

## Construction of coffee transcriptome networks based on gene annotation semantics

Luis F. Castillo<sup>1,2,\*</sup>, Narmer Galeano<sup>3</sup>, Gustavo A. Isaza<sup>1</sup>, Alvaro Gaitan<sup>3</sup>

<sup>1</sup>Universidad de Caldas, Grupo Investigación GITIR, {luis.castillo, gustavo.isaza}@ucaldas.edu.co

<sup>2</sup>Universidad Nacional de Colombia Sede Manizales, Departamento Ing. Industrial

<sup>3</sup>Centro Nacional de Investigaciones de Café (Cenicafé), Manizales, Colombia,  
<http://www.cenicafe.org>,  
{narmer.galeano, alvaro.gaitan}@cafedecolombia.com

### Summary

Gene annotation is a process that encompasses multiple approaches on the analysis of nucleic acids or protein sequences in order to assign structural and functional characteristics to gene models. When thousands of gene models are being described in an organism genome, construction and visualization of gene networks impose novel challenges in the understanding of complex expression patterns and the generation of new knowledge in genomics research. In order to take advantage of accumulated text data after conventional gene sequence analysis, this work applied semantics in combination with visualization tools to build transcriptome networks from a set of coffee gene annotations. A set of selected coffee transcriptome sequences, chosen by the quality of the sequence comparison reported by Basic Local Alignment Search Tool (BLAST) and Interproscan, were filtered out by coverage, identity, length of the query, and e-values. Meanwhile, term descriptors for molecular biology and biochemistry were obtained along the Wordnet dictionary in order to construct a Resource Description Framework (RDF) using Ruby scripts and Methontology to find associations between concepts. Relationships between sequence annotations and semantic concepts were graphically represented through a total of 6845 oriented vectors, which were reduced to 745 non-redundant associations. A large gene network connecting transcripts by way of relational concepts was created where detailed connections remain to be validated for biological significance based on current biochemical and genetics frameworks. Besides reusing text information in the generation of gene connections and for data mining purposes, this tool development opens the possibility to visualize complex and abundant transcriptome data, and triggers the formulation of new hypotheses in metabolic pathways analysis.

## 1 Introduction

Coffee is the second most-heavily traded commodity after oil [1]. Coffee plantations are circumscribed to tropical regions with two major cultivated species: *Coffea arabica* that produces mild coffee, and *Coffea canephora* which is the source of the robusta coffee. However, there

---

\*To whom correspondence should be addressed.

are over 100 different species, most of them growing wild in the African landscape. As in many crops, high throughput technologies applied to explore genetic data as well as bioinformatics developments are key steps to prospect the coffee biodiversity present in germplasm collections. Mining of genomic data enables the identification of genetic features responsible for disease and pest resistance, adaptation to changing environments, yield and cup quality, all important features to be considered in the generation of new varieties that will ensure a sustainable supply of coffee in the years to come and improve market standards for consumers. Genomic analyses include the structural reconstruction of gene models from sequencing data and the identification of molecular architectures that in turn define the function of genes and proteins. Most of this sequence annotation with putative functional data is done through broadly used bioinformatics tools such as BLAST and Interproscan. These tools are the first step towards the annotation of cDNA and protein sequences that belong to the transcriptome of species that have been scarcely studied at the experimental level [2,3,4]. However, this information has to be complemented on a larger scale with the finding of associations among genes in order to construct networks that in a more comprehensive way explain the biochemical pathways that have evolved inside the cells and that ultimately give rise to the phenotypic features that characterize every organism.

In order to represent the relationships between genes and gene products, networks are used ubiquitously throughout biology. They are traditionally constructed by accumulating and connecting wet lab evidence, either from biochemistry or genetics experiments, although transcriptome profiles obtained from expression studies in high throughput genomics are nowadays becoming more widely used. At the end, the biological knowledge is reduced to scientific text, represented as gene annotations related to genomic sequences. Currently, extracting relationships by data mining of text information is of great interest, even more if the process can be done in an automatic way that enables the discovery of new associations. Semantics, with its focus on the study of relations between different linguistic units and compounds, offers an opportunity to the generation of gene networks. By the process of identifying members of a lexicon within the text, syntactic sentence parsers can process large rule sets automatically, and ontology can organize them, allowing the normalization of relationships and permitting inference over them. This approach has been recently used in pharmacogenomics [5] and in the study of gene diseases [6].

To design and implement the ontology, Web Ontology Language (OWL) has been used as a language which allows for formalisms and representations of knowledge. Some of the most important influences in the design of OWL comes from its predecessor DAML OIL in the Logic Descriptive paradigm frameworks and models RDF / XML. The Logic of Description and interpretation of it has a great influence on the design of OWL, particularly in the formalization of the semantics, the choice of languages builders and integration of data types and values, thus OWL DL and OWL Lite can be seen as expressions of the logic description or descriptive. The logic description or Descriptive Logic (DL) is a set of languages that represent knowledge from a domain terminologies of structured and understood formalism, to describe and semantic concepts that denote formulas, relations and logical expressions of first order predicate. Some salient features of the description logic formalisms is the inclusion of descriptive roles, concepts, individuals, formalisms terminological formalisms assertive and abilities to infer new knowledge from technical reasoning. Among its main features that allow you to define your language are the names of concepts, roles, individuals, developers, complex concepts, among others. The Description Logics ancestors come from formalisms as semantic networks and

frames, which were incorporated new features of formal semantics. To define the rules will be use SWRL (Semantic Web Rule Language), which is defined as an expression language based on OWL rules, can write rules expressed as OWL concepts providing reasoning capabilities. A SWRL rule contains an antecedent to describe the body of the rule and consistent with respect to the head of it, each consisting of a set (May be empty) of atoms. Both the body and the head of the rule are positive conjunctions of atoms can be understood as meaning SWRL that if all elements in the antecedent is True, intone the consequent must also be true [7]. SWRL rules will be written in terms of classes, properties, and values of OWL individual.

