



NGS (NEXT GENERATION SEQUENCING-
GƏLƏCƏK NƏSİL SEKVENSERİ) VASİTƏSİLƏ
ƏLDƏ OLUNMUŞ NƏTİCƏLƏRİN ANALİZİ VƏ
SNP-LƏRİN AŞKARLANMASI

By: Salayeva Samirə

- *Ezamiyyə Azərbaycan Respublikasının Prezidenti yanında Elmin İnkişafı Fondunun 2012-ci il üçün müəyyən edilmiş 2-ci qrant müsabiqəsindən (EIF-2012-2(6)) uğurla keçmiş “Azərbaycanın torpaq və bitki genetik müxtəlifliyinin səmərəli istifadəsi istiqamətində fundamental tədqiqatların maddi-texniki bazasının gücləndirilməsi” adlı və Z.İ.Əkpərovun rəhbəri olduğu pilot layihə çərçivəsində 04.02.2014-04.03.2014 tarixlərində Fransanın Bordo şəhərində INRA-nın (The Institut National de la Recherche Agronomique) Meyvə Biologiyası və Patologiya İnstitutunda (UMR-1332 Biologie du Fruit et Pathologie) reallaşdırılmışdır.*

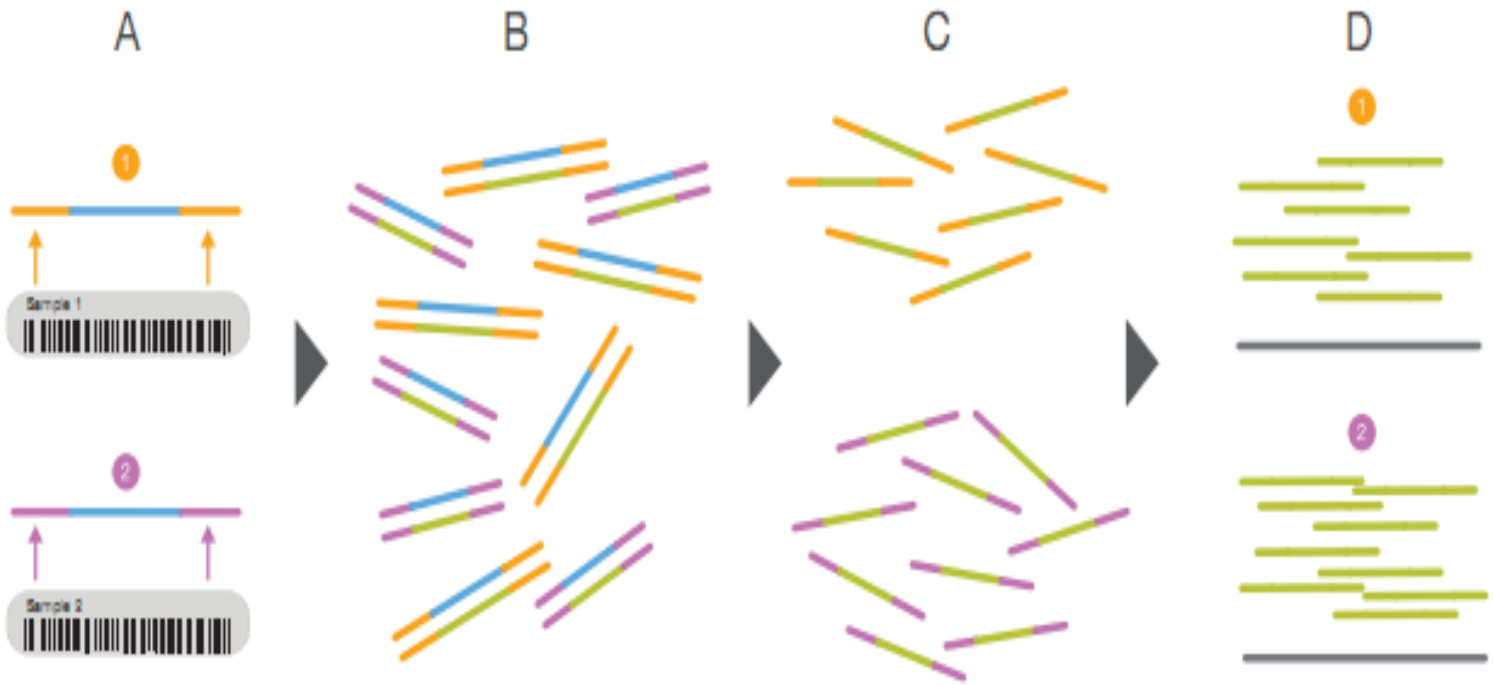
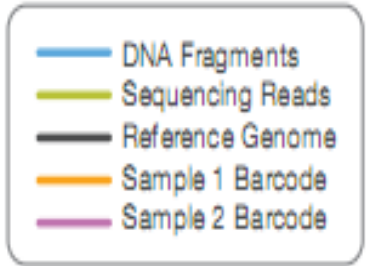


- NGS-dən (next generation sequencing) əvvəlki,

Sanger sequencing əsaslı kapilyar elektroforez texnologiyalarının işləmə prinsipi kiçik DNT fraqmentinin nukleotidlərinin DNT-nin ana zənciri üzərində hər bir fraqmentin yenidən sintez olunması ilə verilən siqnallar əsasında ardıcıl olaraq identifikasiyasına əsaslanır.

NGS Sanger Sequencing-dən fərqlənərək, bir və ya bir neçə DNT fraqmenti ilə məhdudlaşmayaraq, geniş məstəabda, milyon reaksiyaların paralel yerinə yetirilməsi ilə bu prosesi genişləndirir. Bu proses bir sekvensinq nəticəsində yüzlərlə gigabasis (10^9 n.c.) məlumatı istehsal etməklə, bütün genomu əhatə edəcək, DNT-nin böyük parçalarının sürətli oxunmasına imkan verir.

- Bunun üçün vahid DNT genomu (gDNT) kiçik fraqmentlərə ayrılır (kiçik fraqmentlər kitabxanası) və hər bir fraqment eyniliklə (uniform) və dəqiqliklə milyon reaksiyalarla sekvens olunur. Yeni identifikasiya olunmuş əsas sıraları (the newly identified strings of bases) **“read”lər adlandırılır (xam material)**, sonradan onlar nümunə (reference) genom əsasında scaffold, yəni ilişikli qruplar şəklində yenidən qruplaşdırılır (reasssembled). Bu proses **“resequencing”** (yenidən sequencing) adlandırılır. Reference genomun yoxluğu halında readlərin xəritələnməsinə (alignment), yəni sıralanmasına **de novo sequencing** deyilir. Xəritələnməmiş “read” bütün dəsti gDNT-nin hər bir xromosomunun bütün nukleotid ardıcılığını göstərir.



“Read” nədir?

- Fraqmentlərə ayrılmış hər bir DNT parçasının eyniliklə və dəqiqliklə milyon reaksiyalar əsasında oxunması nəticəsində əmələ gəlmiş, yeni identifikasiya olunmuş əsas sıralarıdır.

NGS-də iki növ protokollardan istifadə olunur:

- Cüt sonluqlu sequencing protokolları - **paired-end (PE) sequencing** (bu halda DNT fraqmentlərinin hər iki ucları seqvens olunur, yəni DNT fraqmentlərinin müxtəlif sonluqları R_1 və R_2 read-lər şəklində bir-birinə əks olaraq (uyğun olaraq, $5' \rightarrow 3'$ və $3' \rightarrow 5'$ istiqamətlərində) seqvens olunur);
- Vahid “read”-li sequencing - **single-read sequencing** (bu zaman DNT fraqmentlərinin yalnız bir sonluqdan başlayaraq sekvensinqi həyata keçirilir).



BAŞLANGIÇ NÖQTƏSİ

NGS (Next Generation Sequencing) müxtəlif məqsədlərlə tətbiq olunur; onlardan biri də müəyyən populyasiya daxilində SNP-ləri təyin etmək, müxtəlif pozisiylar (lokuslar) üzrə mövcud olan SNP-lərin cədvəlinin tərtib etmək və populyasiya daxilində öyrənilmiş təsərrüfat əhəmiyyətli (aqronomik) əlamətlərlə SNP-lər arasında ilişikliyi müəyyən etməkdir.

Məs., 93 nümunədən ibarət mədəni ərik bitkisi populyasiyasında müxtəlif pozisiylar (lokuslar) üzrə SNP cədvəlinin tərtib etmək və təsərrüfat əhəmiyyətli iki əlamət: Şarkaya davamlılığı və ağacların çiçəkləmə müddəti əlamətləri ilə müəyyən SNP-lər arasında ilişikliyi təyin etmək.

Sekvensinq GATC Biotech (Almaniyanın Konstanz şəhərində yerləşmiş, dünyanın 40 ölkəsindən çox, 10000 istifadəçisi olan, DNT-nin sekvensi və bioinformatika sahəsində Avropanın aparıcı təchizatçılarından) və Kannapolis (Şimali Korolinada yerləşir) kimi müxtəlif təşkilatlar tərəfindən yerinə yetirilmişdir.

Sekvensinq **Illumina texnologiyası** əsasında icra olunmuşdur.



NƏTİCƏLƏRİN ABRIWG LAYİHƏSİ TİMSALINDA ANALİZİ

Ərik populyasiyasında təsərrüfat əhəmiyyətli
(aqronomik) əlamətlərin xüsusiyyətlərinin SNP-lərin
tədqiqi vasitəsilə öyrənilməsi

Və nəticələrin analizi üçün sequencing-in icra
strukturunun (PIPELINE) tərtibi



SEKVENSIYONUN İCRA STRUKTURU (PIPELINE)



Xam Məlumatlar (Raw Data)
İon Torrent "Read"ləri



Keyfiyyətin yoxlanılması
NGS QC Toolkit



Nümunə (reference) genom
əsasnda xəritələnmə
BWA



Gəndaxili "read"lərin
ekstraksiyası
In-house Python script



Nümunə (reference) genom
əsasnda SNP-lərin təyini
SAMtools - mpileup



Oxunulacaq genomun (məs., ərik genomunun)
SNP cədvəlinin tərtibi



Populyasiya daxilində (məs; ərik
populyasiyasında) SNP-lərin axtarışı
In-house Python script



"Read"lərin dərinliyinin
müəyyənəndirilməsi ilə SNP-lərin filtrasiyası
və allel tezlikləri (In-house Python script)



XAM MƏLUMATLAR (ROW DATA) VƏ ONLARIN İŞLƏNİLMƏSİ-İLKİN MƏRHƏLƏ

- Sekvenserin nəticəsi iki FASTQ formatında yazılmış fayl şəklində əldə olunur. Fayllardan biri yalnız R1 “read”ləri, digəri isə R2 “read”ləri (ID-ləri, ardıcılıqları və PHRED score-ları veriməklə) cəmləyir. Bu fayllar sekvensinqin nəticələrini izah edəcək bioinformatik proqram tərəfindən (məs., CLC) oxunulur.
- **FASTQ** formatlı fayl dedikdə bir read-in ID-si, nukleotid ardıcılığının və səhv nukleotidlərin oxunma ehtimalının (score of probability of mistakes) göstərildiyi fayl nəzərdə tutulur. R1 və R2 cüt sonluqlu “read” eyni ID ilə tanınır, lakin bu ID-nin müxtəlif portallarında read-in R1 və ya R2 paired-end olması göstərilir.
- Nümunə (reference) genom isə **FASTA** formatında əldə olunur; FASTA faylının FASTQ faylından başlıca fərqi orada səhv nukleotidlərin oxunma ehtimalının (score of probability of mistakes) olmamasıdır.



Aşağıda **FASTQ formatında** yazılmış **faylın** parçasını əks etdirən şəkil verilmişdir.

FASTQ faylında **birinci sırada** read-in proqram tərəfindən verilmiş **ID-si** (identification read) göstərilir. Hər bir read-in (o cümlədən, R1 və R2 paired-end-lər üçün) özünəməxsus ID-si olur.



```
@HISEQ:126:C1MD9ACXX:7:1101:2362:1868 1:N:0:CTTGTC
NTGCCGGCGGAGTCCTAAAAGTAACATCCGCCGATCCCTGGTCGGCATCGTTTATGGTTGAGACTAGGACGGTATCTGATCGTCTTCGAGCCCCCAACT
+
#1:DDDDFFHHGDHIIJJJJIBDHDHEIJIIJDEEEFFFCAC@BB:@CBDBBACDDDCDDDDCCCCD@C;B5<@CCDCDCB?8<?C@A?9>@DDDDDD>
```

İkinci sırada read-in **nukleotid ardıcılığı** verilir (cari tədqiqatda istifadə olunmuş read-lər 100 nukleotiddən ibarət ardıcılıqlardır). Ion torrent sekvensirindən əldə olunan readlər isə 400 n.c. uzunluğunda olur.


