RSG-Italy presents

# CASPita

# Critical Assessment of Student Programming in Italy

ISCB REGIONAL Student GROUP
Italy

## International Society for Computational Biology
Established in 1994

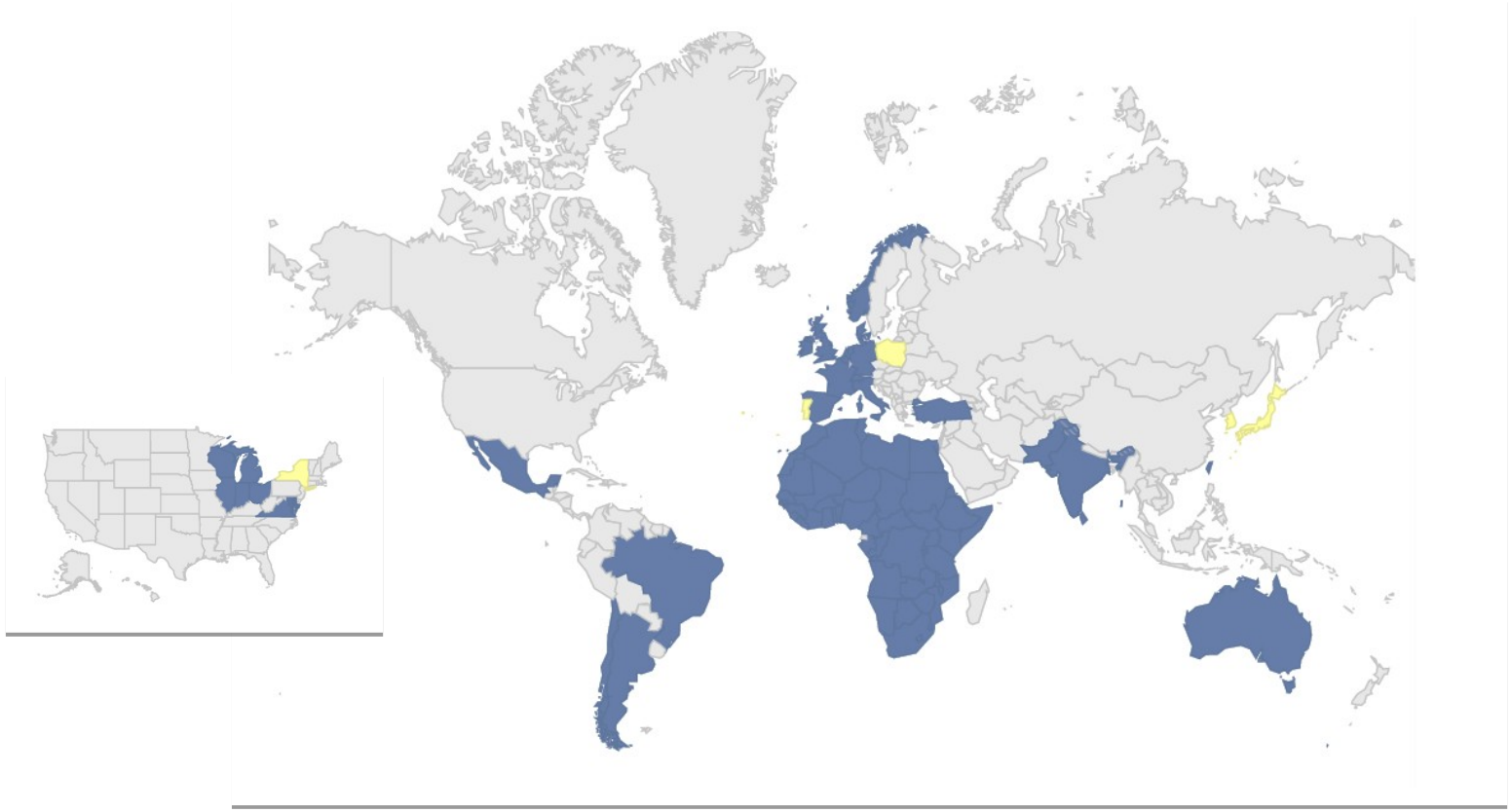
Scholarly society dedicated to advancing the scientific understanding of living systems through computation.