Complex networks can be envisioned in 2 or 3 dimensions using visualization utilities for bioinformatics [8]. The display of existing connections in a wide data set provides a simple way to interpret and understand the gene relationships lying underneath cell metabolic processes, and facilitates the discovery of new associations previously overlook due to data deluge. Among the tools available can be found well-know examples such as Cytoscape, Zentity, VisANT, Pathway Studio and Patika. Cytoscape is widely used in projects of data representation in biological, genomic and proteomic systems Cytoscape ([www.cytoscape.org](http://www.cytoscape.org)). Zentity is used to visualize the existing relations in management process of research documents, and Pathway Studio combines an extensive set of features with a polished graphic user interface.

This document presents a procedure to find explicit and implicit relationships among gene models based on semantic analysis of their corresponding annotations. We applied the method to a highly cured coffee genomics database to construct a small scale prototype of a transcriptome network that could generate hypotheses for further biological validation.

## 2 Recent work in bioinformatics using Semantic Web

One initiative that has been significantly strengthened in the field of Semantic Web Bioinformatics applied to points to the use of standards such as Linked Open Data. The Linked Open Data (LOD) is a community project on the World Wide Web Consortium (W3C), which aims to expand “the network with a shared data by publishing various open datasets as RDF on the Web and by establishing links between the elements of RDF data from different data sources” [9,10]. In this context, many biomedical databases have already been made available (an open diagram cloud-based data is available online [11]). Multiple data sets are derived deBio2RDF, but also some that were built independently. Approaches based on Semantic Web for biomedical data integration have been proposed in some instances in recent years [12,13,14,15]. In the biomedical field, a resource copy has been represented by Bio2RDF [16], a system for integrated access to a large number of biomedical databases through Semantic Web technologies RDF, ie, for data representation and SPARQL (SPARQL Protocol and RDF Query Language) [17] for queries. In [18] and [19] is presented the portal Chem2Bio2RDF a Linked Open Data (LOD) portal for chemical systems biology to facilitate drug discovery. It converts about 25 different datasets in the genes, compounds, drugs, roads, side effects, diseases and RDF triples that links to other bubbles LOD, as Bio2RDF, LODD and DBpedia. The portal is based on D2R server and provides a SPARQL endpoint, but adds a few unique features such as faceted RDF, easy to use from SPARQL query builders, MEDLINE / PubMed service of cross-validation, and monitoring via Cytoscape. More recent efforts like [20] present a website that is aimed at developing community around biological data linked (LOD). This public space provides several services as a collaborative infrastructure in order to stimulate the generation of activity in

**Table 1: Annotated contigs**

Species	Annotated contigs			Contigs With No hits	Total per species
	Annotated as “whole genome shotgun”	Annotation different from “whole genome shotgun”	Subtotal		
<i>C canephora</i>	5892	3336	9228	3929	13157
<i>C arabica</i>	16488	8250	24738	6735	31473
<i>C kapakata</i>	884	655	1539	508	2047
<i>C liberica</i>	6374	3515	9889	1763	11652
Total	29638	15756	45394	12935	58329

the use of biological data linked and therefore contributes to the implementation of the benefits of Web data in this area.

