

#### Issue 5

### **G-IMEx:** A comprehensive software tool for detection of microsatellites from genome sequences

# Suresh Babu Mudunuri<sup>1</sup>, Pankaj Kumar<sup>4</sup>, Allam Appa Rao<sup>2</sup>, Pallam Setty Sanaboina<sup>3</sup>, Hampapathalu Adimurthy Nagarajaram<sup>4,\*</sup>

<sup>1</sup>Department of Computer Science and Engineering, Aditya Engineering College (AEC), Surampalem 533 437, India; <sup>2</sup>Jawaharlal Nehru Technological University (JNTU), Kakinada, 533 003, India; <sup>3</sup>Department of Computer Science and Systems Engineering, Andhra University College of Engineering (AUCE), Visakhapatnam 530 003, India; <sup>4</sup>Laboratory of Computational Biology, Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad 500 001, India; H.A.Nagarajaram - Email: han@cdfd.org.in; Phone: +91-40-24749367; +91-9908209193; Fax: +91-40-24749448; \*Corresponding Author

Received June 28, 2010; Accepted August 25, 2010; Published November 1, 2010

#### Abstract:

Microsatellites are ubiquitous short tandem repeats found in all known genomes and are known to play a very important role in various studies and fields including DNA fingerprinting, paternity studies, evolutionary studies, virulence and adaptation of certain bacteria and viruses etc. Due to the sequencing of several genomes and the availability of enormous amounts of sequence data during the past few years, computational studies of microsatellites are of interest for many researchers. In this context, we developed a software tool called Imperfect Microsatellite Extractor (IMEx), to extract perfect, imperfect and compound microsatellites from genome sequences along with their complete statistics. Recently we developed a user-friendly graphical-interface using JAVA for IMEx to be used as a stand-alone software named G-IMEx. G-IMEx takes a nucleotide sequence as an input and the results are produced in both html and text formats. The Linux version of G-IMEx can be downloaded for free from http://www.cdfd.org.in/imex

Keywords: Bioinformatics Tool, Stand-alone program, Microsatellites, Simple Sequence Repeats, Genomes

#### **Background:**

Microsatellites, also known as Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs), are tandem repetitions of a nucleotide motif of size 1-6 bp. They are distributed in both coding as well as non-coding regions of all known genomes. Because of their polymorphic nature, they are known to play an important role in gene regulation, pathogenesis, bacterial adaptation and in evolution of genomes [1-5]. They are also applied in various fields such as DNA fingerprinting, Paternity studies, Forensics, Evolutionary studies etc. As the sequencing of new genomes is increasing day-by-day, microsatellites of many genomes remain unexplored. Analysis of these microsatellites is important to understand their role in various studies. Computational analysis is a better alternative to the time-consuming and money-intensive traditional wet lab microsatellite studies. A software tool that can extract all types of microsatellites with greater sensitivity and provides flexible options to analyze the repeats detected is the need of the day.

Few tools [6-9] exist in the public domain for extracting microsatellites from genome sequences, but many of them suffer from certain lacunae interms of their features and their efficiency. In the course of our studies on evolution of microsatellites in prokaryotic genomes, we developed a novel algorithm [10] to detect imperfect microsatellites from nucleotide sequences. The algorithm has been implemented in the form of a standalone software with a user-friendly graphical user interface (GUI) called G-IMEx. The present communication gives the details of this software.

