New perspectives of distance in International Innovation Networks. The role of cognitive and of geographic distance in biotechnology

M. Balconi°, G. Palamara#, P.P. Saviotti*, A. Zucchella#

° CIBIE, Università di Pavia, Italy;
Dipartimento di Ricerche Aziendali R.Argenziano, Università di Pavia, Italy
*INRA GAEL, Univérsité Pierre Mendés-France

Corresponding Author : Antonella Zucchella, Center for Research In International Business, University of Pavia, Via S.felice 5, 27100 Pavia, Italy
Mail: antzuc04@unipv.it

Keywords: innovation, knowledge, networks, cognitive distance, geographic distance

Abstract

Distance represents one of the key constructs in IB. This paper analyses the issue of distance from the two perspectives of geographic and cognitive distance in the working of innovation networks on a global scale. While the former is a better explored field, the latter gained much less attention, especially in terms of operationalisation and empirical studies. Many industries rely extensively on innovation networks for the creation of new knowledge. In these networks, cognitive distance -instead of psychic or cultural distance- represents a key driver of knowledge creation, together with geographic distance. Innovation networks are crucial for biotech firms because, due to the growing rate of creation of new knowledge and to its increasing specialization, they need to combine different knowledge bases to create innovative products. Biotech firms rely extensively on collaborations for R&D not just because they need to share risks and costs of long lasting and uncertain research processes, but also because their technological platform (the biotechnological one) is still at an early stage of exploration and exploitation, especially in some fields, where innovation requires the pooling of differentiated knowledge bases (Malerba and Orsenigo, 2001; Pisano, 2007). The object of our research is to explore collaborative innovation processes in the Italian biotech industry, in particular from the point of view of partners' heterogeneity and complementarities, in terms of cognitive and geographic proximity. This contribution describes the study of the co-publications of a sample of Italian biotechnology firms with collaborators situated in a very large number of countries.

1. Introduction

The construct of distance is a constitutive one in international business studies. Traditionally it has been viewed in terms of geographic, cultural and psychic distance. The former represents a first frame not only for IB studies but also for economic geography and innovation studies. Cultural distance aims at identifying clusters of related national cultures along certain attributes (Hofstede, 1980), thus enabling a sort of objective measurement of distance in its cultural dimension between countries/groups of countries. Psychic distance represents a subjective measure of distance because it is related to geographic and cultural distance but also to the firm risk aversion and experiential knowledge cumulated by the firm in the previous domestic/international expansion (Hornell, Vahlne and Wierdesheim-Paul, 1973).

The literature in IB, as well as in management and in innovation studies, recognises that the access to knowledge and the creation of new knowledge frequently takes place in networks and in dyadic relationships. This paper adopts this point of view and aims at a deeper understanding about the role of distance in knowledge creation processes inside innovation networks. These networks can be better approached adopting the perspective of geographic distance together with cognitive distance, which expresses distance between different firms' knowledge bases, rather than between markets or between firms experience and markets.

The main objective of this paper is to shed light on the role played by cognitive and geographic distance (or, inversely, proximity) in biotechnology innovation networks (INs). To this aim we present a study of the innovation network constituted by a sample of 32 Italian biotechnology firms and by the set of 553 organizations co-publishing with them. The specificities of this paper are the following: (i) it describes the first (or one of the very first) measures of cognitive distance ever published for INs and it develops an original methodology to do it; (ii) it focuses on Italian biotechnology firms which are mostly small, privately owned and have few or no patents, which in turn have typically a single applicant.

For this type of firms it is very difficult both to identify the alliances in which they are involved and the field of knowledge they master. Interestingly, a few of these firms publish more than they patent. Thus, in order to discover the network of collaborators and to map the knowledge base of Italian biotechnology firms and of their partners we relied on co-publications. We are aware that networks based on co-publications are not representative of all the networks in which firms can participate. For example, a co-publication is unlikely to involve the same kind of commitment of resources as the joint development of a new drug. However, given the central role which the generation of new knowledge is known to play in innovation networks, we can expect co-publications to involve generally used mechanisms of collaboration and their study to help us gain a general understanding of innovation networks. In spite of the above remarks, the study of co-publications cannot be considered a substitute for a full analysis of innovation networks and in future it will be important to understand the role of co-publications in a complete study of innovation networks.

This paper represents an interesting methodological exercise because it does not rely on patents which are the most common source of information about firms' knowledge bases, but which are not often available for Start Ups (SUs) or for Dedicated Biotechnology Firms (DBFs). Since DBFs are an essential component of biotechnology INs, to provide a method which allows studying them in their generality is an important contribution to our understanding of INs.

In what follows we first briefly summarize the literature on INs, with particular reference to the role played in them by the creation and utilisation of knowledge.

Subsequently we describe the concepts of the knowledge base of the firm, of cognitive and of geographic distance and the methodology we used to measure them. Finally we present our results and discuss their interpretation.

