "NCBI" smo poiskali nukleotidna zaporedja gena *leptin* za več vrst (prašič, govedo, konj, človek, gorila, šimpanz, pes, miš, podgana in oposum) in štiri SNP-je v njih. Nukleotidno zaporedje v okolici najdenih SNP-jev smo primerjali pri več vrstah z bioinformacijskim orodjem "MultAlin". Glede na medvrstno ohranjenost gena *leptin* v območjih s SNP-ji smo za analizo oz. sestavo diagnostičnega testa izbrali nukleotidno zamenjavo timina v citozin (T>C) z imenom rs45431504, ki se nahaja v eksonu 2 gena *leptin* pri prašiču. Ugotovili smo, da je to območje ohranjeno pri desetih vrstah sesalcev (slika 3).

	rs45431504
	T>C
PRASIC	GTCACC GGTT TGGACTT CATTCCCTGGGCTCCATCCTGTCTTGAGT  TGTCCAGATGGACCGAGACCTGGCAGTCTACCAACA
GOVEDO	GTCACT GGTT TGGACTTCATCCCTGGGCTCCACCCCTCTCTGAGT  TGTCCAGATGGACCGAGACATGGCAGTCTACCAACA
KONJ	GTCACT GGTT TGGACTTCATCCCTGGGCTCCACCCCTCTCTGAGT  TGTCCAGATGGACCGAGACATGGCAGTCTACCAACA
CLOVEK	GTCACC GGTT TGGACTTCATCCCTGGGCTCCACCCCTCTGACG  TATCCAGATGGACCGAGACACTGGCAGTCTACCAACA
GORILA	GTCACC GGTT TGGACTTCATCCCTGGGCTCCACCCATCCTGACG  TATCCAGATGGACCGAGACACTGGCAGTCTACCAACA
ŠIMPANZ	GTCACC GGTT TGGACTTCATCCCTGGGCTCCACCCATCCTGACG  TATCCAGATGGACCGAGACACTGGCAGTCTACCAACA
PES	GTCGET GGCT TGGACTTCATCCCTGGGCTCCACCCATCTGAGT  TGTCCAGATGGACCGAGACCTGGCAGTCTACCAACA
MIS	GTCACT GGCT TGGACTTCATCCCTGGGCTCCACCCATCTGAGT  TGTCCAGATGGACCGAGACCTGGCAGTCTACCAACA
PODGANA	GTCACC GGTT TGGACTTCATCCCGGGCTCCACCCATCTGAGT  TGTCCAGATGGACCGAGACCTGGCAGTCTACCAACA
OPOSUM	GTAACT GGCT TGGATTTCATCCCGGGCTTCATCCCTTCAGAGT  TGTCAAGATGGACCGAGACCTGGCAGTCTACCAACA

Slika 3: Medvrstna primerjava odseka nukleotidnega zaporedja gena *leptin*, v katerem se nahaja nukleotidna zamenjava rs45431504 T>C

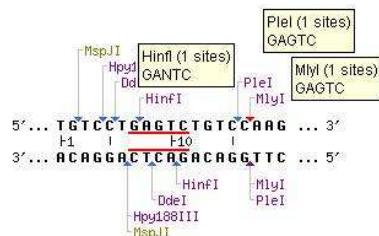
### 6.3.2 Iskanje začetnih oligonukleotidov

Začetni oligonukleotidi (angl. *primers*) so dolgi približno 20 baznih parov (bp) in se komplementarno prilegajo začetku (levi - L) in koncu (desni - D) zaporedja DNA, ki ga želimo pomnoževati z reakcijo PCR. Z bioinformacijskim orodjem "Primer3" smo za analizo zamenjave timina v citozin v genu *leptin* določili začetna oligonukleotida (L: 5' CCAACTATG-TGAGAACAGAC 3' in D: 5' CCAGGGCTGAGGTCCAGCTGC 3') (slika 4).