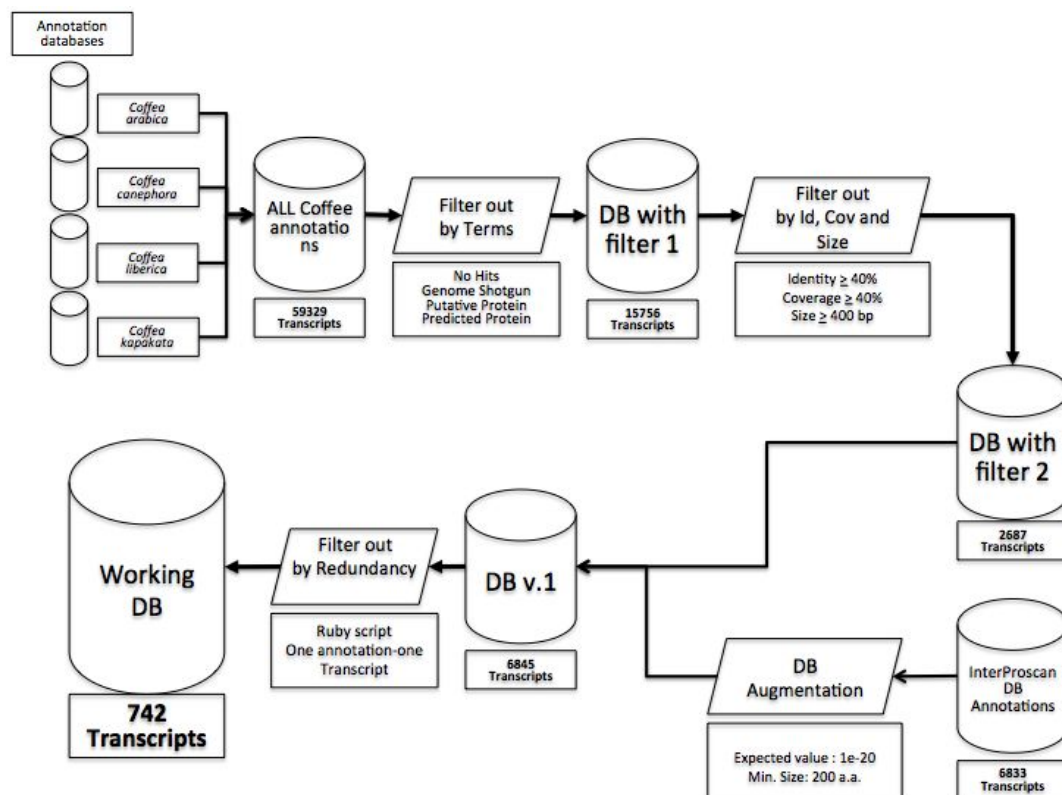
### 3 Methods

Considering annotation text as the raw material for network construction, and due to the dependence on local significant alignments in automatic annotation systems, a filtering pipeline was designed to select sequences comprising global homologies to well known genes. The experimental dataset was built with annotations from selected transcriptome sequences (contigs) from the *Coffea* species *C. arabica*, *C. canephora*, *C. kapakata* and *C. liberica*, publicly available in the Genbank and annotated at the National Coffee Research Center (Cenicafe) [bioinformatics.cenicafe.org]. Out of 58,329 coffee transcriptome entries, those containing as the only description “No hits found”, were eliminated in a first filtering round, followed by the ones annotated as “Genome shotgun”, “Putative protein” or “Predicted protein” (Table 1).

For the remaining 15,756 sequences, a quality filter was set up based on the BLAST output which involved the percentage coverage of the reference sequence, the length of the query sequence involved in the match and the percentage similarity between the two local alignments [21]. Since only a small percentage of the dataset passed the filter, the procedure was also applied to the openly available set of transcripts of the highly annotated model species *Arabidopsis thaliana* [arabidopsis.org] in order to detect any bias against the coffee data. With the purposed of enhancing the data set, Interproscan annotations were included in the analysis, considering the e-value and the length of the match as thresholds. Remaining redundancy after the augmentation of the data collection was eliminated to keep a single representative for each selected contig (Figure 1), ending up with a highly cured working database.

A corpus is considered as a wide range of real examples of language use. The procedure to find associations among selected transcripts was supported on a corpus validated against the term descriptors listed in the Oxford Dictionary for Molecular Biology and Biochemistry [22], and complemented with the terms present in Wordnet [23].

In the project context, the corpus was a set of useful notes, where each entry was linked to an identifier of a contig. A problem may arise when choosing the useful terms associated to the descriptions in the coffee genome, foreseeing this issue it was decided to use the annotations of the sequences that complied with the established quality filter determined above. Flat text files stored the results throughout the process of filtration.



**Figure 1: Pipeline for the construction of a biologically significant annotation data set from a collection of *Coffea* transcripts**

Interproscan annotations were stored as flat files in the CoffeaBank database in the format:

*Id | Id\_contig | id\_interproScan | evaluate | annotation | reference GO (Gene Ontology)*

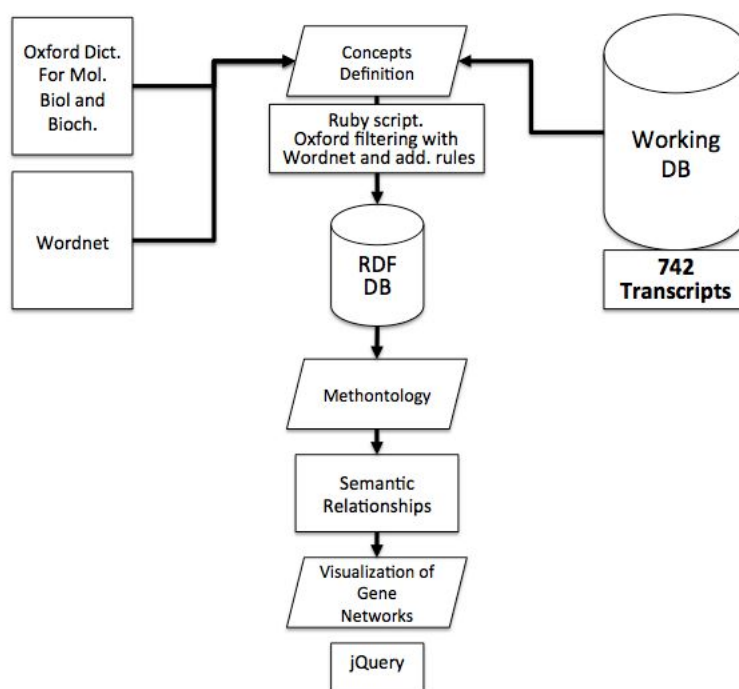
and the BlastX annotation kept the following format:

*Id\_Contig | id\_BLAST | annotation | %C | %I | Contig\_length*

Annotations in the working database were screened using the reference corpus and concepts found in the annotations were related to transcripts identifications using oriented vectors by means of Ruby scripts. Annotation concepts of the experimental dataset were encoded as Resource Description Framework (RDF) triplets that served as entries for semi-automatic ontological analysis with Methontology [24, 25]. Relationships among concepts and transcripts were then visualized using a graph visualization library applying web workers and jQuery<sup>1</sup>, see Figure 2.

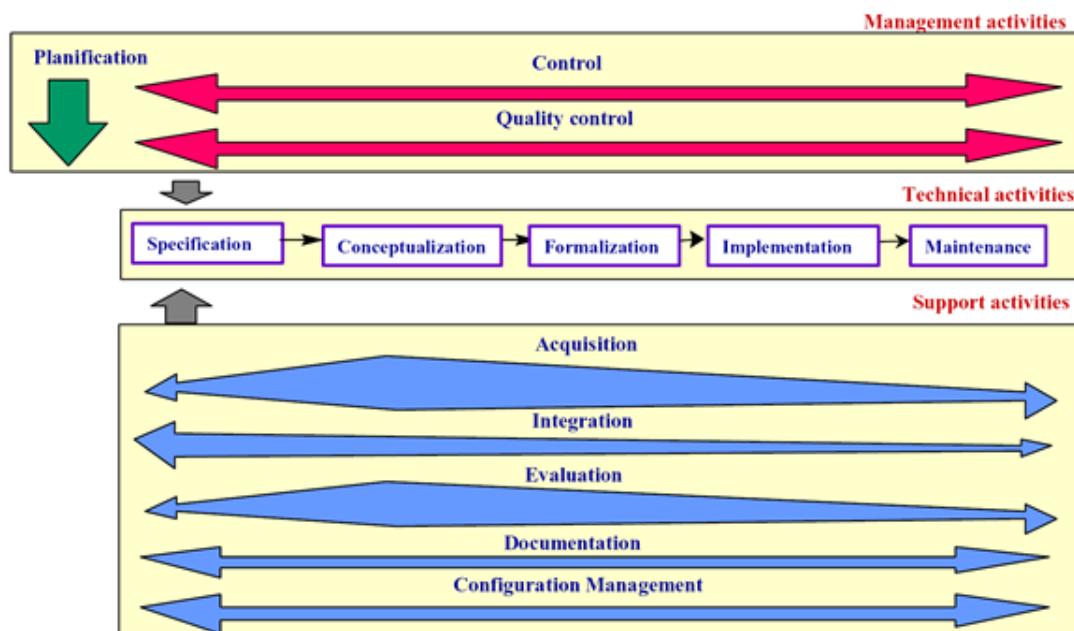
RDF is considered as a semantic markup language, which allows to describe or categorize web resources, a fundamental aspect of what is considered the definition of metadata for information retrieval. RDF is a general method to decompose knowledge into small pieces, with some rules about the semantics or meaning of these pieces. RDF is a formal description of concepts, terms and relationships within a given knowledge domain.

<sup>1</sup><http://www.arborjs.org>



**Figure 2: Workflow followed to build up and visualize a network of transcript annotations based on semantics associations**

Methontology (see Figure 3) provides a user-friendly approach to knowledge acquisition by non knowledge engineers. Ontology specifications goal is to put together a document that covers the ontology's primary objective, purpose, granularity level, and scope. The aim is to identify the set of terms to be represented, their characteristics, and their granularity. When most of the knowledge has been acquired, the ontologist has a lot of unstructured knowledge that must be organized. Conceptualization organizes and structures the acquired knowledge using external representations that are independent of the implementation languages and environments. Specifically, this phase organizes and converts an informally perceived view of a domain into a semiformal specification, using a set of intermediate representations that the domain expert and ontologist can understand. We built a glossary of terms that includes all the terms (concepts, instances, attributes, verbs, and so on) of the relations between the annotations and the Oxford-Wordnet dictionaries. The ontology life cycle schedules the ontology development activities detailed previously, although not all of them are currently considered by the Methontology life cycle. The life cycle is cyclic, based on evolving prototypes [25]. It allows an incremental development of the ontology that enables earlier validation and readjustment. Each cycle starts with the scheduling activity that identifies the tasks to be performed, their arrangement, their temporal extent and the resources they need. After that the development activities are engaged, starting with specification. Simultaneously, the management activities, control and quality assurance, and the support activities, knowledge acquisition, integration, evaluation, documentation and configuration management, are launched. They take place in parallel with the development activities (Figure 3, [25]).