2. Conceptual background

2.1 Innovation networks

The economics of networks approach (Hankasson, 1987, 1990) establishes a relationship between learning, innovation and networks, based on actors, activities and resources. Among resources, the knowledge base is a fundamental one. The construct of knowledge base can be referred to the general (scientific) knowledge base, the industry/product knowledge base and the firm specific knowledge base, which refers not only to technical knowledge but also to the way technology is linked to other activities and processes (Smith, 1995). As a consequence, the knowledge bases of firms tend to be multi-layered and highly firm-specific in their combination and structure. The concept of knowledge base of the firm (KB) on which this contribution rests is that of the 'collective knowledge that firms can use for their productive purposes' (Nesta and Saviotti, 2005). The collective nature is determined by the interactions which various types of knowledge undergo due to the firm structure and organisation. For this reason, the KB of each firm has a degree of specificity and, especially for knowledge intensive sectors, determines firm performance (Nesta and Saviotti 2005, 2006; Fontana and Nesta, 2007). It follows that establishing collaboration with external actors is crucial in order to overcome the limitations of highly specific knowledge bases when the creation of new knowledge is involved.

According to Lundvall (1988) “learning by interacting” is a fundamental pattern through which new knowledge creation and innovation takes place. The relevant interactions for innovation processes may involve the public and private knowledge infrastructure and the production chain (Oerlemans, Meeus and Boekema, 2000).

Innovation networks (INs) involving different firms and/or research organizations (Shan et al., 1994; Ahuja, 2000) are a relatively new phenomenon which emerged significantly since the beginning of the 1980s (Powell, 1990; Pyka and Kuppers, 2001). According to the literature, innovation processes became more and more the result of inter-firm collaborations (Breschi et al., 2003) due to important environmental changes such as an increasing cost of R&D, convergences of technologies, increased global competition and rapid technological change.

Recent innovation theories consider networking as the main locus of new knowledge creation (Freeman, 1991; Pennings and Harianto, 1992) and one of the determinants of competitive advantage of the firm. In a network different knowledge bases are simultaneously present: the diversity among each firm’s knowledge base constitutes an element fostering learning processes and new knowledge creation, and an element characterising the effectiveness of the network itself.

Powell, Koput and Smith-Doerr (1996) suggest that in knowledge-intensive industries, where knowledge bases are complex and dispersed, interfirm cooperation is crucial as innovation takes place in networks of partnerships rather than in individual firms. The degree of knowledge intensity and the turbulence characterizing these industries constitute elements supporting a collaborative innovation pattern.

The biotech industry is widely recognised as one of the main cases of distributed innovation, where research and development take place through collaborations among organisations belonging to different scientific and business areas (Powell and Brantley, 1992; Powell et al., 1996; Chiesa, 2003).

Depending on the desired balance between exploration and exploitation, two main kinds of network are found in biotechnology (March, 1991; Gilsing and Nooteboom, 2005; Rothermael and Deeds, 2004): partnerships between universities, research centres and dedicated biotechnology firms, based mostly on the exploration of knowledge; and partnerships between dedicated biotechnology firms (DBF) and large pharmaceutical companies (in the case of red biotech, the largely prevailing sector), based primarily on the exploitation of knowledge, through licensing out, commercialisation agreements and similar contracts aimed at bringing closer to the market the outcomes of the R&D pipeline.

The innovation pattern in the biotechnological firms is thus characterised by multiple interactions with other organisations, generating INs. The latter show, at the research network level, a partial overlap of inter-organisational ties together with inter-personal (researchers) ties. This work focuses its empirical section on the latter, captured through the co-publications data.

Inter-organisational and inter-personal INs develop in a context which can favour frequent exchanges, in particular of tacit knowledge, and mutual understanding. According to Hakansson (1987), networks tend to be characterised by stability and variety. Stability (repeated interactions over time among partners) enables the access to scarce and valuable resources, while variety supports the exploration of new opportunities of knowledge creation. The economics of innovation networks is thus subject to some trade-offs, around the core issue of proximity. A certain degree of geographic proximity is needed, as well as of similarity in knowledge bases (cognitive proximity) in order to enable and foster knowledge access and sharing. On the other hand, the creation of new knowledge requires a certain variety and "distance" in order to enhance "novel combinations" of knowledge.

2.2. The role of geographic and cognitive proximity in innovation networks

The working of INs rests upon interactions among organisations (learning by interacting), as we mentioned above. Interactions which aim at new knowledge creation rest upon the balance between proximity and distance, both in their geographic and cognitive dimension.

According to Kirat and Lung (1999, p.29) "technological innovation is a process that is based on relationships of proximity, the forms, modes and combinations of which may be quite varied."

Proximity is indeed a complex construct (Lundvall, 1988; Bellet et al. 1993) and it is analysed along different dimensions: geographic, industrial, organisational and institutional proximity. The literature of economic geography and local innovation systems/*milieux innovateurs* has highlighted the first one. It is a fundamental aspect of proximity also in the perspective of this contribution, since in biotechnology - as well as in many other industries - a significant role of spatial embeddedness of innovation processes has been found. Geographic proximity is not in itself a pre-condition for learning by interacting and collaborative innovation, because it needs to be complemented by institutional thickness. The latter is "the combination of factors, including their inter-institutional interaction and synergy, a collective representation by many bodies, a common industrial purpose and shared cultural norms and values" (Amin and Thrift, 1993, p.417).