## ISCB Student Council (SC)
Established in 2004

**VISION**
To become the leading student organization in computational biology
**MISSION**
To develop the next generation of computational biologists

ISCB REGIONAL Student GROUP
Italy

# Active Regional Student Groups

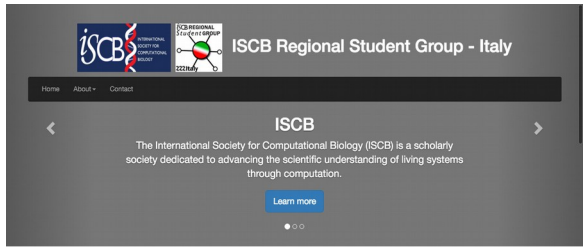


22 countries

ISCB REGIONAL Student GROUP
Italy

- **RSG-ITALY**
- Officially approved in January 2016
- **VISION**
  - To become the leading student organization in computational biology in Italy
- **MISSION**
  - To provide to the Italian students in the fields of computational biology a hub of connections to universities, academic experts and other students
  - To promote the spreading of knowledge within the field
  - To facilitate collaborations among future Italian Bioinformaticians

**Website**
**rsg-italy.iscbsc.org**




ISCB REGIONAL Student GROUP
Italy

# CASPita - Critical Assessment of Student Programming in Italy

Such as the famous CASP, it is a challenge proposed to the community. RSG Italy, as organizer, will give italian bioinformatic students the possibility to test their skills with a bioinformatic puzzle.

**First coding assessment specifically conceived to young programmers within the Bioinformatics and Biology areas**

Supporter and sponsors

ISCB REGIONAL Student GROUP Italy

# CASPita: first call

## Parser for BLAST and PSI-BLAST output

Annotation section

Alignment section, example:

```
 Score = 209 (95.7 bits), Expect = 5.0e-47, Sum P(3) = 5.0e-47
 Identities = 37/93 (39%), Positives = 68/93 (73%)

Query:    62 VGSSLMFLKEGDRVSVRDLSRGLIVDSGNDACVALADYIAGGQRQFVEMMNNYAEKLHLK 121
             +G S +FL+ G+ ++V+++ +G+ + SGNDA VA+A++I+G + +FV+ MN  A++L LK
Sbjct:    98 MGGSQIFLEPGEEMTVKEMLKGIAIASGNDASVAMAEFISGSEEEFVKKMNKKAKELGLK 157

Query:   122 DTHFETVHGLDAPGQHSSAYDLAVLSRAIIHGE 154
             +T F+   GL   G +SSAYD+A++++ ++  E
Sbjct:   158 NTSFKNPTGLTEEGHYSSAYDMAIMAKELLKYE 190

 Score = 161 (73.8 bits), Expect = 5.0e-47, Sum P(3) = 5.0e-47
 Identities = 45/153 (29%), Positives = 73/153 (47%)

Query:   187 VDGLKTGHTSGAGFNLIASAVDGQRRLIAVVMGADSAKGREEEARKLLRWGQQNFTTVQI 246
             VDG+KTG+T  A + L ASA  G  R IAVV GA + K R  +  K+L +    + T  +
Sbjct:   226 VDGVKTGYTGEAKYCLTASAKKGNMRAIAVVFGASTPKERNAQVTKMLDFAFSQYETHPL 285

Query:   247 LHRGKKVGTERIWYGDKENIDLGTEQEFWMVLPKAEIPHIKAKYTLDGKELTAPISAHQR 306
               R + V   ++  G ++ I+L T +   ++  K E  +   K       ++API   Q
Sbjct:   286 YKRNQTVAKVKVKKGKQKFIELTTSEPISILTKKGEDMNDVKKEIKMKDNISAPIQKGQE 345

Query:   307 VGEIELYDRDKQVAHWPLVTLESVGEGSMFSRL 339
             +G + L    + +A  P+   E + +   S L
Sbjct:   346 LGTLVLKKDGEVLAESPVAAKEDMKKAGFISFL 378
```

Annotation section

1. The default BLAST and PSI-BLAST output format is human friendly but not designed to be processed automatically.

2. Among the possible BLAST output format, this is the one comprising the most of the information.

3. A number of tools is available to parse the output in different programming languages, mainly in Perl. However, they usually do not meet all possible needs:

4. For example, an ideal tool could be used both as a module to import it in a script and do additional processing of data, and as a fast format coverter.

5. Some programming languages completely lack an updated tool to parse this kind of output.

The challenge is a simple exercise that however requires a good preliminary documentation and knowledge of the problem.

ISCB REGIONAL Student GROUP Italy

# Evaluation

The submitted proposals will be evaluated by a team composed both of programmers and biologists with limited coding experience.

The evaluation will consider:

- Execution time and memory usage
- Readability and reusability of the scripts
- Readability and completeness of help and documentation
- Program flexibility for errors in files.
- Novelty of program features and functionality

# Applications

The applications will be submitted as an email to rsg-italy@iscbsc.org. The call officially opens **today**. All additional information will be available at rsg-italy.iscbsc.org.

