An analysis of the Brazilian researchers on freshwater fish genetics and reproduction as coauthors in academic papers

Carlos Eduardo M. Viegas da Silva[1, 3], Elisabete Maria Macedo Viegas[2]

[1]Faculdade de Zootecnia e Engenharia de Alimentos (USP), Departamento de Engenharia de Biossistemas, Pirassununga, SP, Brazil.
[2]Faculdade de Zootecnia e Engenharia de Alimentos (USP), Departamento de Zootecnia, Pirassununga, SP.

Abstract

Science is a social activity in some senses: first, in any of its fields science is an accumulated body of knowledge inherited from generations of antecessors by its practitioners, the scientists. In second place, the results of the scientific findings are shared socially and finally, those results to be successfully achieved demands from the researchers a high grade of collaboration, through the formations of a social network. The collaboration between researchers may result in the publication of an academic article. This publication serves as an index of the existence of a social network of collaboration between researchers and may reveal how a specific field of knowledge has been established, who are the researchers that act as facilitators among colleagues as well as to give some clues about the direction the specific field may unfold. The objective of this work was apply techniques of Social Network Analysis on the community formed by the Brazilian researchers, that are granted a CNPq scholarship on research productivity, with an interest in genetics and reproduction of freshwater fishes. Their linkage in co-authoring research papers published by academic journals were examined. For the determination of this community and the academic articles, the existing data of researchers in the Platform Lattes of the CNPq database were extracted through the applicative ScriptLattes, which search for the works published in co-authorship. Later, on these data was applied a software dedicated to the Analysis of Social Networks, and graphs of social network and some of its metrics were obtained. The analysis found the Lattes CVs of 33 researchers who published 1156 articles in co-authorship in 502 different journals. As a result, the analysis shows that these researchers form an academic network with low density, and in twenty periodicals were published 25% of the articles produced in co-authorship by those researchers.

Keywords: Academic Network, coauthorship, fish reproduction scientific collaboration.

Introduction

The communities of scientists have attracted interest, as object of study, once the recognition that the deepening of scientific knowledge depends on the network structures formed by these communities (Verspagen and Werner, 2004), at the same time when large databases with researcher's data became available over the internet. Through their work, their interaction, researchers form social networks by establishing strong ties between them, which are the links between advisors and advisees, or co-authorships in the publication of works resulting from the scientific research work (Granovetter, 1973).

The publication in co-authorship has been used as an indicator of cooperation between researchers (Lima, 2011), but it may be misleading because "not all collaboration results in an article and co-authorship does not always indicate collaboration" (Vanz and Stumpf, 2010). Having this observation in mind, this work starts from the assumption that academic researchers belong to a specialized field of knowledge establishing a social group whose foundation is, like any other social group, the division of labor and cooperation. Their individual members socialize themselves by means of the formation of network structures. In sociology, the theoretical foundations of this view can be found in the pioneering works of Georg Simmel (1858 - 1918).

In this study, the object of interest was the community of scientists formed by Brazilians researchers who were granted a scholarship called productivity in research by the National Research Council (CNPq), having as object of interest the genetics and reproduction of freshwater fish. The starting point of this work is the assumption that when these researchers publish the results of their work in scientific articles written in co-authorship, they give materiality to the existence of a social network formed by researchers interested in genetics and reproduction of fish. This work assumes the hypothesis that, being so, on this network it is possible to apply some of the tools of Social Network Analysis, to understand some of its characteristics, measure its main features and draw conclusions about the use of scientific periodicals as well as evaluate the level of cooperation among researchers.

Objectives

To find the linkage between the CNPq holders of research productivity scholarships in freshwater fish genetics and reproduction by means of their works published as co-authors. To take this linkage as shaping a Social Network and analyze it using graphs to show and find its characteristics and metrics.

---

[3]Corresponding author: carlos.viegas@usp.br
Phone: +55(19)3565-6708

Materials and Methods

The publicly accessible records existent in the Lattes Platform (Curriculum Lattes - CV Lattes) of the National Research Council (CNPq) database (Plataforma Lattes, 2017) were used in this analysis. The CNPq or the authors of this work are not liable for any flaws or inconsistencies in the data collected, since, according to the provisions of the CNPq website, the researchers are full responsible for the data quality fed into the Lattes Platform database. Initially, a search for the Lattes CVs was done on the CNPq website, in the Lattes Platform, filtering the results to identify the *curricula* of doctors who are CNPq productivity scholarship holders, at all levels this scholarship is granted. The search was done by subject, using the keyword "fish reproduction" and resulted in 302 CVs Lattes, which were examined one by one, manually, since many of the CVs selected by the website automatic search are incongruent with the goal of the inquiry, using the mentioned keyword. The manual selection of the initial sampling resulted in the selection of 33 CVs Lattes of researchers whose work is related to genetics and/or reproduction of freshwater fish in Brazil. This selection returned their respective identifiers (id lattes), a code that allows the public access to the academic data of each researcher registered at this CNPq database.

