

Scipio Documentation

Contents

Example searches

- Gene containing a mismatch and a frameshift
- Gene spread on several contigs and containing a gap
- Gene with a 5' exon just encoding a methionine

Expert searches

- Parameters to account for additional/missing bases in predicted exons
- Parameters to identify divergent exons and very short exons
- Parameters to account for low homology at intron borders

Gene containing a mismatch and a frameshift

Organism Fusarium oxysporum f. sp. lycopersici 4287

Genome file  supercontigs v 1.0.0




Genomic DNA subsequence (useful for command-line tests of Scipio)

Query sequence MALVIPYTYIQPCSDQSPDLPQARQSQSSDERTFDPRDPRSNYSLYPLEYLLYCEDCQ
QIRCPRCVNEEVVTTYCPNCLFEVPSSNLRSDGNRCTRSCYQCPVCIGPLQVMETPIEKD
QSHLGADIPGPQYALYCQYCNWTSTEIGIKFDKPNGIHSQLSKINNGGDLKLTAKELKER
RKENPDEPLADSDVDLQYANLKSFYQSQLADTNAASSGISPLNDTTGYGSPAASLSR
IMAMYTGHGHARKRNGPSEVMREALSAEGLKLADLDESAQIKKLHQEGWDATATIQQNL
EQAQEVQRFQDGLRPIPHLLRTRKSKRCSVCRHIISKPENKVTSTRFKIRLVAKSYIPTIT
IKPLNPTAGTVPTTQRPQILEERPLKPLTPHHYIITFKNPLFDGIKVTLATPNSTPGRFS
SKVTILCPQFDIDANTDMWDDALKDDDRDKKRKGEESGQPEAGKIWERGRNWVSIILEV
VPASLRDGGQDKSPLKEDEDILEIPMFVRMEWEPDSQQQDVGAASAKEKDAQERRELAYW
CVLGVGRISHD

Search parameters standard parameters



Resultfile  ScipioResult_Fusarium_oxysporum.yaml

Gene spread on several contigs and containing a gap

Query sequence derived from cDNA.

Organism Bombyx mori str. Dazao

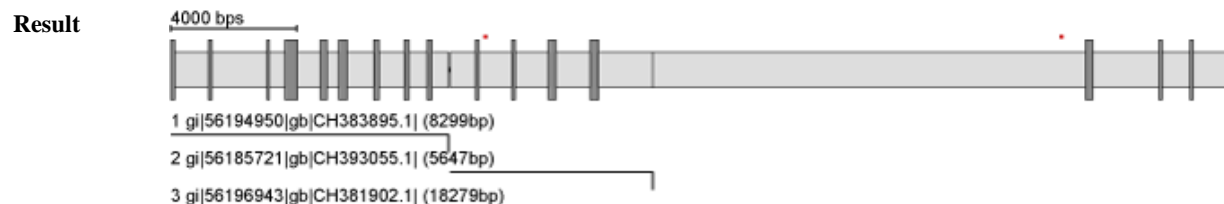
Genome file  supercontigs v 1.0.0



Genomic DNA subsequence (useful for command-line tests of Scipio)

Query sequence MEHSLQHRERVGVQDFVLLEDYRSEAAFIDNLKKRFHENIITYIGNVLISVNPYKNLPI
YTEEKTLYFKKAFFEAPPHVFAIADNAYRSLVYEHREQCILISGESGSGKTEASKKVLE
YIAARTNHLRNVENVKDKLLQSNPLLEAFGNAKTHRNDNSSRFGKYMDIQFNYEGGPEGG
HILNYLLEKSRVVSQMHGERNFHIFYQLLASSDQSLMTHLKLQGRPEAYKYTSDSTSHMS
QRANDQEQFRVVQEAMKVIEIGESEQREIFEIVASVLHLGNVKFVQNDKGYAEILSHDAN
SGNAADLLKVNATALREALTNRTIEARGDVVSTPLDVEQAQYARDALAKAIYDKHFSWL
SRLNSSLAPIEKDAKSSVIGILDIYGFEIFPKNSFEQFCINFCNEKLQQLFIQLTLRQEQ
EEYLREGIEWEPVEYFNNIIICDLIEARHKGIISILDDECLRPGDATDASFLDKLNQHLD
GHQHYKSHRKSDTKTQKLMGRDEFCLVHYAGEVTYNVNTFLEKNNDLLFRDIQSLMASSD
NTIVGCCFKVTFSNREPSYIRCIKPNDFKAPMQFDDKLVS HQVKYLGLMENLRVRRAGFA
YRRTYEAFLERYKCLSAETWPNYRGAARDGVQRLVEALQYEKEEYRMGN TKIFVRFPKTL
FATEDAFQIKKNDIATIIQSRWRGYYLRKRYLRMRNAAIVIQKWVRRFLAQRLRERRRKA
ADVIRAFIKGFITRNGPETPENRRFLGVAKVHWLKRLSAQLPTKLLDLSWPPCPSTCREA
SEELHRLHRAHLARKYRLALSPDDKKQFELKVLAEKIFKYSCEAVKYDRRGYKARARGLL
ASRAALYVLDAGGRRTFRLKHLPLDRLTVVVTNESDSL L LVKVPRDLKKDKGDLIISVT
HLIEALTIVTDYTKKPELIEIVDTRTIAHSLVNGKQGGTIEVTKGTQPAIQRKSGNLLV
VATP