Industrial proximity refers to vertical and horizontal interdependencies among players (Torre, 1993), which establish complementarities and similarities among organisations and favour shared goals, mutual understanding and collaborative patterns of innovation. In biotechnology these vertical and horizontal links among organisations (universities, DBF, big pharma, venture capitalists, hospitals) characterise deeply the innovative activity (Stuart, Ozdemir, Ding, 2007).

Organisational proximity "relies upon a certain consistency in the configuration of relationships between agents and is structured around a common cognitive framework" (Kirat, Lung, 1999, p.30). The need of proximity in terms of cognitive framework recalls the role of similarity/complementarity of the knowledge bases of collaborating organisations.

This contribution highlights in particular geographic and cognitive distance/proximity. The institutional one is to some extent included in the geographic scales proposed, which represent not only spatial representations but also institutional frames. Industrial proximity is captured through the characteristics of collaborating organisations, but we considered more fruitful to focus on the similarities/differences of knowledge bases rather than the generic categories to which organisations belong.

Geographic proximity tends to establish a fertile ground for cognitive interactions which enable the sharing of tacit knowledge, while codified knowledge may also be exchanged /accessed without physical proximity (Nonaka, 1994; Nonaka and Takeuki, 1995).

In fact the transfer of tacit knowledge, which very often characterizes learning processes, is possible only through a social process of personal and face to face interaction between the "owner" and the "recipient" of knowledge (Storper, 1997, Balconi et al., 2007), involving demonstration, practice, exercise and repetition. Thus the co-location of partners is supposed to foster knowledge creation and sharing processes in INs.

Moreover, geographic proximity allows firms localised in the same spatial area not only to gain better and faster infra and inter-organisational learning processes, but also to exploit other localisation advantages, such as ease of cooperation, research speed, agglomeration of technologies and knowledge. The literature on National and Regional Innovation Systems emphasizes the positive relation and feedbacks among institutions, spaces and firms sharing the same location and sharing the same innovation system (Gertler et al., 2000; Lundvall, 1992; Cooke, 2002). Several studies based on this approach and on the agglomeration economies framework have found that knowledge production and innovative activities tend to agglomerate geographically (Asheim, 1996; Breschi 1999; Feldman, 2000; Arundel and Geuna, 2001).

Many Authors have analysed the importance of location for R&D, knowledge creation and exploitation (see Oerlemans et al., 2000, for a complete literature review): geographic proximity to institutions such as universities, public or private research centres (Autio et al., 2004) is critical to innovation acting as learning environments, and this explains why innovative firms tend to cluster geographically (Mowery et al., 1996; Tijssen, 2001; Mowery and Sampat, 2001). Confirmation of this aspect comes from different empirical analysis, studying co-localisation using citations of patent documents (Jaffe, 1989; Jaffe et al., 1993; Acs et al., 2002; Haas, 1995). However, the literature mentioned above also shows that while some factors favour the co-location of collaborating firms and organizations, collaborations may be determined uniquely by the competencies of the potential partners irrespective of the geographic distance, this being particularly the case in science-based industries. As a consequence, in INs local clusters tend be nodes of global networks.

Proximity cannot be considered only in its geographic dimension: being located in the same area is neither a sufficient nor a necessary condition to share and develop knowledge or to create effective INs (Boschma, 2004).

In INs in highly knowledge intensive sectors, of which biotechnology is a very good example, the other crucial type of proximity is likely to be cognitive proximity (Brown and

Duguid, 1991; Nooteboom, 2000), and in particular the proximity among the knowledge bases of the collaborating partners (Nesta and Saviotti, 2005; Pyka and Saviotti, 2005). As we will discover later, cognitive and spatial proximity interact. “Pure” spatial proximity encourages knowledge transfer *only* through closeness of knowledge sources, without any kind of explicit coordination or interaction, or formal relations (Audretsch, 1998; Gordon and Mecann, 2000). However, when the required knowledge sources are located far away from the firms wanting to acquire them, spatial proximity is likely to play a very limited role. Only when the required knowledge is available locally, being located in the same area can contribute to create *cognitive proximity by* facilitating informal relations, thanks to the greater possibilities of face to face contact, thus contributing to create trust and firms interrelations (Harrison, 1992). In these conditions cognitive proximity between organisations leads to a common knowledge base and to the sharing of experiences and technologies, thus enhancing learning processes by the members of the cognitive community (Nooteboom, 2000). However, the geographic distribution of innovative capabilities in very recent and highly knowledge intensive sectors is far from homogeneous. In particular in biotechnology the USA have a considerable superiority with respect to European countries and important poles of competence exist in Canada, Australia, China, India to mention just a few. In these conditions it is clear that not all forms of collaboration will be locally based. We can expect a trade off to be attained in which knowledge that cannot be found locally will be sourced everywhere it is available. The extent of such a trade off is affected by the costs of travelling and of communicating.

According to the above discussion, innovative communities do not necessarily need to be spatially confined, because *the actual power of cognitive proximity is not uniquely determined by spatial closeness, but by accessibility of the required knowledge types and by the emergence of shared understanding* (Coenen et al., 2004, p. 9).