**Figure 3: Methontology life cycle [12]**

If a development cycle is completed, these are the steps that are performed:

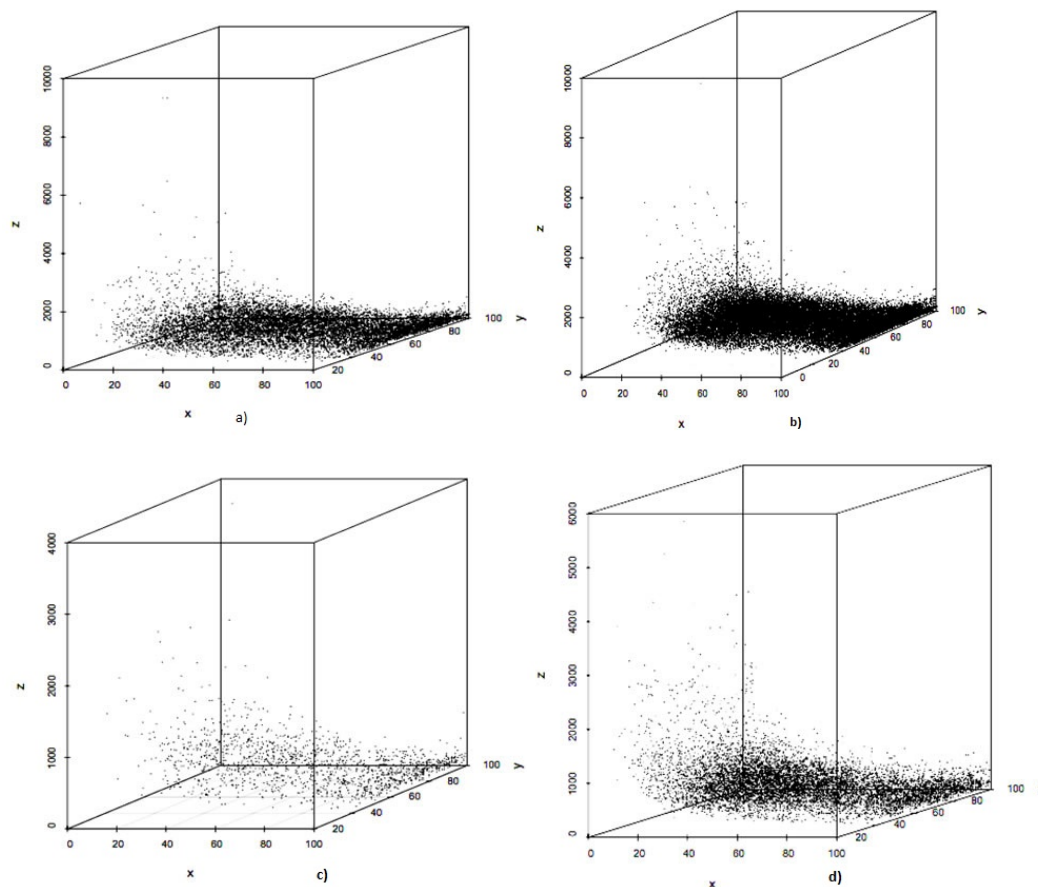
- To specify the prototype.
- To build a conceptual model
- To formalise the conceptual model.
- To implement the formalised conceptual model.
- To maintain the resulting ontology, which might lead to a new development cycle if unsatisfied or new requirements are detected.

## 4 Results and discussion

An experimental set of sequences with biologically relevant annotations was selected from the available coffee transcripts. For the four coffee species included, the distribution of the transcripts according to the threshold parameters defined was the same (Figure 4, a to d), indicating that less than 8 % of sequences (2,687) met simultaneously the three requirements of at least 40 % similarity, 40 % coverage and 400 bp in length, that defined significant confidence in the annotation. A comparison with the transcript set of the model species *Arabidopsis thaliana* showed that 35,5 % (9,024) of the transcripts passed the same filter (data not shown).

This reflects the current situation in gene annotation among plant species, and in other organisms, where many of the genes found by high throughput technologies lack a significant putative annotation due to gene evolution and the lag in functional characterization of proteins





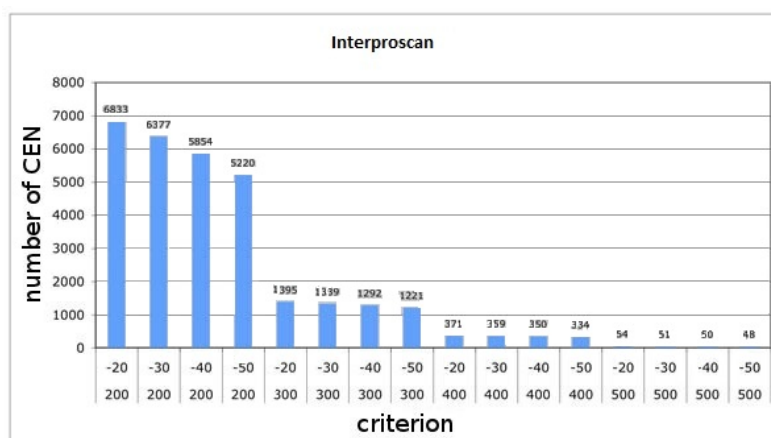
**Figure 4: Distribution of transcripts according to threshold variables associated to BLAST results.** (a). *C. canephora*; (b). *C. arabica*; (c). *C. kapakata*; (d). *C. liberica*; x-axis: percentage coverage, y-axis: percentage identity, z-axis: contig length.

when compared to sequencing. To widen the dataset, sequences displaying e-values smaller than  $e^{-20}$  and with at least 200 bp of matching length in Interproscan (Figure 5) were added, completing a preliminary experimental set up to 6,845 sequences.