N-məlum olmayan, yaxud qeyri-deqiq nukleotidi ifadə edir.



```
@HISEQ:126:C1MD9ACXX:7:1101:2362:1868 1:N:0:CTTGTC
NTGCCGGCGGAGTCCTAAAAGTAACATCCGCCGATCCCTGGTCGGCATCGTTTATGGTTGAGACTAGGACGGTATCTGATCGTCTTCGAGCCCCCAACT
+
#1:DDDDFFHHGDHIIJJJJIBDHDHEIJIIJDEEEFFFCAC@BB:@CBDBBACDDDCDDDDCCCCD@C;B5<@CCDCDCB?8<?C@A?9>@DDDDDD>
```


Nəhayət, **üçüncü sırada PHRED score**-verilir. Şəkildən də məlum olur ki, sekvensinqin PHRED score-su xüsusi kod şəklində verilir ki, bu kod yalnız müvafiq bioinformatik proqram (məs., CLC) tərəfindən başa düşülür və onun neçəyə bərabər olması bu proqram vasitəsilə oxunularaq, istifadəçiyə aydınlaşdırılır.

```
@HISEQ:126:C1MD9ACXX:7:1101:2362:1868 1:N:0:CTTGTC
NTGCCGGCGGAGTCCTAAAAGTAACATCCGCCGATCCCTGGTCGGCATCGTTATGGTTGAGACTAGGACGGTATCTGATCGTCTTCGAGCCCCCAACTT
+
#1:DDDDFFHHGDHIIJJJJIBDHDHEIJIIJDEEEFFCCAC@BB:@CBDBBACDDDCDDDDCCCCD@C;B5<@CCDCDCB?8<?C@A?9>@DDDDDD>
```

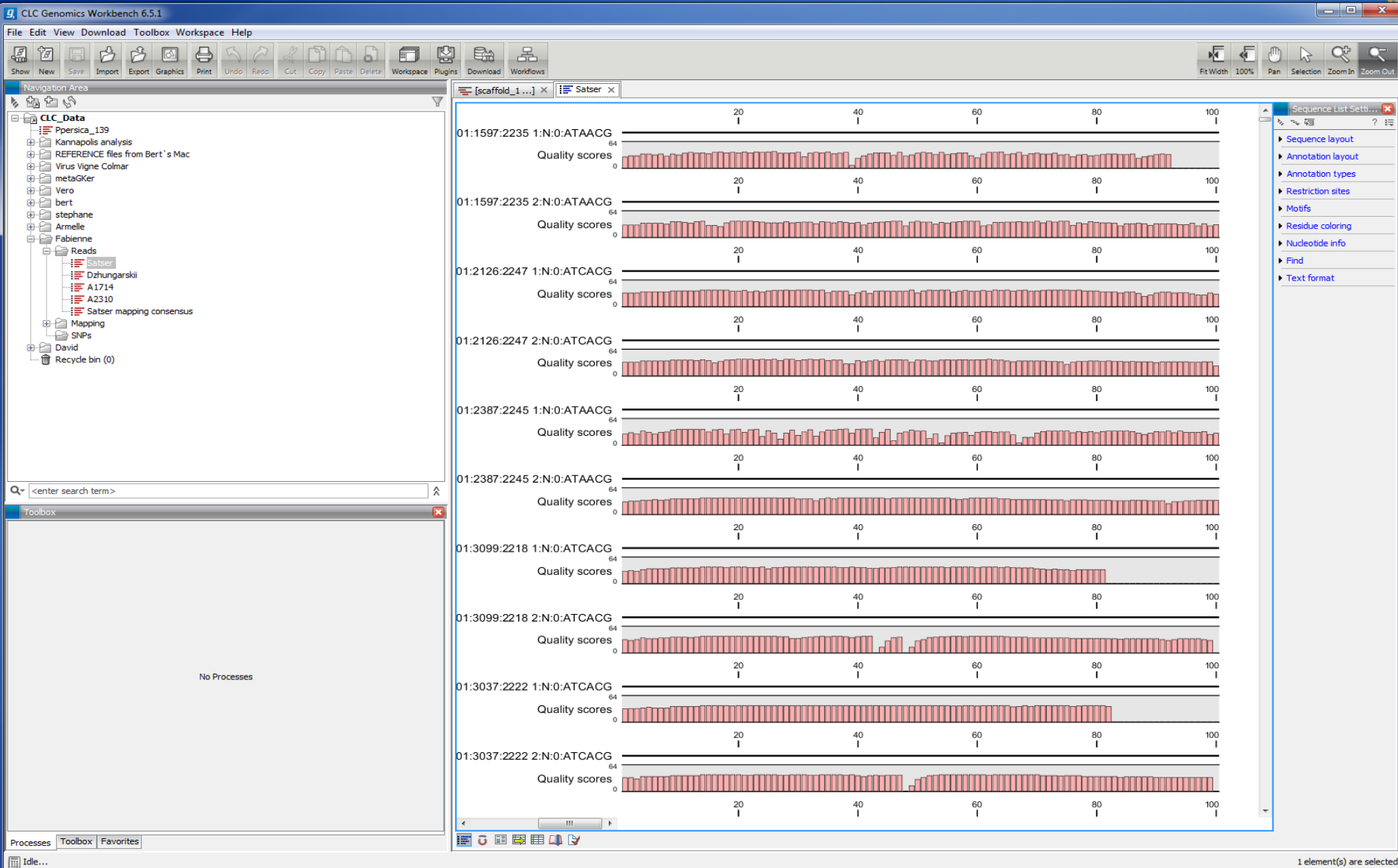


PHRED score probability is for predicting base to be on error - **nukleotidlərin səhv oxunma ehtimalını proqnozlaşdırmaq üçündür.**

PHRED score maksimum 40 qiymətini ala bilər ki, bu da əla nəticə deməkdir. Belə ki, bu halda nukleotidlərin səhv oxunma ehtimalı olduqca aşağıdır. If score is 40, it means we have excellent result. On this case, probability is being low. PHRED score-nun qiyməti 0-40 arasında olur. Score=10, 10 nukleotiddən birində (1/10), score=20, 100-də 1-də (1/100), score=30, 1000 nukleotiddən birində, score=40 olduqda, 10000 nukleotiddən birində səhvin mövcud olduğunu göstərir.



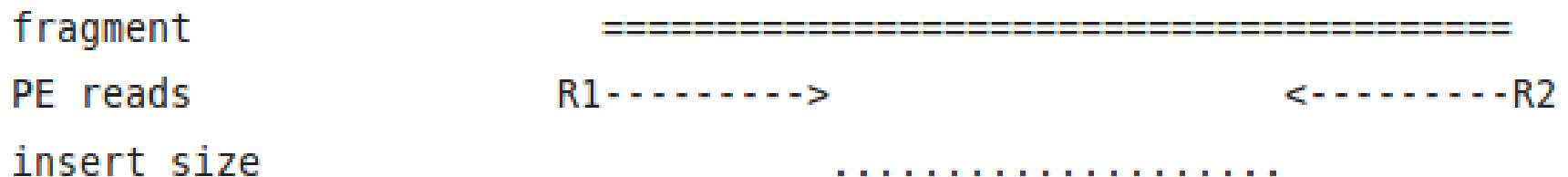
Şəkildə sekvenslər tərəfindən fayl şəklində hər bir read üçün verilmiş informasiyanın **CLC programı** tərəfindən oxunulmuş variantı (reads faylı) əks olunmuşdur. Müşahidə olunduğu kimi, hər bir readin ID ünvanı (adı), ardıcılığı və keyfiyyəti histqram formatında göstərilmişdir.



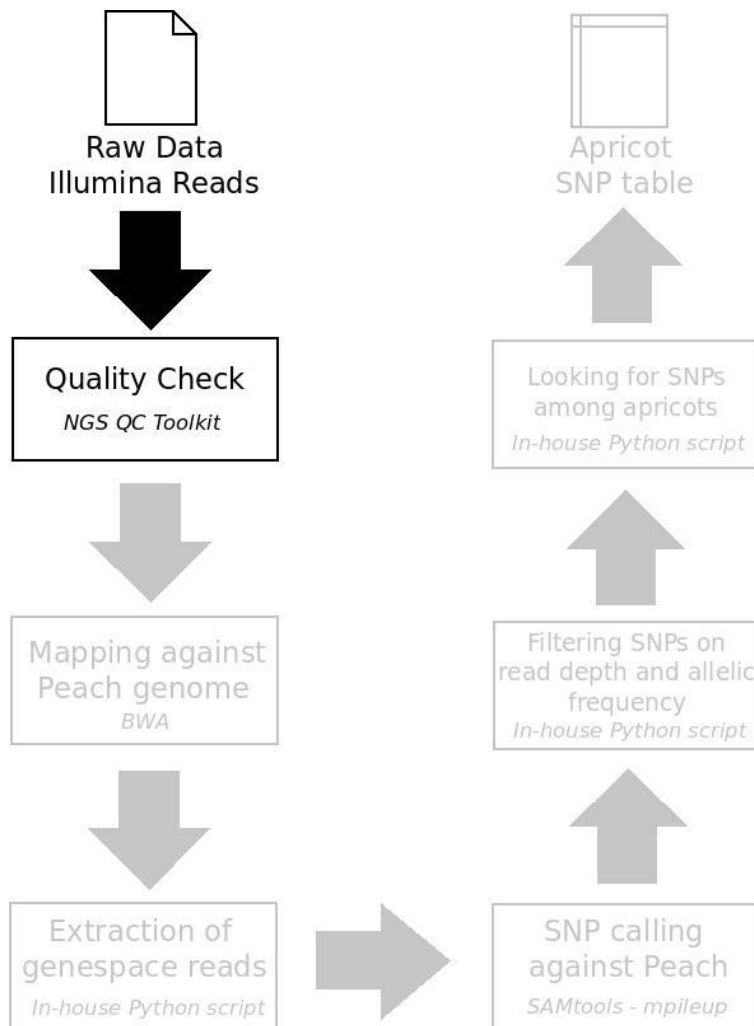
XAM MƏLUMATLAR (ROW DATA) VƏ ONLARIN İŞLƏNİLMƏSİ

Beləliklə,

- “Read”-DNT fraqmentinin müəyyən hissələrinin oxunulmuş DNT ardıcılıqlarıdır;
- Illumina read-lərinin hər biri 100 nc ölçüsündə, Ion Torrent read-lərinin ölçüsü isə 400 nc ölçüsündədir və FASTQ formatında əldə olunur;
- DNT fraqmentlərində hər bir pozisiyanın orta oxunma ölçüsü (average depth of coverage) 15 dəfəyə (15X) bərabər olmuşdur. (Coverage - total number of reads on the map for 1 position)
- Cüt sonluqlu read-lər arası məsafə 100-700 nc ölçüsündə olmalıdır.
- FASTQ faylının ölçüsü 2-3-GB –dır (gigabayt-109).



“Read”lərin keyfiyyətinin yoxlanılması

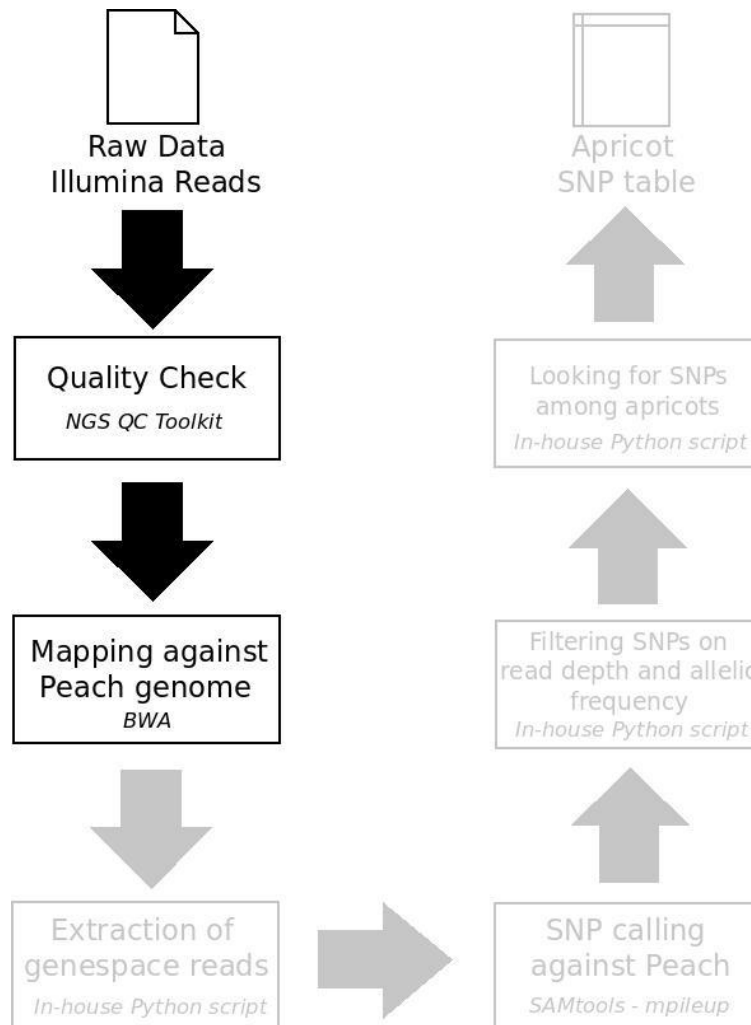


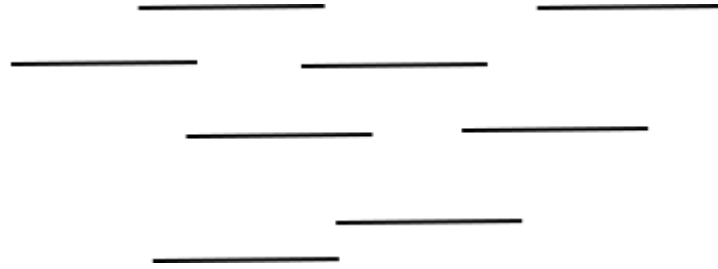
KEYFİYYƏTİN YOXLANILMASI – SEQUENSİNQ NƏTİCƏLƏRİNİN İCRA STRUKTURUNDA II MƏRHƏLƏDİR

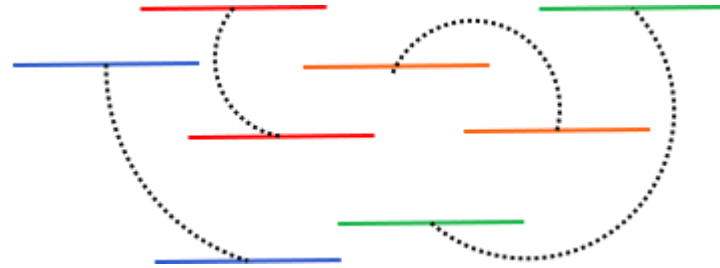
- Bu mərhələdə *NGS QC Toolkit* kompüter proqramının imkanlarından (ödənişsizdir) istifadə olunur; bu kompüter proqramı read-lərin keyfiyyətini yoxlamağa imkan verir.
- PHRED score vasitəsilə read-lərin keyfiyyəti haqqında mühakimə yürüdülmür.
- Bu zaman aşağıdakı kriterilərdən istifadə olunur:
 - PHRED score-nun minimum qiyməti 20 olmalıdır, yəni səhv nukleotid pozisiyalarının ehtimal olunan ehtimalı 1/100-ə bərabər olmalıdır (oxunulmuş DNT fraqmentlərində yalnız hər 100 nukleotiddən biri səhv ola bilər).
 - Read-lərin uzunluğunun 70%-nin PHRED score qiyməti 20 və ondan yüksək olmalıdır.
- Read-lərin uzunluğunun 80-95%-i bu təmizlik keyfiyyətinə (yəni score>20) malik olmalıdır.
- Readlərin 80-95%-i 20-dən yüksək keyfiyyət göstəricisinə malik olduqda read-lər keyfiyyətli (yəni düzgün oxunmuş) hesab olunur.