In the next step a list composed by the id Lattes of the 33 selected researchers was processed by means of the software scriptLattes (Mena-Chalco and César, 2009), which performs the semiautomatic extraction of the Lattes CVs supplied by the list, analyzes them and searches the articles they published as coauthors, according to what each researcher has fed into the Platform Lattes database. Thus, it was possible to determine the articles published by the 33 researchers as coauthors, as well as to identify the periodicals used by them for the publishing, and the frequency each researcher published articles by periodical. On the collected data some tools and techniques of Social Network Analysis were applied, and metrics regarding coauthors were determined, as well as their representation in the form of social network graphs consisting of entities that are connected, or related to each other in some way; and is worth to note that the mathematical theory (Feofiloff *et al.*, 2016) underlying this network analysis will not be addressed here. For the analysis of the academic collaboration network on genetics and fish reproduction, as well as to generate metrics and graphs of this network, it was used the Gephi software (Gephi Consortium, 2017), which is dedicated to the analysis of social networks. This list of names, processed using the ScriptLattes software, resulted in a data set that allowed the production of a graph with the Gephi applicative depicting the collaborative network among researchers as co-authors. The results obtained are shown below.

Results

The extraction of the information contained in the 33 CVs Lattes of the researchers taken for this study resulted 1156 articles published in 502 different journals.

The manual selection of researchers, registered in the Lattes Platform, after filtered by the keyword "fish reproduction", nominated the following researchers:

List of CNPq´s Researchers on Fish Genetics and Reprodction.

Alex Pires de Oliveira Nuñer
Diogo Teruo Hashimoto
Elizabeth Romagosa
Elizete Rizzo
Evoy Zaniboni Filho
Fábio Porto-Foresti
Fausto Foresti
Jayme Aparecido Povh
João Batista Kochenborger Fernandes
Leandro Cesar de Godoy
Luís André Nassr de Sampaio
Luis David Solis Murgas
Luis Fernando Fernandes Marins
Luiz Renato de França
Maria Ines Borella
Maria Iracilda da Cunha Sampaio
Nelson Ferreira Fontoura
Nilo Bazzoli
Renata Guimarães Moreira Whitton
Ricardo Pereira Ribeiro
Ricardo Vieira Rodrigues
Rilke Tadeu Fonseca de Freitas
Roberto Ferreira Artoni
Robie Allan Bombardelli
Robie Allan Bombardelli
Rodrigo Augusto Torres
Ronald Kennedy Luz
Ronaldo Oliveira Cavalli
Rosicleire Veríssimo Silveira
Sathyabama Chellappa
Sergio Ricardo Batlouni
Silvio Ricardo Maurano Peixoto
Tarcízio Antônio Rêgo de Paula
Wilson Francisco Britto Wasielesky Junior