Literature argues that in the process of new knowledge creation the partners involved share the expectation that the higher the cognitive distance between partners, the higher the advantage generated by the alliance (Nooteboom, 2000). That is to say that, also considering the problems linked to cognitive lock in phenomena (Lambooy, 2003), a firm can expect to learn more from an organization having the knowledge it wants to acquire than its same knowledge. However, also the costs and difficulties a firm faces in learning a new type of knowledge tend to rise with a growing cognitive distance. Taking into account both benefits and costs, cognitive distance is considered to have an inverted U shaped relation with innovative performance (Nooteboom 2000, Gilsing et alii 2008). In fact large cognitive distances yield high opportunities for accessing new knowledge and promise a great extent of learning, but at the same time they also have a negative effect on absorptive capacity (Cohendet and Llerena, 1997), and may reach the point to preclude the mutual understanding between the interacting actors.

Thus the literature suggests that the cognitive distance is a likely determinant of the probability of existence and of success of INs in biotechnology (Brown and Duguid, 1991). A firm is likely to be interested in collaboration if the knowledge it needs is so different from the one it has that it cannot develop it alone. Collaborating with another firm having a knowledge base close to the target knowledge is likely to be an effective way of acquiring it. However, large cognitive distances (CDs) cannot be expected to be very frequent, since the cost of learning would rise with a growing CD. Actual measures of CDs are extremely rare as well as empirical tests regarding these issues. Wuyts et al. (2006) measure cognitive distance in terms of technological and organisational distance on a set of inter-organisational alliances in the pharma-biotech and ICT industries. Their analysis supports the above mentioned inverse U-shaped relation.

We argue that the trade-off between the attractiveness of the target knowledge to be acquired and the cost of acquiring it can be expected to differ depending on the objective of the collaboration. For example, we could expect CD to be higher in an alliance formed to create new knowledge than in one aiming for joint production. In the latter case the objective of the alliance is likely to be a reduction in production costs, ease of communication, and thus a low CD, is likely to be privileged by partners.

We can also expect cognitive distance to vary at constant alliance objective in the course of time. As firms learn by collaborating, their CD can be expected to fall. Furthermore, average CDs are likely to be higher in the early years of a radically new technology and to fall as the technology matures.

3. Research design, methodology and data

The main objective of our paper is to study the influence of cognitive and geographic distance on the formation of innovation networks in biotechnology. The choice of these two variables is determined by their relevance in the processes of exploration and new knowledge creation, according to the above outlined literature survey.

In order to investigate how cognitive and geographic distance influence the INs of biotech firms, we develop an exploratory analysis of 2.230 research collaborations realized by a sample of 32 Italian biotech firms with other organizations during the period 1992-2008. The data we use are the scientific publications co-authored by these firms and the other entities. Even though we are aware that this is only a part of the overall innovation network involving biotech firms, we think joint publications reveal an important part of the network, directly linked with the R&D activity of the cooperating actors.

The main methodological issue we had to address was the operationalization of the concept of cognitive distance, a problem which is far from being solved in the literature (Gallaud and Rherrad, 2002).

Operationalising cognitive distance clearly involves two main steps. The first is identifying the main cognitive specialization, or field of competence, of the various actors, and the second is finding a methodology to measure the distance between these fields of competence.

The first step has been particularly demanding, since it required a thorough examination and comparison of various data sources. Table 1 lists the sources we used to build our dataset.

Table 1. Sources of data and information collected

Sources	*Kind of information*
The web site www.Italianbiotech.com	List of Italian biotech firms, their general characteristics, addresses
Data collected by Cresit[1]	Various data on Italian biotech firms
Web sites of the firms	Qualitative data on knowledge bases
Patent descriptions	Information on inventions and on technologies
Science Citation Index Expanded	Information on research collaborations and scientific topics

Note that the initial database comprised 97 biotech firms: missing values, unavailability of data about collaborations or knowledge bases both of the Italian firms and of their collaborators compelled us to use only part of the information, and to limit the analysis to a final sample of 46 firms. Among these, only 32 have published at least one article. These 32 firms belong mostly to the red biotech sector, with only 3 specialised in the green area; they are small or medium sized, and all of them have at least one patent registered at USPTO or WIPTO.

Overall, we distinguished 14 fields of competence, so that, for each of our firms a of their collaborators we could construct a vector, constituted by these 14 components, where the presence of a field of competence is denoted by a one and its absence by a zero (Fig.1). Given the small size and the high level of specialization of most of the firms or of their collaborators, the fields of competence vectors contained few ones and many zeros. The nature of the data we used is important, because it constrained the measure of distance we could use.

In this paper we considered several possible measures of cognitive distance which have been previously proposed in the literature.

Since formally the problem we face is identical to that faced by ecologists attempting to measure the distance of different species, once we substitute technologies for biological species traits, we could apply several of the measures which have been developed by ecologists (Pielou, 1984). All these measures are indexes of similarity or of their converse, dissimilarity or distance. Some of them can only be used for continuous or for highly variable data, while others can also be used for presence or absence data such as the ones we have. Furthermore, even amongst the types of measures which can be used for presence or absence data, some are inappropriate for cases in which most of the data points are zero, as in our case. Another measure of similarity often used in studies of innovation has been introduced by Jaffe (1986). This measure is based on the assumption that two technologies are similar to the extent that they can be combined with the same third technology. Unfortunately, this type of measure cannot be used due to the nature of our data.