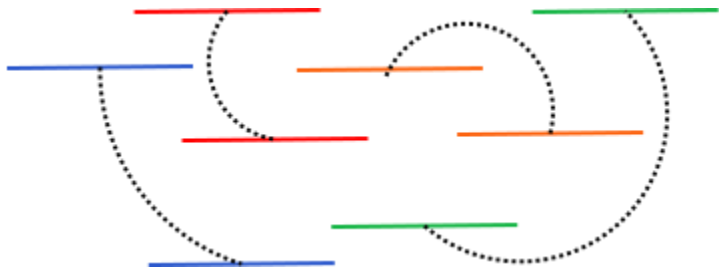


READLƏRİN XƏRİTƏLƏNMƏSİ VƏ YA DÜZLƏNDİRİLMƏSİ (ALIGNEMENT OF READS)





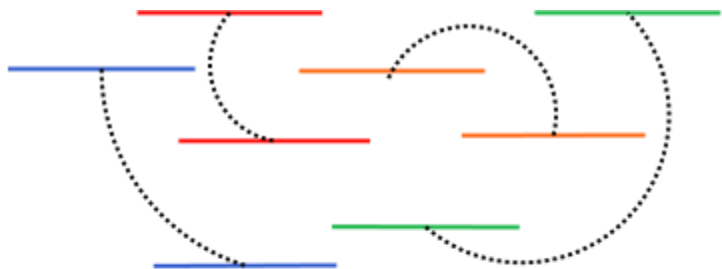




+

Reference genome





+

Reference genome



Hər bir DNT fraqmenti 15X, yəni 15 təkrar nəticəsində oxunmalıdır (**Sequencing coverage**).

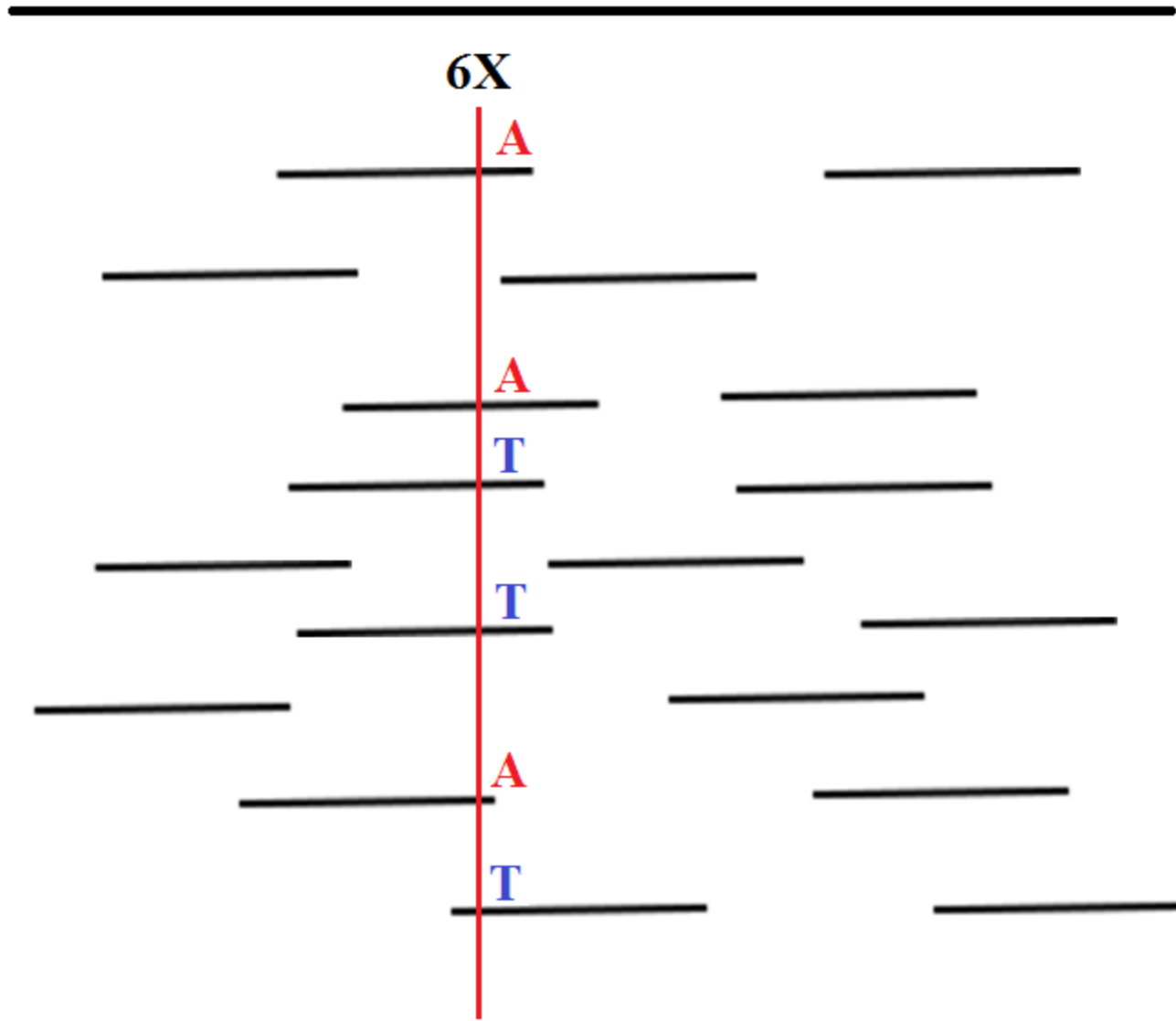
Lakin bu fraqmentlər başdan axıra qədər 15 dəfə oxunmur. Bir fraqmentin ayrı-ayrı hissələri, ~100 nc olmaqla, R1 və R2 cüt sonluqlu readlər kimi oxunur, sonra bir-birlərinə uyğun R1 və R2 readlər tanınır, onların və 100-700 nc ölçüsündə olan onlar arasındakı məsafənin ardıcılığı readlər düzləndirildikdən (alignment) sonra, readlərin bir-birlərinin üzərini qapası (overlapping) hesabına dəqiqləşdirilir.

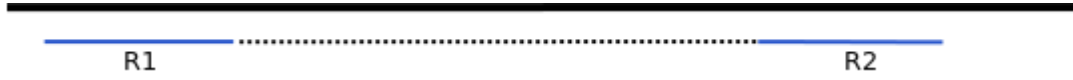
Hər bir fraqment üçün 15X, yəni 15 dəfə oxunmadan istifadə olunsa da, fraqmentdə hər bir pozisiya ən azı 6X (**mapping coverage**) əsasında tanınmalıdır. Hər bir pozisiyanın 6X vasitəsilə müəyyənləşdirilməsi homo-və heteroziqotluğu müəyyənləşdirmək, yaxud SNP-i təyin etmək üçün mütləqdir.

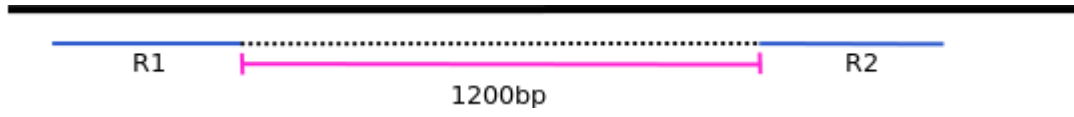
Bütün bu prosedurlar nümunə (reference) genom, onun nukleotid ardıcılığı əsasında icra olunur.

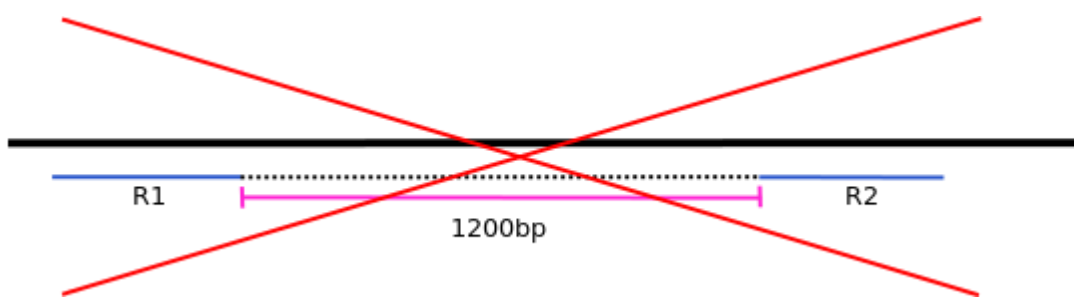


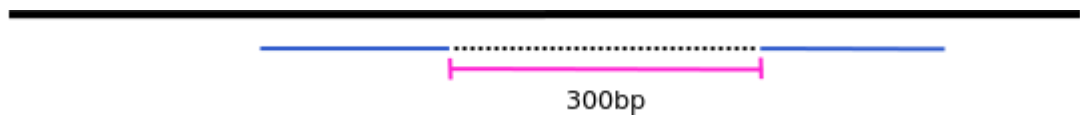
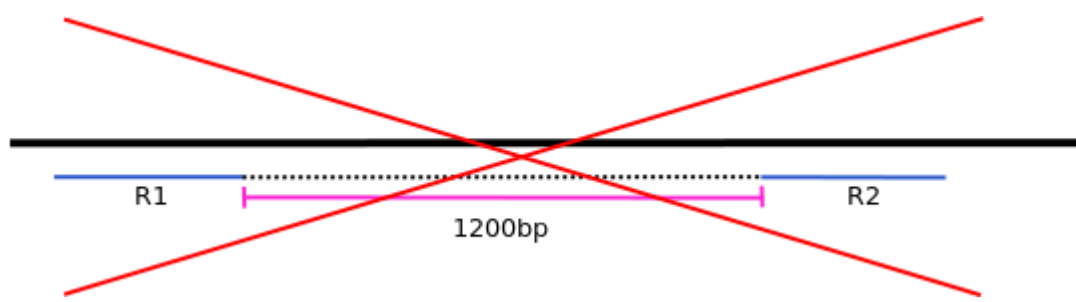
Reference genome

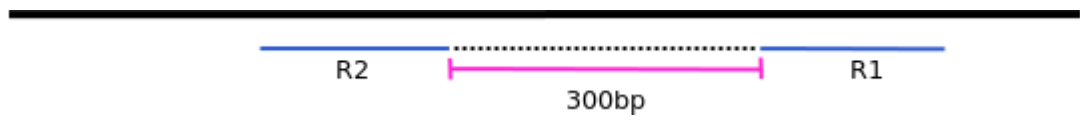
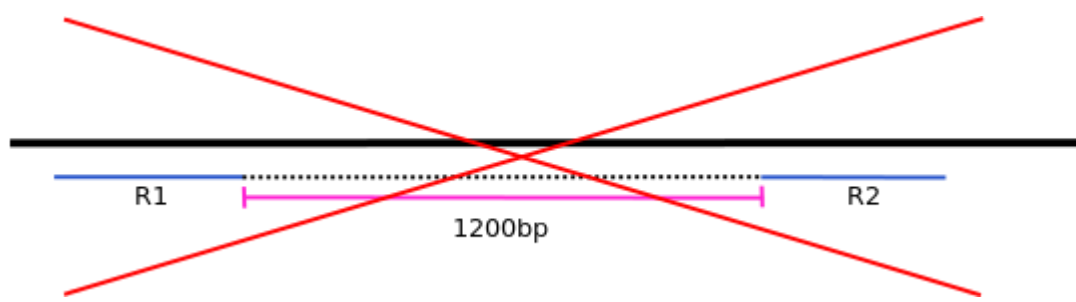


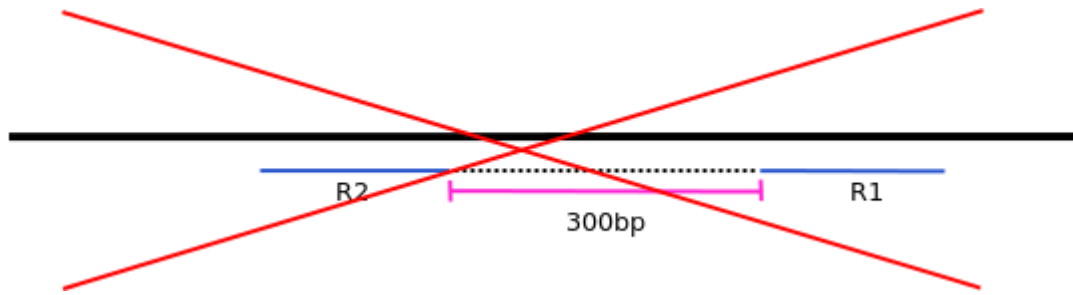
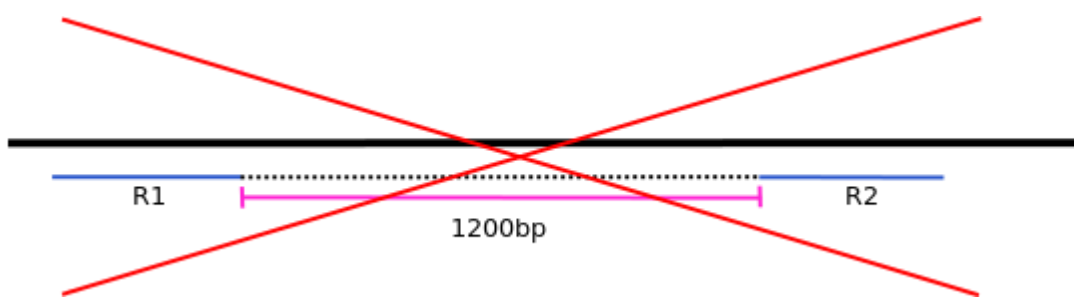