The graph of Figure 1 is a representation of how the Brazilian researchers, who are granted a productivity research scholarship by CNPq, in the field of genetics and reproduction of freshwater fish, collaborate with each other through the coauthoring of papers in scientific journals. In this representation, each point is a researcher and each line represents a co-authoring interaction. The size of the letter type used in the researcher's name is proportional to the amount of researcher's co-author role with the other colleagues (their degree) and the thickness of the lines linking one researcher to another is proportional to the quantity of works published in co-authorship between these researchers. Due to their degree of collaboration, measured by the number of works published with others, each researcher has a different impact on the set of collaborations of the social network of coauthors, which is measured by their collaboration rank (Page *et al.,* 1999). Below is a table (Tab. 1), produced by the applicative ScriptLattes, which reports this metric for each researcher involved.
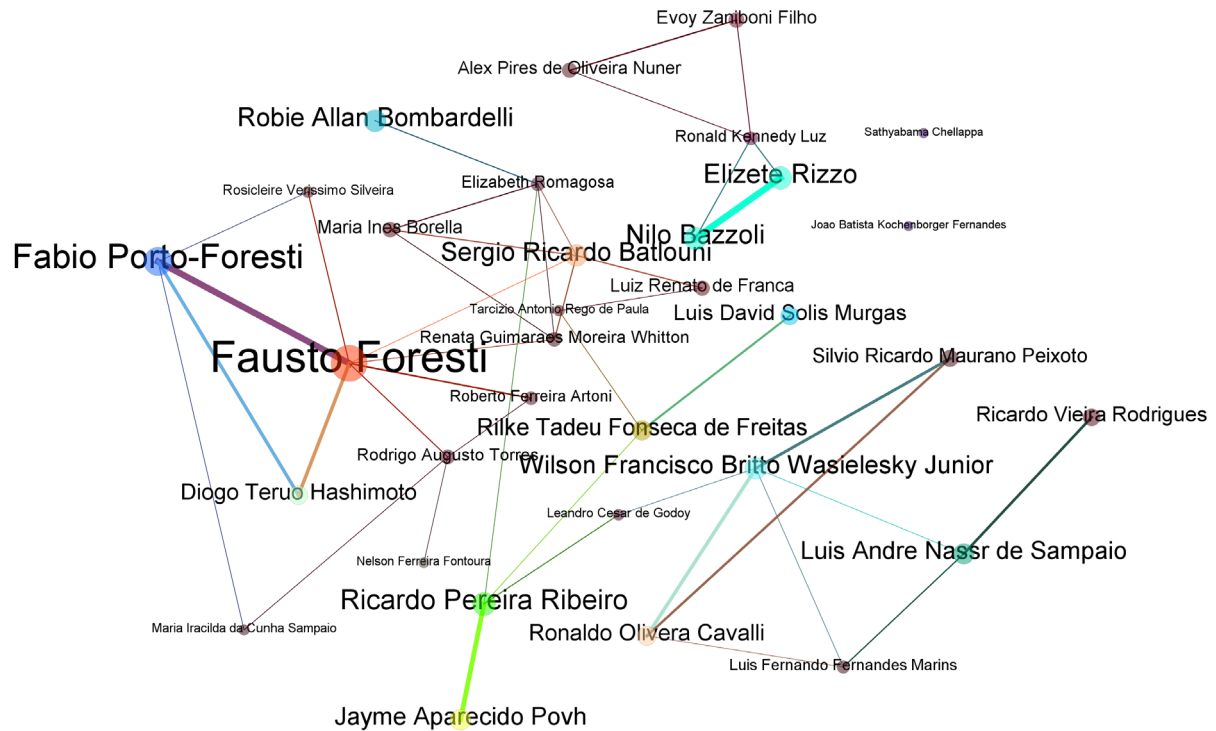
Figure 1. Graph of the collaboration network in publishing academic papers among co-authors researchers.

Table 1. Researchers classified in decreasing order by collaboration rank.

Names	Collaboration Rank
Fausto Foresti	3.0
Fábio Porto-Foresti	2.16
Ricardo Pereira Ribeiro	1.56
Nilo Bazzoli	1.51
Elizete Rizzo	1.51
Robie Allan Bombardelli	1.45
Robie Allan Bombardelli	1.45
Sergio Ricardo Batlouni	1.41
Wilson Francisco Britto Wasielesky Junior	1.37
Jayme Aparecido Povh	1.31
Luís André Nassr de Sampaio	1.3
Rilke Tadeu Fonseca de Freitas	1.19
Ronaldo Olivera Cavalli	1.13
Diogo Teruo Hashimoto	1.11
Luis David Solis Murgas	1.05
Ricardo Vieira Rodrigues	0.95
Silvio Ricardo Maurano Peixoto	0.91
Evoy Zaniboni Filho	0.74
Maria Ines Borella	0.72
Luiz Renato de França	0.71
Alex Pires de Oliveira Nuñer	0.69
Rodrigo Augusto Torres	0.63
Renata Guimarães Moreira Whitton	0.63
Elizabeth Romagosa	0.58
Ronald Kennedy Luz	0.55
Roberto Ferreira Artoni	0.53
Luis Fernando Fernandes Marins	0.48
Rosicleire Veríssimo Silveira	0.33
Leandro Cesar de Godoy	0.29
Tarcízio Antônio Rêgo de Paula	0.28
Maria Iracilda da Cunha Sampaio	0.25
Nelson Ferreira Fontoura	0.22
Sathyabama Chellappa	0.15
João Batista Kochenborger Fernandes	0.15

