BioNLP Shared Task 2013 – An overview of the Genic Regulation Network Task

Robert Bossy, Philippe Bessières, Claire Nédellec

Unité Mathématique, Informatique et Génome
Institut National de la Recherche Agronomique
UR1077, F78352 Jouy-en-Josas, France
forename.name@jouy.inra.fr

Abstract

The goal of the Genic Regulation Network task (GRN) is to extract a regulation network that links and integrates a variety of molecular interactions between genes and proteins of the well-studied model bacterium *Bacillus subtilis*. It is an extension of the BI task of BioNLP-ST'11. The corpus is composed of sentences selected from publicly available PubMed scientific abstracts. The paper details the corpus specifications, the evaluation metrics, and it summarizes and discusses the participant results.

1 Introduction

The Genic Regulation Network (GRN) task consists of (1) extracting information on molecular interactions between genes and proteins that are described in scientific literature, and (2) using this information to reconstruct a *regulation network* between molecular partners in a formal way. Several other types of biological networks can be defined at the molecular level, such as metabolisms, gene expressions, protein-protein interactions or signaling pathways. All these networks are closely interconnected. For example, a gene codes for a protein that catalyzes the transformation of small molecules (metabolites), while the expression of the gene and its related regulation is controlled by other proteins.

The concept of biological networks is not new. However, the development of new methods in molecular biology in the past twenty years has made them accessible at the level of an organism as a whole. These new methods allow for the design of large-scale experimental approaches with high throughput rates of data. They are then used to build static and dynamic models that represent the behavior of a cell in the field of Systems Biology (Kitano, 2002; de Jong, 2002). In this context, there has recently been a considerable focus on "biological network inference", that is to say the process of making inferences and predictions about these networks (D'haeseleer, *et al.*, 2000). Therefore, it is expected that Information Extraction (IE) from scientific literature may play an important role in the domain, contributing to the construction of networks (Blaschke *et al.*, 1999). IE also plays a role in the design and the validation of large-scale experiments, on the basis of detailed knowledge that has already been published.

2 Context

Extracting molecular interactions from scientific literature is one of the most popular tasks in IE challenges applied to biology. The GRN task adds a supplementary level that is closer to the biological needs: the participant systems have to extract a regulation network from the text that links and integrates basic molecular interactions.

The GRN task is based on a series of previous challenges in IE that started with the LLL challenge in 2005 (Nédellec, 2005). The LLL corpus is a set of sentences of PubMed abstracts about molecular interactions of the model bacterium *Bacillus subtilis*. Originally, the LLL task defined a unique binary genic interaction relation between proteins and genes. Since then, it has evolved to include the description of interaction events in a fine-grained representation that includes the distinction between transcription, different types of regulations and binding events, as proposed by (Manine *et al.*, 2009). This new schema better captures the complexity of regulations at the molecular level. Entities other than genes and proteins were introduced, such as DNA sites (*e.g.* transcription promoter sites, transcriptional regulator binding sites). We proposed the Genic Interaction task (Bossy *et al.*, 2012) in the BioNLP'11 Shared Task with a full re-annotation of the LLL corpus that follows this schema. The GRN task in

Proceedings of the BioNLP Shared Task 2013 Workshop, pages 153–160,
Sofia, Bulgaria, August 9 2013.

BioNLP-ST'13 builds on this corpus and includes annotation improvements and extensions that are detailed below.

3 Task description

The BioNLP-ST 2013 GRN task consists of the automatic construction of the regulation network that can be derived from a set of sentences. As usual in relation extraction tasks, the GRN corpus includes text-bound annotations. However the extraction target is the network, which is a structure with a higher level of abstraction. GRN thus also provides an explicit procedure to derive a network from a set of text-bound annotations.
The GRN annotation is stacked in four successive levels of annotation:

1. **Text-bound entities** represent genes, proteins and aggregates (families, complexes). Some entities directly relate to a gene and are given a unique gene identifier corresponding to a node of the network. These entities are hereby called *genic named entities*.
2. **Biochemical events and relations** are molecular-level events (*e.g.* transcription, binding) and detailed knowledge on relationships between entities (*e.g.* promoter of gene, regulon membership).
3. **Interactions** denote relations between entities and events and relations. Interactions are the first abstract annotations; they are the key to the construction of the network arcs.
4. Finally, the **Genic Regulation Network** is derived from the Interactions and from the identifiers of the named genic entities.