Slika 4: Nukleotidno zaporedje gena *leptin* z izbranimi začetnima oligonukleotidoma (pod-črтано) in SNP-jem (v oglatih oklepajih)

### 6.3.3 Določanje restrikcijskih encimov

Restriktivne endonukleaze so encimi, ki režejo dvojno (dsDNA; angl. *double-stranded DNA*) ali enojno vijačnico DNA (ssDNA; angl. *single-stranded DNA*) na specifičnih mestih. Z bioinformacijskim orodjem "NEBCutter" (v 2.0) smo našli tri restriktivne encime *HinfI*, *PleI*, *MlyI* s prepoznavnim mestom na območju SNP-ja. Restriktivni encim *HinfI* prepozna zaporedje nukleotidov GANTC (N je A, T, C ali G), *PleI* in *MlyI* pa prepoznavata zaporedje GAGTC. Omenjena prepoznavna mesta so znotraj območja s SNP-jem, zato je te encime možno uporabiti za restriktivno analizo (slika 5).

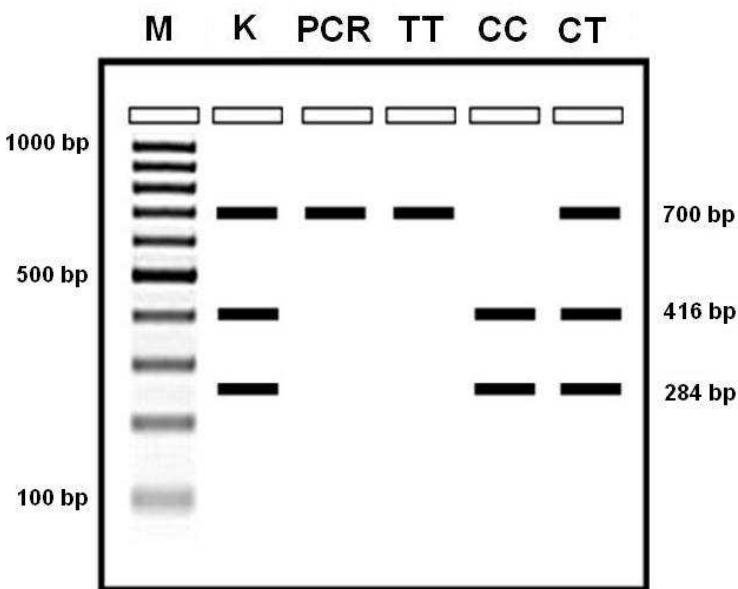


Slika 5: Restriktijski encimi *HinfI*, *PleI*, *MlyI* za identifikaciju nukleotidne zamenjave T>C u genu *leptin*

### 6.3.4 Pričakovani rezultati restrikcijske analize na agaroznem gelu

Po restriktivni analizi rezultate preverimo na 1-odstotnem agaroznem gelu, s čimer dočimo genotip nekega osebka. Encim nukleotidno zaporedje cepi v primeru zaporedja GAGTC, v primeru zaporedja GAGTT pa cepitve ni. Označevalec molekulskih mas (M) na gel nanesemo za pomoč pri odčitavanju rezultatov. Za kontrolo vzamemo vzorec z znanimi

dolžinami odsekov ter necepljen produkt PCR. V našem primeru so pričakovane dolžine 700 baznih parov (bp) (v primeru necepljenega produkta PCR in homozigota TT) ter po cepljenju z restriktijskim encimom 416 bp in 284 bp (v primeru homozigota CC). Heterozigot CT ima eno verigo DNA z zaporedjem GAGTC ter drugo z zaporedjem GAGTT, zato na gelu pričakujemo odseke vseh treh dolžin (700 bp, 416 bp in 284 bp; slika 6).