Thus, amongst all the available distance measures we chose the one called Percentage Remoteness (PR), which is the complement of Ruzicka's similarity index (RI). According to Pielou (1984, pp. 43-44 and 55-57) this measure has the advantages of (i) being usable for presence, absence data and (ii) not being adversely affected by the presence of few ones and many zeros in the data. The PR measure is calculated by first computing Ruzicka's similarity index and then its complement to 100. To calculate Ruzicka's similarity index we need to compute the minimum (MIN) and maximum (MAX) for each component of the technology

[1] Cresit is a research center of Insubria University, which publishes one of the most complete and indepth report on biotech in Italy in collaboration with FarmaIndustria, Assobiotech (Italian national associations of pharma and biotech industries) and Blossom and Associati (Italian consultancy firm).

vectors representing the knowledge bases of the collaborating partners (Fig 1 and equations 1 and 2).

	KB_1	KB_2	MIN	MAX
T_1	0	1	0	1
T_2	1	0	0	1
T_3	0	0	0	0
T_4	0	0	0	0
T_5	1	1	1	1
			ΣMIN = 1	ΣMAX = 3

Figure 1. Example of steps in the calculation of Ružička's similarity index (RI) and of percentage remoteness (PR)

In the examples of figure 1, the technology vectors representing the knowledge bases of two firms, KB_1 and KB_2, contain five component technologies (T_1-T_5). In the KB vectors the number one indicates the presence of a technology in the KB of the firm and zero its absence.

Equation (1) is Ružička's index of similarity RI.

$$RI = 100 \times \frac{\sum_{i=1}^{s} \min(x_{i1}, x_{i2})}{\sum_{i=1}^{s} \max(x_{i1}, x_{i2})} \qquad (1)$$

Equation (2) is the calculation of PR, percentage remoteness.

$$PR = 100 - RI \qquad (2)$$

Measuring geographic proximity among the collaborating entities does not present particular methodological problems. We have operationalised it through a variable whose values are based on the closeness of collaborators to the location of the firms. More precisely, we have distinguished five different values:

- 0 if the firm (F) and the collaborating organisation (C) are located in the same region of Italy;
- 1 if F and C are located in the same macro area (North-West, North East, Centre, South, according to ISTAT classification), but in different regions;
- 2 if F and C are located in Italy, but in different macro-areas;
- 3 if C is located in Europe;

– 4 if C is located in the rest of the world.

4. Preliminary results on exploration networks

4.1. The network: the nodes, the links and their value

F have published 764 articles with 553 collaborating institutions (table 2). Since various firms have the same collaborators, the number of relationships (links) realised by F (928) is higher than the total number of partners. The number of collaborations, where a collaboration is a co-publication of an F firm with any co-author, is even higher (2.230), since most of the relations are repeated (2,4 times on average).

Table 2. Main data on the network of collaborations

Total # of articles published by F	764
Total # of collaborators C (# of nodes)	553
Total # of relationships (# of links)	928
Total # of collaborations (# of links*value of each link)	2.230
# of articles per firm : Average	23,9
Median	8
Modal value	2
Min. value	1
Max. value	200
# of collaborators (C) per firm : Average	29,03
Median	15,00
Modal value	15,00
Min. value	1
Max. value	171
# of times a collaboration is repeated between F and C : Average	2,40
Median	1,00
Modal value	1,00
Min. value	1,00
Max. value	69,00

The collaborating institutions are different kinds of organisations worldwide (34 countries in total): universities (35%), hospitals (27%), research institutions including science parks (19%) and firms (18%), mainly of big or medium size. We consider the network of co-authoring institutions (C from now on) a network for exploration.

Obviously, the importance of the different kinds of collaborating institutions varies if we consider, instead of the sheer number of entities belonging to each category, the number of relationships or rather the number of collaborations that the entities of each category realize with F. Thus, when the number of relationships is considered, the role of universities results augmented, since they spawn 43,5% of total links, mainly at the expense of firms (table 3 and fig. 2).

Table 3. Weight of the various types of collaborating institutions in the relationships, collaborations and repeated collaborations with the Italian biotech firms of the sample

	# of relationships	Shares %	# of collaborations	Shares %	# of repeated collaborations	Shares %
Firms	115	12,4%	244	10,9%	165	9,9%
Research institutes	179	19,3%	400	17,9%	286	17,2%
Hospitals	231	24,9%	471	21,1%	318	19,2%
Universities	403	43,5%	1.115	50,0%	890	53,6%
Total	928	100,0%	2.230	100,0%	1.659	100,0%