Search parameters standard parameters



Resultfile  ScipioResult_Bombyx_mori.yaml

Gene with a 5' exon just encoding a methionine

Organism Magnaporthe grisea 70-15

Genome file  supercontigs v 6.0.0



Genomic DNA subsequence (useful for command-line tests of Scipio)

Query sequence

MGITRRGKDKAAAGQAVAGGASGGRARPKKATFETSKKKDVGVSDDLTLSSKVSNEAINEN
LQKRFEGREIYTYIGHVLVSVNPFRLGIYTDQVLDSYKGKNRLEMPHVFAIAESAYYN
MKAYKDNQCVIISGESGAGKTEAAKRIMQYIASVSGGDSTDIQIKDMVLATNPLLESFG
NAKTLRNNNSSRFQKYLQIHFNVSVEPVGADITNYLLEKSRVVGQITNERNFHIFYQFTK
GASEHYRQMFGIQKPETYIYTSRSKCLDVDGIDDLAEFQDTLNAKVVIGLSQEEQDSVFR
ILAAILWTGNLVFREDDEGYAAVTDQSVVEFLAYLLEVDPPQLIKAITIRILTPRSGEVI
ESPANVAQAMATRDALAKSLYNNLFDWIVERINQSLKARQPTSNSVGILDIYGFEIFEKN
SFEQLCINYYNEKLQQIFIQLTLKAEQDEYAREQIKWTPIKYFDNKIVCDLIESVRPPGV
FSALKDATKTAHADPAACDRTFMQSVNGMSNAHLIPRQGSFIKHAYAGDVAYTVDGITDK
NKDQLLKGLLGMFQVSQNPFLHTLFPNQVDQDNRKQPPTAGDRIRTSANALVETLMKCQP
SYIRTIKPNENKSPTEYNVPNVLHQIKYLGLQENVRIIRAGFAYRQSFEKFVDRFFLLSP
ATSYAGEYTWQGSYEAQVQILKDTQIPQEEWQMGVTKAFIKSPETLFALEHMRDRYWHN
MATRIQRMWRAYLAYRAESATRIQTFWRKKRTGAEYLQLRDHGHVRLQGRKERRRMSILG
SRRFIGDYLGINASSGPGAHIRNAIGIGSNEKTVFSCRGILEAKFGRSSKASPRILIVT
NSKFYVVAQMLVNGQVQITAEKAIPLGAIKFIGASSSRDDWFSLGVGSPQEADPLLNCVL
KTEMFTQMERVMPGGFNLKIGDSIEYAKKPGKMQVVKVLKDSPNPVDFYKSGAVHTQQGE
PPNSVSRPTPKGKPVPPRITRGLIRPGGPNGRPARGTTNRTPPRPGGASASAVASRP
VPQAQPPAQAQVAASIPVRTQQSQSTSSASVRAPPPPPPAAPPAAKAKIMAKVLYDFAGQK
ENEMSIKEGDLIEIVQKENNGWWLAKSGNQQAQWVPAAYVEEQKQAPPPVAASRPPPPAPP
AANGKNKPLPPAKRPAAGKKPASLQPRDSGMSLNGSDGSRSTPTPSLGNLADALLARK
QAMAKKDDDDDDW

Search parameters standard parameters

Result



Resultfile  ScipioResult_Magnaporthe_grisea.yaml

Parameters to account for additional/missing bases in predicted exons

Organism Phytophthora sojae P6497

Genome file  supercontigs v 2.0.0




Genomic DNA subsequence (useful for command-line tests of Scipio)