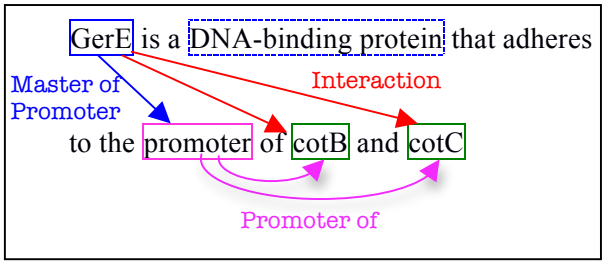

Figure 1. Example of annotated sentence.

Levels 1, 2 and 3 were obtained by a manual annotation of the GRN corpus sentences by a domain expert. Level 4 was automatically computed from the lower level annotations. The training corpus was provided to the participants with level 1, 2 and 3 annotations. The algorithm to compute the next level was described and implemented as a script and made available to the participants during the training stage of the challenge.
The test corpus was provided with only level 1 annotations (entities). The participants submitted their prediction either as a set of Interactions (level 3) or directly as a network (level 4). This setting allows the participants to train systems that work at different levels of abstraction.
Submissions in the form of *Interactions* are translated into a *Genic Regulation Network* using the algorithm provided during the training stage. The evaluation of each submission is carried out by comparing the predicted network with the reference network. The reference network is itself computed from the gold level 1, 2 and 3 annotations of the test sentences.
The following subsections describe the four annotation levels. The full annotation schema that specifies the constraints on event and relation arguments can be found on the task web page[1].

3.1 Text-bound entity types

Text-bound entities come in three kinds: event trigger words, genic entities and entity aggregates. Trigger words are of type *Action*, they serve as anchors for events.
Genic entities represent mentions of biochemical objects of the bacteria cell. Genic entity types include *Gene*, *mRNA*, *Promoter*, *Protein* and *Site*. Finally aggregates denote composite objects of the bacteria cell. Aggregate types are:

- *GeneFamily*: homologous gene families.
- *Operon*: operons *sensu* prokaryotes.
- *PolymeraseComplex*: RNA polymerase complexes, either the core complex alone, or bound to a sigma factor.
- *ProteinComplex*: protein complexes formed by several proteins that bind together.
- *ProteinFamily*: homologous protein families.
- *Regulon*: regulons, *sensu* prokaryotes.

3.2 Biochemical events and relation types

Biochemical events and relations represent the knowledge of cellular mechanisms at the molecular level. There are three types of events:

- *Transcription_by* represents the transcription event by a specific RNA

1 https://sites.google.com/site/bionlpst2013/tasks/gene-regulation-network

polymerase. Its agent is usually a *PolymeraseComplex*.

- *Transcription_from* represents the transcription from a specific site or promoter.
- *Action_Target* is a generic bio-molecular event.

The relation types represent three major genetic regulation patterns in bacteria: promoter activation, regulons and binding to specific DNA sites. Two types of relations specifically denote mechanisms that involve promoters:

- *Promoter_of* is a relation between a gene (or operon) and its promoter.
- *Master_of_Promoter* relation represents the control of the transcription from a specific promoter by a proteic entity (*Protein*, *ProteinComplex* or *ProteinFamily*).

Two other relation types represent the function of regulons:

- *Member_of_Regulon* relation denotes the membership of a genic entity to a regulon.
- *Master_of_Regulon* relation represents the control of the activity of an entire regulon by a protein.

