Innovation Networks and Gatekeepers of Canadian Biotechnology Clusters [±]

Andrea Schiffauerova and Catherine Beaudry[*]

École Polytechnique de Montréal

Abstract

This paper studies the importance of collaboration networks for the innovation creation in Canadian biotechnology. Using information contained in 3550 biotechnology patents of 4569 inventors we construct Canadian biotechnology innovation network and describe the collaborative behaviour of the inventors. We find that most of the collaborative activity which involves Canadian inventors takes place inside biotechnology clusters. We examine the structural properties of the local collaboration subnetwork of each cluster and relate them to the efficiency in knowledge diffusion and innovation creation in that cluster. We then investigate the network architecture of inter-cluster collaborations by examining the cooperation ties among inventors who are directly or indirectly interconnected in network components. We find that the cluster-based and component-based collaboration spaces overlap to a certain extent, but differ in their structure. Gatekeepers are Canadian inventors who by bridging over the two spaces enable the nurturing of clusters with fresh knowledge originating outside. We propose a number of indicators measuring an inventor's importance as a procurer of external knowledge and find that usually only around 10%-20% of all inventors are responsible for the inflow of the external knowledge to the cluster.

Keywords: innovation, collaboration, knowledge networks, network structure, knowledge transmission, gatekeeper, intermediary, patents, biotechnology, clusters, Canada

---

[±] Beaudry acknowledges financial support of the Social Science and Humanities Research Council of Canada. We acknowledge helpful comments from Jorge Niosi and Nathalie de Marcellis-Warin as well as from the InnovaRisQ research team. Ahmad Barirani provided research assistance. None of these, however, are responsible for any remaining errors.

[*] **Corresponding author**: Département de mathématiques et de génie industriel, École Polytechnique de Montréal, C.P. 6079, succursale Centre-ville, Montréal (Québec) Canada, H3C 3A7, catherine.beaudry@polymtl.ca

1. Introduction

As an alternative to the three classical locations where an innovation takes place (which are non-profit institutions, profit-seeking firms and the minds of individual inventors), Allen (1983) introduced the concept of collective invention. The key to understanding the phenomenon of collective invention is in the exchange and free circulation of knowledge and information within groups of socially connected (but often competing) agents rather than in the inventive efforts of particular firms or individuals. Open sharing of information thus results in fast knowledge accumulation and high invention rates[1]. A concept of collective invention is convenient for describing the dynamics of knowledge diffusion through various innovation networks. The network of innovators is an inter-personal network of individual innovators, who collaborate and exchange information in order to produce innovations and scientific knowledge. Social network analysis[2] can be used to analyze the way these innovators are interconnected. Patent documents and bibliometric data can thus be exploited to map the complex web of social ties among innovators and to construct innovation networks.

A number of studies have used the network approach to examine knowledge flows and their implications. Newman constructed networks of collaboration based on databases of scientific papers (2001a, 2001b, 2001c, 2001d). Breschi and Lissoni (2003, 2004), Balconi *et al.* (2004), Singh (2005), Fleming *et al.* (2006), Gauvin (1995), Mariani (2000) or Ejermo and Karlsson (2006) adopted the co-inventorship of patents as an appropriate device to derive maps of social relationships between inventors and to build their networks. Another line of research related to innovation networks involves theoretical simulation studies. Network models which simulate knowledge diffusion through the network were built by Cowan and Jonard (2003, 2004), Morone and Taylor (2004) and Cowan *et al.* (2007).

---

[1] See examples of collective invention in Saxenian (1994), Dahl and Pedersen (2004), Schrader (1991), von Hippel (1987) or Lamoreaux and Sokoloff (1997).

[2] Social network analysis is the mapping and measuring of relationships and flows between people, groups, organizations, computers or other information/knowledge processing entities. The nodes in the network are the people or groups, whereas the links show the relationships or flows between the nodes. Social network analysis provides both a visual and a mathematical analysis of complex human systems (Krebs, 2006).

Some researchers adopted the network approach and included geographical aspects into their models. Gittelman (2006) argued that the geography of research collaborations has distinct impacts on firms' scientific contribution and their inventive productivity. Beaucage and Beaudry (2006) and Schiffauerova and Beaudry (2008b) compared the major Canadian biotechnology clusters in terms of their innovation network structures and of the distinct geographical collaborative patterns on knowledge generation and innovation production.

This paper aims at understanding the role of collaboration networks in the creation of innovation in Canadian biotechnology clusters. Our research will examine the diffusion of knowledge through the network of Canadian biotechnology inventors constructed from the patent co-inventorship data. The construction of the network will allow us to derive the collaborative behaviour of the inventors within and between clusters. The special focus is on the network architecture, its role in the knowledge generation and thus in the growth of high technology clusters in Canada. Finally, the points of interaction between the intra-cluster and inter-cluster collaboration spaces are explored. The paper is organised as follows: section 2 introduces the methodology and data used in this study, section 3 presents the results of the network analyses related to the collaboration within and between clusters, section 4 explores the points of interaction between the cluster-based and component-based collaboration spaces and finally section 6 concludes.

2. Methodology

The patent database used for the empirical analysis is the United States Patents and Trademarks Office (USPTO) database. This is the only patent database which provides the geographical location of the residence for each inventor (unlike the Canadian Intellectual Property Office database (CIPO) or the European Patent Office (EPO)). The use of the USPTO database instead of the CIPO may introduce a bias in the data, but we consider it minimal, since Canadian inventors usually patent both in Canada and in the US. For example, in 1998 and 1999 out of all the patent applications submitted by Canadian biotechnology firms worldwide, the majority (36%) was delivered to the USPTO, followed by 28% to the CIPO, 21% to the EPO and the balance of 16% to other offices[3] (Statistics Canada, 2001). The population of Canada is relatively small and as a consequence, building a viable industry based on domestic sales alone may prove difficult. In addition, because of the long development cycles for biotechnology products (typically 10 years for a single product), access to large markets is needed to ensure an adequate return on investment (Strachan, 1995). As a result, Canadian biotechnology companies prefer to protect their intellectual property in the USA. The much larger and easily accessible American biotechnology market offers great potential to Canadian biotechnology firms. An analysis of the Canadian patents registered at the USPTO should hence provide a realistic picture of Canadian biotechnology innovation.

Biotechnology encompasses several different research technologies and several fields of application. A Statistics Canada study (Rose, 2000) has shown that different interpretations of the meaning of biotechnology can result in differences in the results of biotechnology surveys. One of the initial tasks for us was therefore to select a clear and practical definition of biotechnology. We have opted to base our USPTO search strategy on the OECD definition of biotechnology, which is based on the group of carefully selected International Patent Codes (IPC)[4]. The OECD has carried out an extensive consultation (including the work conducted by Statistics Canada, which shares similar definitions of biotechnology (Munn-Venn and Mitchell, 2005), to develop the definitions of biotechnology techniques, and the validation showed that the definition appears to capture a significant proportion of biotechnology patents. It might not be complete and may include some patents with non-biotechnology techniques, however, errors are likely to be small (OECD, 2005).

---

[3] Note that firms may have submitted patents regarding the same invention to a number of patent offices at the same time. For instance, some triadic patents are registered at the USPTO, CIPO and EPO.

[4] The OECD definition of biotechnology patents covers the following IPC classes: A01H1/00, A01H4/00, A61K38/00, A61K39/00, A61K48/00, C02F3/34, C07G(11/00, 13/00, 15/00), C07K(4/00, 14/00, 16/00, 17/00, 19/00), C12M, C12N, C12P, C12Q, C12S, G01N27/327, G01N33/(53*, 54*, 55*, 57*, 68, 74, 76, 78, 88, 92).

An automated extraction program was used to collect the required information[5] from the biotechnology patents. All the biotechnology patents registered before March 31, 2007 were included. According to the above definition, there are around 100 000 biotechnology patents registered at the USPTO. We have thus created a patent database, which contains all the patents in which at least one inventor resides in Canada and which comprises 3550 patents. The total numbers of Canadian patents registered at the USPTO each year found by the aforementioned search strategy largely correspond with the other authors' data: for example with the results of Statistics Canada (2001) or with the study of Rasmussen (2004). We have nevertheless noticed substantial differences in the findings of other researchers. These were usually caused by the choice of different search strategies such as keywords in the patents' names and abstracts (as in Niosi and Bas, 2001) or by the decision to use a rather narrow biotechnology definition (as in Beaucage and Beaudry, 2006).

This extraction allowed the construction of a database of inventors, which includes all biotechnology inventors who are Canadian residents or at least once patented with a Canadian resident in biotechnology (and registered their patent at the USPTO). Our work involves a quantitative analysis of the information extracted from the patents. Furthermore, biotechnology innovation networks were created by mapping the collaborations of inventors of particular biotechnology patents. We used the concept of social network analysis defined above to create connections between inventors and to construct innovation networks. The social network analysis program PAJEK was used to build the networks from the extracted patent information. An analysis of these collaborative networks enabled us to describe their structural properties and to understand the collaborative behaviour of the inventors inside or outside Canadian biotechnology clusters.

Since the obtained patent data span over a period of 30 years, we have assumed that once inventors collaborate on one patent they continue to be in contact afterwards and are

---

[5] Extracted information necessary for the research leading to this paper includes the patent number and the inventors' names and their addresses.

able to exchange knowledge acquired long after the patent has been granted. This allows us to disregard the time of collaboration and consider all links among inventors in the network as simultaneously active.

3. Canadian biotechnology collaboration network

The network of Canadian biotechnology inventors which we created includes 4569 vertices (representing inventors) and 9731 edges (representing collaborative relations[6]). Based on the residences of inventors we have identified 12 Canadian biotechnology clusters[7]. 20% of inventors reside in the Toronto cluster, 15% in the Montreal cluster and 9% in the Vancouver cluster. Only a very small portion of inventors (less than 1%) residing in Canada lives outside the defined clusters and around 31% of the innovators in our sample reside outside the Canadian borders.[8] The following sections explore the structures of the collaboration networks within and between clusters.

3.1 Intra-cluster collaborations

It has been suggested and empirically supported that firms in clusters are more innovative (Baptista and Swann, 1998; Beaudry, 2001; Beaudry and Breschi, 2003; Beaudry and Swann, forthcoming). The companies collocated in a close geographical proximity enjoy numerous benefits[9], among which the most discussed in the context of biotechnology innovation are knowledge spillovers[10]. Biotechnology knowledge is largely tacit, which limits knowledge diffusion over long distances. In fact, the

---

[6] Each collaborative relation (also called a tie or a link) represents a connection between a pair of inventors, which involves one or more instances of co-invention of a biotechnology patent.

[7] A cluster is defined in this study as a geographically continuous region active in biotechnology (as measured by the patent production).