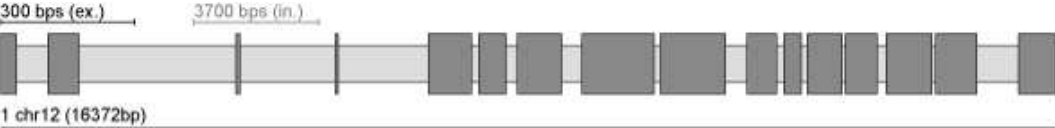
Query sequence MTSEGDVVDDVCVRYACKCGHLAPVSSLFYSETCEKLCRLPGCSVEEFESYYCGNLLVNLPSKEASM
YQNRSSRCFSCPTCENALSTAFHENEQRFFFLCAHCRWESLSVGLADDDPDALVMTAIRERQAAHED
IFQALHSHYSTLSSSPANGGAAGVLSASSGAPFGRSSLQLLADSMKELQREHQMKKFRLQRMAEMGG
WKYDQALAKVEEKERWLMEQRREHQWPELSMQLAALSKQGTSYGEESPEKTREVLRLNSQKCDMGEVS
TLHQRLNPLDQAREADQLFSPRPLLRVKRTWRCVESIERGTAGILVKPQISPMMSGDSSLPVSASWFK
KANLAAQYVPVVTFRQLPYRVGEGGGVECILLVENPLDDAIRITFHSTPTASEDSNENGYVDFQDLTP
IIVGPYEDPNLADAFIDDEPPFGANGDEHNTMLLQATRNLIKIKLPLSVNAATTIPISLTSARFIMDT
EKFDDEDANEVIENSLLSVPVVITAPIIKIGD

Search parameters Min. Score = 0.2; Min. Identity = 60%; Max. Mismatch = "∞"; Gap to Close = 15 aa; Min. Intron Length = 34 na



Resultfile  ScipioResult_Phytophthora_sojae.yaml

Parameters to identify divergent exons and very short exons

Organism	Homo sapiens
Genome file	 chromosome_assembly1 v 1.0.0
	 Genomic DNA subsequence (useful for command-line tests of Scipio)
Query sequence	MADPKYADLPGIARNEPDVYETSDLPEDDQAEFDFAFAQELEELTSTSVEHIIVNPNAAYD KFKDKRVGTKGLDFSDRIGKTKRTGYESGEYEMLGEGLGVKETPQQKYQRLLHEVQELTT EVEKIKTTVKESATEEKLTPVLLAKQLAALKQQLVASHLEKLLGPDAAINLTDPDGALAK RLLQLLEATKNSKGGSGGKTTGTPPDSSLVTYELHSRPEQDKFSQAAKVAELEKRLTELE TAVRCDQDAQNPLSAGLQGACLMETVELLQAKVSALDLAVLDQVEARLQSVLGKVNEIAK HKASVEDADTQSKVHQLYETIQRWSPIASTLPELVQRLVTIKQLHEQAMQFGQLLTHLDT TQQMIANSLKDNTLLTQVQTTMRENLATVEGNFASIDERMKKLGK
Search parameters	Exhaust Align Size = 15000 bp
Result	 <p>1 chr12 (16372bp)</p> <p>For clarity introns have been scaled down by a factor of 13.27</p>
Resultfile	 ScipioResult_Homo_sapiens.yaml

Parameters to account for low homology at intron borders

Organism Thielavia heterothallica ATCC 42464

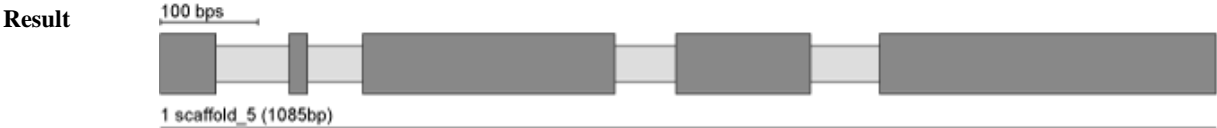
Genome file  supercontigs v 1.0.0




Genomic DNA subsequence (useful for command-line tests of Scipio)

Query sequence MSSHKAIVSSFVEGAPPGELSDVVADIKALTISSPGLVNELGPAFQKYNEEQFTTVKLPG
GSQPVIISHSALGNTRYDLESSCSFAFDHITQKASAVESYVPEGDGVDLAKSILKGLS
TYVNEHFTNAAYGAYPIENGSKIAIIIVANKYSPNFWNGRWRSLYIFNPASSALEGSIK
VDVHYYEDGNVRLLTNKSATATVPSGTGAAIAKEISAGERKYQEELNRSFTSLSEGAFFKA
LRRQLPVTRQKIEWDKVASYRLGQDIGGGSSRR

Search parameters Max. Move Exon = 6 aa



Resultfile  ScipioResult_Thielavia_heterothallica.yaml