This classification of the researchers allowed to select the top ten Brazilian researchers with the highest degree of collaboration as coauthors, in the publication of works related to genetics and reproduction of fish, among the other 33 listed and to determine the academic journals where those publications were made. Processing the data with the Gephi platform, it was possible to obtain the graph shown below (Fig. 2).
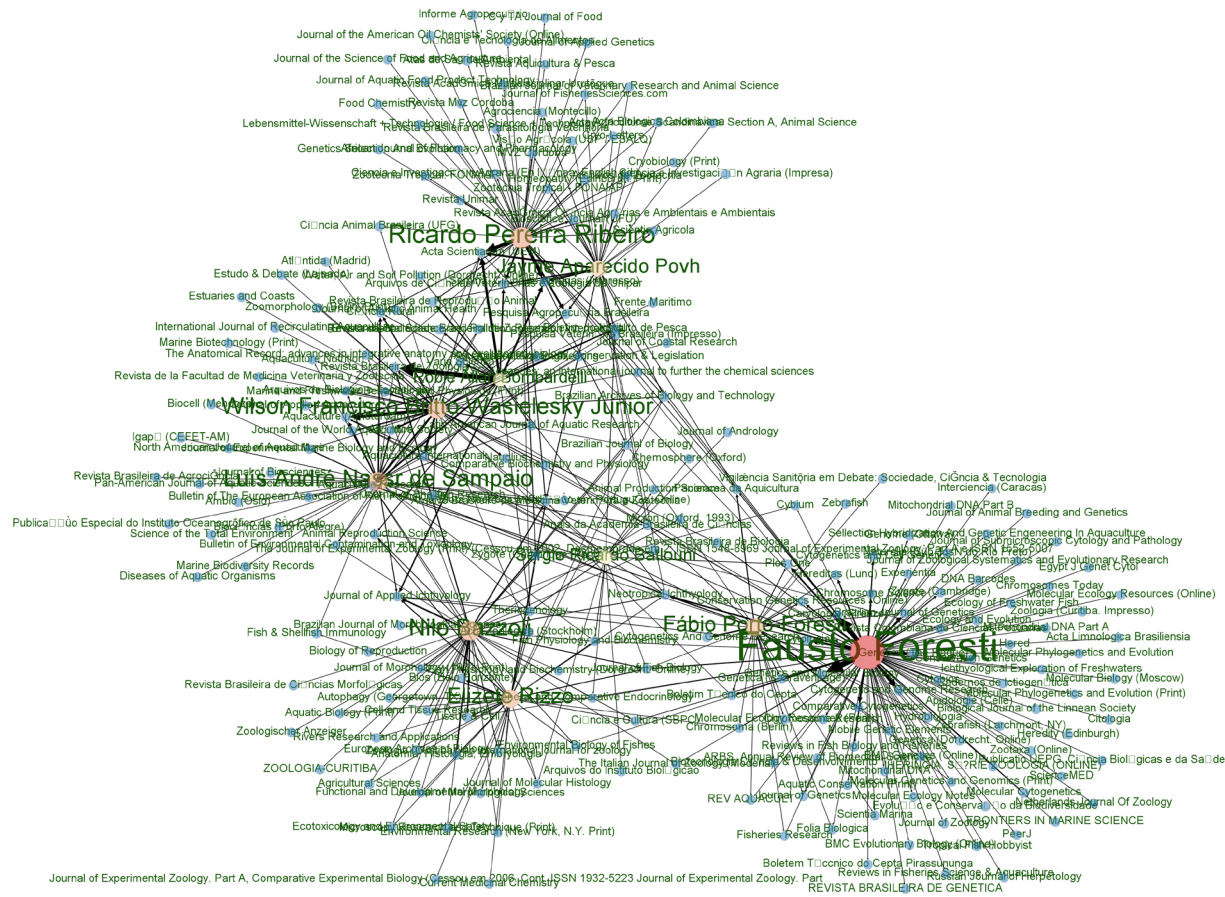

Figure 2. Graph of scientific network formed by journals and most collaborative authors.

In this graph the green lines represent a link between a researcher and a journal, and the size of the font type used in the researcher's name is proportional to the number of publications made by that researcher. The green dots represent different journals and the line thickness is proportional to the number of articles the researcher published in that periodical. The next graph (Fig. 3) represents the same relationship in an inverted way and describes which journals were the ones that received the largest amount of article publications by the ten researchers with the greatest degree of collaboration among them. The font type size used in the names of journals is proportional to the number of articles published in each of them and the thickness of the lines is proportional to the number of articles published per researcher.

The Table 2 shown below lists forty journals used by the ten researchers with the highest degree of collaboration as coauthors in genetics and reproduction of freshwater fish, for publishing theirs works. These journals belong to the universe of 502 journals listed. The entrance degree (indegree) measures the number of articles published by these researchers in the periodical in question and the Page Rank means is an index to measure the number of times the periodical was used, relative to all other periodicals.

Using the Gephi software, another analysis was applied on the collaboration network formed by the group of 33 researchers, this time with the purpose of analyzing the general characteristics of this network (Scott, 2011). The results are shown in Figure 4 and in Table 3. Figure 4 represents the degree of centrality of intermediation between the members of the network (betweenness centrality), which is a metric that shows the extent into a member of a social network is connected to the other components of the network, acting as a "bridge" between them.