Finally two types are used to represent relations that are common to different regulation mechanisms:

- *Bind_to* relation represents the binding of a proteic entity to a site on the chromosome.
- *Site_of* relation denotes the belonging of a chromosomal site to a genic entity such as a gene or a promoter.

3.3 Interaction types

Interaction relations are labeled with one of six types grouped into a small hierarchy following two axes: *mechanism* and *effect*. The hierarchical levels are figured here by the text indentations.

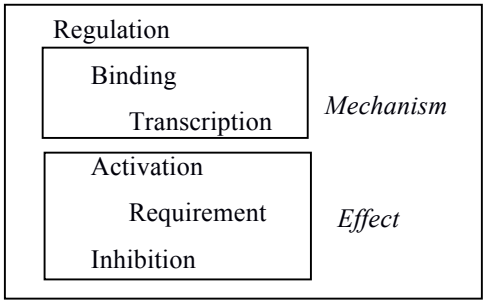

Figure 2. Types of Interaction relations

The *Binding* and *Transcription* types specify the mechanism through which the agent regulates the target. In a *Binding Interaction*, the agent binds to the target; this includes Protein-DNA binding and excludes Protein-Protein binding mechanisms. In a *Transcription Interaction*, the agent affects the transcription of the target.
The *Activation*, *Requirement* and *Inhibition* types specify the effect of the agent on the target. In an *Activation Interaction*, the agent increases the expression of the target. In a *Requirement Interaction*, the agent is necessary for the expression of the target. In an *Inhibition Interaction*, the agent reduces the expression of the target.
The *Regulation* type is the default type: in such interactions, neither the mechanism nor the effect is specified.

3.4 Genic Regulation Network inference algorithm

The genic regulation network corresponding to a corpus is inferred from the set of *Interaction* relations. The network presents itself as a directed labeled graph where nodes represent gene identifiers and edges represent gene interactions. The inference is done in two steps: the resolution of *Interaction* relations and the removal of redundant arcs.

Step 1: Resolution of Interaction relations
The agent and the target of an *Interaction* relation are not necessarily genic named entities. They can be secondary events or relations, another *Interaction*, or auxiliary entities (*e.g.* Promoter). The resolution of an *Interaction* aims to look for the genic named entity in order to infer the node concerned by the network edge. The resolution of *Interaction* arguments is performed using the rules specified below. These rules express well-known molecular mechanisms in a logical manner:

1. If the agent (or target) is a genic named entity, then the agent (or target) node is the gene identifier of the entity. If the entity does not have a gene identifier, then it is not a genic named entity and there is no node (and thus no edge).
2. If the agent (or target) is an event, then the agent (or target) node is the entity referenced by the event.
3. If the agent (or target) is a relation, then the agent (or target) of both arguments of the relation are nodes.

4. If the target is a *Promoter* and this promoter is the argument of a *Promoter_of* relation, then the target node is the other argument of the *Promoter_of* relation. *i.e.* if A interacts with P, and P is a promoter of B, then A interacts with B.
5. If the agent is a *Promoter* and this promoter is the argument of a *Master_of_Promoter* relation, the agent is the other argument of the *Master_of_Promoter* relation. *i.e.* if A is the master of promoter P, and P interacts with B, then A interacts with B.

The resolution of *Interaction* arguments consists of a traversal of the graph of annotations where these rules are applied iteratively. Event and relation arguments are walked through. *Promoter* entities are handled according to rules 4 and 5.
If the resolution of the agent or the target yields more than one node, then the *Interaction* resolves to as many edges as the Cartesian product of the resolved nodes. For instance, if both the agent and the target resolve to two nodes, the *Interaction* relation resolves into four edges.
Edges are labeled with the same set of types as the *Interactions*. Each edge inherits the type of the *Interaction* relation from which it has been inferred.