Filtering out redundant descriptions yielded a workable database of 742 unique annotations. This sizable reduction may be due in part to the presence of large and highly characterized gene families that share similar function but with specialized spatial, developmental or environmental expression patterns. That is the case, for instance, of protein kinases (253), cytochrome p450 (225) or peroxydases (107), just to name a few. In fact, 50% of the preliminary database was represented by only 80 entries in the workable database. The oriented vector visualization (Figure 6) reflects this situation of association of multiple contigs to common annotation terms, where for some concepts there are numerous transcripts linked, while in infrequent cases (14 %) there were concepts unique to one or a few transcript annotations.

A complex network of transcript relationships can be constructed based on the method applied (Figure 7) that still requires a manual examination to verify the biological relevance of the connections. Until additional refinements are taken on the system, perhaps integrating some quantification in the significance of the associations found, care must be taken accepting the inferences that are being established and additional check outs are mandatory.





**Figure 5: Distribution of transcripts according to threshold variables associated to Interproscan results.**

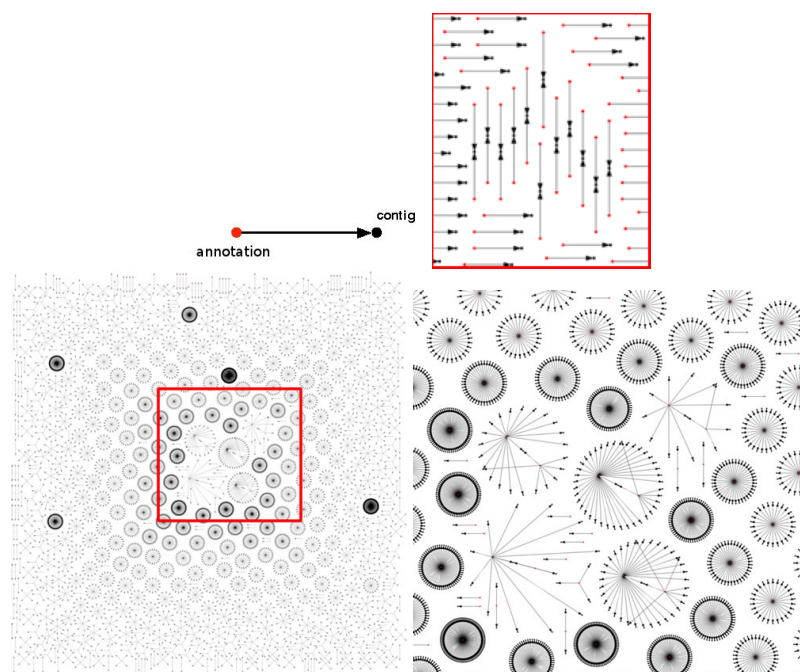
In one example of true-positive associations (Figure 8a), the concept “peroxidase” is displayed in association to hemoproteins and oxidoreductases. Hemoproteins are defined as proteins to which an iron-porphyrin compound is linked in a stoichiometric manner. Within the biochemical framework, peroxidases are considered among the group of hemoproteins, in the basal catalysis of the reaction of hydrogen peroxidase. Similarly, a group of oxidoreductases act as peroxidases in the role of electron acceptor on peroxide. An additional association could be inferred involving the hemoprotein and oxidoreductase concept, which is a valid enzyme class known as NADPH:hemoprotein oxidoreductase, absent in the workable database.

Other relationships have unclear meaning, which could be considered as false-positives until further evidence is gathered. An instance of this group can be considered the case of serpins (Figure 8b), a group of enzymes defines as serine protein inhibitors, that appears linked to protein phosphatases on a first level, and to storage proteins on a second level. Bibliographic searches cannot support significant structural or functional associations among these concepts. For both cases, the system designed provides a new layer of information that integrates the results from traditional automatic annotations, such as BLAST and Interproscan, and allows the formulation of a new series of hypotheses in the understanding of transcriptomes that must be complemented with gene expression, such as microarrays and RNAseq, and proteomics experiments.

The proposed model was applied using Methontology to generate high-level ontological specification for identifying conceptual relationships based on the genome annotations selected (see Figure 9) . This phase corresponds to the preliminary design of the ontology lifecycle and it is currently under construction to complement the ontology looking for integration with Gene Ontology to identify new gene relationships from axioms, inferences and reasonings.