#### Methodology:

The algorithmic details of IMEx have been reported elsewhere [10]. For

ISSN 0973-2063 (online) 0973-8894 (print) Bioinformation 5(5): 221-223 (2010) 221

sequence and looks for two consecutive exact repeat units or two alternate exact repeat units and considers them as the 'candidate' microsatellite repeat tract. The 'candidate' tract is expanded on both sides by allowing few mismatches in each individual repeat unit ('k' – imperfection limit / repeat unit) such that the percentage of imperfection of the entire tract does not cross the threshold set by the user. The expansion is also terminated if a repeat unit with more than 'k' mismatches is encountered. The program further collates and clusters equivalent microsatellite repeats into families. It also has an option to identify compound microsatellites, which are regions containing more than one microsatellite tract separated by a certain distance as defined by the user.

the sake of continuity we reiterate the method. IMEx scans the input

#### Software Requirements:

G-IMEx has been developed on the Linux platform and requires preinstalled C and Java (for graphical interface). An ideal environment for running G-IMEx would be a latest Fedora or other Linux distribution with a gcc compiler (version 3.4 or higher), Java version (1.6 or higher) and any browser software.

#### Input options:

G-IMEx offers several options for identification, extraction, collation, clustering and reporting of microsatellites from an input DNA sequence in FASTA format. The software can handle large sequences such as genomes easily and is comparatively faster than many other tools. Users can set the limits for repeat size, repeat number, repeat type and imperfection level. In addition users can set levels (0 to 4) for clustering of equivalent microsatellites i.e., those

**Software** 

www.bioinformation.net

Issue 5

microsatellites which are close to each other sequentially. There is also an option to use the core IMEx program in batch mode for scanning multiple sequences.

#### **Output options:**

G-IMEx creates a folder with the name of the input sequence file and the results are stored in two formats – html and text. The text format of results is optional and separate directories are created for text and html results. The output includes a well-formatted summary table file with information such as the repeating motif, repeat number, imperfection %, tract size, nucleotide composition and protein information (if it falls in coding region) etc. Along with the information about the microsatellite extracted, its corresponding alignment with its perfect repeat counterpart is also produced automatically in a separate alignment file which facilitates

analysis of mutational events in a microsatellite tract. **Figure 1** shows the snapshot of the GUI and the result pages of G-IMEx.

#### **Future Work:**

The current version of G-IMEx is available only for Linux users. Efforts are underway to develop versions compatible to Windows and Macintosh systems.

#### Acknowledgements:

The authors acknowledge the support from the lab members of computational biology and the core grant of CDFD. SBM would like to thank Mr. Priyatosh Mishra for his valuable help during the development of the graphical interface and the management of AEC for providing necessary facilities.