[8] For the detailed analysis of the inventors and other characteristics of the Canadian biotechnology clusters see Schiffauerova and Beaudry (2008).

[9] Supply-side benefits were defined by Krugman (1991) as labour market pooling, availability of immediate inputs and knowledge spillovers, while Baptista and Swann (1998) who surveyed demand-side benefits indicated that the major ones are strong local demand, market share gain, decreased search costs, and exploitation of local information flows.

[10] Localized knowledge spillovers are defined as knowledge externalities bounded in space that allow companies operating nearby key knowledge sources to introduce innovations at a faster rate than rival firms located elsewhere (Breschi and Lissoni, 2001).

transmission of tacit information and knowledge spillovers is usually associated with face-to-face contact. Collaboration among the inventors working in clusters is thus encouraged by the benefits of acquiring knowledge which the subjects located within short geographical distance spill over. Indeed we observe that 60% of collaboration activities in Canadian biotechnology is carried out within clusters.

This section of the paper analyzes these local collaborations carried out entirely within clusters. We have divided the Canadian biotechnology innovation network into geographically based subnetworks, where each subnetwork strictly includes inventors who reside in one particular cluster, while excluding the ones that do not. Out of cluster inventors are therefore eliminated for the time being. For each of the subnetworks created in this manner several network characteristics were calculated. Table 1 presents some of these properties. The remaining part briefly discusses several of the basic structural properties of the network and explains the indicators used in this paper to measure them. We show how these characteristics could be related to efficiency in the knowledge diffusion among the inventors within the clusters and suggest the possible impact on innovation creation in the cluster.

Collaboration characteristics in the subnetworks

As Table 1 shows, 18-50% of collaborative relations between pairs of inventors involve repetitive instances of collaboration. Inventors in Toronto and Calgary tend to pursue collaborative relations with the same partners more often than inventors in Montreal, Vancouver or Edmonton. In Halifax, half of the collaborative ties of the local inventors include repetitive collaborative relationships. The strongest collaboration link in the network, i.e. the most frequently repeated collaborative relation, concerns two inventors in Toronto. They repeated their collaboration 60 times, i.e. their joint research resulted in 60 patents. In smaller clusters, the maximum number of repeated collaborations is lower. It is still relatively low for the clusters of Montreal (11) and Vancouver (10), where on average, innovative activities involve slightly more co-inventors who collaborate with each other less often.

Table 1: Structural properties of the cluster-based subnetworks

Cluster *	**TRT**	**MTL**	**VAN**	**EDM**	**CAL**	**SAS**	**WIN**	**KIN**	**OTT**	**QUE**	**HAL**	**SHE**	**Network**
Number of inventors	927	698	411	210	91	147	77	94	224	127	33	26	4569
COLLABORATION CHARACTERISTICS													
Number of collaborating pairs	1120	1027	568	334	91	259	54	96	343	155	20	10	9731
% of repeated collaborations	43%	36%	37%	37%	41%	28%	19%	33%	36%	18%	50%	20%	36%
Max number of repeated collaborations	60	11	10	14	16	8	3	10	19	7	5	3	60
FRAGMENTATION													
# of components	342	218	134	67	39	34	44	38	70	44	20	18	894
Size of the 1st largest component	98	109	38	49	15	54	7	8	75	11	6	3	579
Size of the 1st largest as % of all	11%	16%	9%	23%	16%	37%	9%	9%	33%	9%	18%	12%	13%
Ratio 2nd/1st largest	0,36	0,31	0,47	0,43	0,60	0,67	0,86	0,75	0,15	0,82	0,67	1,00	0,32
Average component size	2,71	3,20	3,07	3,13	2,33	4,32	1,75	2,47	3,20	2,89	1,65	1,44	5,11
Share of components formed by 50% of inventors	13%	11%	15%	11%	18%	6%	25%	24%	9%	21%	30%	33%	10%
Isolates as % of inventors	19%	15%	16%	17%	24%	13%	36%	18%	17%	15%	39%	46%	4%
STRUCTURAL COHESION													
Subnetwork density	0,003	0,004	0,007	0,015	0,022	0,024	0,018	0,022	0,014	0,019	0,038	0,031	0,001
Average degree	2,42	2,94	2,76	3,18	2,00	3,52	1,40	2,04	3,06	2,44	1,21	0,77	4,26
CENTRALIZATION OF SUBNETWORK													
Degree centralization	0,05	0,02	0,06	0,08	0,11	0,15	0,06	0,04	0,06	0,04	0,13	0,05	0,01
Betweenness centralization	0,008	0,011	0,005	0,019	0,011	0,074	0,002	0,003	0,068	0,003	0,010	0,000	0,009
CENTRALITY OF VERTICES													
Max degree centrality	51	16	27	20	12	25	6	6	16	8	5	2	66
Max closeness centrality	0,05	0,05	0,07	0,13	0,14	0,22	0,09	0,07	0,09	0,07	0,18	0,12	0,03
Max betweenness centrality	0,008	0,011	0,005	0,019	0,011	0,076	0,002	0,003	0,070	0,003	0,010	0,000	0,009
GEODESIC DISTANCES													
Subnetwork diameter	9	11	5	7	5	6	3	3	11	4	2	1	17
Average distance among reachable vertices	3.26	4.27	1.98	2.50	1.75	2.75	1.25	1.31	4.95	1.48	1.23	1.00	6.55
Average distance in the largest component	3.70	4.89	2.30	2.73	1.74	2.72	1.24	1.61	5.22	1.73	1.40	1.00	7.09
Max reach	97	108	37	48	14	53	6	7	74	10	5	2	578
CLIQUISHNESS													
Average egocentric density	0,44	0,56	0,57	0,55	0,29	0,64	0,32	0,47	0,59	0,55	0,24	0,23	0.71

* TRT	…Toronto	MTL	…Montreal	VAN	…Vancouver	EDM	…Edmonton	CAL	…Calgary	SAS	…Saskatoon
WIN	…Winnipeg	KIN	…Kingston	OTT	…Ottawa	QUE	…Quebec	SHE	…Sherbrooke	HAL	…Halifax

Fragmentation of the subnetworks

A component is defined as the maximal connected subnetwork (Wasserman and Faust, 1994). It is a part of the network which includes a maximum number of vertices which are all directly or indirectly connected by links. The *largest component* in absolute value is found in the Montreal cluster, even though Toronto has almost twice as many inventors. In proportion of the cluster sizes, the largest components of Saskatoon, Ottawa and Edmonton, regroup a greater share of the cluster's inventors (37%, 33% and 23 %, respectively). Proportionately more inventors in these clusters are thus directly or indirectly interconnected and easily exchange scientific knowledge. The *second largest components* in Montreal and Toronto are of similar sizes, with that of Vancouver being much smaller. The cluster of Vancouver is in general more fragmented than the other two. In Saskatoon, even the second largest component is composed of proportionately many inventors. In smaller clusters the larger components are in general of comparatively smaller sizes than in the larger clusters and the size of the second largest component is much more closer to the size of the first largest (in Sherbrooke they even are exactly the same size). The larger components in Saskatoon and in Ottawa are of relatively greater sizes than in other clusters: 50% of all the cluster's inventors are part of only around 6% (in Saskatoon) or 9% (in Ottawa) of components.

The *average component size is* fairly small for all the clusters (around 2-3 inventors). As hinted from the previous paragraphs, Saskatoon, which comprises components of a large relative size, scores the highest on the average number of interconnected inventors (4.32 inventors). The second rank is occupied by Montreal and Ottawa (both have on average 3.2 connected inventors), but Toronto has a mean of only 2.71 inventors in a component. Many inventors have collaborators outside their clusters or outside Canada that contribute to linking indirectly inventors from the same cluster. As explained previously, we consider only cooperation based on close personal contacts, which are limited by geographical distance.

The *counts of isolate vertices* are proportionately comparable for the three large clusters (15%-19% of all the vertices) and relatively high for the smaller clusters (e.g., in Sherbrooke almost half of the inventors are isolated). In general, around 23% of inventors in the subnetworks do not have any collaborator in the cluster and work in geographical isolation.

Structural cohesion of the subnetworks

Structural cohesion refers to the degree to which vertices are connected among themselves. The most common measure of cohesion is the *density of a network*, which is the number of existing lines in the network expressed as a proportion of the maximum number of possible lines. Table 1 shows the subnetwork densities for each cluster. It is evident that for networks of smaller sizes the density is higher and vice versa. Even though density is an indicator often used in social network analysis, it is more suitable to compare networks of the similar sizes, since density is inversely related to network size. De Nooy *et al.* (2005) explain that this is because the number of possible lines increases rapidly with the number of vertices, whereas the number of social ties, which each person can maintain is limited. Therefore we measured the density by the average degree of a network. The degree of a vertex is the number of lines that are directly connected to the vertex (Wasserman and Faust, 1994). It represents the number of direct collaborators with whom an inventor cooperated on at least one patent. The more co-inventors the inventors have, the tighter is the network structure. The *average degree of a network* then denotes the average of the degrees of all vertices and in fact it also shows the average number of co-inventors in each subnetwork, which we discussed earlier. Accordingly, the innovation subnetworks in the clusters of Saskatoon (3.52), Edmonton (3.18) and Ottawa (3.06) are the densest and Montreal (2.94) and Vancouver (2.76) are still relatively dense. Inventors in these clusters have direct or indirect access to a larger amount of knowledge and a greater number of inventors. Consequently the possibility for two inventors to get in touch through a chain of personal acquaintances is higher as well.

Centrality of vertices

The centrality of a vertex indicates whether the position of an individual inventor within the subnetwork is more central or more peripheral. Inventors that are more central have better access to knowledge and better opportunities to spread information. We measured three indicators of the vertex centrality: degree centrality, closeness centrality and betweenness centrality. The simplest definition of centrality is the *degree centrality of a vertex*, which is in fact equal to the degree of the vertex defined above. Inventors in more central positions in the subnetwork are those directly connected to more other inventors and thus have more sources of knowledge at their disposal. *Closeness centrality of a vertex* is the number of other vertices divided by the sum of all distances between the vertex and all others (de Nooy *et al.*, 2004). It measures the centrality of an inventor in terms of his closeness to other inventors, i.e., his ability to interact with all the others and consequently the speed of his access to all the knowledge in the subnetwork. Finally, *betweenness centrality of a vertex* is defined as a proportion of all shortest distances between pairs of other vertices that include this vertex (de Nooy *et al.*, 2004). An inventor is more central if a lot of shortest paths between pairs of other inventors in the subnetwork have to go through him. Betweenness centrality is therefore based on the inventor's importance to other inventors as an intermediary and it measures his control over the interactions between other inventors and thus over the flow of knowledge in the subnetwork. We measured degree centrality, closeness centrality and betweenness centrality for all the inventors in our database. We thus expect that inventors who occupy the most central positions in the subnetworks will be the most influential and probably the most prolific. Table 1 below shows the maximal centralities in each subnetwork. The most connected inventor lives in Toronto (he has 51 co-inventors), but Montreal's most connected inventor has only 16 direct collaborators. Other well connected inventors are located in Vancouver (27 co-inventors) and Saskatoon (25 co-inventors). In fact, the Saskatoon's most central inventor occupies the most central location based on all three centrality measures. Based on closeness centrality, there is a very central inventor in Halifax and another one in Ottawa if measured by betweenness centrality.