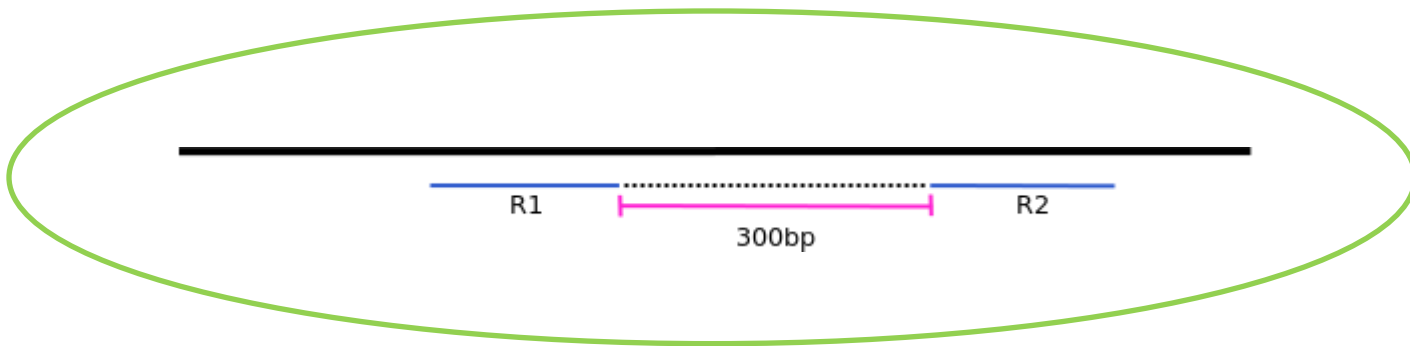
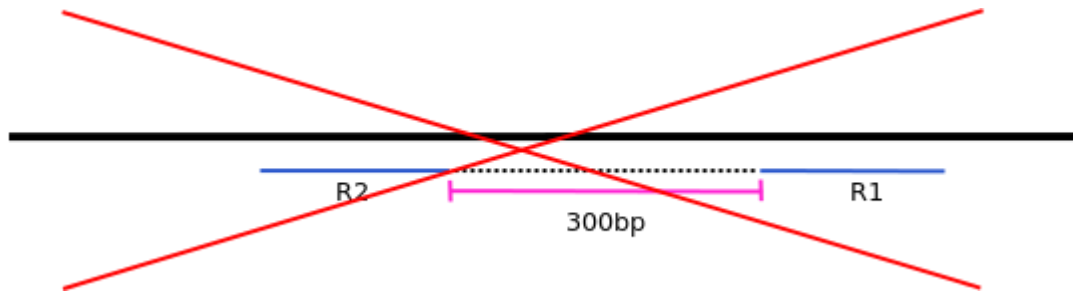
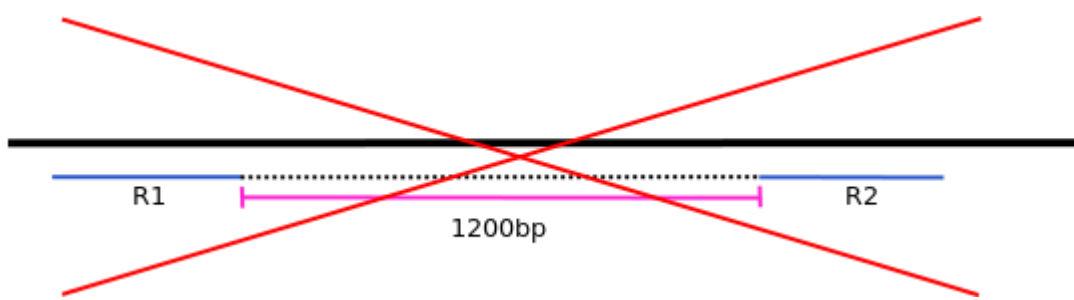












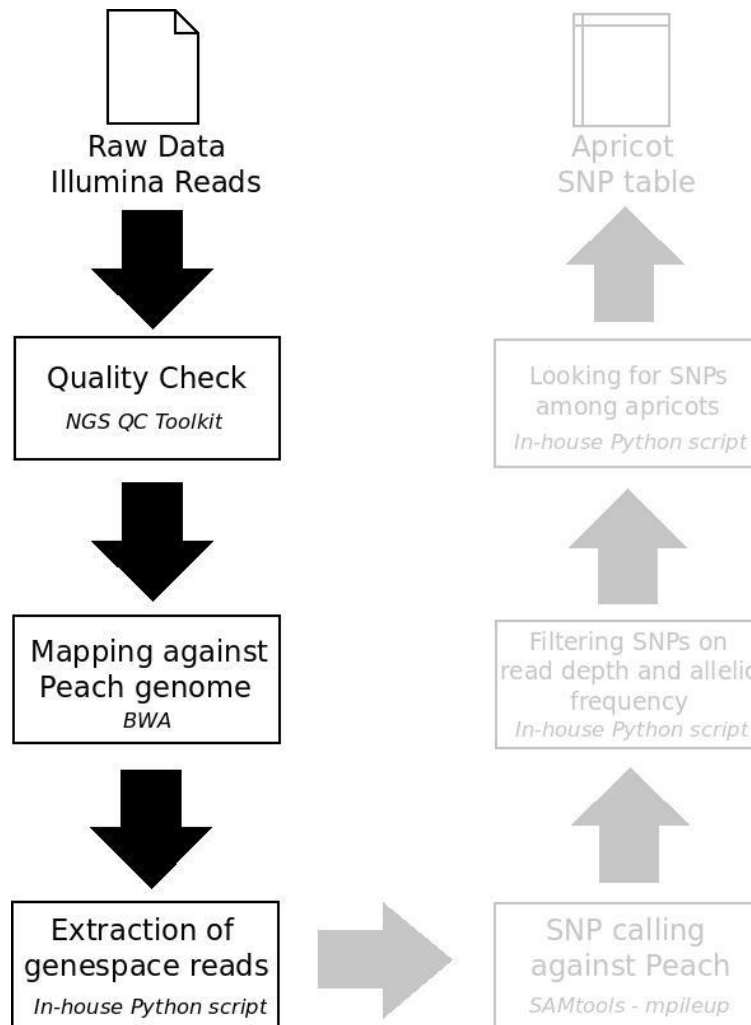
READ-LƏRİN XƏRİTƏLƏNMƏSİ (ALIGNEMENT)

Beləliklə,

- **Xəritələnmə (Alignment)** nümunə (**referens**) genom əsasında aparılmışdır (*de novo sequencing* deyil, resequencing adlanır). Cari tədqiqat işində nümunə genom kimi şaftalı genomundan istifadə edilmişdir; nümunə genom müxtəlif axtarış bazaları, məs., **NCBI** (National Centre of Biotechnology Information), **GDR** (Genome Database for Rosaceae), **TAIR** (The Arabidopsis Information Resource), **PGSB** (Plant Genome and Systems Biology-for Barley data) vasitəsilə gətirilə bilər və **FASTA** faylı şəklində verilir.
- **Xəritələnmə (Alignment)** və ya ardıcıl qruplaşdırma **BWA** kompüter programı əsasında aparılmışdır, pulsuzdur və dos sistemi kimi əmrlərlə (command) işləyir, həmçinin **CLC** vasitəsilə icra oluna bilər.
- Readlərin 50-60%-i nümunə (reference) genomu uyğun qruplaşdırılmalıdır .



GENLƏR ÜZƏRİNDƏ FOKUSLANMA

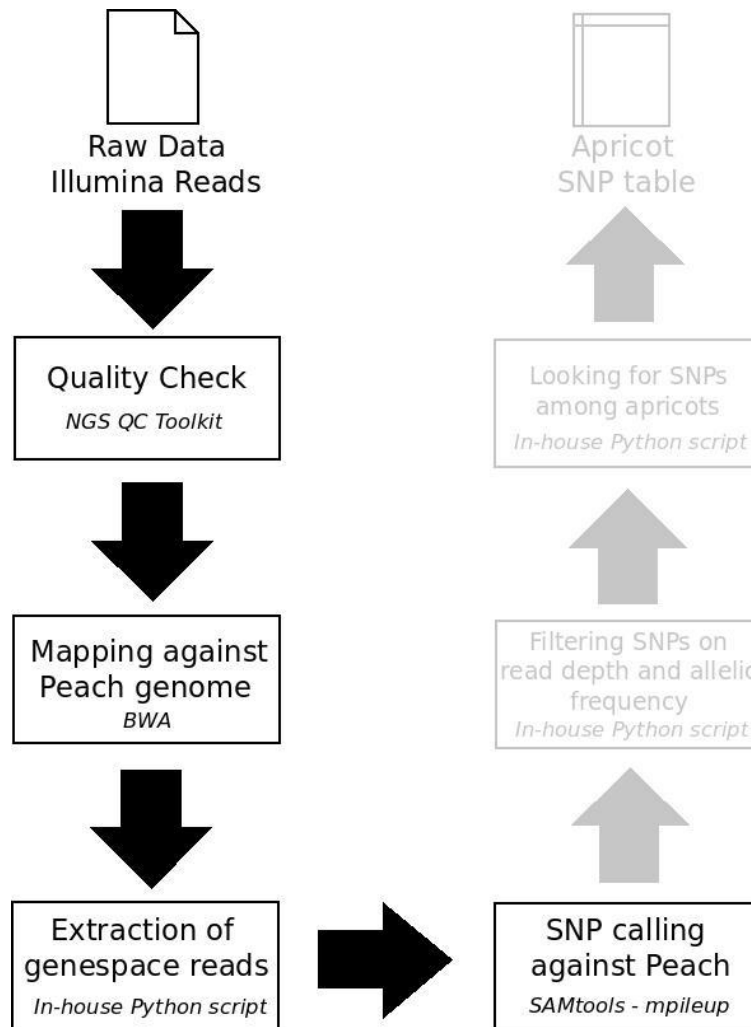


GENLƏR ÜZƏRİNDƏ FOKUSLANMA

- Genlər haqqında annotasiya informasiyası GFF fayl ölçüsündə saxlanılır və müxtəlif mənbələrdən, məs., GDR-dan (Genome Database for Rosaceae) alınır.
- Annotasiya əsasında gen regionlarına uyğun read-lərin ekstraksiyası (digərlərindən ayrılması) aparılır.
- Read-lərin 20-28%-i genlərdaxili sahələrə uyğun olmalıdır.

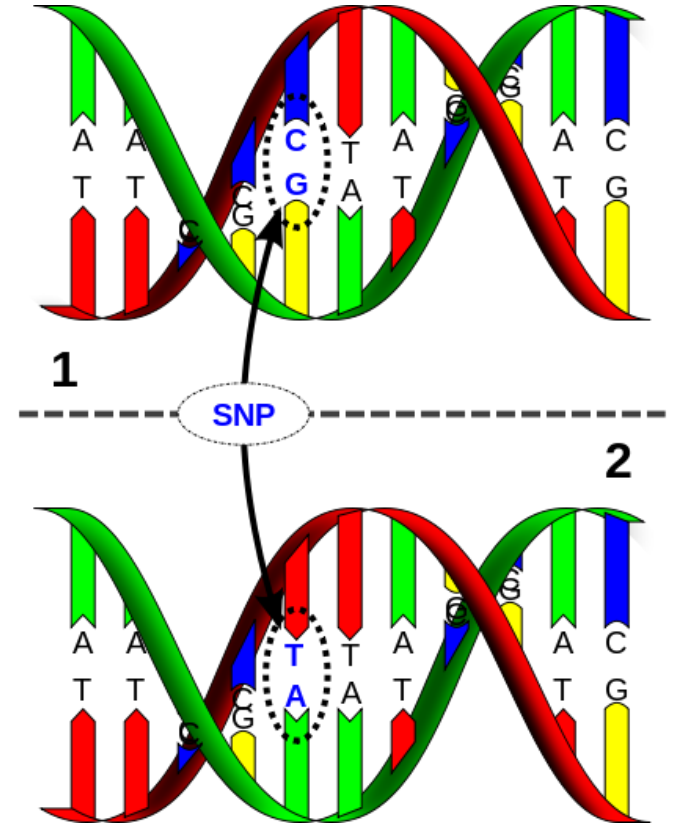


SNP-LƏRİN PROQNOZLAŞDIRILMASI

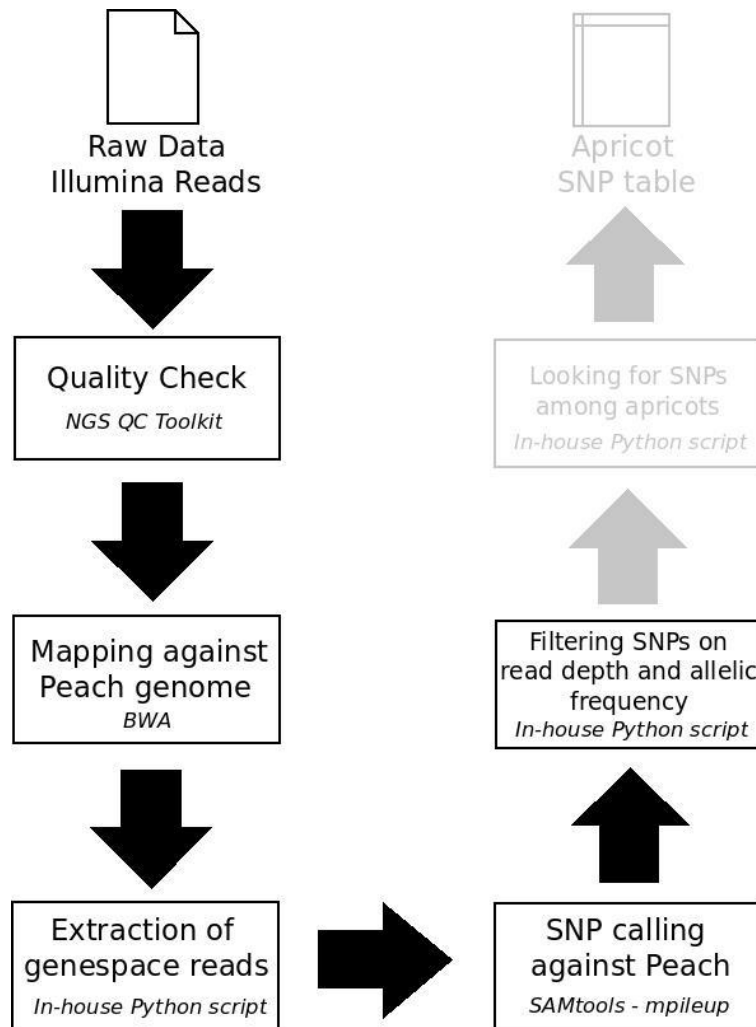


SNP-LƏRİN PROQNOZLAŞDIRILMASI

- **SNP** : Single-Nucleotide Polymorphism – vahid nukleotid polimorfizmləridir
- Şaftalı genomunda read-lərin düzləndirilməsi və ya xəritələnmə üçün mühüm olan addımlar
- **SAMtools** vasitəsilə proqnozlaşdırma
- Şaftalı və ərik genomları arasında gendaxili regionlarda 1.8-2 milyon SNP mövcuddur.