The data set, collected from the Lattes Platform, to locate the co-authorship in academic articles produced by researchers granted by CNPq with a scholarship, through the keyword "fish reproduction", brought some biases using the search tools available on the CNPq website. One bias, the most serious of the data extraction, was to list many names of researchers who do not work with animal reproduction at all, or who work with reproduction of other animals but not with the reproduction of fish properly. Another bias resulted in the collection of names of researchers who work with genetics of fish, but not properly with their reproduction. These biases had to be adjusted manually and, therefore, it is more convenient to analyze the results obtained in terms of coauthors in genetics and fish reproduction.
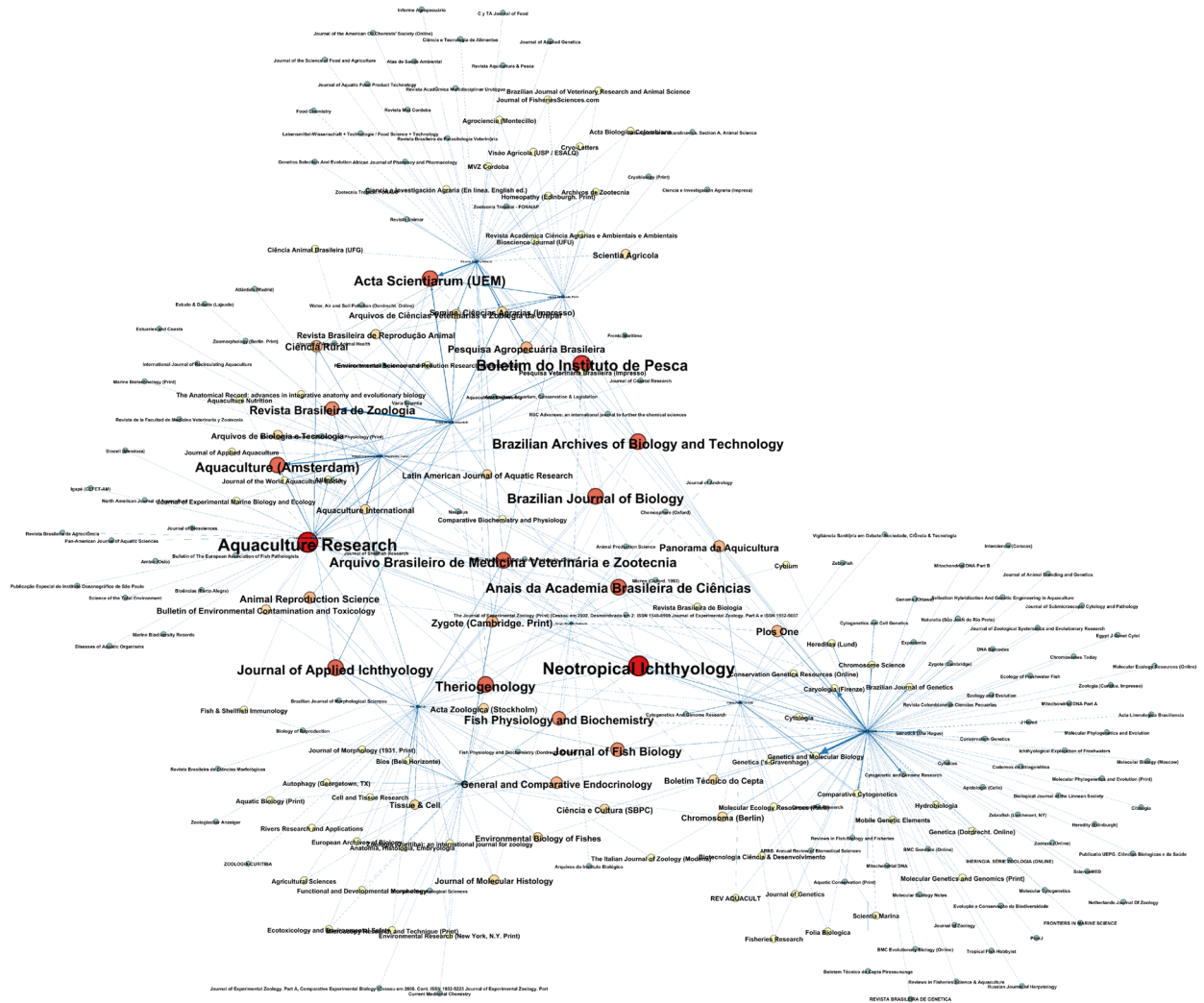
Figure 3. Graph of scientific network formed by authors and the most used journals.