Centralization of the subnetworks

Contrary to centrality, which refers to positions of individual inventors, centralization characterizes an entire network. A highly centralized network has a clear boundary between the center and the periphery. The center of a centralized network allows more efficient transmission of knowledge, which consequently spreads fairly easily in highly centralized networks. A network is hence more centralized if centralities of the vertices vary substantially. Centralization of a network is defined as the variation in the degree centrality of vertices, divided by the maximum degree variation which is possible in a network of the same size (de Nooy *et al.*, 2004). Similarly as with centrality, there are three main measures of network centralization: degree centralization, closeness centralization and betweenness centralization. *Degree centralization of a network* is based on the variation in degree centrality of vertices in a network. The Saskatoon, Halifax and Calgary subnetworks show the highest degree centralization scores. Analogous to degree centralization, *closeness centralization of a network* is based on the variation in closeness centrality of vertices in the network. This however could be measured only in connected network (i.e., where all vertices are directly or indirectly connected), since otherwise there are no paths between some vertices and thus it is impossible to compute the distances between them. Our subnetworks are not connected and therefore we could not measure closeness centralization. Finally, *betweenness centralization of a network* is based on the variation in betweenness centrality of vertices in the network. Betweenness centrality can be computed even in unconnected networks. The results are shown in Table 1. It is again Saskatoon and Ottawa, which previously showed the highest maximal betweenness centralities of the vertices and now score the highest in betweenness centralization of all the subnetworks as well.

Geodesic distances in the subnetworks

A shortest path between two vertices is referred to as geodesic. The geodesic distance is then the length of a geodesic between them, which depends on the number of intermediaries needed for an inventor to reach another inventor in the subnetwork. A short path length in innovation networks should improve knowledge production and knowledge diffusion (Cowan and Jonard, 2004; Fleming *et al.*, 2004), since knowledge can move to the different parts of a network more quickly and spread rapidly among inventors. Moreover, as Cowan and Jonard (2004) suggest, decreased path length will cause knowledge to degrade less by bringing new sources of ideas and perspectives from farthest parts of the network to the inventors.

The longest geodesic in a network (the longest shortest path) is called the *diameter of a network*. It quantifies how much apart are the two farthest vertices in a network and it is a rough indicator of the effectiveness of a network in connecting pairs of inventors. In general, the diameters seem to be fairly long when compared to the overall size of the components. This suggests a quite low connectedness in our subnetworks. An indicator of the *average distance of a network*, which denotes an average of all the distances of all the vertices in the subnetwork, is a more global measure of efficiency in communication. Nevertheless, the distance between two unconnected vertices is not defined (does not exist) and the average distance hence could be measured only in fully interconnected networks. We therefore carried out two measurements – first, we calculated the average distance only between reachable vertices (i.e., directly or indirectly connected) and second, the average distance in the largest subnetwork component. Both of these measures show similar results as the subnetwork diameter. The largest diameter and average distances are found in the Montreal and Ottawa clusters, where an inventor transmitting the knowledge needs on average around 4 other intermediaries to reach another inventor within the same component, but he may need as many as 10 intermediaries in the worst case. The exchange of knowledge is much easier in Vancouver and obviously also in many other smaller clusters. However, since small or highly disconnected networks should obviously yield low geodesic scores, it is necessary to evaluate this measure more globally – while considering how many inventors could be reached within the cluster. The *reach of a vertex* is defined as the number of vertices that can be reached from this particular vertex. Table 1 shows the maximal reach for each subnetwork, i.e. the maximum number of reachable inventors within a subnetwork. Evidently, more inventors could be directly or indirectly reached in larger networks. In the Montreal subnetwork 108 inventors can reach each other, while in the larger Toronto cluster it is only 97 inventors who are connected among themselves. The knowledge can spread easily among 74 inventors in Ottawa and 53 inventors in Saskatoon, which are smaller clusters, but only among 37 inventors in much larger Vancouver cluster. The clusters with lower maximal reach are likely to be more disconnected and thus show lower scores of geodesic distances, whereas the clusters with highest numbers of reachable vertices are more connected and should show higher geodesic distances. The exception seems to be Saskatoon for which, with a relatively high maximal reach, does not show a very high average shortest distance. This is indicative of a network structure which enables more efficient knowledge diffusion.

Cliquishness in the subnetworks

Cliquishness is a property of a local network structure which refers to the likelihood that two vertices that are connected to a specific third vertex are also connected to one another. Cliquish networks have a tendency towards dense local neighbourhoods, in which individual inventors are better interconnected with each other. Such networks exhibit a high transmission capacity, since a great amount of knowledge could be diffused rapidly (Burt, 2001). Moreover, a high degree of cliquishness in an innovation network supports friendship and trust-building, and hence facilitates collaboration between innovators. Uzzi and Spiro (2005) and Schilling and Phelps (2007) argue that higher cliquishness enhances system performance and knowledge diffusion. However, Cowan and Jonard (2003) point out the existence of negative effects of cliquishness stemming from the loss due to repetition, as the knowledge exchanged in highly cliquish neighbourhoods is often redundant. Moreover, empirical findings of Fleming *et al.* (2006) confirm the negative impact of the higher cliquishness in the network on innovative productivity. The role of a high degree of cliquishness in the innovation production is still not obvious and the optimal degree will apparently depend on a variety of factors.

In this paper we measured the degree of local cliquishness for each vertex with egocentric density of a vertex, which is the fraction of all pairs of the immediate neighbours of a vertex that are also directly connected to each other, and then we calculated the *average egocentric density of a subnetwork*. The highest values for cliquishness are found in Saskatoon, Ottawa, Vancouver and Montreal. Cliquishness is quite comparable among the larger subnetworks, only Toronto shows much lower cliquishness. The subnetworks of the smaller sizes are less cliquish. Our results are, however, not in agreement with Newman (2001a) who found that the degree of network cliquishness in biomedicine is much lower than in other fields (clustering coefficient of 0.066), which he explained by the differences in social organization between biomedical and other research communities. The values for his other databases quite correspond to our results. The differences are probably caused by the distinct kinds of the studied networks (as we mentioned before, he created his networks based on the co-authorship of the scientific articles and not the co-inventorship of patents).

We can conclude that in order to enhance the efficiency of each network in terms of knowledge diffusion, the network should be cohesive (which means that inventors are closely interconnected), cliquish (which fosters trust and close collaboration), it should have a long reach within large components (which enables bringing fresh and non-redundant knowledge from distant locations) and it should have a centralized structure (which supports fast knowledge transmission). As Table 1 shows, the closest to these properties is the Saskatoon subnetwork. It is the densest, most cliquish and most centralized of all Canadian biotechnology clusters. It has on average the largest components and lowest share of isolates of all clusters. Despite the great size of the components, the diameter is still only of an average size. Inventors from both the Ottawa and Edmonton clusters also benefit greatly from quite large components and fairly dense, relatively cliquish and rather centralised biotechnology subnetworks. The long geodesic distances however make it more difficult to bring new knowledge fast to all researchers. In contrast, we found that the structural properties of the subnetworks of Calgary, Quebec and Toronto are not very suggestive of efficient knowledge transmission and innovation generation. Both the Calgary and Quebec intra-cluster subnetworks are quite sparse and consist of the components of rather small sizes, suggesting great disconnectedness among inventors. Calgary, however, is quite centralized, which supports a more efficient transmission of knowledge, but it has a high share of researchers working in geographical isolation. Quebec is fairly cliquish and hence better interconnected. In both clusters, relatively short geodesic distances increase the speed of the knowledge transmission. The subnetwork of the Toronto cluster is rather sparse, neither very cliquish nor centralized, and comprises components of relatively small sizes, many of which are completely isolate inventors. The Montreal cluster, on the other hand, contains relatively large components through which knowledge has to travel large distances. It is also denser and more cliquish than the Toronto one; researchers seem to be more interconnected and knowledge could still be diffused more rapidly. The subnetwork structure of the Vancouver cluster is somewhere in between the two previous patterns. It is denser than the Toronto subnetwork and quite cliquish, but comprises smaller components and thus involves shorter geodesic distances. The structural network properties are quite diverse within the individual clusters and their exact role in knowledge creation and innovation generation still remains to be determined.

3.2 Inter-cluster Collaborations

Collaboration among clusters is based on co-inventorship through the network components which commonly span several clusters. All inventors in a component are directly or indirectly interconnected and it is thus supposed that they all collectively contribute to the innovation process. The attachment of inventors to their local environment is considered as secondary and the innovation network is analyzed regardless of the inventors' place of residence. Canadian biotechnology inventors are grouped into 894 components, which suggests that the network is quite fragmented and that inventors are not highly interconnected. In terms of the number of vertices, the largest component (Component #1) includes 579 inventors, the second one (Component #33) consists of 185 inventors and the third (Component #9), of 175 inventors. There are few large components (10% of components include around 50% of inventors); most of the components however are relatively small. As a consequence, the average number of inventors in a component is relatively small (5.11). This is attributable to the fact that around 22% of all the components (195 components) are isolates.

The structure and main characteristics of the 30 largest components in our network are shown in Table 2. It is obvious that most components consist of inventors residing in several distinct clusters. This is particularly true for the largest components, where inventors are geographically spread over the entire country and abroad (Components C1 or C2).