				IMEx Mode:	Advanced 👻	(out	resh@loca	alhost IM	EX2.0]\$ ./	inex					
The	Sequence file should be	in FASTA or Plain Fo	ormat												
: U	Jpload a file from you	ur directory						IMEX	IMPERFEC	T MICROSA	TELLITE EX	TRACTOR			
	В	rowse		Sam	ple Sequence				••••••	•••••					
00	Cut and Paste your se	equence				ENTI	ER THE IN	NPUT FILE	NAME (Ex.	:/home/su	resh/seq1.	fna, sample	e.txt) :NO	C_000913.f	na
AC GC T/ AT CC GT Set 1 I Typ	gi   49175990   ref   N GCTITITITITITITI CAGCAACTGGTTACCTG ATAGGCATAGCGCACAG TTAGCCATAGCGCCACAG CCCGCACCTGACAGTGCG TTCCGCGGTACATCAGT TTCGGCGGTACATCAGT the parameters here: nclude Flanking regio e of Repeat: [mperf Main parameter: inum Beneat Number	TTCATTCFGACTGC CCGTGAGTAAATTA ACAGATAAAAATTA ATTACCACAGGTAAA GGCTTITTTTTCG GGCAAATGCAGAAC ONN 10 V V Ge Fect V ® Repea S Mono	AACGGGCAATATG AAAATITTATTGAC CAGAGTACACAAC GGGTGCGGGCGGA ACCAAAGGTAACG GTTTTCTGCGTGT Inerate Alignmen It Size: Ali	TCTCTGTGTGGATT/ TTAGGTCACTAAAACGAT TTCAGTGAAACGCAT CGCGTACAGGAAACGA AGGTAACAACGAT GGCGTACAGGAAACCATG TGCCGATATTCTGG/ nts Generate ' C Enter a Pat ri Tetra P	AAÂAAAAGAGT ACTTTAACCAA TTTAACCACAC ACAGAAAAAAG CGAOTGTTGAA AAAGCAATGCC Text output files :tern: :enta Hexa	ENTI HOM DI [] TETI FETI REXJ GENJ	ER THE RE 0(>=2):15 =2]:7 [>=2]:5 RA(>=2):4 TA(>=2):2 ER THE FI ERATE ALI ERATE TEN	EPEAT NUM 5 4 3 LANKING SU IGNMENTS KT OUTPUT	BER FOR EA EQUENCE LE ? [YES(Y/y S ? [YES(Y	NGTH [0-20 1)/300(N/n)]	] 19	1.	J ıp		
Min	imum Repeat Number	r: 15	7 5	4 3	3							ta. A secondaria da	7		
Imp	erfection Limit/repea	at unit: 1	1 - 1	<b>▼</b> 2 <b>▼</b> 2	▼ 3 ▼	HAX	INCH DIST	TANCE ALL	OWED BETWE	EN ANY 2	SSRs (dMAX	) [0-50]:1	0		
	Jpload Coding info file Extract Compound Mic	and the second sec	Bre	owse			1.Le		andardizat		Repeat (	Consensus:	TAG		
- E	1	crosatellites d between two SS	Rs (dMax) 10 ((				1.Le 2.Le 3.Le 4.Co	evel I Sta evel II S evel III : omplete /		ion tion ation	Occurre	nces:	TAG Start	End	Stand. Leve
	ixtract Compound Mic Max. distance allowe Level of Standardizat	crosatellites d between two SS	Rs (dMax) 10 ((	0-50)		ENTI	1.Le 2.Le 3.Le 4.Co ER STANDA	evel I Sta evel II S evel III : omplete / ARDIZATION	andardizat tandardiza Standardiz Full Stan	ion tion ation dardizatio	or Repeat	nces:	Start	End 83681	Stand. Leve
- E	ixtract Compound Mic Max. distance allowe Level of Standardizat	crosatellites d between two SS tion: <mark>No Standar</mark>	Rs (dMax) 10 (	0-50)		ENTI	1.Le 2.Le 3.Le 4.Co ER STANDA	evel I Sta evel II S evel III : omplete / ARDIZATION	andardizat tandardiza Standardiz Full Stan N Level:4	ion tion ation dardizatio	or Repeat	nces: No. Repeat	Start		
	xtract Compound Mid Max. distance allowed Level of Standardizat	crosatellites d between two SS tion: <mark>No Standar</mark>	Rs (dMax) 10 (	0-50)	End	ENTI	1.Le 2.Le 3.Le 4.Co ER STANDA	evel I Sta evel II S evel III : omplete / ARDIZATION	andardizat tandardiza Standardiz Full Stan N Level:4	ion tion ation dardizatio	or Repeat 1	nces: No. Repeat (TAG) <sub>4</sub> (ATG) <sub>4</sub>	Start 83670	83681	LO
	xtract Compound Mid Max. distance allowed Level of Standardizat	crosatellites d between two SS tion: No Standar Clear Values	Rs (dMax) 10 ( dization Extract Micro	0-50) v	End 32776	enti exti	1.Le 2.Le 3.Le 4.Co ER STANDA RACT MICS	avel I Sta avel II S avel III S amplete / ARDIZATION ROSATELLI	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL	ion tion sation dardizatio , SIZES (1	Occurre Repeat t 2 5	nces: No. Repeat (TAG) <sub>4</sub> (ATG) <sub>4</sub> (CAT) <sub>4</sub>	Start 83670 130322	83681 130333	L0 F
	Extract Compound Mic Max. distance allower Level of Standardizat	crosatellites d between two SS tion: No Standard Clear Values Regations	Rs (dMax) 10 (i dization Extract Micro Tacl-size	0-50) y satellites Start		SNT SXT	1.Le 2.Le 3.Le 4.Co ER STANDA RACT MICS	avel I St. svel II S svel III S omplete / ARDIZATION ROSATELLI	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL	ion tion ation dardizatio , SIZES (1	CCCUTTE Repeat 1 5 16 19	nces: No. Repeat (TAG) <sub>4</sub> (ATG) <sub>4</sub> (CAT) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893	83681 130333 420697 544904	L0 F L3 L3