# Prizes

- First prize:

**Amazon coupon (of the value of 200€)**

- Special prize:

**Reserved place in an ELIXIR course (in 2017)**

## Publication of the results

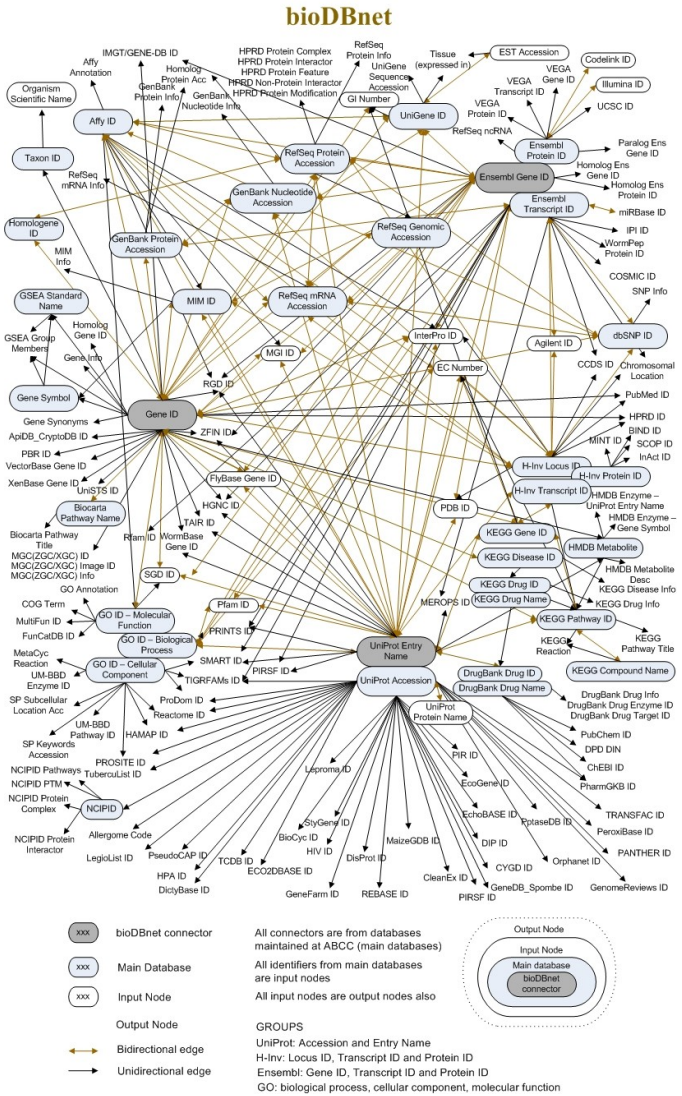
Assessment results will be published on the RSG-Italy Facebook page, website and on Twitter. All participants will be notified by email. Winner(s) will receive a specific notice.

The top ranking programs will be announced and shared with the community (in github), so that all those that have and will have similar problems will benefit from them.

# New calls are coming...

... Inter-database mapping?

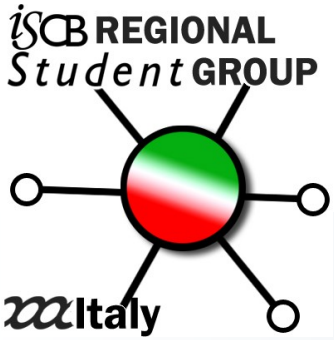
... Text-mining?


ISCB REGIONAL Student GROUP
Italy

## Long term goals

- Propose at least one run of challenges per year.
- Discuss possible challenges on social networks, in order to approach problems of interest for the community.
- Creation of a set of simple, well commented and re-usable parsing and processing tools and algorithms.
- Promote the dialogue and discussion on biological problems from the programming point of view and promoting the interaction among young researchers.



# Other RSG-Italy initiatives

## Reaching italian Universities…

- List all italian Universities with Bioinformatics related courses (sources: CINECA, AlmaLaurea, others)
- List all PHD courses with interest in Bioinformatics
- Collect a mailing list of all Institutions interested in Bioinformatics related seminars and events to publicize them.

**… and the european ones.**

- Propose the same initiative to european RSG in order to collect informations about the whole Europe.

ISCB REGIONAL Student GROUP
Italy

## Reaching the students…

- Collect Bioinformatics students opinion about the courses they attended.