Slika 6: Shematski prikaz reakcije PCR-RFLP za analizo gena *leptin* z enim od restriktijskih encimov *HinfI*, *PleI* ali *MlyI* s prikazanimi dolžinami restriktijskih odsekov. PCR: necepljen produkt PCR, K: kontrola, M: označevalec molekulskih mas

#### 6.4 Zaključki

V prispevku smo za prikaz sestave diagnostičnega testa izbrali gen *leptin*, ki je bil predhodno že povezan z nalaganjem maščobe pri prašiču. Z metodami *in silico* smo za omenjeni gen prikazali razvoj molekularno-genetskega diagnostičnega testa. S podatkovno zbirko ‐Ensembl‐ smo pridobili nukleotidno zaporedje tega gena in informacije o njegovi genetski variabilnosti (SNP-jih). Na osnovi nukleotidnega zaporedja smo ugotovljali ohranjenost gena *leptin* med vrstami sesalcev in za izvedbo reakcije PCR-RFLP z računalniškimi orodji poiskali začetne oligonukleotide in restriktijske encime. Sestavljeni diagnostični test je sedaj možno uporabiti v laboratoriju. Opravljene analize obdelamo s pomočjo statističnih metod. S primerjavo genotipa s fenotipskimi lastnostmi lahko določimo vlogo posameznega gena oz. mutacij znotraj gena.

## 6.5 Viri

- Ahima R., Saper C., Flier J., Elmquist J. 2000. Leptin regulation of neuroendocrine systems. *Front. Neuroendocrin.*, 21: 263–307.
- Barb C., Hausman G., Houseknecht K. 1999. The brain-pituitary-adipocyte axis: role of leptin in modulating neuroendocrine function. *J. Anim. Sci.*, 77: 1249–1257.
- Cameron N., Penman J., McCullough E. 2000. Serum leptin concentration in pigs selected for high or low daily food intake. *Genet. Res., Camb.*, 75: 209–213.
- Conde L., Vaquerizas J.M., Ferrer-Costa C., de la Cruz X., Orozco M., Dopazo J. 2005. PupasView: a visual tool for selecting suitable SNPs, with putative pathological effect in genes, for genotyping purposes. *Nucleic Acids Res.*, 33: W501–W505.
- Corpet F. 2010. Multalin.  
<http://multalin.toulouse.inra.fr/multalin/> (12. dec. 2010)
- EMBL-EBI in Wellcome Trust Sanger Institute 2010. Ensembl release 60.  
<http://www.ensembl.org> (12. dec. 2010)
- Heiman M. 1997. Webcutter 2.0.  
<http://bio.lundberg.gu.se/cutter2/> (12. dec. 2010)
- Houseknecht K., Portocarrero C. 1998. Leptin and its receptors: regulators of whole-body energy homeostasis. *Domest. Anim. Endocrin.*, 15: 457–475.
- National Center for Biotechnology Information 2010a. Ncbi.  
<http://www.ncbi.nlm.nih.gov> (12. dec. 2010)
- National Center for Biotechnology Information 2010b. Ncbi-mapviewer.  
<http://www.ncbi.nlm.nih.gov/projects/mapview/> (24. dec. 2010)
- New England Biolabs 2003. Nebcutter v2.0.  
<http://tools.neb.com/NEBcutter2/> (12. dec. 2010)
- Ramsay T., Yan X., Morrison C. 1998. The obesity gene in swine: sequence and expression of porcine leptin. *J. Anim. Sci.*, 76: 484–490.
- Rozen S., Skaletsky H. 2010. Primer3.  
<http://frodo.wi.mit.edu/primer3/> (12. dec. 2010)
- Stachowiak M., Mackowski M., Madeja Z., Szydlowski M., Buszka A., Kaczmarek P., Rubis B., Mackowiak P., Nowak K., Switonski M. 2007. Polymorphism of the porcine leptin gene promoter and analysis of its association with gene expression and fatness traits. *Biochem. Genet.*, 45: 245–253.
- Switonski M., Stachowiak M., Cieslak J., Bartz M., Grzes M. 2010. Genetics of fat tissue accumulation in pigs: a comparative approach. *J. Appl. Genet.*, 51(2): 153–168.

Zhang Y., Proenca R., Maffei M., Barone M., Leopold L., Friedman J. 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature*, 372: 425–432.