## 5 Conclusions and future work

A semantic tool immersed in a biological environment was constructed to enable the coding and interchange of annotation data in a set of transcripts of an organism. Relationships found with this tool must resemble metabolic pathways already described in the scientific literature, but can also produce new and uncovered associations that have to be confirmed with wet-lab

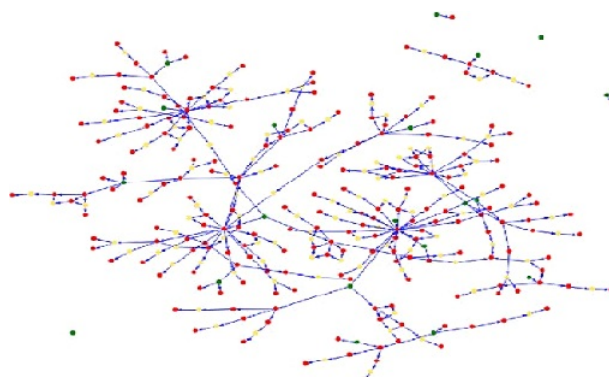


**Figure 6: Vector representation of transcript-to-concept associations using the workable database from coffee. Out of 742 transcripts, only 105 had unique associations, reflecting a significant presence of gene families in the database**

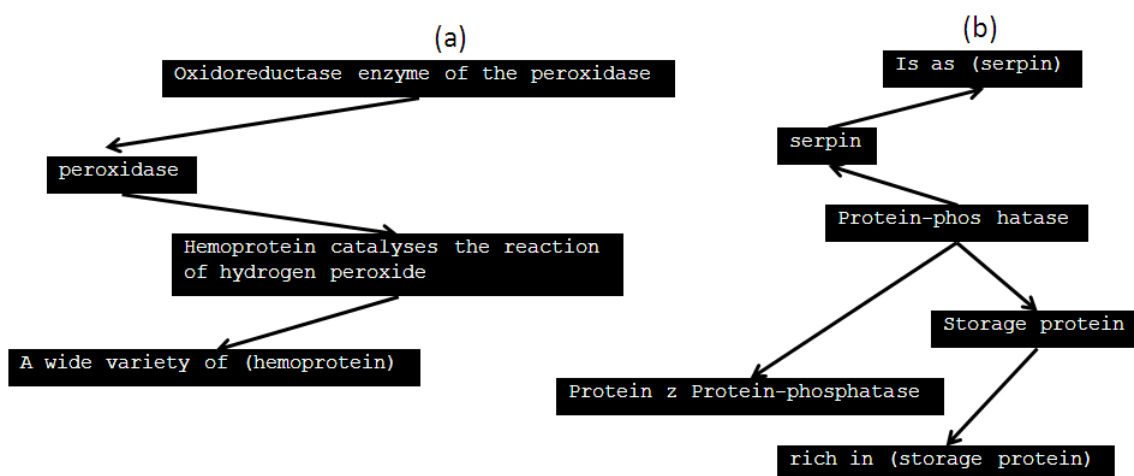
data.

In the short term, alternative displays centered on the transcripts ID (rather than in the concepts themselves) must be developed to ease the visualization in highly cluttered networks, as well as the generation of semantic similarity scores that provide deeper quantitative support for the evaluation of link strength. The structured data must be compatible to current systems to be reused to be mined with other metadata and complement the annotation dossier of a genomic project.

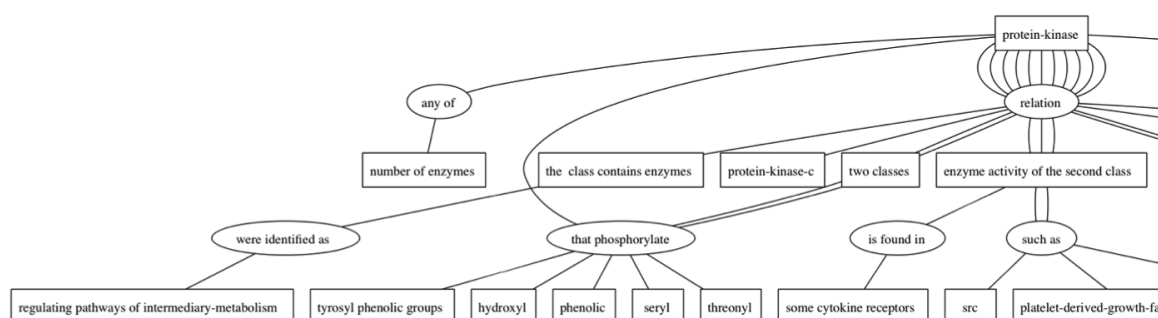
The application will be implemented using semantic web services technology. The visualization module will be connected with Bio2RDF, this framework creates and provides machine understandable descriptions of biological entities using the RDF/RDFS/OWL Semantic Web languages. The Bio2RDF network is a loosely coupled set of RDF databases which can re-



**Figure 7: Complex network of coffee transcripts constructed on RDF associations from transcript annotations. Green: Contig annotation; Yellow: Relational Concept; Red: Annotation concept**



**Figure 8: Detailed networks used for biological validation of concept relationships. (a). True-positive relationship around the concept of peroxidase. (b). False-positive relationship around the concept of protein phosphatase.**



**Figure 9: Partial High Level Ontology from Coffee Annotations**

spond to queries for the RDF versions of particular records on bioinformatics databases that they have information about. To enable this network to function reliably, some redundancy needs to be available, and queries need to be efficient.

## Acknowledgements

This work was supported by the Universidad de Caldas, the National Coffee Research Center (Cenicafé) and the Colombian Ministry of Agriculture.

## References

- [1] E. Hartwick. The Cultural Turn in Geography: A New Link in the Commodity Chain. *In* Encounters and Engagements between Economic and Cultural Geography, vol. 104 of GeoJournal Library, pages 39–46, 2012.
- [2] W. Li, J. Feng, T. Jiang. Workshop: Transcriptome assembly from RNA-Seq data: Objectives, algorithms and challenges. *In* IEEE 1st International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), page 271, 2011.