	Extract Compound Mic Max. distance allower Level of Standardizat	crosatellites d between two SS tion: No Standar Clear Values Rerations 5	Rs (dMax) 10 (r dization Extract Micro Tract-size 10	o-50) v satellites Start 32767	32776	ENTI SXTT 0%	1.Le 2.Le 3.Le 4.Co ER STANDA RACT MICS A 0.0	evel I Sta evel II Sta maplete / ARDIZATION ROSATELLIT	andardizat tandardiza Standardiza Standardiz Full Stan N Level:4 TES OF ALL G 50.0	ion tion dardization sIZES (1	Occurre Repeat I 2 5 16	nces: No. Repeat (TAG) <sub>4</sub> (ATG) <sub>4</sub> (CAT) <sub>4</sub> (ATC) <sub>4</sub>	Start 83670 130322 420686	83681 130333 420697 544904	L0 F L3
	xtract Compound Mic Max. distance allowed Level of Standardizat GC TAG	crosatellites d between two SS tion: <u>Into Standar</u> Clear Values <u>Berations</u> 5 4	Rs (dMax) 10 (( diration Extract Micro Tract-size 10 12	0-50) 	32776 83681	ENTI EXTI	1.Le 2.Le 3.Le 4.CC ER STANDA RACT MICS 0.0 33.3	avel I St. avel II St. avel II St. avel II St. avel II St. avel II St. avel II St. avel I St. avel II St.	andardizat tandardiza Standardiza Full Stan N Level:4 TES OF ALL 50.0 33.3	ion tion ation dardizati SIZES (1 50.0 0.0	CCCUTTE Repeat 1 5 16 19	No.         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (CAT) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893	83681 130333 420697 544904 1463704	L0 F L3 L3
S.No	xtract Compound Mid Max. distance allower Level of Standardizat GC TAG GGC	crosatellites d between two SS tion: <u>Ino Standar</u> Clear Values <u>Rerations</u> 5 4 4 4	Rs (dMax) 10 (d dization Extract Micro Text-size 10 12 12	0-50) satellites Start 32767 93670 92844	32776 83681 92855	ENTI EXTI 0% 0%	1.Le 2.Le 3.Le 4.CC ER STANDA RACT MICE 0.0 33.3 0.0	vel I St. vvel II St. smplete / REDIATION ROSATELLI 0.0 33.3 0.0	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL G 50.0 33.3 66.7	ion tion ation dardizati 50.0 0.0 33.3	Occurre Repeat 1 2 5 16 19 42	Inces:         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890	83681 130333 420697 544904 1463704 1516901	L0 F L3 L3 L3
. E	Xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG	crosatellites d between two SS tion: <u>Ino Standar</u> Clear Values Berations 5 4 4 5	Rs (dMax) 10 (d dization Extract Micro Tract-size 10 12 12 15	0-50) Satellites Start 32767 83670 92844 105358	32776 83681 92855 105372	ENT EXT 0% 0% 0%	1. L4 2. L6 3. L4 4. C0 ER STANDJ ERCT MICS A% 0.0 33.3 0.0 0.0	vel I St. vvel II St. omplete / REDIATION ROSATELLI 0.0 33.3 0.0 0.0	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL 50.0 33.3 66.7 66.7	ion tion ation dardizati. , SIZES (1 0.0 0.0 33.3 33.3	Occurre Repeat I 5 16 19 42 43 46	Incest         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890           1720637	83681 130333 420697 544904 1463704 1516901 1720648	L0 F L3 L3 L3 F F
. E	xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG ATG	crosatellites d between two SS tion: <u>Fio Standar</u> Clear Values <u>Benutions</u> 5 4 4 5 4	Rs (dMax) 10 (d dization Extract Micro Tract-size 10 12 12 15 12	0-50) Satellites Start 32767 83670 92844 105358 130322	32776 83681 92855 105372 130333	ENTI EXTI 0% 0% 0% 0%	1. Lo 2. Lo 3. Lo 4. CC ER STANDA RACT MICE 0.0 33.3 0.0 0.0 33.3	svel I St.           svel II St.           svel II St.           smplete /           amplete /           ROSATELLY           0.0           33.3           0.0           33.3           0.0           33.3	andardísat standardísa Standardís Full Stan N Level:4 TES OF ALL CS 50.0 33.3 66.7 66.7 33.3	ion tion ation dardizati SIZES (1 50.0 0.0 33.3 33.3 0.0	Occurre Repeat I 5 16 19 42 43	Incest         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890	83681 130333 420697 544904 1463704 1516901 1720648	L0 F L3 L3 L3 F