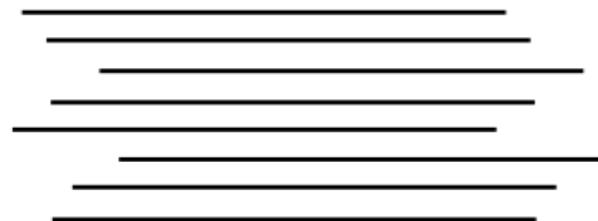
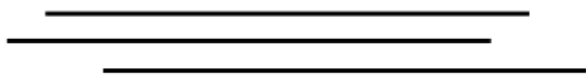


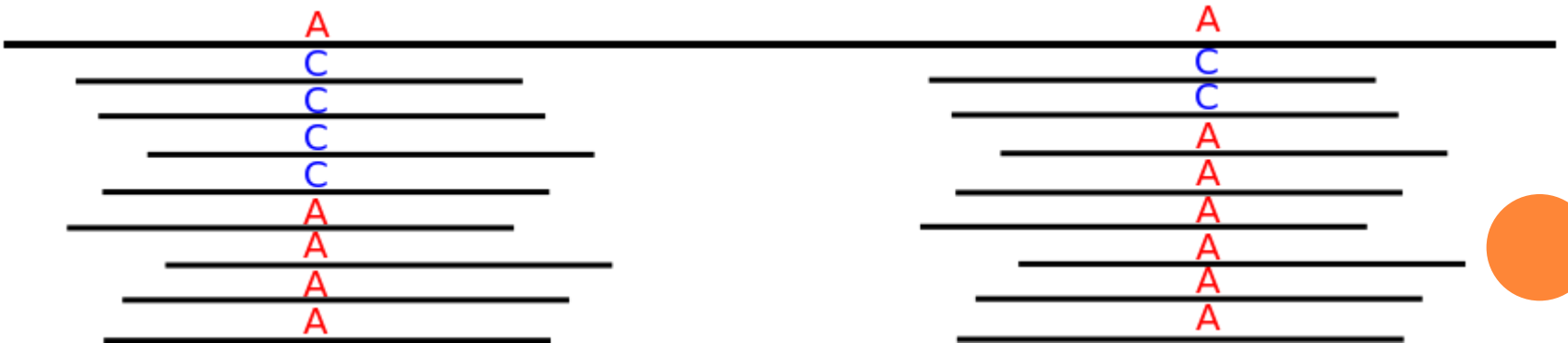
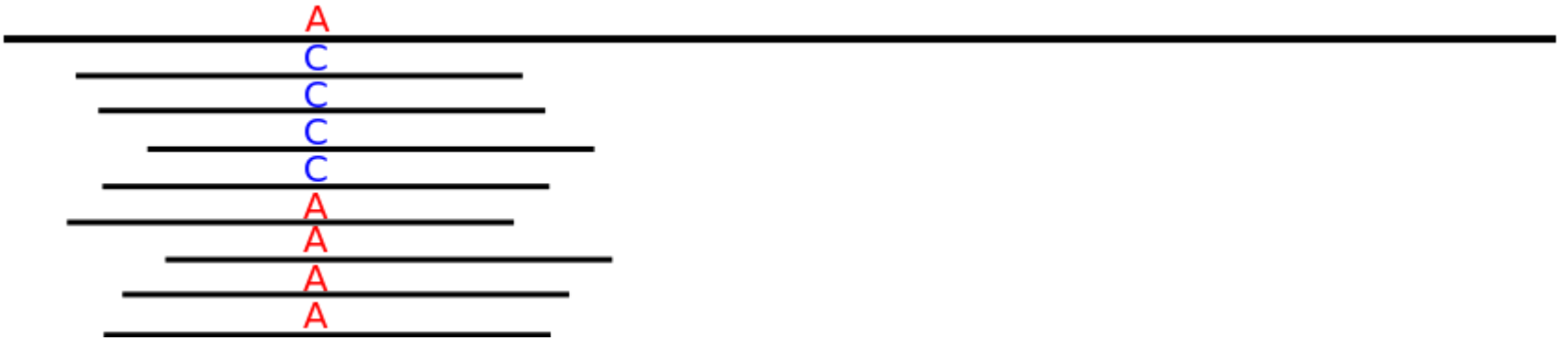
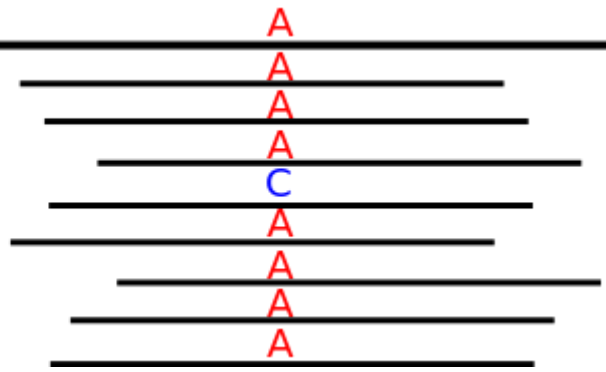
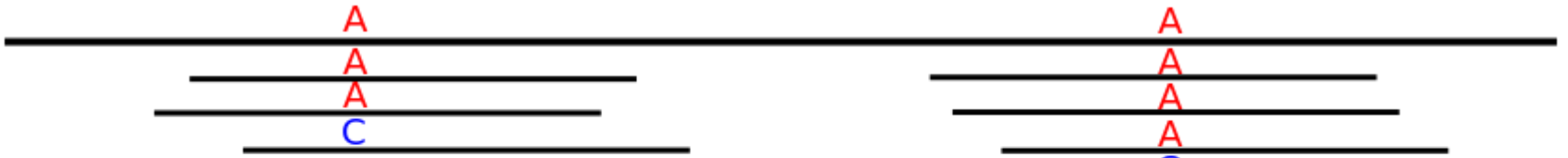
SNP-LƏRİN FİLTRASIYASI



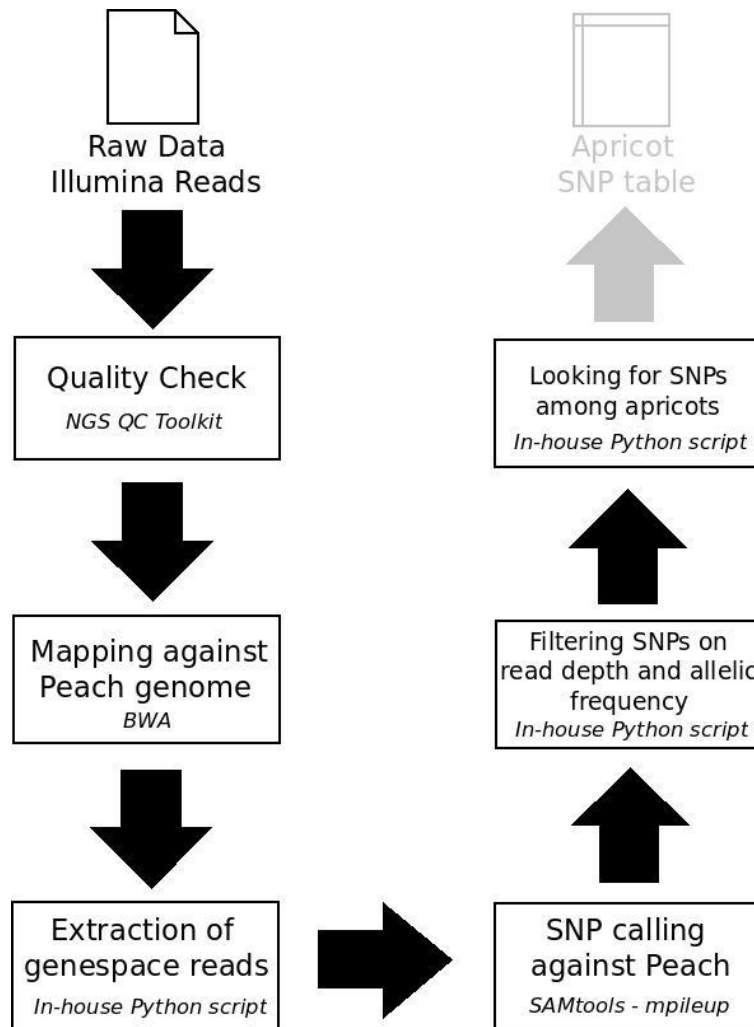
SNP-LƏRİN FİLTRASIYASI

- Ətraflı proqnozlaşdırma
- Filtrasiya aşağıdakı kriterilər əsasında aparılır:
 - Minimum təkrarların miqdarı (Minimum depth) müəyyənləşdirilir; bu məqsədlə 8X, 6X, 4X təkrarların nəticələri müqayisə olunur
 - Təyin edilmiş SNP-nin təsdiqi üçün alternativ allel variantına malik read-lərin sayı ümumi readlərin 35%-ni təşkil etməlidir. Yəni əgər hər hansı pozisiya 8 dəfə oxunmuşdursa (8X və ya coverage mappeng 8-ə bərabərdir, yəni 8 readlə oxunmuşdur), onda bu pozisiya üzrə SNP-nin mövcudluğunu (yəni homo- və heteroziqotluğu, yaxud, nümunə genomla müqayisədə nöqtəvi mutasiyanın baş verməsini) əminliklə söyləmək üçün, onlardan ən azı 3-ü (yəni 8-in 35%-i) alternativ allelə malik variant olmalıdır, əks təqdirdə, təyin olunmuş SNP seqvensinqin səhvi kimi qəbul olunur.





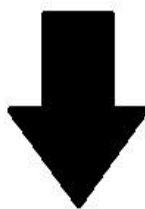
SNP-LƏRİN TƏYİNİ



SNP-LƏRİ NECƏ TƏYİN ETMƏLİ?

A	1	2	3	4	5	6	7
Peach	A	T	C	C	G	A	T
Apricot1	A	A	C	C	T	A	T

B	1	2	3	4	5	6	7
Peach	A	T	C	C	G	A	T
Apricot2	G	T	C	C	T	T	T



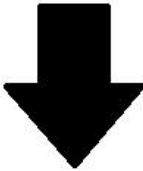
Apricot1	A	A	C	C	T	A	T
Apricot2	G	T	C	C	T	T	T



SNP-LƏRİ NECƏ TƏYİN ETMƏLİ?

A	1	2	3	4	5	6	7
Peach	A	T	C	C	G	A	T
Apricot1	A	A	C	C	T	A	T

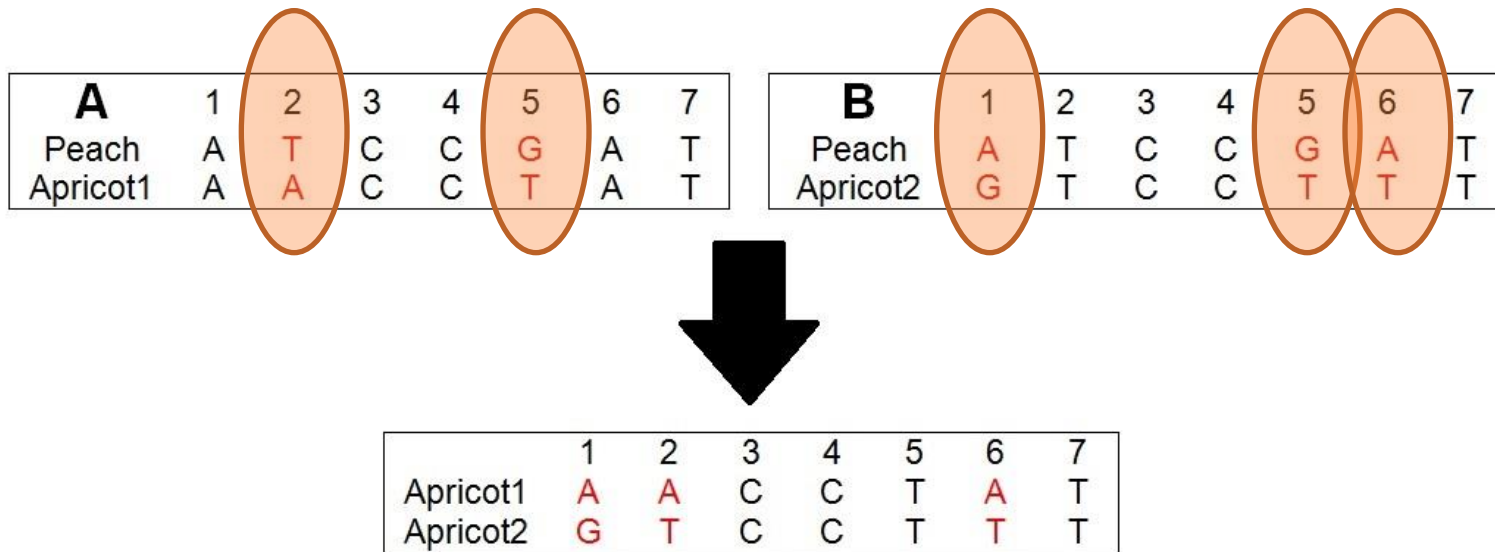
B	1	2	3	4	5	6	7
Peach	A	T	C	C	G	A	T
Apricot2	G	T	C	C	T	T	T