Table 2: Main characteristics and composition of the 30 largest components in the Canadian biotechnology innovation network

Component #	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**	**C7**	**C8**	**C9**	**C10**	**C11**	**C12**	**C13**	**C14**	**C15**
# of inventors	579	185	175	78	50	44	39	36	30	29	27	27	27	24	23
# of patents	606	155	139	70	32	70	31	50	30	6	12	15	65	1	12
Patents/inventor	1.05	0.84	0.79	0.9	0.64	1.59	0.79	1.39	1.00	0.21	0.44	0.56	2.41	0.04	0.52
Number of the component’s inventors in each cluster															
Toronto	154	16	8	16		35		5	22	1		25	22		1
Montreal	13	2	112	35	4		34								8
Vancouver	55	10	2		38			1			11		1		
Edmonton	50	1						2	1					1	
Calgary	20							9			3				
Saskatoon	54	40							2						
Winnipeg	1			7											
Kingston	9														
Ottawa	17	77	1	6											1
Quebec	9	2	1	1											1
Halifax		1						1							
Sherbrooke	2	2													
out-of-cluster	25	5		4								1			
abroad	170	29	51	9	8	9	5	18	5	18	13	1	4	23	12

Table 2: continued

Component #	**C16**	**C17**	**C18**	**C19**	**C20**	**C21**	**C22**	**C23**	**C24**	**C25**	**C26**	**C27**	**C28**	**C29**	**C30**
# of inventors	23	22	20	19	18	18	18	18	17	17	17	17	16	15	15
# of patents	16	10	10	37	7	14	10	8	13	7	8	12	22	5	8
Patents/inventor	0.7	0.45	0.5	1.95	0.39	0.78	0.56	0.44	0.76	0.41	0.47	0.76	1.38	0.33	0.53
Number of the component’s inventors in each cluster															
Toronto	8		1	18	1	2	1	6				1	15		
Montreal	5	19	12			7	12				12				1
Vancouver					3	5				1					
Edmonton					1				10						
Calgary						1			1					2	
Winnipeg										1					
Kingston					4				1						5
Ottawa												7			
Quebec	1														
Halifax									1						
out-of-cluster	2		1						1	1					
abroad	7	3	6	1	9	3	5	12	3	14	5	9	1	13	9

Some components (C6, C7, C19, C23 or C28) present intra-cluster cooperation within Canada, but also include some foreign cooperation relations. In fact, all of these 30 largest components include at least one foreign collaborator. Some of these “international” components consist of a majority of foreign inventors with only one or two Canadians (Components C10 or C14). These are probably much larger foreign networks in which a few Canadian inventors participate. For instance, Component C10 is based on collaboration on one single patent and is composed of 24 inventors; out of which 23 are foreign and only one is Canadian. Understandably, these mostly foreign components also show very low ratios of patents per inventor.[11]

Let us now turn to the network characteristics of the 30 largest network components (Table 3). Four largest components usually show higher cohesion but lower centralization than smaller components. They obviously also have larger geodesic distances but higher maximal reach, since it takes longer for the knowledge to travel all over the large component but it can reach many more other inventors. Striking exceptions to this pattern are two medium-sized components, in which all inventors (Component C14) or almost all inventors (Component C10) are connected to each other, since they have all collaborated with each other on all their patents (or almost all for Component C10).

A comparison of the structural properties with the cluster-based subnetworks (Table 1) reveals that the component-based subnetworks (Table 3) are denser, more centralized and present more cliquishness, but they also have greater diameters. This should not be surprising as the cluster-based subnetworks in facts consist of smaller parts of components separated by the cluster of residence of its inventors. Collaboration within components is thus probably more efficient because higher structural cohesion of subnetworks indicates closer interconnectedness of inventors; higher cliquishness fosters trust and close collaboration, and higher centralization supports fast knowledge transmission. In contrast, the cluster-based subnetworks show smaller diameters due to the high structural fragmentation.

---

[11] We are well aware of the fact that concentrating on inventors of Canadian patents may miss some much larger North American or even worldwide network which might link (indirectly) some of the components obtained. Since our focus in on Canadian cluster gatekeepers, however, this does not constitute an obstacle to our study.

Table 3: Structural properties of the component-based subnetworks (see Appendix for explanation description of the structural properties)

Component Number	***C1***	***C2***	***C3***	***C4***	***C5***	***C6***	***C7***	***C8***	***C9***	***C10***	***C11***	***C12***	***C13***	***C14***	***C15***
Number of inventors	579	175	185	78	50	44	39	36	29	30	27	27	27	24	23
COLLABORATION CHARACTERISTICS															
Number of collaborating pairs	2057	560	517	185	167	105	83	92	336	89	61	44	46	276	87
% of repeated collaborations	45%	40%	29%	38%	46%	43%	40%	82%	82%	48%	16%	16%	54%	100%	62%
Max number of repeated collaborations	60	11	19	9	8	9	3	12	6	6	2	2	32	1	5
STRUCTURAL COHESION															
Subnetwork density	0,01	0,04	0,03	0,06	0,14	0,11	0,11	0,15	0,83	0,20	0,17	0,13	0,13	1,00	0,34
Average degree	7,11	6,40	5,59	4,74	6,68	4,77	4,26	5,11	23,17	5,93	4,52	3,26	3,41	23,00	7,57
CENTRALIZATION OF SUBNETWORK															
Degree centralization	0,10	0,10	0,10	0,12	0,45	0,37	0,24	0,27	0,19	0,26	0,27	0,16	0,48	0,00	0,27
Closeness centralization	0,19	0,19	0,16	0,22	0,49	0,31	0,37	0,34	0,28	0,36	0,32	0,23	0,52	0,00	0,42
Betweenness centralization	0,55	0,36	0,49	0,43	0,53	0,53	0,54	0,47	0,01	0,48	0,45	0,51	0,77	0,00	0,57
CENTRALITY OF VERTICES															
Max degree centrality	66	24	24	14	28	20	13	14	28	13	11	7	15	23	13
Max closeness centrality	0,24	0,29	0,24	0,32	0,62	0,48	0,51	0,50	1,00	0,57	0,51	0,39	0,63	1,00	0,71
Max betweenness centrality	0,56	0,38	0,51	0,48	0,56	0,56	0,59	0,51	0,02	0,52	0,51	0,60	0,81	0,00	0,59
GEODESIC DISTANCES															
Subnetwork diameter	17	12	14	10	7	6	7	7	2	5	6	8	6	1	3
Average distance	7,09	5,31	6,58	4,78	2,75	3,10	3,18	3,10	1,17	2,57	2,90	3,65	2,68	1,00	1,98
Max reach	57	174	184	77	49	43	38	35	28	29	26	26	26	23	22
CLIQUISHNESS															
Average egocentric density	0,80	0,81	0,79	0,75	0,80	0,57	0,82	0,77	0,94	0,80	0,83	0,69	0,69	1,00	0,94

Table 3: ***continued***

Component number	***C16***	***C17***	***C18***	***C19***	***C20***	***C21***	***C22***	***C23***	***C24***	***C25***	***C26***	***C27***	***C28***	***C29***	***C30***
Number of inventors	23	22	20	19	18	18	18	18	17	17	17	17	16	15	15
COLLABORATION CHARACTERISTICS															
Number of collaborating pairs	53	43	46	46	41	37	39	38	40	53	54	62	41	40	48
% of repeated collaborations	34%	7%	9%	59%	7%	70%	23%	0%	18%	32%	67%	19%	24%	5%	88%
Max number of repeated collaborations	3	3		26	2	3	3	1	5	2	5	12	12	2	4
STRUCTURAL COHESION															
Subnetwork density	0,21	0,19	0,24	0,27	0,27	0,24	0,25	0,25	0,29	0,39	0,40	0,46	0,34	0,38	0,46
Average degree	4,61	3,91	4,60	4,84	4,56	4,11	4,33	4,22	4,71	6,24	6,35	7,29	5,13	5,33	6,40
CENTRALIZATION OF SUBNETWORK															
Degree centralization	0,32	0,42	0,67	0,63	0,49	0,26	0,57	0,85	0,59	0,34	0,54	0,62	0,37	0,71	0,30
Closeness centralization	0,36	0,33	0,73	0,68	0,57	0,27	0,65	0,91	0,68	0,51	0,60	0,74	0,39	0,83	0,49
Betweenness centralization	0,58	0,62	0,71	0,67	0,68	0,44	0,70	0,85	0,69	0,50	0,36	0,50	0,42	0,71	0,50
CENTRALITY OF VERTICES															
Max degree centrality	11	12	16	15	12	8	13	17	13	11	14	16	10	14	10
Max closeness centrality	0,58	0,53	0,86	0,86	0,74	0,49	0,81	1,00	0,84	0,76	0,89	1,00	0,65	1,00	0,78
Max betweenness centrality	0,63	0,68	0,73	0,69	0,71	0,53	0,72	0,85	0,71	0,53	0,38	0,51	0,48	0,71	0,53
GEODESIC DISTANCES															
Subnetwork diameter	4	5	3	3	5	7	3	2	4	3	3	2	4	2	3
Average distance	2,49	2,76	1,96	1,89	2,17	2,89	2,01	1,75	1,95	1,93	1,68	1,54	2,20	1,62	1,85
Max reach	22	21	19	18	17	17	17	17	16	16	16	16	15	14	14
CLIQUISHNESS															
Average egocentric density	0,82	0,82	0,84	0,83	0,75	0,68	0,83	0,90	0,88	0,85	0,85	0,94	0,66	0,95	0,91

This means that the paths are shorter and knowledge can travel faster in cluster-based subnetworks, but because of the smaller maximal reach, knowledge will finally be acquired by much less inventors. It is not unexpected that the transmission of knowledge through the network is more efficient if there are no geographical barriers and all the interconnected inventors could freely and frequently cooperate regardless of the distance between them. In reality, however, this is not usually the case. Even though we observe that collaboration of Canadian inventors with non-local partners are very common in biotechnology, most inventors, in fact, local intra-cluster collaborative relations are more frequent (Beaudry and Schiffauerova, 2008b). Biotechnology inventors in Canada do take the geographical distance into consideration when searching for partners. Since our cluster-based subnetworks consist of the local fragments of the components-based subnetworks, let us now find the key individuals who link both these spaces.

4. In a search of the gatekeepers

The last part of the paper involves both cluster-based and component-based subnetworks and searches for the bridges between them. Here, we aim to understand exactly how knowledge travels among clusters through the component channels. We look for the inventors who bridge over the two spaces and thus enable the nurturing of biotechnology clusters with fresh knowledge originating outside. Since these inventors stand at the gate through which external knowledge enters clusters, we shall call them *gatekeepers* or *intermediaries*.

First we roughly categorize all Canadian inventors residing in the studied clusters based on the nature of each inventor's connections with other inventors. Three categories of inventors are established: internal inventor, external inventor and intermediary. *An internal inventor* has only intra-cluster connections, i.e. no collaboration partner outside the cluster. *An external inventor*, on the other hand, does not participate in any intra-cluster cooperation, since all of his links are directed out of the cluster. Even if he physically resides in the cluster he has no contacts there and any external knowledge which he acquires remains on the cluster's border. None of the internal or external inventors can thus contribute to the actual knowledge transmission between clusters; *an intermediary* however maintains both intra-cluster and inter-cluster connections and as such, his existence is instrumental to delivering fresh outside knowledge to the cluster. Out of 3065 inventors residing in Canadian clusters 31% (936 inventors) are such intermediaries. The remainder of this section evaluates these intermediaries in their role of procurers of nonlocal knowledge to the cluster.