The co-authored production of scholarly articles in the subject analyzed has an expressive concentration on just a single researcher, who participate in about 30% of the total published articles, which perhaps can be explained for this researcher being a senior one with extensive academic life and his CV Lattes reports the bibliographic production throughout a whole academic career. On the other hand, it is also significant the dispersion in terms of the number of journals used by researchers for the publication of their works, totaling 502 journals with the most diverse characteristics. One way to evaluate this dispersion is to check the number of articles published per periodical. Taking in decreasing order the first three journals most searched after by researchers, the figure is 24 articles for the most wanted, 21 for the second place and 19 articles for the third one. However, the last 311 less-wanted journals received only one article for publication, meaning that in twenty journals there was a concentration of 25% of the total articles published by these researchers. It is also worth noting that among the forty most wanted journals a presence of Brazilian journals alongside with non-Brazilian journals, which may suggest that several Brazilian journals are vehicles that have a good reputation among researchers in the field.

Table 3 (above) shows the positioning metrics of the researchers, taken the general set of the collaboration network, that is, this measure considers the importance of each researcher in the collaboration network as a whole. From this table it is possible to notice that although the third author (in the order) has a bibliographic production more than four times superior to the first one, the first in the rank has a greater degree of intermediation between the coauthors. The PageRank index roughly measures the number of times a page (e.g.: the number of times a website page is visited) is accessed; and for this analysis, this index is correlated to the bibliographic production (the quantity of articles published), which does not capture the impact of the researcher in terms of his/her academic collaboration for the whole network of researchers. In order to examine this feature, it is necessary to determine its centrality of intermediation, which, in the case of the network examined, represents the position of one researcher as an intermediary among other researchers, that is, even if researcher A does not have an article published as co-author with researcher C, it may exist author B who is a co-author with A and C, acting as a bridge between them. This is an important metric for locating

agents who play a key role in any social network and in the case examined, shows the network of co-authors with a different dimension than the rough number of paired collaborations between single members. This measurement is illustrated by the graph of Figure 4, produced from the data of Table 3, which highlights other researchers than those represented in the graph of Figure 3.

Table 2. List of forty journals used by the ten researchers with the highest degree of collaboration as coauthors in genetics and reproduction of freshwater fish.

#	Journal	indegree	degree	pageranks
1	Neotropical Ichthyology	24	24	0.003373
2	Aquaculture Research	21	21	0.003267
3	Arquivo Brasileiro de Medicina Veterinária e Zootecnia	19	19	0.003016
4	Journal of Applied Ichthyology	19	19	0.002876
5	Boletim do Instituto de Pesca	18	18	0.003586
6	Revista Brasileira de Zoologia	17	17	0.003597
7	Acta Scientiarum (UEM)	16	16	0.003903
8	Aquaculture (Amsterdam)	16	16	0.003382
9	Brazilian Archives of Biology and Technology	15	15	0.002575
10	Theriogenology	15	15	0.002351
11	Journal of Fish Biology	14	14	0.002454
12	Brazilian Journal of Biology	14	14	0.002265
13	Fish Physiology and Biochemistry	13	13	0.002437
14	Anais da Academia Brasileira de Ciências	13	13	0.002216
15	Ciência Rural	13	13	0.002524
16	Pesquisa Agropecuária Brasileira	12	12	0.002573
17	Plos One	12	12	0.00236
18	Animal Reproduction Science	12	12	0.002352
19	Revista Brasileira de Reprodução Animal	11	11	0.00236
20	Journal of the World Aquaculture Society	10	10	0.002301
21	Boletim Técnico do Cepta	9	9	0.002153
22	Aquaculture International	9	9	0.002619
23	Genetics and Molecular Biology	8	8	0.002776
24	Tissue & Cell	8	8	0.002185
25	Zygote (Cambridge. Print)	8	8	0.00245
26	General and Comparative Endocrinology	8	8	0.002244
27	Hydrobiologia	7	7	0.001976
28	Revista Brasileira de Biologia	7	7	0.00214
29	Scientia Agricola	7	7	0.001916
30	Ciência Animal Brasileira (UFG)	7	7	0.001961
31	Atlântica	7	7	0.002331
32	Conservation Genetics Resources (Online)	6	6	0.002137
33	Environmental Biology of Fishes	6	6	0.002022
34	Comparative Biochemistry and Physiology	6	6	0.002082
35	Caryologia (Firenze)	5	5	0.002138
36	Comparative Cytogenetics	5	5	0.001976
37	Journal of Genetics	5	5	0.001863
38	Panorama da Aquicultura	5	5	0.002064
39	Zebrafish (Larchmont, NY)	5	5	0.001945
40	Ciência e Cultura (SBPC)	5	5	0.001918