Step 2: Removal of redundant arcs
In this step, edges with the same agent, target and type are simplified into a single edge. This means that if the same *Interaction* is annotated several times in the corpus, then it will resolve into a single edge. This means that the prediction of only one of the interactions in the corpus is enough to reconstruct the edge.
Moreover, *Interaction* types are ordered according to the hierarchy defined in the preceding section. Since the sentences are extracted from PubMed abstracts published during different periods, they may mention the same *Interaction* with different levels of detail, depending on the current state of knowledge. For a given edge, if there is another edge for the same node pair with a more specialized type, then it is removed. For instance, the edges *(A, Regulation, B)* and *(A, Transcription, B)* are simplified into *(A, Transcription, B)*. Indeed the former edge conveys no additional information in comparison with the latter.

4 Corpus description

The GRN corpus is a set of 201 sentences selected from PubMed abstracts, which are mainly about the sporulation phenomenon in *Bacillus subtilis*. This corpus is an extended version of the LLL and BI (BioNLP-ST'11) corpora. The additional sentences ensure a better coverage of the description of the sporulation. An expert of this phenomenon examined the regulation network derived from the annotation of the original sentences, and then manually listed the important interactions that were missing. We selected sentences from PubMed abstracts that contain occurrences of the missing pairs of genes. In this way, the genic interaction network is more complete with respect to the sporulation. Moreover, the publications from which the sentences are extracted cover a wider period, from 1996 to 2012. They represent a diverse range of writing styles and experimental methods. 42 sentences have been added, but 4 sentences were removed from the BI sentences because they described genic interactions in bacteria other than *Bacillus subtilis*. The distribution of the sentences among the training, development and test sets has been done in the following way:

- Legacy sentences belong to the same set as in previous evaluation campaigns (LLL and BI).
- Additional sentences have been randomly distributed to training, development and test sets. The random sampling has been constrained so that the proportion of different types of interactions is as much as possible the same as in the three sets.

The GRN task does not include the automatic selection by the participant methods of the relevant sentences, which are provided. With regards to a real-world application, this selection step can be achieved with good performance by sentence filtering, as demonstrated by Nédellec *et al.* (2001), by using a Naive Bayesian classifier. Moreover, the corpus contains sentences with no interaction.
Tables 1 to 3 detail the distribution of the entities, relations and events in the corpus. They are balanced between the training and test sets: the test represents between a quarter and a third of the annotations. Table 1 details the entity frequency and their distributions by type. Column 5 contains the contribution of each entity type to the total. Genes and proteins represent two thirds of the entities, since they are the main actors in genic interactions. It is worth noting that the high number of promoters and polymerase complexes is specific to bacteria

where the biological mechanisms are detailed at a molecular level.

Entity	#	Train+Dev	Test
Gene	199	70%	30%
GeneFamily	2	50%	50%
mRNA	1	100%	0%
Operon	33	67%	33%
PolymeraseComplex	62	71%	29%
Promoter	63	73%	27%
Protein	486	65%	35%
ProteinComplex	7	100%	0%
ProteinFamily	18	78%	22%
Regulon	14	79%	21%
Site	32	78%	22%
Total	917	68%	32%

Table 1. Entity distribution in the GRN corpus.

Table 2 details the distribution of the biochemical events and relations (level 2). The most frequent event is *Action Target*. *Action Target* links, for instance, *Transcription by* and *Transcription from* events to the target gene.

Event/Relation	#	Train+dev	Test
Action target	226	68%	32%
Bind to	9	78%	22%
Master of Promoter	60	80%	20%
Master of Regulon	13	85%	15%
Member of Regulon	12	92%	8%
Promoter of	47	72%	28%
Site of	24	75%	25%
Transcription by	86	71%	29%
Transcription from	18	78%	22%
Total	495	72%	28%

Table 2. Distribution of the biochemical events and relations in the GRN corpus.

Finally, Table 3 details the distribution of the *Interaction* relations (level 3). The distribution among *Interaction* relations is more uniform than among entities and molecular events. The frequency of the Transcription relation is much higher than Binding, which is not surprising since transcription is the major mechanism of regulation in bacteria, while binding is rare. Conversely, the relative frequency of relations among Effect types of relations is balanced.