The most obvious evaluation criterion is based on the amount of knowledge intermediaries bring into the cluster, which here corresponds to *the number of direct sources of external knowledge* to which each intermediary is connected. Table 4 shows the average number of inter-cluster links (or inter-lines in the fifth column) for intermediaries in each cluster, which corresponds to the amount of potential knowledge an average intermediary delivers to his cluster. Moreover, the average number of links (or average degree), including both intra-cluster (within the cluster) and inter-cluster (between clusters) that are connected to the intermediaries in each cluster, is displayed in the fourth column. This measure indicates how well an average intermediary is interconnected in general. Furthermore, we have grouped the intermediaries based on the number of their inter-cluster links the results of which are provided in the last four columns of the same table. Around 70% of all intermediaries collaborate with only 1 or 2 out-of-cluster partners and are thus connected to only 1 or 2 channels through which they can introduce external knowledge into the cluster. An intermediary with a low number of external connections could be extremely important for the cluster as a transmitter of external knowledge, since this depends on his position in the network, but also on the size of the out-of-cluster component to which he is connected.

In order to evaluate the positions of the intermediaries in the network we use the notion of *betweenness centrality***.** Since this measure does not distinguish between the place and direction of knowledge transmission (whether the inventor serves as an important intermediary mainly among the inventors from the same cluster or he is indeed instrumental in the external knowledge transfer to the inventors in the cluster), it cannot fully capture an inventor's strategic position as an external knowledge procurer. At this point we thus use betweenness merely to filter out intermediaries whose betweenness is zero, since any external knowledge which is transmitted through such an inventor is redundant. For instance, imagine an inventor *i* connected to the same exact inventors as at least one other inventor *j* in the component (a co-author on all the same patents as *i* and hence who transmits exactly the same knowledge as the original inventor *i*). If inventor *j* has collaborated on a single additional patent without inventor *i*, then there is at least one other intermediary in the cluster which has exactly the same connections as the original inventor *i* plus at least one additional connection leading to other inventors. The obtained betweenness of the original inventor *i* will thus equal zero. Betweenness in fact measures how the disappearance of an inventor would alter the shortest paths and connectedness between all other inventors. Since the disappearance of inventors with zero betweenness would neither reduce the amount of external knowledge which enters the cluster nor the speed at which it enters (no shortest path would get longer), they are considered redundant and hence excluded from further analysis. After this filtering process, only around half the intermediaries (434 or 14% of all Canadian inventors within clusters) are retained.[12]

Table 4: Inter-lines analysis for all intermediaries

	Number of gate-keepers	*as % of all*	*Average degree*	*Average number of interlines*	*Number of intermediaries with:*			
					1-2 inter-lines	*3-5 inter-lines*	*6-9 inter-lines*	*10 or more*
Toronto	247	27%	6.5	2.4	**187 (76%)**	44 (18%)	9 (4%)	7 (3%)
Montreal	244	35%	6.0	2.3	**174 (71%)**	53 (22%)	15 (6%)	2 (1%)
Vancouver	101	25%	5.7	2.2	**79 (78%)**	11 (11%)	6 (6%)	5 (5%)
Edmonton	92	44%	**7.3**	2.7	52 (57%)	27 (29%)	**13 (14%)**	
Calgary	35	38%	5.9	**3.3**	21 (60%)	7 (20%)	**5 (14%)**	**2 (6%)**
Saskatoon	36	24%	**8.8**	**3.2**	24 (67%)	4 (11%)	**6 (17%)**	**2 (6%)**
Winnipeg	27	35%	3.9	1.8	24 (89%)	3 (11%)		
Kingston	20	21%	4.8	2.4	13 (65%)	6 (30%)	1 (5%)	
Ottawa	97	43%	6.9	2.7	60 (62%)	27 (28%)	8 (8%)	2 (2%)
Quebec	29	23%	4.6	1.6	25 (86%)	4 (14%)		
Halifax	5	15%	5.0	1.8	3 (60%)	2 (40%)		
Sherbrooke	3	12%	3.3	2.0	2 (67%)	1 (33%)		
ALL	**936**	**31%**	**6.3**	**2.4**	**664 (71%)**	**189 (20%)**	**63 (7%)**	**20 (2%)**

Performing once again the interlines analysis exclusively for the non-redundant intermediaries yields Table 5 which can be compared with the previous results including all intermediaries (in Table 4). The comparison suggests that most redundant intermediaries have a very low number of ties to external knowlegde sources As the percentage of intermediaries with only 1 or 2 connections outside the cluster dropped from around 70% to about 50%. This shows

---

[12] Even though for the purpose of the analysis, we do not consider the redundant intermediaries, they are nevertheless important in the regional system of innovation, as the information can “enter” the cluster from a number of sources. The reason to ignore these redundant gatekeepers for the moment will become apparent in the latter part of the paper when we consider the importance of such intermediaries as providers of outside knowledge to the cluster.

that non-redundant intermediaries are usually better interconnected with out-of-cluster collaborators. A proportionally much greater amount of non-redundant intermediaries with many direct sources of external knowledge (6 or more inter-lines) is found in the clusters of Saskatoon (35%) and Calgary (25%), whereas in the big clusters of Toronto, Montreal and Vancouver, almost 90% of all outside knowledge is brought into the clusters by less connected non-redundant intermediaries (1-5 inter-lines). In fact, this is already detectable in the analysis of all intermediaries in Table 4, but the exclusion of the redundant gatekeepers made this observation more pronounced. In Saskatoon and Calgary, gatekeepers have the highest average number of inter-lines. Furthermore, intermediaries from Saskatoon present the highest average degree (for both redundant and non-redundant intermediaries). The intermediaries from these two cities therefore seem to be better interconnected with their external innovation environment than those in other clusters. These observations are however not surprising in the light of the integratedness of Saskatoon researchers within the two largest components identified in Table 2, 94 out of the 147 inventors of Saskatoon collaborate within these two components. Very few inventors are present in the remaining 28 largest components (only two inventors in component C9).

Table 5: Inter-line analysis for non-redundant intermediaries only

	Number of gate-keepers	*as % of all*	*Average degree*	*Average number of interlines*	*Number of intermediaries with:*			
					1-2 inter-lines	*3-5 inter-lines*	*6-9 inter-lines*	*10 and more*
Toronto	124	13%	9.0	3.2	**74 (60%)**	**36 (29%)**	8 (6%)	6 (5%)
Montreal	111	16%	8.1	3.0	**56 (50%)**	**42 (38%)**	11 (10%)	2 (2%)
Vancouver	39	9%	7.7	2.6	**26 (67%)**	**8 (21%)**	3 (8%)	2 (5%)
Edmonton	41	**20%**	9.9	3.1	17 (41%)	18 (44%)	6 (15%)	
Calgary	24	**26%**	7.1	**4.0**	12 (50%)	6 (25%)	**4 (17%)**	**2 (8%)**
Saskatoon	20	14%	**12.6**	**4.6**	9 (45%)	4 (2%)	**5 (25%)**	**2 (10%)**
Winnipeg	5	6%	4.6	1.8	4 (80%)	1 (2%)		
Kingston	10	11%	5.9	2.6	6 (60%)	3 (3%)	1 (10%)	
Ottawa	45	20%	9.5	3.4	21 (47%)	16 (36%)	6 (13%)	2 (4%)
Quebec	9	7%	7.1	1.9	6 (67%)	3 (33%)		
Halifax	3	9%	5.7	2.3	1 (33%)	2 (67%)		
Sherbrooke	3	12%	3.3	2.0	2 (67%)	1 (33%)		
ALL	**434**	**14%**	**8.6**	**3.2**	**234 (54%)**	**140 (32%)**	**44 (10%)**	**16 (4%)**

Table 6: First 25 non-redundant intermediaries with highest inter-line count showing values of all importance indices and other network properties