Finally, the collaborative academic network examined has a density equal to 8.1% as measured by Gephi software. To understand this metric, it is enough to know that a network in which all its components are linked to each other, has a density equal to 100% and, therefore, the network under analysis has a low density, meaning that the collaboration between these researchers in the publication of academic articles is less than their potential for collaboration. The meaning of this metric is reinforced by two other measurements: the diameter of this network is 9 and its average length is 3.8 (measurements obtained by the Gephi). The diameter of a network is the maximum distance that separates two of its components, measured by the number of network members between them. As a network (social or otherwise) may have a diameter of hundreds (or more) of components, the network in question is not large. The average length is the metric that measures the average minimum distance for one member to reach another member (necessarily, it is smaller than its diameter) and the figure reinforces the small size of the network and emphasizes its low density.
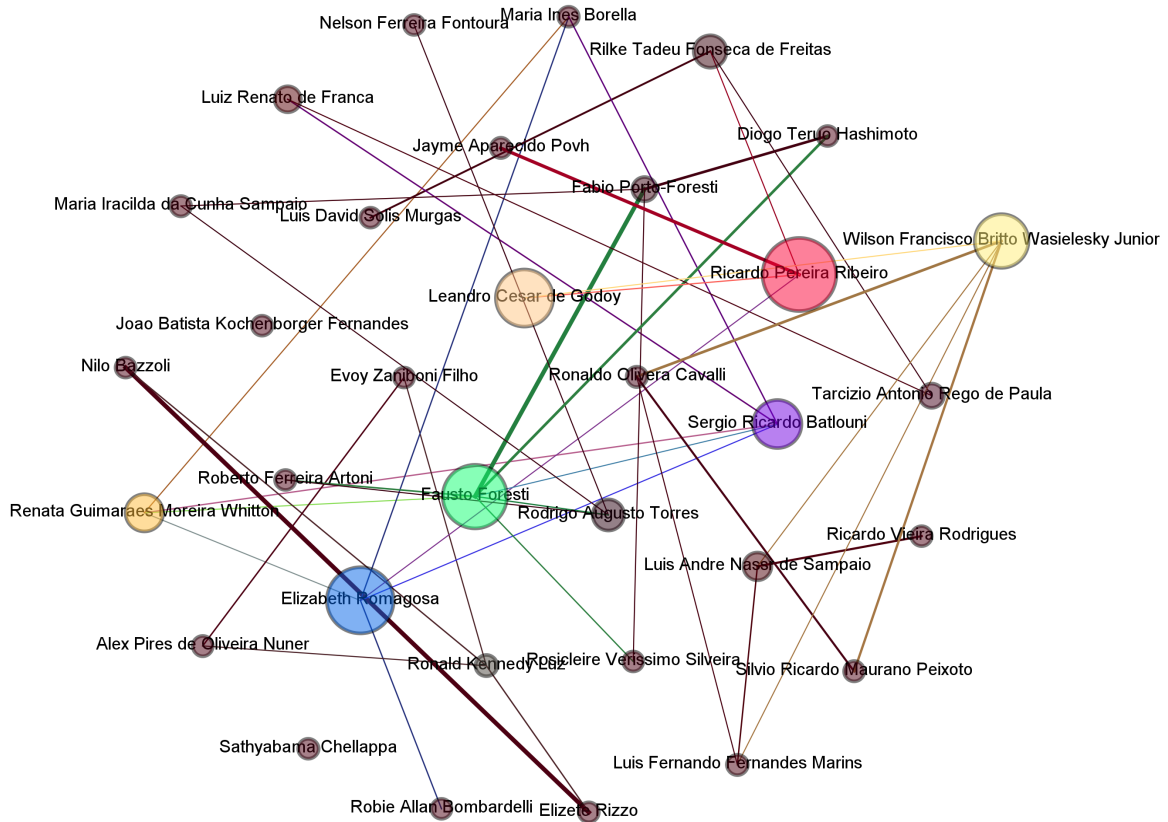
1250

Anim. Reprod., v.14, (Suppl.1), p.1245-152. 2017

Figure 4. Graph of researchers betweenness centrality.

Table 3. Researchers ordered by the centrality of intermediation (betweenness centrality).