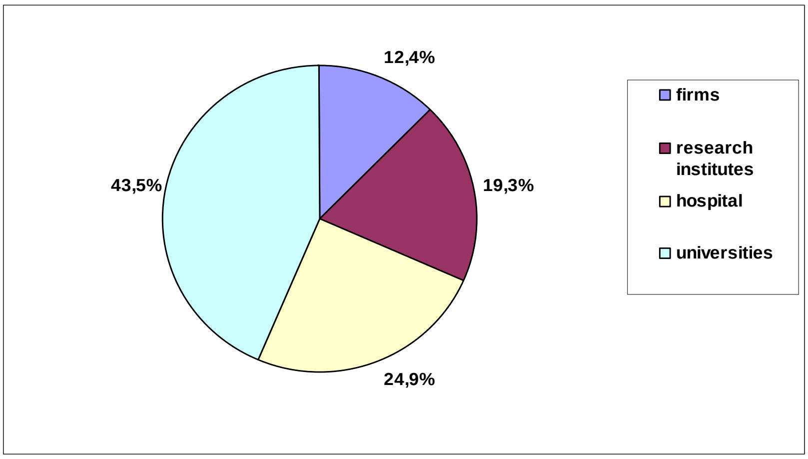

Figure 2. Weight of the various types of collaborating institutions in the relationships with the Italian biotech firms of the sample

Repeated relations amount to 74,4% of collaborations (fig. 3). Most collaborations are repeated between 2 and 5 times (34% of the total), but also the share of "strong ties", that is of collaborations repeated more than 10 times, is relevant (25%). According to the literature on INs, recurring ties are crucial in order to create a solid network.

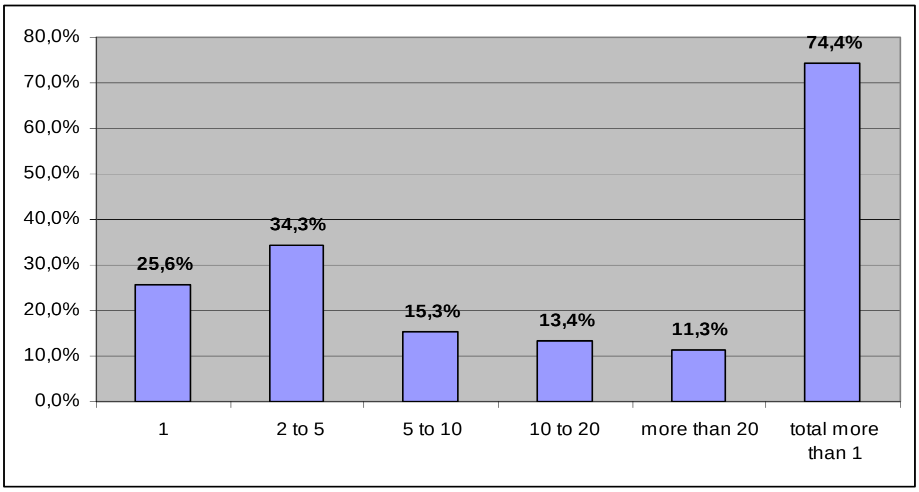

Figure 3. Distribution of collaborations according to the number of times they are repeated

Focusing again on universities, we find that 53.6% of F's repeated collaborations involve universities (table 3), a percentage much higher than with other types of collaborators. Moreover, as table 4 shows, about 30% of universities' collaborations are repeated more than 10 times.

Thus, we have a clear evidence of the major role played by universities in the creation and dissemination new knowledge in collaboration with biotech companies. Not only the number of universities collaborating with F is higher than that of the other types of entities, but their collaboration is also both wider (more links) and more intense (more repeated links). It is also possible to observe that some universities play a polar role in the network of F, "guiding" the firms inside the network (Aygodan and Lyon, 2004), such as University Federico II of Naples, University San Raffaele of Milan, University Statale of Milan, University La Sapienza of Rome and University of Siena.

Table 4. Distribution of the various types of C according to the number of collaborations

# of collaborations	Firms	Research institutes	Hospitals	Universities	Total
1	32,4%	28,5%	32,5%	20,2%	25,6%
2 to 5	31,1%	38,5%	36,7%	32,5%	34,3%
5 to 10	4,9%	21,5%	7,9%	18,6%	15,3%
10 to 20	10,2%	4,0%	8,3%	19,6%	13,4%
More than 20	21,3%	7,5%	14,6%	9,1%	11,3%
Total	100,0%	100,0%	100,0%	100,0%	100,0%

4.2.Cognitive distance

In the great majority of the cases firms and their co-publishing partners have a very high cognitive distance (CD). The mean CD is 88,4 measured on a scale of 100 and the most common value of CD is 100, meaning that the co-publishing partners do not share any cognitive field. While this might be considered evidence that the advantage for each partner of accessing a different specialization greatly outweighs the costs of communicating between partners, such an interpretation would not be warranted without taking into account the factors which could have affected our measurements. We know that our results could have been affected by (i) the method used to measure CD, (ii) the way in which the cognitive fields constituting the knowledge bases of the co-publishing partners are classified, (iii) the fact that the expected CD for co-publications is not necessarily the same as for other types of collaboration.

As we previously pointed out, there are several possible measures of cognitive distance. Not all of these measures are suitable for every data set. The measure we chose (PR) is suitable for a data set in which the KB vector(s) of each collaborating partner tell us about the presence or absence within the KB of the partner of the set of specializations characterising altogether the group of collaborators studied. Furthermore, we know that in its previous use this distance measure turned out not to be adversely affected by the presence of many zeros (absence) and of few ones (presence) in the data (Pielou, 1984). Although we cannot have the absolute certainty that the cognitive distances we measure are the 'true' ones, we can still expect that the high values we generally find are not an artefact of our method: a simple visual inspection of the data matrix displaying the competences possessed by all the co-publishing partners show that in the vast majority of the cases they have no competence in common. Thus, we consider the result obtained a realistic representation of the sample studied.