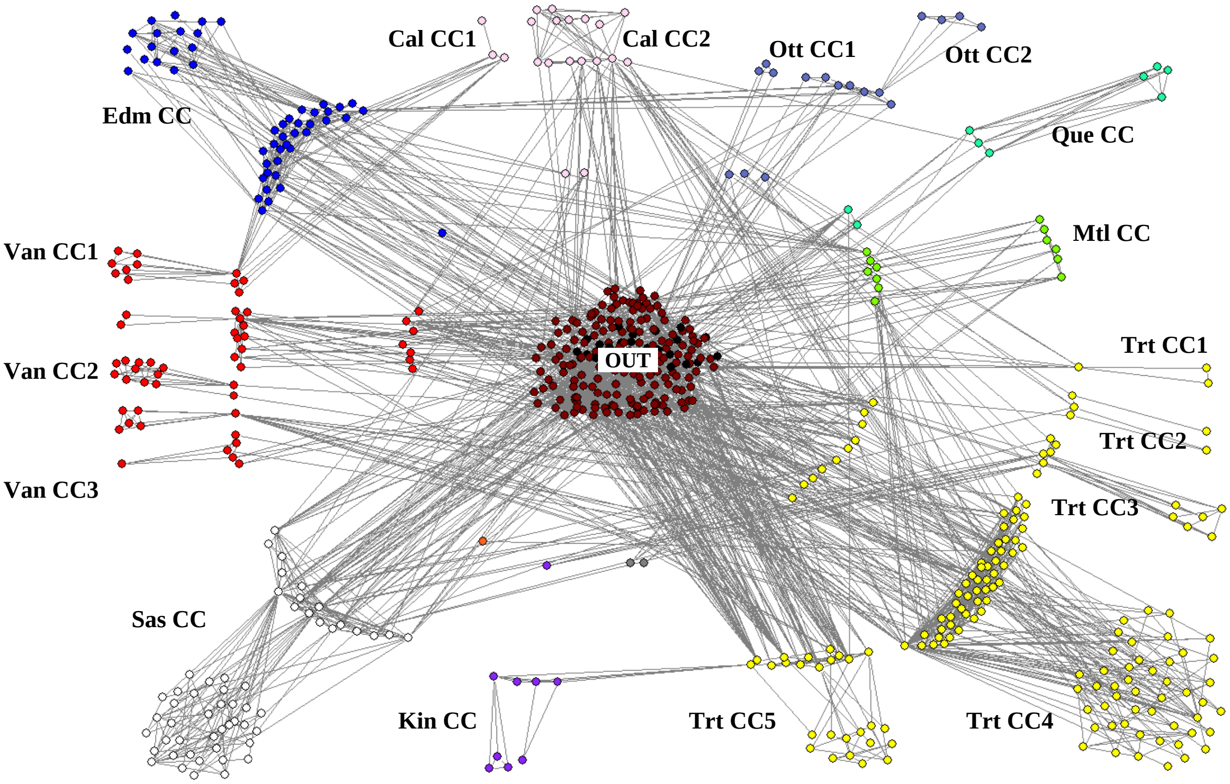
Gate-keeper ID[*]	*Rank by inter lines*	*Inter lines count* (I_i)	*Inter lines value*	*Degree*	*Between-ness* *(1000xB$_i$)*	*Patents by the C-C* (P_{cc})	*C-C size*	***Importance*** *inventor for C-C* *(I$_i$/I$_{cc}$)*	*C-C for cluster* *(P$_{cc}$/P$_{cluster}$)*	*C-C for Canada* *(P$_{cc}$/P$_{Canada}$)*	***Gatekeeper`s Importance Index*** ($GII_i^{cluster}$) *Index value*	*Rank in cluster*	*Rank in Canada*	(GII_i^{Canada}) *Index value*	*Rank in Canada*
TRT $_1$	***1***	29	81	33	1,63	55	25	24,37%	4,05%	1,55%	0,0161	***4thin TRT***	***12***	0,0062	**8**
TRT $_2$	***2***	21	54	28	2,71	55	25	17,65%	4,05%	1,55%	0,0194	***3rdin TRT***	***11***	0,0074	**6**
CAL $_1$	***3***	16	94	25	8,94	64	17	33,33%	30,48%	1,80%	0,9085	***1stin CAL***	***1***	0,0537	**1**
TRT $_3$	***4***	15	59	66	7,21	254	110	8,62%	18,72%	7,15%	0,1163	***1stin TRT***	***5***	0,0444	**2**
MTL $_1$	***5***	14	27	19	0,39	117	112	9,09%	15,70%	3,30%	0,0056	***1stin MTL***	***33***	0,0012	**21**
CAL $_2$	***6***	13	14	14	0,01	5	2	76,47%	2,38%	0,14%	0,0001	***11thin CAL***	***155***	0,0000	**195**
TRT $_4$	***7***	12	38	28	0,66	254	110	6,90%	18,72%	7,15%	0,0085	***8thin TRT***	***23***	0,0032	**14**
TRT $_5$	***7***	12	12	17	0,01	8	6	70,59%	0,59%	0,23%	0,0000	***49thin TRT***	***195***	0,0000	**154**
MTL $_2$	***9***	11	11	19	0,13	117	112	7,14%	15,70%	3,30%	0,0014	***9thin MTL***	***68***	0,0003	**51**
SAS $_1$	***9***	11	21	33	3,30	80	54	13,10%	51,28%	2,25%	0,2219	***2ndin SAS***	***3***	0,0097	**4**
SAS $_2$	***9***	11	16	36	5,13	80	54	13,10%	51,28%	2,25%	0,3443	***1stin SAS***	***2***	0,0151	**3**
TRT $_6$	***9***	11	23	17	5,30	55	25	9,24%	4,05%	1,55%	0,0199	***2ndin TRT***	***10***	0,0076	**5**
VAN $_1$	***9***	11	12	22	1,00	8	12	42,31%	2,00%	0,23%	0,0085	***2ndin VAN***	***22***	0,0010	**22**
OTT $_1$	***14***	10	15	14	0,10	18	6	43,48%	5,94%	0,51%	0,0025	***11th in OTT***	***52***	0,0002	**70**
OTT $_2$	***14***	10	59	16	0,01	13	7	34,48%	4,29%	0,37%	0,0001	***29th in OTT***	***169***	0,0000	**191**
VAN $_2$	***14***	10	25	15	0,27	7	6	100,00%	1,75%	0,20%	0,0048	***4th in VAN***	***35***	0,0005	**35**
CAL $_3$	***17***	9	14	21	2,06	64	17	18,75%	30,48%	1,80%	0,1179	***2ndCAL***	***4***	0,0070	**7**
KIN $_1$	***17***	9	9	12	0,01	4	4	56,25%	4,12%	0,11%	0,0002	***4th in KIN***	***129***	0,0000	**205**
MTL $_3$	***17***	9	13	24	0,38	117	112	5,84%	15,70%	3,30%	0,0035	***3rdin MTL***	***42***	0,0007	**30**
SAS $_3$	***17***	9	38	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***9th in SAS***	***85***	0,0000	**123**
SAS $_4$	***17***	9	46	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***11th in SAS***	***85***	0,0000	**125**
SAS $_5$	***17***	9	38	19	0,05	80	54	10,71%	51,28%	2,25%	0,0025	***6th in SAS***	***51***	0,0001	**85**
SAS $_6$	***17***	9	38	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***10th in SAS***	***85***	0,0000	**123**
TRT $_7$	***17***	9	58	12	0,01	15	5	60,00%	1,11%	0,42%	0,0001	***43rd in TRT***	***176***	0,0000	**135**
TRT $_8$	***17***	9	9	17	0,44	254	110	5,17%	18,72%	7,15%	0,0043	***11th in TRT***	***39***	0,0016	**18**

[*] Gatekeeper ID is based on the cluster of the inventor's residence and his rank according to the number of inter-lines. (Whereas the ranking in the 12th column is based on the values of $GII_i^{cluster}$):

TRT$_{\#}$ …Inventor of rank # in Toronto
MTL$_{\#}$ …Inventor of rank # in Montreal
VAN$_{\#}$ …Inventor of rank # in Vancouver
KIN$_{\#}$ …Inventor of rank # in Kingston
CAL$_{\#}$ …Inventor of rank # in Calgary
SAS$_{\#}$ …Inventor of rank # in Saskatoon
OTT$_{\#}$ …Inventor of rank # in Ottawa

Table 6 provides a list of the 25 non-redundant intermediaries with the highest number of direct sources of outside knowledge and orders them according to the number of their inter-cluster links. An inventor from Toronto (**TRT $_1$**) has the highest number of direct external sources (29). The sum of the value of all his inter-lines is 81, i.e. the inventor has collaborated with 29 external collaborators on 81 occasions. The next column shows the degree of a vertex, which is the sum of all his links, including both inter-cluster and intra-cluster. The inventor **TRT $_1$** has only four additional links within the cluster (his degree is 33), which means that all the external knowledge which he acquires flows further into the cluster only through 4 of his colleagues from the cluster. Since not all inventors in the clusters are interconnected within the cluster itself, we do not know how many of them benefit from the external knowledge introduced by any particular intermediary. These indicators do not allow the measurement of whether an inventor is alone in bringing external knowledge to these cluster inventors or whether there are others contributing to this task (which would make his contribution less critical). Moreover, we are not able to assess how much innovative potential this knowledge may create. Therefore, we have developed several measures to help answer these questions. In order to evaluate the importance of each inventor for the transmission of external knowledge and to assess the external innovative potential delivered by him to other inventors in the cluster we have introduced a *Gatekeeper's Importance Index (GII)* both for the cluster and for Canada.

First, let us start with the definition necessary for understanding the concept: A *Cluster-Component group of inventors (C-C group)* is a group of inventors residing in a Canadian cluster who are all directly or indirectly interconnected within the cluster. In a great majority of components, the C-C groups were created as a simple intersection between the clusters and the components, however - particularly in the 4 largest components - many inventors residing in the same cluster and being part of the same component are not directly connected within the cluster and end up in different C-C groups. Figure 1 illustrates the position of the three types of inventors of Component C1. In the centre of the figure is the largest group of inventors in this component, which is composed mainly of foreigners but also of some Canadian inventors residing outside clusters.



The vertices of different shades of grey indicate the inventors residing in different clusters.

Edm CC	…Edmonton C-C group	Mtl CC	…Montreal C-C group
Van CC#	…Vancouver C-C groups	Que CC	…Quebec C-C group
Sas CC	…Saskatoon C-C group	Ott CC#	…Ottawa C-C groups
Kin CC	…Kingston C-C group	Ca CC#	…Calgary C-C groups
Trt CC#	…Toronto C-C groups	OUT	…foreigners or Canadians outside clusters

Figure 1: Component C1 with all created C-C groups

It is fairly obvious that it is these predominantly foreign inventors who are interconnecting all other Canadian inventors in this component. Many of the inventors within the component do not have any other connection among themselves except through the foreign inventors. Canadian inventors residing in clusters are depicted here in three concentric circles around the core of foreigners and out-of-cluster inventors. The inner circle is composed of external inventors, which do not have any direct connections with their fellow inventors from the cluster, but indirectly through out-of-cluster and foreign inventors. Each of these external inventors actually constitutes a separate C-C group (those formed by the external inventors are neither indicated in the figure nor discussed further). In the middle circle are located the inventors connected to those residing both outside and inside the cluster – these are the intermediaries or gatekeepers.

The rest of the inventors - placed in the outer circle (on the periphery of the figure) - are internal cluster inventors connected only to intermediaries or among themselves. The C-C groups of Edmonton, Saskatoon and Kingston were created by the simple cluster-component intersection and there is thus only one C-C group for each cluster in this component. However, many inventors in other clusters had to be separated, notably in Toronto and Vancouver where they ended up in 5 different C-C groups in each cluster, since the only connections existing between them are through inventors residing outside clusters.

The *Gatekeeper's Importance Indices (GIIs)* are based on the measurement of the importance of each intermediary as a source of external knowledge for the C-C group to which he takes part and the importance of this C-C group either for the cluster or for Canada. The two *GIIs* are defined as:

$$GII_i^{cluster} = \frac{I_i}{I_{cc}} \times \frac{P_{cc}}{P_{cluster}} \times B_i \times 1000$$

$$GII_i^{Canada} = \frac{I_i}{I_{cc}} \times \frac{P_{cc}}{P_{Canada}} \times B_i \times 1000$$

where:

- $GII_i^{cluster}$… *Gatekeeper's Importance Index for Cluster* for inventor i
- GII_i^{Canada}… *Gatekeeper's Importance Index for Canada* for inventor i
- I_i…the number of inter-cluster links of the inventor i
- I_{CC}…the sum of all inter-cluster links of the C-C group cc (which includes inventor i)
- P_{cc}…the sum of all the patents invented or co-invented by at least one inventor from the C-C group cc (which includes the inventor i)
- $P_{cluster}$…the sum of all the patents authored or co-authored by all the inventors in the cluster in which the inventor i resides
- P_{Canada}… the sum of all the patents authored or co-authored by all the inventors residing in Canadian clusters
- B_i…betweenness centrality of the inventor i

The first term of the product in both indices captures the importance of the inventor as a source of external knowledge for the C-C group. It measures the number of inter-links connected to each inventor (I_i) as a share of all the inter-links entering the given C-C group of inventors (I_{CC}). Since we disregard time in this analysis and thus assume that all links are active simultaneously, we could also assume that the amounts of external knowledge incoming by each such channel are equal even if the values of the links in fact differ (No matter how many times the collaboration between two inventors took place the total amount of the knowledge exchanged between them remains the same). The second term of $GII_i^{cluster}$ evaluates the importance of each C-C group for the cluster based on the innovative productivity of that group. The patents which are authored or co-authored by at least one of the C-C group inventors are added for each group and divided by the sum of all the patents invented or co-invented by at least one of the inventors from the cluster ($P_{cluster}$). The last importance measure, which constitutes the second term of GII_i^{Canada} evaluates the importance of each C-C group for Canada and is based on the innovative productivity of the group as well. It also counts the number of patents which have been created within the C-C group of which the given inventor is part and expresses that number as a share of the total innovative production in all Canadian clusters (P_{Canada}). The last term of the product in both indices measures the betweenness of the inventor (B_i) and indicates how well is the inventor interconnected in general[13]. This involves an overall evaluation of his network position which goes far beyond the external channels: it takes into consideration his other connections inside the cluster, the connections of all the inventors to whom he is connected and the positions of all the other inventors in the component from which he can indirectly gather knowledge or to whom he can deliver it. The resulting products are called Gatekeeper's Importance Indices and measure an inventor's importance as a procurer of external knowledge for the cluster ($GII_i^{cluster}$) or for Canada (GII_i^{Canada}) based on the share of innovative production to which he thereby contributes.