**Survey “Studying Bioinformatics”**

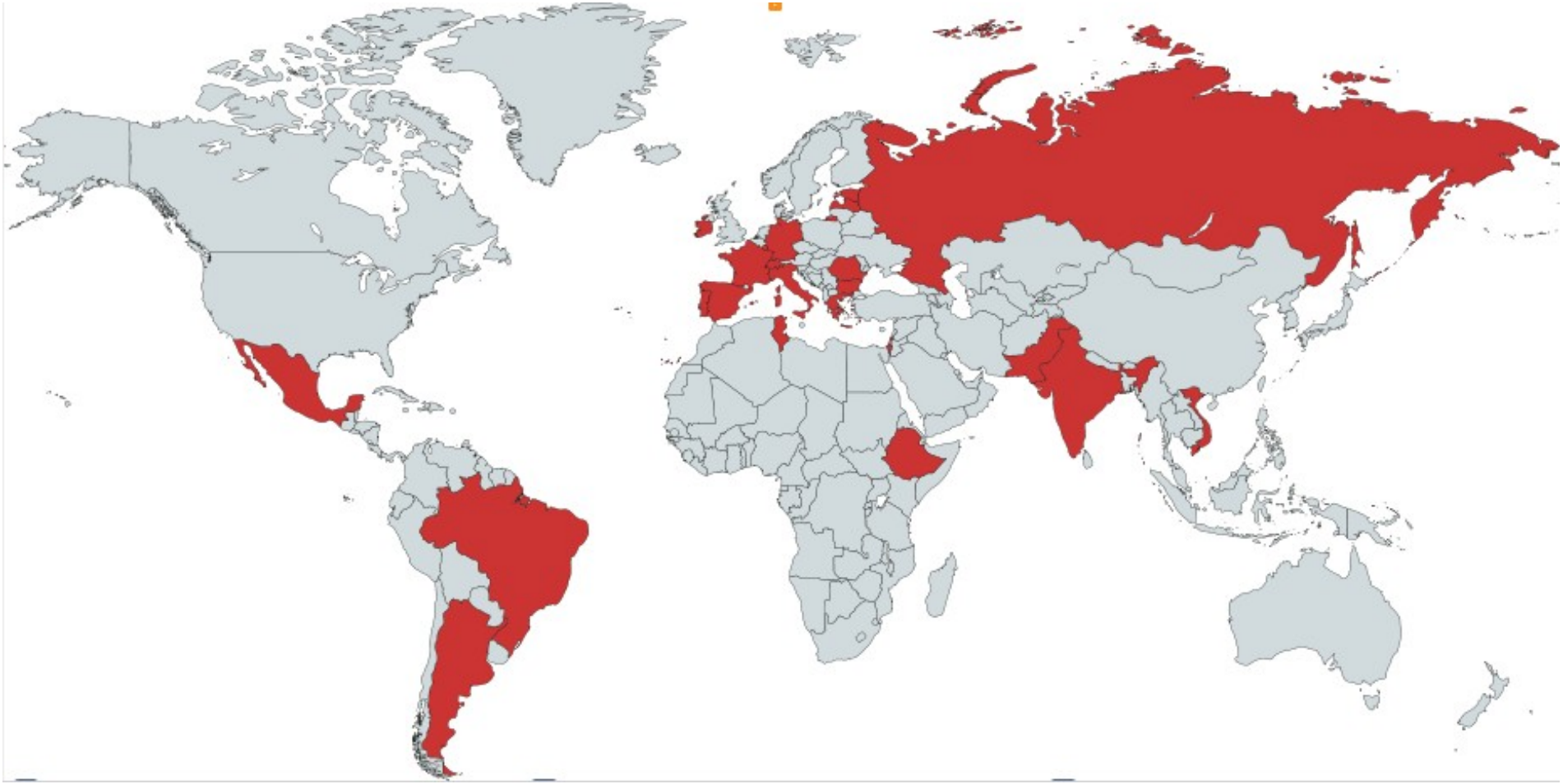
- Collect Bioinformatics students sugestions for new candidates.

**Survey “Bioinformatics training”**

- Collect biologist and biotechnologist opinions about Bioinformatics.

**Survey “What is Bioinformatics?”**

ISCB REGIONAL Student GROUP
Italy

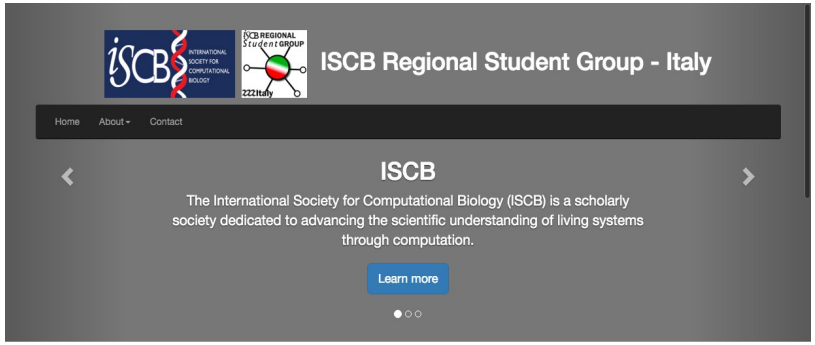


ISCB REGIONAL Student GROUP
Italy

**307 answers**
**from 23 different countries, 68 Universities**

# Thank you for your attention

## Your questions, comments and suggestions are welcome!




rsg-italy@iscbsc.org

**rsg-italy.iscbsc.org**

ISCB REGIONAL Student GROUP
222Italy