Another source of influence on the measured values of cognitive distance is the system used to classify fields of competence. Any such classification system is by definition hierarchical, in the sense that it includes competences at different levels of aggregation. Within each field of competence we can usually identify several competencies at a lower level of aggregation. We can expect cognitive distances and costs of communicating specific knowledge to depend strongly on the level of aggregation used. We can expect cognitive distances and communication costs to rise with a growing level of aggregation. To put it differently, the cognitive distances within a group of technological fields at a given level of aggregation (intra-group distances) can be expected to be generally smaller than the distances between two groups of technological fields at a higher level of aggregation (inter-group distances). For example, if two potential partners having competencies in biotechnology and in electronics attempt to collaborate they are likely to face much higher barriers than two partners having competencies in two different classes of biotechnology. We can observe that all the competencies included in our sample are medical ones, sharing a non negligible part of concepts and theories. Furthermore, most of the co-publishing firms in our sample are highly specialized and their KB contains a very small number of competencies. Even in the case of large or very large co-publishing organizations, such as universities or hospitals, the collaboration occurs with a very small subset of the organization having very specialized competencies. Thus, in general we can expect that the very high cognitive distances we observe occur for a relatively low level of aggregation. Our co-publishing partners can share a lot of knowledge and differ in a very limited subset of their KB. We could say that the lower the level of aggregation at which we measure cognitive distance the more *local* this measure is, in the sense that it indicates the relative values of the cognitive distances *within* a group of fields of knowledge at a low level of aggregation. If we wanted to find an absolute measure of cognitive distance encompassing all levels of aggregation we would need to calibrate it with respect to the maximum possible cognitive distance between any pair of cognitive fields or subsets of knowledge. Such a measure is for the moment impossible to carry out. The local measure of cognitive distance we propose is still useful since many technological alliances occur by combination of different but not too different fields of specialization.

The third factor potentially affecting the measured values of CD is the type of collaboration. A co-publication does not necessarily involve a very heavy or highly irreversible commitment although repeated co-publications indicate a more durable relationship. In the case of biotechnology firms, co-publications are clearly not the final outcome of their activity. We can envisage co-publications being used as a way of learning previous to the realization of a project involving some marketable outcome or as the side result of the search for such a marketable outcome. Especially in the former case we can expect the shared experiences and knowledge required for a successful collaboration to be more limited than for the joint realization of a marketable outcome. In other words, we can expect the average cognitive distance involved in co-publishing to be different and possibly higher than for the joint creation of a new drug or of a new plant variety. Of course, this is more an hypothesis to be tested than an accepted result which can be used as an explanation of new findings. However, it is highly plausible for the average CD observed for co - publications to differ from those which can be observed for other types of collaboration. At the moment we do not have the data to test such an hypothesis, but we hope to acquire them in the foreseeable future.

In summary, the dominance of large CDs in our sample of co-publications is likely to reflect the high degree of specialization of co-publishing partners and the high degree of 'local' differentiation of their knowledge, which is compatible with a very large extent of shared knowledge which allows them to communicate across the cognitive distance observed. This would be possible if the CD measures were intra-group, within the same field of knowledge rather than CDs between two very different and highly aggregate fields. It is possible that the high observed values of CD are specific to co-publications as opposed to other types of collaboration. The last sentence cannot be considered a result but rather an interesting hypothesis emerging in the interpretation of our results and worth of further investigation.

The cognitive distances observed seem to vary slightly with the frequency of co-publication and with the type of partner. Repeated collaborations are more frequent than occasional ones for very high cognitive distances but less frequent for low cognitive distances (Table 5 and Fig 4). For intermediate values of cognitive distances repeated and occasional collaborations are almost equally frequent. Observed cognitive distances seem to be affected also by the type of partner. The percentage of high CDs is very high for all types of partners but it grows from 70% to 80% in the order (Firms-Hospitals-Research Institutes-Universities). In particular Research Institutes and Universities have 80% of their co-publications corresponding to the maximum cognitive distance. Although our paper is a first exploration of these issues, and as a consequence our interpretations are rather tentative, the order in which the percentage of high cognitive distances varies is not implausible. We can expect the role of cognitive distances to vary according to the type of collaboration. For example, we would expect to find higher cognitive distances for alliances aimed at creating new knowledge than for those which produce jointly a new pharmaceutical product. If this were the case, and we have to bear in mind that the study of cognitive distances is an extremely new subject especially for what concerns their empirical analysis, then it is not too surprising to find that Universities and Research Institutes tend to have higher cognitive distances than Hospitals and Firms since we can expect the creation of new knowledge to be a more frequent objective in the alliances involving the first two types of partners than the last two ones.