- [3] L. Jing, M. K. Ng and Y. Liu. Construction of Gene Networks With Hybrid Approach From Expression Profile and Gene Ontology. *IEEE Transactions on Information Technology in Biomedicine*, 14(1):107–118, 2010.
- [4] H. Zheng, F. Azuaje and H. Wang. seGOsa: Software environment for gene ontology-driven similarity assessment. *In IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 539–542, 2010.
- [5] A. Coulet, N. H. Shan, Y. Garten, M. Musen and R. B. Altman. Using text to build semantic networks for pharmacogenomics. *J Biomed Inform.*, 43(6):1009–1019, 2010.
- [6] R. Jiang, M. Gan and P. He. Constructing a gene semantic similarity network for the inference of disease genes. *BMC Systems Biology*, 5(Suppl 2):S2, 2011.
- [7] M. J. O'Connor, R. Shankar, C. Nyulas, S. Tu and A. Das. Developing a Web-Based Application using OWL and SWRL. *In AAAI Spring Symposium*, Stanford, CA, USA, 2008.
- [8] M. Suderman and M. Hallet. Tools for visually exploring biological networks. *Bioinformatics*, 23(20):2651–2659, 2007.
- [9] T. Berners-Lee, T. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer and D. Sheets. Tabulator: Exploring and Analyzing Linked Data on the Semantic Web. *In Proceedings of the 3rd International Semantic Web User Interaction Workshop*, 2006.
- [10] Linked Data-Connect Distributed Data across the web, <http://linkeddata.org/>.
- [11] The Linking Open Data cloud diagram, <http://lod-cloud.net/>.
- [12] S. Stephens, D. LaVigna, M. DiLascio and J. Luciano. Aggregation of bioinformatics data using Semantic Web technology. *Web Semantics: Science, Services and Agents on the World Wide Web*, 4(3):216–221, 2006.
- [13] L. Dhanapalan and J. Y. Chen. A case study of integrating protein interaction data using semantic web technology. *International Journal of Bioinformatics Research and Applications*, 3(3):286–302, 2007.
- [14] H. F. Deus, R. Stanislaus, D. F. Veiga, C. Behrens, I. I. Wistuba, J. D. Minna, H. R. Garner, S. G. Swisher, J. A. Roth, A. M. Correa, B. Broom, K. Coombes, A. Chang, L. H. Vogel and J. S. Almeida. A Semantic Web Management Model for Integrative Biomedical Informatics. *PLoS ONE*, 3(8):2946, 2008.
- [15] A. Miles, J. Zhao, G. Klyne, H. White-Cooper and D. Shotton. OpenFlyData: an exemplar data web integrating gene expression data on the fruit fly *Drosophila melanogaster*. *Journal of Biomedical Informatics*, 43(5):752–761, 2010.
- [16] F. Belleau, M. A. Nolin, N. Tourigny, P. Rigault and J. Morissette. Bio2RDF: Towards a mashup to build bioinformatics knowledge systems. *Journal of Biomedical Informatics*, 41(5):706–716, 2008.
- [17] SPARQL Query Language for RDF, <http://www.w3.org/TR/rdf-sparql-query/>

- [18] B. Chen, Y. Ding, H. Wang, D. J. Wild, X. Dong, Y. Sun, Q. Zhu and M. Sankaranarayanan. Chem2Bio2RDF: A Linked Open Data Portal for Systems Chemical Biology. *In* IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), pages 232–239, 2010.
- [19] B. Chen, X. Dong, D. Jiao, H. Wang, Q. Zhu, Y. Ding and D. J. Wild. Chem2Bio2RDF: a semantic framework for linking and data mining chemogenomic and systems chemical biology data. *BMC Bioinformatics*, 11:255, 2010.
- [20] M. J. García-Godoy, I. Navas-Delgado and J. Aldana-Montes. Bioqueries: a social community sharing experiences while querying biological linked data. *In* Proceedings of the 4th International Workshop on Semantic Web Applications and Tools for the Life Sciences (SWAT4LS '11), pages 24–31, 2012.
- [21] L. Verschelde, M. C. Dos Santos, T. Deray, B. Smith and W. Ceusters. Ontology-assisted database integration to support natural language processing and biomedical data-mining. *Journal of Integrative Bioinformatics*, 1(1):1, 2004.
- [22] Z. Langari and F. W. Tompa. Subject classification in the Oxford English Dictionary. *In* IEEE International Conference on Data Mining (ICDM), pages 329–336, 2001.
- [23] C. Fellbaum. WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA, 1998.
- [24] M. F. Lopez, A. Gómez-Perez, J. P. Sierra and M. P. Sierra. Building a Chemical Ontology Using Methontology and the Ontology Design Environment. *IEEE Intelligent Systems and their Applications*, 14(1):37–46, 1999.
- [25] M. F. Lopez, A. Gómez-Perez, M. D. Rojas Amaya. Ontology's crossed life cycles. *In* Knowledge Engineering and Knowledge Management, LNAI 1937, pages 65–79, 2000.