Interaction	#	Train+dev	Test
Regulation	80	65%	35%
Inhibition	50	66%	34%
Activation	49	67%	33%
Requirement	35	66%	34%
Binding	12	75%	25%
Transcription	108	74%	26%
Total	334	69%	31%

Table 3. Distribution of the Interaction relations in the GRN corpus.

5 Annotation methodology

A senior biologist, who is a specialist of *Bacillus subtilis* and a bioinformatician, a specialist of semantic annotation, defined the annotation schema. The biologist annotated the whole corpus, using the BI annotations as a starting point. The bioinformatician carefully checked each annotation. They both used the AlvisAE Annotation Editor (Papazian *et al.*, 2012) that supported their productivity due to its intuitive visualization of dense semantic annotations. *Subtiwiki* provided the identifiers of genes and proteins (Flórez *et al.*, 2009). *Subtiwiki* is a community effort that has become the reference resource for the gene nomenclature normalization of *Bacillus subtilis*. Other genic named entities, like operons, families or protein complexes, were given an identifier similar to their surface form. Several annotation iterations and regular cross-validations allowed the annotators to refine and normalize these identifiers.

The consistency of the annotations was checked by applying the rules of the network inference procedure that revealed contradictions or dangling events. The biologist double-checked the inferred network against his deep expertise of sporulation in *Bacillus subtilis*.

6 Evaluation procedure

6.1 Campaign organization

The same rules and schedule were applied to GRN as the other BioNLP-ST tasks. The training and development data were provided eleven weeks before the test set. The submissions were gathered through an on-line service, which was active for ten days. We took into account the final run of each participant to compute the official scores. They were published on the BioNLP-ST web site together with the detailed scores.

6.2 Evaluation metrics

The predictions of the participating teams were evaluated by comparing the reference network to the predicted network that was either submitted directly, or derived from the predicted *Interactions*. Since the genic named entity annotations are provided with their identifier, the network nodes are fixed. Therefore, the evaluation consists of comparing the edges of the two networks. Their discrepancy is measured using the *Slot Error Rate* (SER) defined by (Makhoul *et al.*, 1999) as:

$$SER = (S + D + I) / N$$

where:

- S is the number of substitutions (*i.e.* edges predicted with the wrong type)
- D is the number of deletions (false negatives)
- I is the number of insertions (false positives)
- N is the number of arcs in the reference network.

The SER has the advantage over F_1, namely it uses an explicit characterization of the substitutions. (Makhoul *et al.,* 1999) demonstrates that the implicit comprehension of substitutions in both recall and precision scores leads to the underestimation of deletions and insertions in the F score. However, we compute the Recall, Precision and F_1 in order to make the interpretation of results easier:

$$Recall = M / N$$
$$Precision = M / P$$

where:

- M is the number of matches (true positives).
- P is the number of edges in the predicted network.

Matches, substitutions, deletions and insertions are counted for each pair of nodes. The genic regulation network is an oriented graph, thus the node pairs *(A,B)* and *(B,A)* are handled independently. For a given node pair *(A,B)*, the number of exact matches (M) is the number of edges with the same type in the prediction as in the reference. The number of substitutions, deletions and insertions depends on the number of remaining edges. We name q and r, the number of remaining edges between two nodes A and B in the prediction and the reference respectively:

- $S = min(q, r)$
- if $q > r$, then $I = q – r$, $D = 0$
- if $q < r$, then $I = 0$, $D = r – q$

In other words, edges from the prediction and the reference are paired, first by counting matches, then by maximizing substitutions. The remaining edges are counted either as insertions or deletions depending if the extra edges are in the prediction or reference, respectively.

The values of $\mathcal{S}$, $\mathcal{D}$, $\mathcal{I}$ and $\mathcal{M}$ for the whole network are the sum of S, D, I and M on all the node pairs.