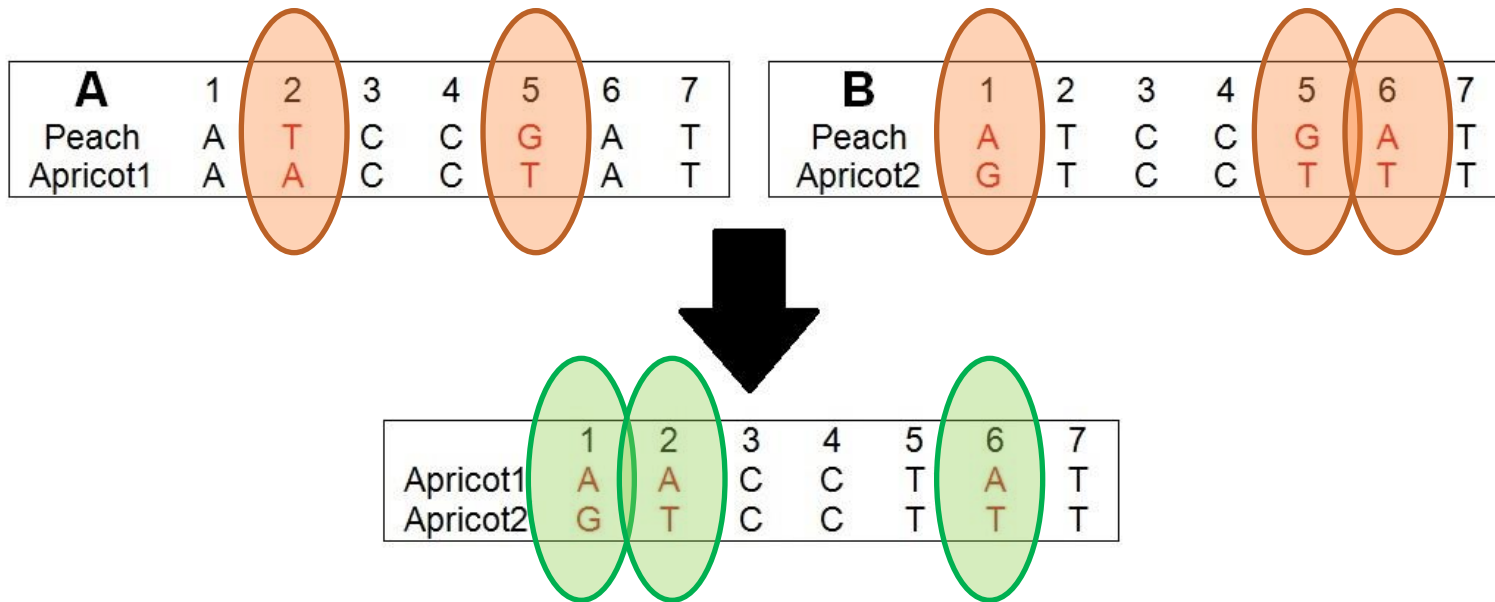
Apricot1	A	A	C	C	T	A	T
Apricot2	G	T	C	C	T	T	T



SNP-LƏRİ NECƏ TƏYİN ETMƏLİ?



SNP-LƏRİ NECƏ TƏYİN ETMƏLİ?

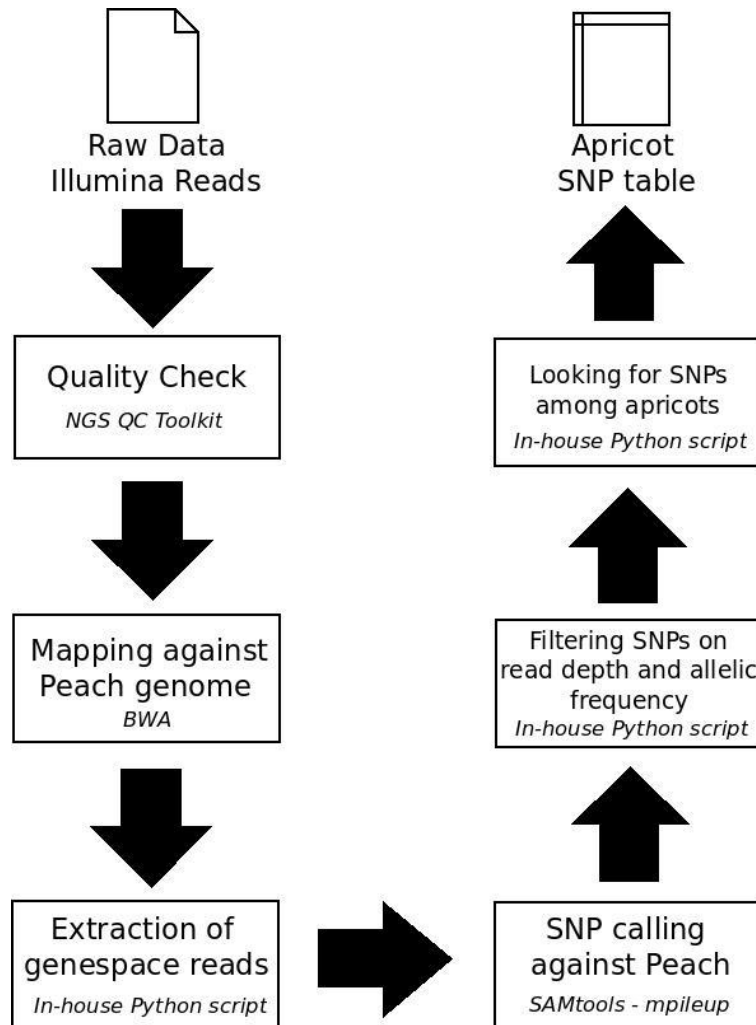


ƏLDƏ OLUNMUŞ MƏLUMATLARIN ÜMUMİLƏŞDİRİLMƏSİ

- Çoxlu sayda nümunə analiz edilməlidir (ilkin olaraq, müxtəlif praymerlərlə)
- Onlar bir-birləri ilə 2-2 müqayisə olunmalıdır;
- Hər bir lokus və hər bir genotip üçün SNP-lərin cədvəli tərtib olunmalıdır. Məs.,

	1	2	3	4	5	6	7	93-cü nümunə
1-ci pozisiya	A	A	A	A/T	T	A	T		
2-ci pozisiya	T	T	A	T/A	A	NA			
3-cü pozisiya və s.	CG	C	C	C	C	C/G			
- Coverage mapping 8-dir, yəni hər bir pozisiya ən azı 8 read vasitəsilə müəyyənləşdirilmişdir;
- Sequencing nəticəsində çoxlu sayda məs., 2909906 lokus və ya pozisiyada SNP öyrənilə bilər, yəni populyasiya daxilində göstərilən miqdarda lokuslar üzrə müxtəliflik aşkar edilə bilər və onlardan az miqdarı, məs., 220,779 lokusun SNP-ləri tamamilə dəqiqləşdirilə, yəni hər birinin nukleotid növü bütün genotiplərdə aşkar oluna, digər 2689127 lokusda isə hər hansı genotip və genotiplərdə müəyyən mövqə tam dəqiqləşdirilmədiyindən, onlar analizdən kənarlaşdırıla bilər.

DAİRƏ QAPANIR! SEQUENCING-İN NƏTİCƏLƏRİNİN İCRA STRUKTURUNUN TAMAMLANMASI



ƏLDƏ OLUNMUŞ MƏLUMATLARIN İSTİFADƏSİ

- **GWAS – Genome Wide Association Studies – Genom üzrə geniş assosiasiya olunmuş tədqiqatlar**
- **Şarkaya davamlılıq və ağacların çiçəkləmə müddəti ilə ilişikli lokusların təyini**
- **QTL analizi vasitəsilə əldə olunmuş məlumatlarla müqayisəsi**



***NGS nəticələrinin CLC kompüter
proqramı vasitəsilə işlənilməsi***





Navigation Area

- CLC_Data
 - Ppersica_139
 - Kannapolis analysis
 - REFERENCE files from Bert's Mac
 - Virus Vigne Colmar
 - metaGker
 - Vero
 - bert
 - stephane
 - Armelle
 - Fabienne
 - David
 - Recycle bin (0)

Q- <enter search term>

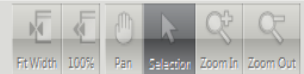
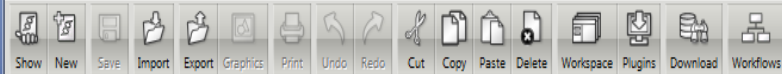
Toolbox

No Processes

Processes | **Toolbox** | Favorites

Quick Start

- Import data
- New sequence
- Read tutorials



Navigation Area

Standard Import... Ctrl+I

Tracks...

- Roche 454...
- Illumina...
- SOLID...
- Fasta...
- Sanger...
- Ion Torrent...
- SAM/BAM Mapping Files...
- Satser mapping consensus
- Mapping
 - A2310 mapping
 - A2310 mapping summary report
 - Satser mapping
 - Satser mapping summary report
 - A1714 mapping
 - A1714 mapping summary report
 - Dzhungarskii mapping
 - Dzhungarskii mapping summary report
- SNPs
- David
- Carole
 - Reference
 - Reads
 - TSS_SRR492389_1 (paired)
 - Fei_0_SRR095695_1 (paired)
 - TS1_SRR492388_1 (paired)
 - Mapping
- Recycle bin (0)

Q <enter search term>

Quick Start

- Import data
- New sequence
- Read tutorials

No Processes

Processes | Toolbox | Favorites

Idle...

MAPPING TABLE – MAPPING (XƏRİTƏLƏNMƏ) CƏDVƏLİ

The screenshot displays the CLC Genomics Workbench 6.5.1 interface. The main window shows a mapping table for scaffold_1. The top part of the mapping view features a gene model for scaffold_1 with two mRNA transcripts labeled "Gene mRNA" and "mRNA". Below this, the "Consensus" sequence is shown with a length of 20236. The "Coverage" plot displays individual reads as horizontal lines with colored arrows indicating mismatches. The x-axis of the coverage plot is labeled with genomic coordinates: 16,259,500, 16,260,000, 16,260,500, and 16,261,000. The left sidebar shows a navigation tree under "CLC_Data" with various project folders and mapping reports. The right sidebar contains the "Read Mapping Settings" panel, which is currently open. The settings are organized into sections: "Read layout" (Compactness: Packed, Gather sequences at top, Show sequence ends, Show mismatches, Disconnect pairs, Packed read height: Medium, Find Conflict), "Sequence layout" (No spacing, Numbers on sequences, Relative to: 1, Numbers on plus strand, Lock top sequence, Hide labels, Lock labels, Sequence label: Name, Matching residues as dots), "Annotation layout", "Annotation types", "Residue coloring", "Alignment info", "Nucleotide info", "Find", and "Text format". The bottom status bar shows "Processes", "Toolbox", and "Favorites" tabs, with "Idle..." selected.

XƏRİTƏLƏNMƏNİN STATİSTİK CƏDVƏLİ – MAPPING SUMMARY REPORT

CLC Genomics Workbench 6.5.1

File Edit View Download Toolbox Workspace Help

Show New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plugins Download Workflows

Fit Width 100% Pan Selection Zoom In Zoom Out

Navigation Area

- virus_vignie_Corimar
- metaGker
- Vero
- stephane
- Armelle
- Fabienne
 - Reads
 - Satsler
 - Dzhungarskii
 - A1714
 - A2310
 - Satsler mapping consensus
 - Mapping
 - A2310 mapping
 - A2310 mapping summary report
 - Satsler mapping
 - Satsler mapping summary report
 - A1714 mapping
 - A1714 mapping summary report
 - Dzhungarskii mapping
 - Dzhungarskii mapping summary report
 - SNPs
- David
- Carole
 - Reference
 - Reads
 - TSS_SRR492389_1 (paired)
 - Fel_0_SRR095695_1 (paired)
 - TS_1_SRR492388_1 (paired)
 - Mapping
 - Recycle bin (0)

Search: <enter search term>

Toolbox

No Processes

A2310 mapping...