. E	xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG ATG TGGC	crosatellites d between two SS tion: No Standar Clear Values Becations 5 4 4 5 4 5 4 3	Rs (dMax) 10 (d dization Extract Micro Tract-size 10 12 12 15 12 12 12 12	0-50) Satellites Start 32767 83670 92844 105358 130322 182464	32776 83681 92855 105372 130333 182475	ENTT EXTT 0% 0% 0% 0% 0%	1. Le 2. Le 3. Le 4. Co ER STANDJ RACT MICS 0.0 33.3 0.0 0.0 33.3 0.0	Test         O           0.0         33.3           0.0         33.3           0.0         33.3	andardizat tandardizat Standardiz Full Stan N Level:4 TES OF ALL TES OF ALL 50.0 33.3 66.7 66.7 33.3 50.0	ion tion dardizati 50.0 0.0 33.3 33.3 0.0 25.0	Occurre Repeat I 5 16 19 42 43 46	Incest         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890           1720637	83681 130333 420697 544904 1463704 1516901 1720648 1832842	L0 F L3 L3 L3 F F
S.No	xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG ATG TGGC GCA	crosatellites d between two SS tion: No Standar Clear Values Berations 5 4 4 5 4 5 4 4 3 4	Rs (dMax) 10 (d diration Extract Micro Tract-size 10 12 12 12 15 12 12 12 12 12	0-50)  Satellites  Start  32767  83670  92844  105358  130322  182464  220481	32776 83681 92855 105372 130333 182475 220492	ENTT EXT 0% 0% 0% 0% 0%	1. Le 2. Le 3. Le 4. CC ER STANDA ER STANDA 0.0 33.3 0.0 33.3 0.0 33.3 0.0 33.3	T         Constraint           0.0         33.3           0.0         33.3           0.0         0.0	andardizat tandardiza Standardiz Pull Stan N Level:4 50.0 33.3 66.7 66.7 33.3 50.0 33.3	ion tion dardizati sIZES (1 50.0 0.0 33.3 33.3 0.0 25.0 33.3	Cocurre Repeat 1 2 5 16 19 42 43 46 52	Incest         Repeat           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890           1720637           1832831           1886583	83681 130333 420697 544904 1463704 1516901 1720648 1832842 1886594	L0 F L3 L3 F F F
5.No	xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG ATG GGC GCA GT	crosatellites d between two SS tion: bio Standar Clear Values 5 4 4 5 4 4 5 4 3 4 5	Rs (dMax) 10 (r dtration Extract Micro Tract-ske 10 12 12 12 15 12 12 12 12 12 12 12 12 12	0-50)  Satellites  Start  32767  83670  92844  105358  130322  182464  220481  226041	32776 83681 92855 105372 130333 182475 220492 226050	ENTI SXT 0% 0% 0% 0% 0% 0%	1.Le 2.Le 3.Le 4.Cc ER STANDA RACT MICE 0.0 33.3 0.0 33.3 0.0 33.3 0.0 33.3 0.0	Tel         State           0.0         33.3           0.0         33.3           0.0         33.3           0.0         50.0	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL 50.0 33.3 66.7 66.7 33.3 50.0 33.3 50.0	ion tion dardizati 50.0 0.0 33.3 33.3 0.0 25.0 33.3 0.0	Coccurre Repeat 2 5 16 19 42 43 46 52 53 54	Incest         Repost           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890           1720637           1832831           1886583           1940236	83681 130333 420697 544904 1463704 1516901 1720648 1832842 1886594 1940247	L0 F L3 L3 L3 F F F L3 L3
	xtract Compound Mic Max. distance allower Level of Standardizat GC TAG GGC CGG ATG TGGC GCA GT TTGA	crosatellites d between two SS tion: No Standar Clear Values 5 4 4 5 4 4 5 4 3 4 5 3	Rs (dMax) 10 (d dtration Extract Micro Tract-size 10 12 12 12 12 12 12 12 12 12 12 12 12 12	0-50) Satellites Start 32767 83670 92844 105358 130322 182464 220481 226041 226041 246315	32776 83681 92855 105372 130333 182475 220492 226050 246326	ENTT SXTT 0% 0% 0% 0% 0% 0% 0%	1.1.4 2.1.6 3.1.6 3.1.6 4.00 ER STANDA RACT MICS 0.0 33.3 0.0 33.3 0.0 33.3 0.0 33.3 0.0 25.0	Tel         State           mplete         /           weil II State         /           mplete         /           weil II State         /           0.0         33.3           0.0         0.0           33.3         25.0           0.0         50.0           50.0         50.0	andardizat tandardiza Standardiz Full Stan N Level:4 TES OF ALL 0 50.0 33.3 66.7 66.7 33.3 50.0 33.3 50.0 25.0	ion tion dardizati sIZES (1 50.0 0.0 33.3 33.3 0.0 25.0 33.3 0.0 0.0	Coccurre Repeat 1 2 5 16 19 42 43 46 52 53	Incest         Repost           (TAG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub> (ATG) <sub>4</sub> (ATC) <sub>4</sub> (ATC) <sub>4</sub>	Start           83670           130322           420686           544893           1463693           1516890           1720637           1832831           1886583	83681 130333 420697 544904 1463704 1516901 1720648 1832842 1886594 1940247	L0 F L3 L3 L3 F F F L3