7 Results

7.1 Participating systems

Five systems participated in GRN:

- University of Ljubljana (Slovenia) (Žitnik *et al.*, 2013),
- K.U.Leuven (Belgium) (Provoost and Moens, 2013),
- IRISA-TexMex (INRIA, France) (Claveau, 2013),
- EVEX (U. of Turku / TUCS, Finland and VIB / U. of Ghent, Belgium) (Hakala *et al.*, 2013),
- TEES-2.1 (TUCS, Finland) (Björne and Salakoski, 2013).

Participant	SER	Recall	Precision
U. of Ljubljana	0.73	34%	68%
K.U.Leuven	0.83	23%	50%
TEES-2.1	0.86	23%	54%
IRISA-TexMex	0.91	41%	40%
EVEX	0.92	13%	44%

Table 4. Final evaluation of the GRN task. Teams are ranked by SER. S: Substitutions, D: Deletions, I: Insertions, M: Matches.

Table 4 summarizes the scores by decreasing order. The scores are distributed between the best SER, 0.73 achieved by the University of Ljubljana, 20 points more than the lowest at 0.92. For all systems, the number of insertions is much lower than the number of deletions, except for IRISA-TexMex.

The substitutions correspond to the edges that were predicted with the wrong type. In order to reveal the quality of the predictions with regards to the edge types, we calculated two alternate SERs. The results are displayed in Table 5.The *SER Network Shape* is obtained by erasing the type of all of the edges in the reference and predicted networks, as if all edges were of the *Regulation* type. The *SER Network Shape* measures the capacity of the systems to reconstruct the unlabeled shape of the regulation network. The *SER Effect* is obtained by erasing the mechanism types of all edges only, as if *Binding* and *Transcription* edges were of type *Regulation*. The *Effect* edges are kept unchanged. The *SER Effect* measures the quality of the predictions for valued networks that only contain *Effect* edges.

Participant	SER	SER Shape	SER Effect
U. of Ljubljana	0.73	0.60	0.74
K.U. Leuven	0.83	0.64	0.83
TEES-2.1	0.86	0.74	0.84
IRISA-TexMex	0.91	0.51	0.87
EVEX	0.92	0.79	0.91

Table 5. Scores obtained by erasing edge types (*Network Shape*) or mechanism types (*Effect*).

The *SER Network Shape* is significantly better for all systems, but the impact is dramatic for IRISA-TexMex and K.U. Leuven, showing that the typing of relations may be the major source of error. The *SER Effect* does not differ significantly from the original score. We deduce from the comparison of the three scores that the types that are the hardest to discriminate are effect types. This result is interesting because *Effect* labels are in fact the most valuable for systems biology and network inference studies.

U. of Ljubljana and TEES-2.1 submissions contained level 2 and 3 predictions (interactions and biochemical events). IRISA provided only predictions at level 3 (interactions only). K.U. Leuven and EVEX directly submitted a network. The performance of the systems that use annotations of level 2 confirms our hypothesis that a significant part of the interactions can be deduced from low-level events.

7.2 Systems description and result analysis

All systems applied machine-learning algorithms with linguistic features that were stems or lemmas, POS-tags and parses, most of them being provided by the BioNLP supporting resources. With the exception of K.U. Leuven, all systems used dependency paths between candidate arguments. However different ML algorithms were used, as shown in Table 6.

Participant	ML algorithm
U. Ljubljana	Linear-chain CRF
K.U.Leuven	SVM (Gaussian RBF)
TEES-2.1	$SVM^{multiclass}$ (linear)
IRISA-TexMex	kNN (language model)
EVEX	SVM (TEES-2.1)

Table 6. ML algorithms used by the participants.

Beyond syntactic parses and ML algorithms, the participant systems combined many different sources of information and processing, so that no definitive conclusion on the respective potential of the methods can be drawn here.

8 Conclusion

The GRN task has a strong legacy since the corpus is derived from LLL. Moreover, the GRN task has advanced a novel IE setting. We proposed to extract a formal data structure from successive abstract layers. Five different teams participated in the task with distinct strategies. In particular, we received submissions that work on all proposed abstraction levels.