1 Mapping summary report

1.1 Summary statistics

	Count	Percentage of reads	Average length	Number of bases	Percentage of bases
References	202	-	1,125,010.43	227,252,106	-
Mapped reads	3,654,517	90.13%	100.00	365,451,700	90.13%
Not mapped reads	400,421	9.87%	100.00	40,042,100	9.87%
Reads in pairs	2,085,726	51.44%	205.42	208,572,600	51.44%
Broken paired reads	1,568,791	38.69%	100.00	156,879,100	38.69%
Total reads	4,054,938	100.00%	100.00	405,493,800	100.00%

1.2 Distribution of read length

Length	Count
100	4,054,938

1.3 Distribution of mapped read length

Length	Count
100	3,654,517

1.4 Distribution of un-mapped read length

Length	Count
100	400,421

1.5 Paired reads distance distribution

Distance

Number of reads

Basepairs

Report Settings

- Table of Contents
 - 1 Mapping summary report
 - 1.1 Summary statistics
 - 1.2 Distribution of read length
 - 1.3 Distribution of mapped read length
 - 1.4 Distribution of un-mapped read length
 - 1.5 Paired reads distance distribution
 - Text format

SNP CƏDVƏLİ

CLC Genomics Workbench 6.5.1

File Edit View Download Toolbox Workspace Help

Show New Save Import Export Graphics Print Undo Redo Cut Copy Paste Delete Workspace Plugins Download Workflows

Navigation Area

- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) un-mapped re
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) un-mapped re
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping sum
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping vari
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping vari
- combined showy_nonshowy
- peach sequences
- chloroplast sequences
- INRA_Dupouy_130125_Run_Sample_7326_CCGTCC_008_R1_001 (paired) mapping
- INRA_Dupouy_130125_Run_Sample_7326_CCGTCC_008_R1_001 (paired) un-mapped re
- INRA_Dupouy_130125_Run_Sample_7326_CCGTCC_008_R1_001 (paired) un-mapped re
- INRA_Dupouy_130125_Run_Sample_7326_CCGTCC_008_R1_001 (paired) mapping sum
- INRA_Dupouy_130125_Run_Sample_7326_CCGTCC_008_R1_001 (paired) mapping vari
- Sample_S7315_130820HSeq_Run_TAGCTT_L005_R1_001 (paired) mapping-1 (Variants)
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping (Vari
- Sample_S7315_130820HSeq_Run_TAGCTT_L005_R1_001 (paired) mapping-1 (Variants)
- Sample_S7323_130820HSeq_Run_GGTACG_L006_R1_001 (paired) mapping (Variants)
- Sample_S7318_130820HSeq_Run_GTTTCG_L006_R1_001 (paired) mapping (Variants)
- Sample_S7324_130820HSeq_Run_GAGTGG_L006_R1_001 (paired) mapping (Variants)
- Sample_S7317_130820HSeq_Run_GTGCC_L006_R1_001 (paired) mapping (Variants)
- INRA_Dupouy_130125_Run_Sample_7325_CCGTCC_008_R1_001 (paired) mapping (Vari
- INRA_Dupouy_130125_Run_Sample_7325_ATGTCA_008_R1_001 (paired) mapping (Vari
- Sample_S7315_130820HSeq_Run_TAGCTT_L005_R1_001 (paired) mapping (Variants)
- Sample_S7316_130820HSeq_Run_GGCTAC_L005_R1_001 (paired) mapping (Variants)

stephane
Armelle
Fabienne
David
Recycle bin (0)

Q <enter search term>

No Processes

Processes Toolbox Favorites

INRA_Dupouy_1... x

Rows: 25,445 Variants

Mapping	Refere...	Type	Length	Refere...	Allele	Linkage	Zygotity	Count	Cover...	Freque...	Forwar...	Avera...	Overla...	Coding...	Amino ...
scaffold_1_m...	227	SNV	1	T	G		Heterozy...	22	36	61.11	0.30	36.70			
scaffold_1_m...	298	SNV	1	G	A		Homozyg...	27	38	71.05	0.33	35.85			
scaffold_1_m...	323	SNV	1	A	T		Homozyg...	31	44	70.45	0.42	36.45			
scaffold_1_m...	327	SNV	1	T	A		Homozyg...	32	46	69.57	0.50	36.81			
scaffold_1_m...	331	SNV	1	A	G		Homozyg...	31	45	68.89	0.45	37.52			
scaffold_1_m...	376	SNV	1	T	C		Homozyg...	30	58	51.72	0.33	38.00			
scaffold_1_m...	387	SNV	1	T	G		Homozyg...	46	61	75.41	0.43	38.43			
scaffold_1_m...	401	SNV	1	G	C		Homozyg...	47	62	75.81	0.48	37.21			
scaffold_1_m...	422	SNV	1	C	T		Heterozy...	24	68	35.29	0.42	37.25			
scaffold_1_m...	451	SNV	1	T	C		Homozyg...	46	61	75.41	0.43	37.13			
scaffold_1_m...	557	SNV	1	C	A		Heterozy...	28	57	49.12	0.31	37.00			
scaffold_1_m...	573	SNV	1	A	C		Homozyg...	37	55	67.27	0.30	36.59			
scaffold_1_m...	582	SNV	1	A	C		Heterozy...	25	48	52.08	0.28	37.96			
scaffold_1_m...	590	Deletion	1	T	-		Heterozy...	25	47	53.19	0.24	37.28			
scaffold_1_m...	709	SNV	1	T	C		Heterozy...	10	25	40.00	0.10	33.70			
scaffold_1_m...	2221	SNV	1	A	C		Heterozy...	9	25	36.00	0.10	36.20			
scaffold_1_m...	2333	SNV	1	C	T		Heterozy...	16	33	48.48	0.50	38.94			
scaffold_1_m...	2468	SNV	1	A	G		Heterozy...	10	21	47.62	0.20	36.20			
scaffold_1_m...	2482	SNV	1	T	G		Heterozy...	12	27	44.44	0.25	35.67			
scaffold_1_m...	4450	SNV	1	C	A		Homozyg...	9	9	100.00	0.11	37.22			
scaffold_1_m...	4538	Deletion	1	C	-		Homozyg...	18	18	100.00	0.22	35.78			
scaffold_1_m...	4565	Deletion	1	T	-		Homozyg...	15	15	100.00	0.33	37.73			
scaffold_1_m...	4571	Insertion	1	-	A		Homozyg...	16	16	100.00	0.31	38.38			
scaffold_1_m...	4612	SNV	1	T	C		Homozyg...	16	16	100.00	0.38	37.69			
scaffold_1_m...	5494	SNV	1	G	A		Heterozy...	6	17	35.29	0.33	37.33			
scaffold_1_m...	5519	SNV	1	G	T	1	Heterozy...	9	23	39.13	0.33	37.50			
scaffold_1_m...	5520	SNV	1	C	T	1	Heterozy...	9	23	39.13	0.33	37.50			
scaffold_1_m...	5528	SNV	1	C	T		Heterozy...	10	25	40.00	0.40	37.70			
scaffold_1_m...	5541	SNV	1	G	A		Heterozy...	11	28	39.29	0.45	36.73			
scaffold_1_m...	5600	SNV	1	A	G		Heterozy...	15	32	46.88	0.50	38.12			
scaffold_1_m...	5605	SNV	1	G	A		Heterozy...	15	32	46.88	0.50	37.25			
scaffold_1_m...	5615	SNV	1	T	C		Heterozy...	13	27	48.15	0.43	38.21			
scaffold_1_m...	5642	SNV	1	G	A	2	Heterozy...	8	17	47.06	0.38	39.44			
scaffold_1_m...	5643	SNV	1	A	T	2	Heterozy...	8	17	47.06	0.38	39.44			
scaffold_1_m...	5655	SNV	1	G	A		Heterozy...	7	18	38.89	0.43	37.43			
scaffold_1_m...	5683	SNV	1	G	A		Homozyg...	22	22	100.00	0.32	34.32			
scaffold_1_m...	5711	SNV	1	G	C		Heterozy...	13	26	50.00	0.38	37.69			
scaffold_1_m...	5715	SNV	1	A	G		Heterozy...	13	26	50.00	0.38	37.69			
scaffold_1_m...	5720	SNV	1	A	G		Heterozy...	12	26	46.15	0.33	38.17			
scaffold_1_m...	5723	SNV	1	G	C		Heterozy...	14	28	50.00	0.36	36.64			
scaffold_1_m...	5808	SNV	1	C	T		Heterozy...	10	26	38.46	0.50	38.10			
scaffold_1_m...	5872	SNV	1	G	A		Heterozy...	9	25	36.00	0.22	38.33			
scaffold_1_m...	5989	SNV	1	C	T	3	Heterozy...	14	34	41.18	0.36	37.50			
scaffold_1_m...	5990	SNV	1	G	C	3	Heterozy...	14	34	41.18	0.36	37.50			
scaffold_1_m...	6023	SNV	1	T	C		Heterozy...	20	42	47.62	0.45	37.95			
scaffold_1_m...	6098	SNV	1	C	T	4	Heterozy...	13	34	38.24	0.46	37.15			
scaffold_1_m...	6099	SNV	1	A	G	4	Heterozy...	13	34	38.24	0.46	37.15			
scaffold_1_m...	6100	SNV	1	T	G	4	Heterozy...	13	34	38.24	0.46	37.15			
scaffold_1_m...	6202	SNV	1	T	G		Heterozy...	20	49	40.82	0.35	35.00			
scaffold_1_m...	6354	SNV	1	G	A		Heterozy...	14	30	46.67	0.50	37.64			
scaffold_1_m...	6537	SNV	1	T	C		Heterozy...	17	37	45.95	0.47	34.12			
scaffold_1_m...	6617	SNV	1	G	A		Heterozy...	19	41	46.34	0.42	38.47			
scaffold_1_m...	6640	SNV	1	A	C		Heterozy...	14	33	42.42	0.29	37.93			
scaffold_1_m...	6675	SNV	1	A	T		Heterozy...	16	33	48.48	0.44	37.44			
scaffold_1_m...	6730	SNV	1	A	T		Heterozy...	17	39	43.59	0.18	38.00			
scaffold_1_m...	7209	SNV	1	C	T		Heterozy...	12	31	38.71	0.15	36.69			

Table Settings

Column width: Automatic

Show column

- Mapping
- Reference Position
- Type
- Length
- Reference
- Allele
- Linkage
- Zygotity
- Count
- Coverage
- Frequency
- Forward/reverse balance
- Average quality
- Overlapping annotations
- Coding region change
- Amino acid change

Select All

Deselect All

1 element(s) is selected

*Bioinformatik programların tətbiqi ilə
izah olunmuş NGS nəticələrinin
statistik analizi*



Mapping table – Mapping cædvæli

D:\PROJET ABRIWGP\Polymorphisme\LINK\LINK_INPUT_NewPipelines\Seuil2_avechétérozygotes_lastpipeline\genotype_8X_vcf_v2_74acc_sharka_complete_SNPs_2all_2v2.map - Notepad++

Fichier Edition Recherche Affichage Encodage Langage Paramétrage Macro Exécution TextFX Compléments Documents ?

genotype_8X_vcf_v2_74acc_sharka_complete_SNPs_2all_2v2.ped genotype_8X_vcf_v2_74acc_sharka_complete_SNPs_2all_2v2.map

```
1 scaffold_1 294 47077 0.9
2 scaffold_1 423 49900 0.9
3 scaffold_1 424 49903 0.9
4 scaffold_1 523 57824 0.9
5 scaffold_1 532 57887 0.9
6 scaffold_1 533 57894 0.9
7 scaffold_1 646 72836 0.9
8 scaffold_1 750 90055 0.9
9 scaffold_1 751 90067 0.9
10 scaffold_1 801 90754 0.9
11 scaffold_1 895 91585 0.9
12 scaffold_1 896 91588 0.9
13 scaffold_1 897 91595 0.9
14 scaffold_1 898 91600 0.9
15 scaffold_1 900 91605 0.9
16 scaffold_1 1099 101643 0.9
17 scaffold_1 1108 101903 0.9
18 scaffold_1 1111 101944 0.9
19 scaffold_1 1118 102145 0.9
20 scaffold_1 1135 102646 0.9
21 scaffold_1 1136 102648 0.9
22 scaffold_1 1449 142672 0.9
23 scaffold_1 1450 142707 0.9
24 scaffold_1 1666 147995 0.9
25 scaffold_1 1672 148067 0.9
26 scaffold_1 1784 150689 0.9
27 scaffold_1 1878 155185 0.9
28 scaffold_1 1879 155190 0.9
29 scaffold_1 1880 155205 0.9
30 scaffold_1 1894 155349 0.9
31 scaffold_1 1904 155450 0.9
32 scaffold_1 1905 155451 0.9
33 scaffold_1 1906 155456 0.9
34 scaffold_1 1907 155482 0.9
35 scaffold_1 1908 155483 0.9
36 scaffold_1 1909 155486 0.9
37 scaffold_1 1910 155496 0.9
38 scaffold_1 1912 155498 0.9
39 scaffold_1 1913 155500 0.9
40 scaffold_1 1915 155511 0.9
41 scaffold_1 1918 155525 0.9
```

Normal text file length : 5149220 lines : 188307 Ln : 1 Col : 1 Sel : 0 UNIX ANSI INS

démarrer Explorer... Tinn-R R Console (32-bit) Microsoft Excel ... FlintGarciaAnnu... Microsoft Powe... D:\PROJET ABR... FR 15:14

Diversity results from PLINK software – PLINK kompüter programının tətbiqi ilə genetik müxtəlifliyin araşdırılması

14-02-14-lastpipeline.xlsx - Microsoft Excel

Accueil Insertion Mise en page Formules Données Révision Affichage

Insérer une fonction Somme automatique Utilisé(s) récemment Financier Logique Texte Date et heure Recherche et référence Maths et trigonométrie Plus de fonctions

Gestionnaire de noms Définir un nom Utiliser dans la formule Créer à partir de la sélection Noms définis

Repérer les antécédents Repérer les dépendants Supprimer les flèches Afficher les formules Vérification des erreurs Évaluation de formule Fenêtre Espion Options de calcul