Figure 1: Snapshots of Graphical User-Interface and Results Pages.

## **Bioinformation**

### Volume 5

www.bioinformation.net

Issue 5



#### **References:**

- [1] AV Belkum et al. Microbiol MolBiol Rev 62(2):275 (1998) [PMID: 9618442].
- P Martin et al. Proc Natl. Acad Sci USA 102(10):3800(2005) [PMID:15728391]
- [3] ER Moxon et al. Curr Biol 4(1):24 (1994) [PMID: 7922307]
- [4] VB Sreenu et al. BMC Genomics 7:78 (2006) [PMID: 16603092]
- [5] R Kofler et al. BMC Genomics 9:612 (2008) [PMID: 19091106]
- [6] G Benson Nucleic Acids Res 27:573(1999) [PMID: 9862982]
- [7] R Kofler et al. Bioinformatics 23: 1683 (2007) [PMID: 17463017]
- [8] M LaRota et al. BMC Genomics 6: 23(2005) [PMID: 15720707]
- [9] http://www.genomics.ceh.ac.uk/msatfinder.
- [10] SB Mudunuri & HA Nagarajaram *Bioinformatics* **23**(10):1181(2007) [PMID: 17379689]

#### Edited by P. Kangueane

Citation: Mudunuri et al. Bioinformation 5(5): 221-223 (2010)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.