This shows that Information Extraction implementations have reached a state of maturity, which allow for new problems to be addressed quickly. The performances are promising, yet some specific problems have to be addressed, like the labeling of edges.

Acknowledgments

This work was partially supported by the Quaero programme funded by OSEO (the French agency for innovation).

References

Jari Björne, Tapio Salakoski. 2013. TEES 2.1: Automated Annotation Scheme Learning in the BioNLP 2013 Shared Task. In *Proceedings of the BioNLP 2013 Workshop*, Association for Computational Linguistics.

Christian Blaschke, Miguel A. Andrade, Christos Ouzounis, Alfonso Valencia. 1999. Automatic Extraction of Biological Information From Scientific Text: Protein-Protein Interactions. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology (ISMB 1999)*, 60-67.

Robert Bossy, Julien Jourde, Alain-Pierre Manine, Philippe Veber, Erick Alphonse, Marteen van de Guchte, Philippe Bessières, Claire Nédellec. 2012. BioNLP Shared Task - The Bacteria Track. *BMC Bioinformatics*. 13(Suppl 11):S3.

Vincent Claveau. 2013. IRISA participation to BioNLP-ST 2013: lazy-learning and information retrieval for information extraction tasks. In *Proceedings of the BioNLP 2013 Workshop*, Association for Computational Linguistics.

Patrik D'haeseleer, Shoudan Liang, Roland Somogyi. 2000. Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics*. 16(8):707-726.

Lope A. Flórez, Sebastian F Roppel, Arne G Schmeisky, Christoph R Lammers, Jörg Stülke. 2009. A community-curated consensual annotation that is continuously updated: the *Bacillus subtilis* centred wiki SubtiWiki. *Database (Oxford)*, 2009:bap012.

Kai Hakala, Sofie Van Landeghem, Tapio Salakoski, Yves Van de Peer and Filip Ginter. 2013. EVEX in ST'13: Application of a large-scale text mining resource to event extraction and network construction. In *Proceedings of the BioNLP 2013 Workshop*, Association for Computational Linguistics.

GenBank. http://www.ncbi.nlm.nih.gov/

Hidde de Jong. 2002. Modeling and simulation of genetic regulatory systems: a literature review. *J. Computational Biology*, 9(1):67-103.

Hiroaki Kitano. 2002. Computational systems biology. *Nature*, 420(6912):206-210.

John Makhoul, Francis Kubala, Richard Schwartz and Ralph Weischedel. 1999. Performance measures for information extraction. In *Proceedings of DARPA Broadcast News Workshop,* Herndon, VA, February.

Alain-Pierre Manine, Erick Alphonse, Philippe Bessières. 2009. Learning ontological rules to extract multiple relations of genic interactions from text. *Int. J. Medical Informatics*, 78(12):31–38.

Claire Nédellec, Mohamed Ould Abdel Vetah, Philippe Bessières. 2001. Sentence filtering for information extraction in genomics, a classification problem. *Practice of Knowledge Discovery in Databases (PKDD 2001),* 326-337.

Claire Nédellec. 2005. Learning Language in Logic - Genic Interaction Extraction Challenge" in *Proceedings of the Learning Language in Logic (LLL05)* workshop joint to *ICML'05*. Cussens J. and Nédellec C. (eds). Bonn.

Frédéric Papazian, Robert Bossy and Claire Nédellec. 2012. AlvisAE: a collaborative Web text annotation editor for knowledge acquisition. *The 6th Linguistic Annotation Workshop* (The LAW VI), Jeju, Korea.

Thomas Provoost, Marie-Francine Moens. 2013. Detecting Relations in the Gene Regulation Network. In *Proceedings of the BioNLP 2013 Workshop*, Association for Computational Linguistics.

Slavko Žitnik, Marinka Žitnik, Blaž Zupan, Marko Bajec. 2013. Extracting Gene Regulation Networks Using Linear-Chain Conditional Random Fields and Rules. In *Proceedings of the BioNLP 2013 Workshop*, Association for Computational Linguistics.