---

[13] It is in part for the calculation of these indices that we decided to ignore the redundant gatekeepers.

Table 6, which presents the importance measures for 25 intermediaries with the highest number of direct external sources, contains all the importance indices as well. Here are few examples which show how to interpret the measures: inventor **TRT_1** has the greatest count of inter-cluster collaboration links and contributes to around 24% of all the potential external knowledge input flowing into his C-C group (i.e. the percentage of TRT_1 interlinks with respect to the total number of interlinks of the cluster). The C-C group's share of the patent production represents around 4% of the cluster's production and around 1.5% of the total Canadian patent production . The final *Gatekeeper's Importance Indices*, which also take into account his network position, place inventor **TRT_1** in 8th position for his importance in the cluster and to the 12th position for his importance in Canada. Within his own Toronto cluster, he is the 4th most important inventor in terms of his function as intermediary of external knowledge. The inventor **CAL_2** brings over 76% of external knowledge into the C-C group; this group however does not contribute significantly to the overall patent production in the cluster (2.4%) and even less in Canada (0.1%). Furthermore, even though **CAL_2** has 13 direct sources of knowledge outside the cluster his C-C group inside the cluster is actually formed only by him and one additional inventor and his betweenness score is very low. In spite of the high number of external sources to which he has a direct access, the importance of such intermediary is quite negligible and he ranks very low in both in cluster and in Canada. Similar situation could be observed for the inventors **TRT_1**, **OTT_2**, **KIN_1** and **TRT_7**. These intermediaries utilize relatively many direct sources of external knowledge for themselves, but they do not transfer the knowledge to many fellow inventors inside their own clusters. It would seem that these gatekeepers act in fact as ambassadors of knowledge from their own clusters to the outside world.

Four inventors with the highest scores of $GII_i^{cluster}$ in Canada are from the Saskatoon and Calgary clusters, which points out towards the crucial role played by these intermediaries in their own cluster. Table 7 presents the average number of the importance indices for all the inventors acting as intermediaries for the cluster. It shows that the average scores of $GII_i^{cluster}$ for Calgary (0.04) and Saskatoon (0.03) are much higher than that of any other cluster. The situation changes slightly when the average importance indices for Canada (GII_i^{Canada}) are calculated. Inventors from Toronto gain significantly in importance as gatekeepers for Canada (10 out of the first 20 intermediaries with the highest GII_i^{Canada} are from Toronto). Moreover, Table 7 shows that the average scores of GII_i^{Canada} are highest in Saskatoon (0.0026), Calgary (0.0013) and Toronto (0.0008), while it is much lower for other intermediaries.

Table 7: Average values of the indices of importance for the gatekeepers in each cluster

Cluster	*Importance of intermediary for the C-C*	*Importance of the C-C for cluster*	***Gatekeeper's importance index for cluster***	*Importance of the C-C for Canada*	***Gatekeeper's importance index for Canada***
Toronto	29.45%	6.16%	0.00204	**2.35%**	**0.00078**
Montreal	28.30%	5.01%	0.00030	1.05%	0.00006
Vancouver	39.61%	3.01%	0.00097	0.34%	0.00011
Edmonton	15.35%	18.47%	0.00510	1.36%	0.00037
Calgary	**39.57%**	9.74%	**0.04334**	0.58%	**0.00256**
Saskatoon	17.57%	**27.63%**	**0.02993**	1.21%	**0.00132**
Winnipeg	43.33%	6.49%	0.00046	0.14%	0.00001
Kingston	51.91%	4.95%	0.00029	0.14%	0.00001
Ottawa	11.51%	**29.94%**	0.00148	**2.56%**	0.00013
Quebec	34.55%	7.97%	0.00267	0.28%	0.00009
Halifax	58.33%	15.74%	0.00004	0.16%	0.00000
Sherbrooke	100.00%	12.64%	0.00003	0.10%	0.00000
Average	**28.36%**	**10.52%**	**0.00522**	**1.47%**	**0.00050**

Figure 2 displays both the absolute numbers and relative proportions of inventors in each cluster allocated to the categories of inventors based on their importance as procurers of external knowledge for the cluster $GII_i^{cluster}$. *Internal* and *external* inventors do not participate in the transmission of external knowledge to the cluster, since they lack either the connection outside or inside their cluster. These inventors constitute the majority of all inventors in all clusters (60%-80% for most of the clusters). Inventors which do maintain both intra-cluster and inter-cluster collaborations, but do not serve as indispensible intermediaries for other inventors are *redundant intermediaries*. As it was described, such intermediary brings a redundant external knowledge to the cluster, since not only would his disappearance not reduce the amount of external knowledge which enters the C-C group but it would not even make the shortest paths for that transmission longer. These inventors could be still productive and thus considered important creators of biotechnology innovation (even star scientists), but they are quite useless as external knowledge procurers. Around 15%-20% of inventors in most of the clusters are such intermediaries.

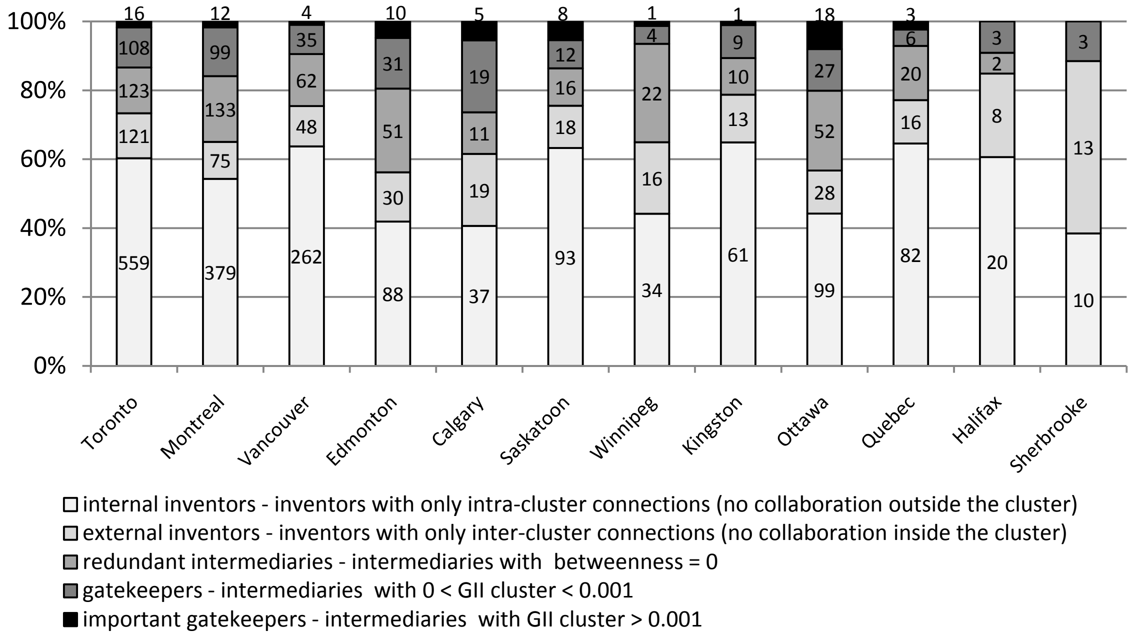


Figure 2: Numbers and relative proportions of inventors in the clusters categorized according to their importance as intermediaries

The remainder of the inventors are considered to be the *gatekeepers*. These are the intermediaries which do introduce non-redundant knowledge to the cluster and contribute thereby to the innovative potential of other inventors in the cluster. The highest percentage of gatekeepers among the cluster's inventors is found in Calgary (26%), Edmonton (20%) and Ottawa (20%), whereas Vancouver (9%) and the small clusters (6%-12%) have the lowest shares. However, the levels of contribution differ significantly among the gatekeepers themselves and therefore we have designated any gatekeeper with $GII_i^{cluster}$ of at least 0.001 as an ***important gatekeeper*.** Quite high percentages (around 60%) of all gatekeepers are considered to be important gatekeepers in the clusters of Saskatoon and Ottawa, but also in Quebec (30%), Edmonton (24%) and Calgary (20%), however in the greatest clusters of Toronto, Montreal and Vancouver only around 10%-13% of all gatekeepers are important gatekeepers for the cluster. (The number of the important gatekeepers in Ottawa is higher than their count in Toronto even in absolute terms.) The smallest clusters have none or very few of important gatekeepers.

Let us briefly summarize the findings of this section concerning clusters. This analysis has shown that the proportions of gatekeepers among inventors is highest among the Calgary, Edmonton and Ottawa inventors. The clusters of Calgary and Saskatoon benefit from relatively many quite important and well-interconnected intermediaries with numerous direct sources of external knowledge and even from a couple of gatekeepers which are of extreme importance for the cluster's innovative productivity. On the other hand, in the greater clusters of Toronto, Montreal and Vancouver the number of gatekeepers is proportionally lower; most of them are not of a very high importance for the cluster and also have a relatively low number of connections outside the cluster. The relative contribution of the Toronto inventors to the total Canadian biotechnology innovation production is however much more important.

Most of the network components (758 components, which represents 85% of all components) do not involve any gatekeeper. These are either components with only internal and external inventors (often single-inventor components or isolates) or components where all the inventors are connected to each other (each inventor is an intermediary who absorbs outside knowledge, but does not transmit it any further, since all of his colleagues have access to the the same knowledge sources, i.e.they are all redundant intermediaries.). As for the components with gatekeepers (136 components, or 15% of the total), over half of them involve only one gatekeeper for the entire component. In this case there is one C-C group within the component where all external knowledge could be transferred to the group only through a single intermediary. If there are any other C-C groups within such component they consist either only of an external inventor or only of redundant intermediaries. Almost half (44%) of the 434 gatekeepers are part of the largest four components. This highlights a critical role played by the large components in the introduction of new knowledge to the cluster. Figure 1 illustrated the collaboration among inventors within the largest component in the Canadian biotechnology network (Component C1, which involves 24% of all gatekeepers). It shows that inventors within the same cluster may not in fact be connected within the cluster and a foreign or out-of-cluster inventor is necessary to transmit knowledge between them. Within the same cluster and component there are groups working completely separately and the short geographical distance between them does not seem to play a role when seeking the collaboration partners. This allows us to make some conjectures about the position of the Canadian biotechnology network in the worldwide biotechnology innovation network. Many Canadian inventors who now seem to be disconnected may in fact be part of the same international component in the worldwide biotechnology innovation network. The complete Canadian biotechnology network would then be in fact much less fragmented than we see it now and there may exist one giant Canadian biotechnology network component, which would comprise a great majority of all the inventors as suggested by Newman (2001a). Furthermore, if we extend this theory further, most biotechnology inventors in the world might in fact be united in one giant international component where they all indirectly collaborate, share their knowledge and create collective inventions.