Id	Label	Bibliografic Production	Articles in Journals	Betweeness Centrality	Page Ranks
13	Fausto Foresti	1171	334	271169	96597
16	Ricardo Pereira Ribeiro	268	153	329973	51543
8	Sergio Ricardo Batlouni	105	30	173723	50348
2	Elizete Rizzo	386	86	0	48594
3	Nilo Bazzoli	411	107	0	48594
27	Diogo Teruo Hashimoto	181	38	0	46829
24	Wilson Francisco Britto Wasielesky Junior	623	146	208669	43747
21	Jayme Aparecido Povh	215	76	0	43097
18	Luis Andre Nassr de Sampaio	381	93	48387	39722
22	Ronaldo Olivera Cavalli	326	78	1008	37997
15	Rilke Tadeu Fonseca de Freitas	422	213	77621	37968
5	Elizabeth Romagosa	173	83	288642	34755
25	Silvio Ricardo Maurano Peixoto	270	69	0	32721
30	Maria Ines Borella	108	37	0	30989
31	Ricardo Vieira Rodrigues	126	26	0	29219
1	Renata Guimaraes Moreira Whitton	214	56	107527	23367
23	Luiz Renato de Franca	338	129	36962	22212
4	Evoy Zaniboni Filho	401	91	0	21657
6	Ronald Kennedy Luz	413	74	8065	20653
17	Alex Pires de Oliveira Nuner	197	34	0	20246
26	Robie Allan Bombardelli	355	70	0	15344
29	Luis Fernando Fernandes Marins	202	80	2016	14233
10	Roberto Ferreira Artoni	243	134	0	14154
20	Rosicleire Verissimo Silveira	216	22	0	9175
28	Leandro Cesar de Godoy	106	24	229839	9093
14	Tarcizio Antonio Rego de Paula	204	90	25874	8945
12	Maria Iracilda da Cunha Sampaio	317	159	2016	7441
9	Fabio Porto-Foresti	423	79	24194	7246
11	Nelson Ferreira Fontoura	99	58	0	6757
0	Joao Batista Kochenborger Fernandes	193	83	0	4792
19	Sathyabama Chellappa	447	130	0	4792
7	Luis David Solis Murgas	481	135	0	3347
33	Rodrigo Augusto Torres	121	44	69556	1849

In conclusions, the analysis of the social network, formed by the Brazilian researchers in genetics and fish reproduction, granted with a scholarship in research productivity by CNPq, through their collaboration in coauthoring the publication of academic papers, reveals that there are some opportunities for improvement in terms of their collaboration. Initially, it is possible to say that the low density of this network (8.1%) suggests that it is possible to increase the collaboration and measure it through the works published by coauthors. Although the authors' survey found 33 names, there is a strong concentration of authorship in a few names. An increase in the academic partnerships would lead to an increase in the density of the network, meaning a more cohesive social group, and perhaps a more integrated scientific community. On the other hand, there is a significant dispersion in the application of papers to be published by the academic journals; as a result, three hundred different journals have received only one application for publication. Among those most in demand, there are some Brazilian journals, which suggests that an increase in the applications of academic papers in these journals could lead to an improvement in their international ranking and to strengthen the collaboration between authors.

Acknowledgements

References

Feofiloff P, Kohayakawa Y, Wakabayashi Y. 2016. Uma Teoria Sucinta à Teoria dos Grafos. Available in http://www.ime.usp.br/~pf/teoriadosgrafos. Accessed on 13-jan-2016.

Gephi Consortium. 2017. Computer Program. Available from: https://www. gephi.org. Accessed on 08-11-2017.

Granovetter M. 1973. The Strength of Weak Ties. American Journal of Sociology. 78(6):1360-1380.

Lima MY. 2011. Coautoria na produção científica do PPGGeo/UFRGS: uma análise de redes sociais. Ciência da Informação**,** 40(1):38-51.

Mena-Chalco JP and Cesar-Jr RM. 2009. scriptLattes: An open-source knowledge extraction system from the Lattes platform. Journal of the Brazilian Computer Society, 15(4):31-39

Page L, Brin S, Motwani R, Winograd T. 1999. *The PageRank Citation Ranking: Bringing Order to the Web.* Stanford, CA: Stanford InfoLab. 17 pp. (Technical Report, 1999-66).

Plataforma Lattes [database on internet]. 2017. Available on: www.lattes.cnpq.br. Accessed on: November, 8,2017.

Scott J. 2011. Social Network Analysis. SAGE Publications Ltd., London.

Vanz SAS, Stumpf IRC. 2010. Colaboração científica: revisão teórico-conceitual. Perspectivas em Ciência da Informação, 15(2):42-55.

Verspagen B, Werker C. 2004. Keith Pavitt and the Invisible College of the Economics of Technology and Innovation. Research Policy**,** 33:1419-1431

1252

Anim. Reprod., v.14, (Suppl.1), p.1245-1252. 2017