Bibliothèque de fonctions

Audit de formules

O16 fx 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	CHR	SNP	bp	TEST	A1	A2	GENO	O(HET)	E(HET)	P	MAF	NCHROBS	Fis	-0.13632961	CHR	SNP	bp	TEST	A1	A2	GENO	O(HET)	E(HET)	P	MAF	NCHROBS
2	1	294	47077	ALL	2	1	1/0/73	0	0.02666	0.006803	0.01351	148	1	0.29744615	1	895	91585	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5	
3	1	423	49900	ALL	1	2	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	0.20059444	1	896	91588	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5	
4	1	424	49903	ALL	1	2	1/0/73	0	0.02666	0.006803	0.01351	148	1	1	2302	182885	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
5	1	523	57824	ALL	1	2	0/7/67	0.09459	0.09012	1	0.0473	148	-0.04960053	1	2322	183272	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
6	1	532	57887	ALL	1	2	0/21/53	0.2838	0.2435	0.3391	0.1419	148	-0.16550308	1	2324	183297	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
7	1	533	57894	ALL	1	2	0/72/2	0.973	0.4996	7.86E-19	0.4865	148	-0.94755805	1	2328	183365	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
8	1	646	72836	ALL	2	1	1/0/73	0	0.02666	0.006803	0.01351	148	1	1	2329	183366	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
9	1	750	90055	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	5159	313953	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
10	1	751	90067	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	5261	320315	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
11	1	801	90754	ALL	2	1	0/43/31	0.5811	0.4123	0.0001716	0.2905	148	-0.40941062	1	9299	442910	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
12	1	895	91585	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5	148	-1	1	9304	442942	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
13	1	896	91588	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5	148	-1	1	9423	444160	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
14	1	897	91595	ALL	1	2	0/20/54	0.2703	0.2337	0.3411	0.1351	148	-0.15661104	1	10264	496981	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
15	1	898	91600	ALL	2	1	0/73/1	0.9865	0.4999	3.59E-20	0.4932	148	-0.97339468	1	10266	497027	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
16	1	900	91605	ALL	1	2	0/44/30	0.5946	0.4178	0.0001408	0.2973	148	-0.42316898	1	10267	497028	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
17	1	1099	101643	ALL	1	2	0/17/57	0.2297	0.2033	0.5843	0.1149	148	-0.12985735	1	10268	497046	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
18	1	1108	101903	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	10651	514158	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
19	1	1111	101944	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	10855	525227	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
20	1	1118	102145	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	10873	525340	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
21	1	1135	102646	ALL	2	1	0/2/72	0.02703	0.02666	1	0.01351	148	-0.01387847	1	12508	595642	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
22	1	1136	102648	ALL	1	2	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	12511	595708	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
23	1	1449	142672	ALL	2	1	0/17/57	0.2297	0.2033	0.5843	0.1149	148	-0.12985735	1	12512	595709	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
24	1	1450	142707	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	12917	647920	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
25	1	1666	147995	ALL	2	1	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	12925	647981	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
26	1	1672	148067	ALL	2	1	4/14/56	0.1892	0.2531	0.04467	0.1486	148	0.25246938	1	13152	649522	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
27	1	1784	150689	ALL	1	2	0/1/73	0.01351	0.01342	1	0.006757	148	-0.00670641	1	13175	649677	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
28	1	1878	155185	ALL	1	2	0/2/72	0.02703	0.02666	1	0.01351	148	-0.01387847	1	13178	649704	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
29	1	1879	155190	ALL	2	1	0/60/14	0.8108	0.4821	4.99E-10	0.4054	148	-0.68180875	1	13386	651887	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
30	1	1880	155205	ALL	2	1	0/56/18	0.7568	0.4704	2.92E-08	0.3784	148	-0.60884354	1	13422	652181	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
31	1	1894	155349	ALL	1	2	0/52/22	0.7027	0.4558	1.02E-06	0.3514	148	-0.54168495	1	16280	758420	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		
32	1	1904	155450	ALL	1	2	0/55/19	0.7432	0.467	4.70E-08	0.3716	148	-0.59143469	1	16314	758770	ALL	1	2	0/74/0	1	0.5	8.83E-22	0.5		

Prêt seuilà2 seuilà35%

démarrer Seuilà2_ave... 2011-HORV... Seuilà2_ave... Seuilà2_ave... Tinn-R R Console (... Microsoft Ex... FlintGarciaA... FR 15:13

Resequencing və analiz olunmuş rüşeym plazması materialının genetik strukturu (Frappe software)



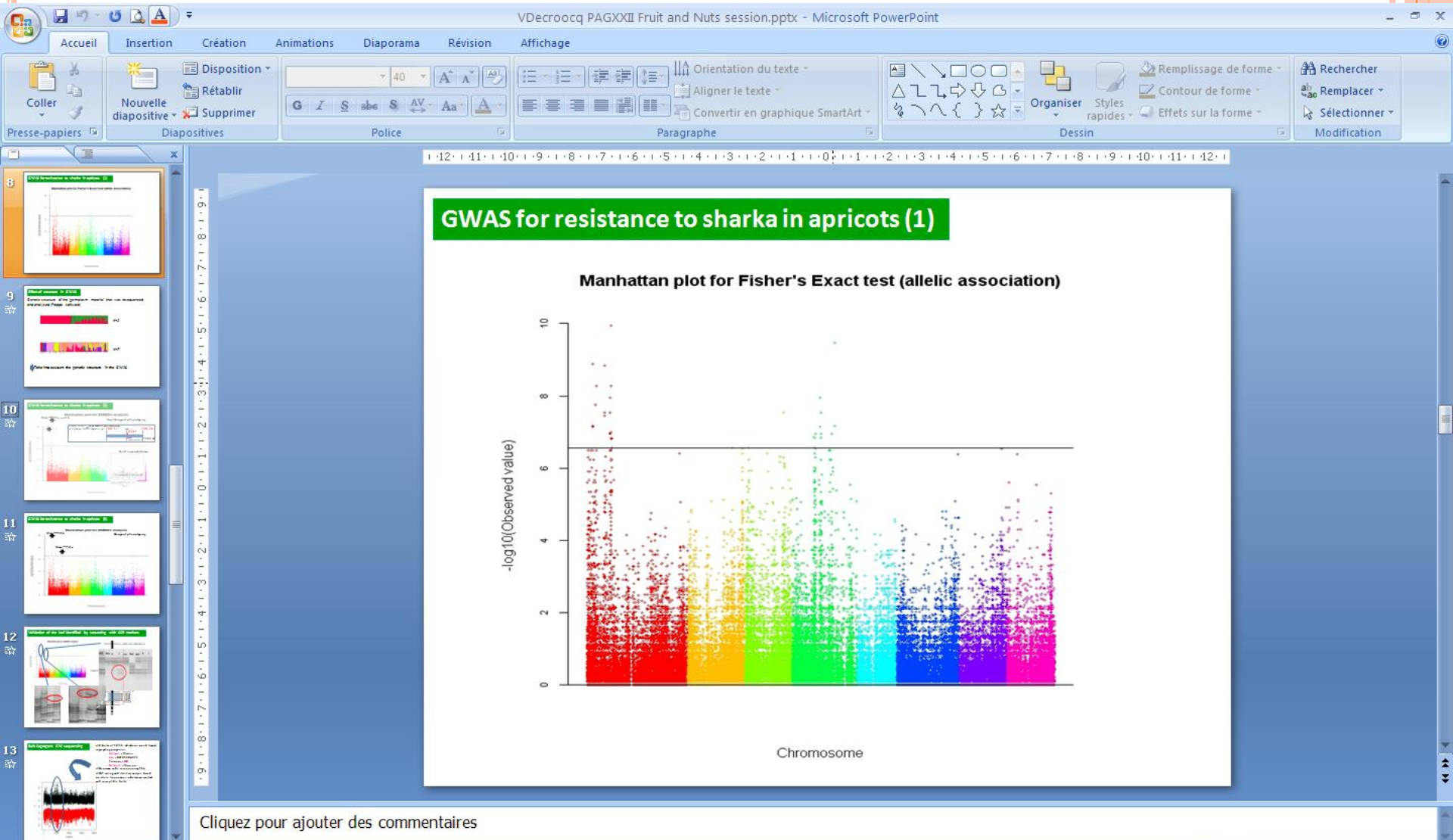
$K=2$



$K=7$

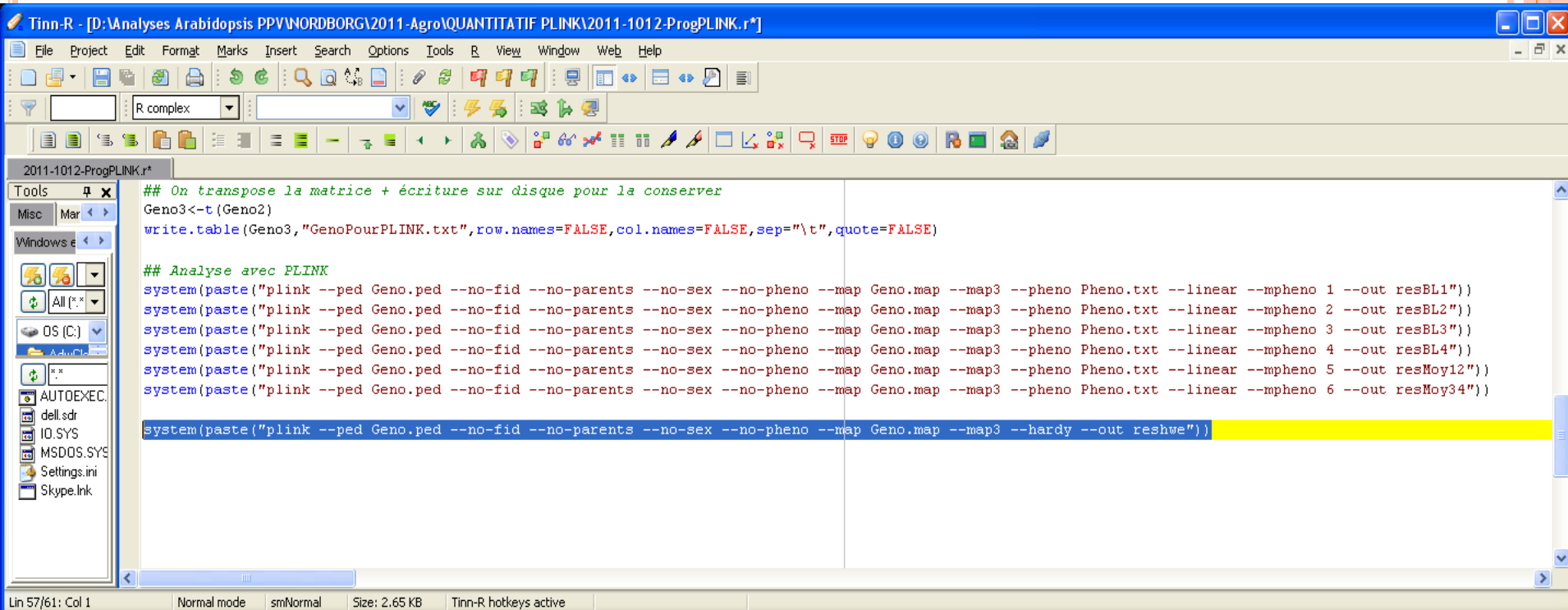


GWAS analysis (association between phenotypic and genotypic variations) – GWAS analizi (fenotipik və genotipik variasiyalar arasında assosiasiyalar)



Cliquez pour ajouter des commentaires

R software – R kompjüter programı

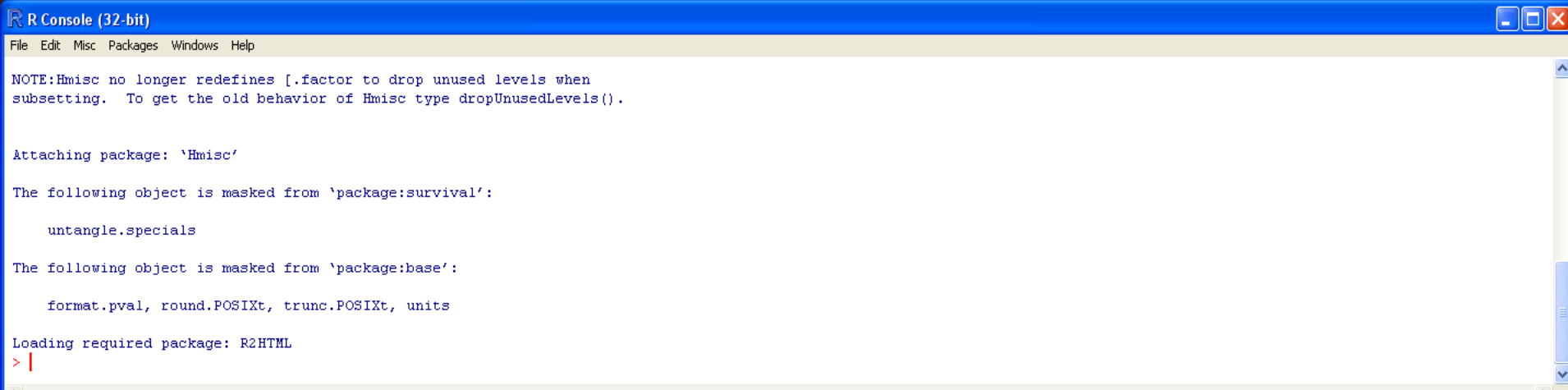


The screenshot shows the Tinn-R editor interface. The title bar reads "Tinn-R - [D:\Analyses Arabidopsis PPV\NORDBORG\2011-Agro\QUANTITATIF PLINK\2011-1012-ProgPLINK.r*]". The menu bar includes File, Project, Edit, Format, Marks, Insert, Search, Options, Tools, R, View, Window, Web, and Help. The toolbar contains various icons for file operations and editing. The main editor window displays the following R code:

```
## On transpose la matrice + écriture sur disque pour la conserver
Geno3<-t(Geno2)
write.table(Geno3,"GenoPourPLINK.txt",row.names=FALSE,col.names=FALSE,sep="\t",quote=FALSE)

## Analyse avec PLINK
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 1 --out resBL1"))
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 2 --out resBL2"))
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 3 --out resBL3"))
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 4 --out resBL4"))
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 5 --out resMoy12"))
system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --pheno Pheno.txt --linear --mpheno 6 --out resMoy34"))

system(paste("plink --ped Geno.ped --no-fid --no-parents --no-sex --no-pheno --map Geno.map --map3 --hardy --out reshwe"))
```



The screenshot shows the R Console (32-bit) window. The title bar reads "R Console (32-bit)". The menu bar includes File, Edit, Misc, Packages, Windows, and Help. The console displays the following output:

```
NOTE:Hmisc no longer redefines [.factor to drop unused levels when
subsetting. To get the old behavior of Hmisc type dropUnusedLevels().

Attaching package: 'Hmisc'

The following object is masked from 'package:survival':

  untangle.specials

The following object is masked from 'package:base':

  format.pval, round.POSIXt, trunc.POSIXt, units

Loading required package: R2HTML
> |
```

Diqqətinizə görə təşəkkür edirəm!