5. Conclusions

In this paper we have studied social networks of inventors in which a tie between two actors represents a co-inventorship of one or more patents. Drawing from the lists of inventions on the USPTO website, we have created a patent database and constructed an innovation network for all the registered patents in the field of biotechnology in which at least one inventor or co-inventor resides in Canada. The principal objective of our research was to investigate and to describe the collaborative behaviour of Canadian biotechnology researchers with a special focus on the geographical and network aspects.

The results show that most of the biotechnology collaborative activity which involves Canadian inventors takes place within Canadian clusters. We examined the structural network properties of each cluster and related them to the likely efficiency of each subnetwork in the knowledge diffusion and the innovation creation. We observed that the network should be cohesive (which means that the inventors are closely interconnected), cliquish (which fosters trust and close collaboration), it should have a long reach within large components (which enables bringing fresh and non-redundant knowledge from distant locations) and it should have a centralized structure (which supports fast knowledge transmission). However, the structural network properties are quite diverse within the individual clusters and their exact role in knowledge creation and innovation generation still remains to be determined.

The examination of the network structure has revealed that many inventors from the same cluster may also be part of the same network component. The bulk of smaller components are entirely contained within one cluster, larger components however usually encompass several clusters. Moreover, most of the larger or medium-sized components include some foreign cooperation relations as well. We argue that these foreign inventors are extremely important in connecting Canadian inventors from different clusters together (or even from the same cluster - particularly in the largest components), which makes their presence critical for the transmission of knowledge between Canadian inventors. We conjectured that if all biotechnology patents in the world were included in the analysis, the Canadian biotechnology network would be less fragmented and most of the inventors would in fact be a part of one giant international biotechnology innovation component in which all inventors indirectly collaborate, share their knowledge and create collective inventions.

We have also investigated the points of interaction between the cluster-based and components-based subnetworks. In order to understand exactly how the knowledge travels among the clusters through the channels of the components we have searched for gatekeepers - inventors who bridge over the two spaces and thus enable the nurturing of biotechnology clusters with fresh knowledge originating outside. For the purpose of identification of these gatekeepers we have developed indicators, which measure each inventor’s importance as a procurer of external knowledge for the cluster (or for Canada) based on the share of innovative production to which he thereby contributes. Only around 10%-20% of all inventors in most clusters were identified as gatekeepers and are responsible for the inflow of external knowledge to the cluster.

This paper represents another step towards the understanding of the influence of knowledge networks on the innovative activities of inventors located within high technology clusters. We intend to continue exploring the exact role played by networks and their importance in the chain of knowledge creation. One avenue for further research is the examination of the intermediary role of star scientists: Are the most productive inventors also the best procurers of external knowledge for the cluster? Moreover, we plan to gather knowledge about the affiliation for each gatekeeper in order to better understand from which environment they arise: academics, industrial or governmental inventors. Finally, an addition of all the worldwide biotechnology patents would allow to see the networks in their entirety and to gain a full picture of innovation production in Canadian biotechnology.

References

Allen, R.C. (1983). Collective invention. Journal of Economic Behaviour and Organization, 4, 1-24.

Balconi, M., Breschi, S., and Lissoni, F. (2004). Networks of inventors and the role of academia: An exploration of Italian patent data. Research Policy, 33, 127-145.

Baptista, R. and Swann, P. (1998). Do firms in clusters innovate more? Research Policy, 27, 525-540.

Beaucage, J-S. and Beaudry, C. (2006). The Importance of Knowledge Networks within Canadian Biotechnology Clusters. International Schumpeter Conference, Sophia-Antipolis

Beaudry, C. (2001). Entry, growth and patenting in industrial clusters. International Journal of Economics of Business, 8, 3, 405-436.

Beaudry, C. and Breschi, S. (2003). Are firms in clusters really more innovative? Economics of Innovation and New Technology, 12, 4, 325-342

Beaudry and Swann (2008 forthcoming) Firm-level growth in industrial clusters: a bird’s eye view of the United Kingdom, Small Business Economics.

Breschi, S. and Lissoni, F. (2003). Mobility and social networks: localized knowledge spillovres revisited. CESPRI Working Papers, 142.

Breschi, S. and Lissoni, F. (2004). Knowledge networks from patent data: Methodological issues and research targets. CESPRI Working Papers, 150.

Burt, R. (2001). Bandwidth and echo: Trust, information, and gossip in social networks . In: A. Casella and J.E. Rauch (Eds.), Networks and Markets: Contributions from Economics and Sociology New York: Russel Sage Foundation.

Cantner, U. and Graf, H. (2006). The network of innovators in Jena: An application of social network analysis. Research Policy, 35, 463-480.

Cockburn, I.M. and Henderson, R.M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics, 46(2), 157-182.

Cowan, R. and Jonard, N. (2003). The dynamics of collective invention. Journal of Economic Behaviour and Organization, 52, 513-532.

Cowan, R. and Jonard, N. (2004). Network structure and the diffusion of knowledge. Journal of Economic Dynamics and Control, 28, 1557-1575.

Cowan, R., Jonard, N. and Ozman, M. (2004). Knowledge dynamics in a network industry. Technological Forecasting and Social Change, 71, 469-484

Cowan, R., Jonard, N., and Zimmermann, J.-B. (2007). Bilateral collaboration and the emergence of innovation networks. Management Science, 53, 7, 1051-1067

de Nooy, W., Mrvar, A. and Batagelj, A. (2005). Exploratory Social Network Analysis with Pajek. Cambridge: Cambridge University Press

Dahl, M.S. and Pedersen, C.O.R. (2004). Knowledge flows through informal contacts in idustrial clusters: myth or reality? Research Policy, 33, 1673-1686.

Ducor, P. (2000). Intellectual property: Coauthorship and coinventorship. Science, 289(5481), 873-875.

Ejermo, O. and Karlsson, C. (2006). Interregional inventor networks as studied by patent coinventorships. Research Policy, 35, 412-430.

Fleming, L., King, C.IV., and Juda, A. (2006). Small worlds and innovation. SSRN Working Papers.

Gauvin, S. (1995). Networks of innovators: Evidence from Canadian patents. Group Decision and Negotiation, 4, 411-428.

Gittelman, M. (2006). Does Geography Matter for Science-Based Firms? Epistemic Communities and the Geography of Research and Patenting in Biotechnology. DRUID Summer Conference on Knowledge, Innovation and Competitiveness: Dynamics of Firms, Networks, Regions and Institutions, Copenhagen, Denmark.

Krebs, Valdis. (2006). Social network analysis. Available on-line: http://www.orgnet.com/sna.html

Krugman, P. (1991). Geography and Trade. Cambridge, MA: The MIT Press.

Lamoreaux, N.R. and Sokoloff, K.I. (1997). Location and technological change in the American glass industry during the late nineteenth and early twentieth century. NBER Working Papers, 5938.

Mariani, M. (2000). Networks of inventors in the chemical industry. MERIT Research Memorandum.

Morone, P. and Taylor, R. (2004). Knowledge diffusion dynamics and network properties of face-to-face interactions. Journal of Evolutionary Economics, 14, 327-351.

Munn-Venn, T and Mitchell, P (2005). Biotechnology in Canada: A Technology Platform for Growth, the report from the Conference Board of Canada. Available on-line: http://www.agwest.sk.ca/biotech/documents/115-06-Biotechnology%20in%20Canada.pdf

Newman, M.E.J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. Physical Review, 64, 016131.

Newman, M.E.J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Physical Review, 64, 016131.

Newman, M.E.J. (2001c). Clustering and preferential attachment in growing networks. Physical Review, 64(025102).

Newman, M.E.J. (2001d). The structure of scientific collaboration networks. Proceedings of National Academy of Sciences, 98(2), 404-409.

Newman, M.E.J., Watts, D.J., and Strogatz, S.H. (2002). Random graph models of social networks. Proceedings of the National Academy of Sciences, 99, 2566-2572.

Niosi J. and Bas, T.G. (2001). The Competencies of Regions – Canada's Clusters in Biotechnology. Small Business Economics, 17(1-2), 31-42.

OECD (2005).A Framework for Biotechnology Statistics. Available on-line: http://www.oecd.org/dataoecd/5/48/34935605.pdf

Putsch, F. (2006). Analysis and modeling of science collaboration networks. Working Paper.

Rasmussen, B. (2004). An Analysis of The Biomedical Sectors in Australia and Canada in a National Innovation Systems Context, Centre for Strategic Economic Studies, Victoria University of Technology, Working Paper Series, No. 21. Available on-line:

http://www.cfses.com/documents/pharma/21-Biomedical_Sect_Aust_&_Can_Innovation_Rasmussen.pdf

Rose, A. (2000). A Challenge for Measuring Biotechnology Activities, The Economics and Social Dynamics of Biotechnology, 2000.

Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, MA: Harvard University Press.

Schiffauerova, A. and Beaudry, C. (2008). Innovation in Canadian Biotechnology Clusters

Schilling, M.A. and Phelps, C.C. (2007). The impact of large-scale network structure on firm innovation. Management Science, 53, 7, 1113-1126

Schrader, S. (1991). Informal technology transfer between firms: Cooperation through information trading. Research Policy, 20, 153-170.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion. Management Science, 51(5), 756-770.

Statistics Canada (2001). Practices and Activities of Canadian Biotechnology Firms: Results from the Biotechnology Use and Development Survey - 1999. Available on-line: http://www.statcan.ca/english/research/88F0006XIE/88F0006XIB2001011.pdf

Strachan, G. (1995). The impact of regulations on the growth and evolution of Canadian commercial biotechnology. Current Opinion in Biotechnology 1995, 6:261-263

Uzzi, B. and Spiro, J. (2005). Collaboration and creativity: The small world problem. American Journal of Sociology, 111(2), 447-504.

von Hippel, E. (1987). Cooperation between rivals: Informal know-how trading. Research Policy, 16, 291-302.

Wasserman, S. and Faust, K. (1994). Social network analysis. Cambridge: Cambridge University Press.