Table 5. Cognitive distance and number of repeated relationships

Cognitive Distance	Not repeated relationships		Repeated relationships		2 to 5 times		6 to 10 times		More than 10 times		Total	
Value	N.	%	N.	%	N.	%	N.	%	N.	%	N.	%
0	36	6,3%	14	3,9%	8	2,8%	4	8,7%	2	7,1%	50	5,4%
33,3	3	0,5%	-	-	-	-	-	-	-	-	3	0,3%
50	43	7,5%	19	5,3%	17	6,0%	0	0,0%	2	7,1%	62	6,7%
66,7	13	2,3%	16	4,5%	11	3,9%	3	6,5%	2	7,1%	29	3,1%
75	16	2,8%	4	1,1%	4	1,4%	-	-	-	-	20	2,2%
80	25	4,4%	12	3,4%	10	3,5%	2	4,3%	-	-	37	4,0%
83,3	2	0,4%	-	-	-	-	-	-%	-	-	2	0,2%
87,5	9	1,6%	1	0,3%	1	0,4%	-	-	-	-	10	1,1%
100	424	74,3%	291	81,5%	232	82,0%	37	80,4%	22	78,6%	715	77,0%
Total	571	100%	357	100%	283	100%	46	100,0%	28	100%	928	100%

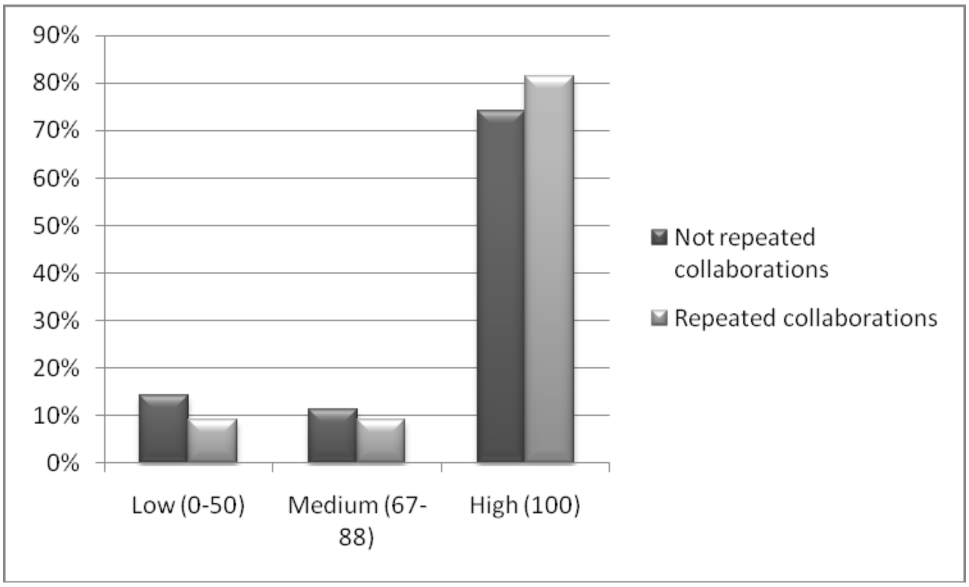

Figure 4. Cognitive distance among F and C in repeated and in not repeated relations

The values of cognitive distance that emerge in this investigation seem quite high with respect to what one should have expected from the literature on optimal cognitive distance, according to which a median value between 0 and 100 should have prevailed (Nooteboom, 2000; Brown and Duguid, 1991, Cohendet and Llerena, 1997). The above results can be justified by two types of considerations. First, the network we are examining is not the full network of collaborations established by our sample of firms. It comprises only those relations among actors who conduct research together, and are thus co-authors of the resulting publication. In order to conduct research together, the individuals of the institutions involved in general need to understand each other quite well. A high absorptive capacity of the knowledge of the collaborators is a precondition of these types of collaboration. Moreover, we have to consider that biotechnology is a scientific field characterised by a high specialisation of actors and of their research projects. One should than expect partners' selection in biotechnology research networks to be based on the sharing of the same experiences and technologies, in order to benefit from the results of learning processes realized by other members of the epistemic community. As we already pointed out, the cognitive distances we measure are 'local' or intra-group, in the sense that they tell us the *relative* values of the differences in knowledge of partners the KBs of which are located within the same field of knowledge. This means that the collaborating partners can still have high observed cognitive distances while sharing a large part of their KB and thus having a reasonable absorption capacity for each other's knowledge. Thus, we can expect to observe large cognitive distances if the collaborations occur over a range of knowledge which is large enough to justify the investment required for its study but not so large as to involve very low absorption capacities of the partners. Furthermore, the risk of lock in might be of very little weight in the production of new knowledge. Clearly, the choice of partners in networks for exploitation might follow quite a different logic.

Table 6. Cognitive distance and number of relations by kind of collaborating institutions

Cognitive distance	Firms		Hospitals		Research insitutions		Universities	
Values	N. relations	%	N. relations	%	N. relations	%	N. relations	%
0	7	6,1%	9	3,9%	13	7,3%	21	5,2%
33,3	2	1,7%	1	0,4%	-	-	-	-
50	10	8,7%	17	7,4%	11	6,1%	24	6,0%
66,7	4	3,5%	10	4,3%	6	3,4%	9	2,2%
75	3	2,6%	6	2,6%	2	1,1%	9	2,2%
80	5	4,3%	16	6,9%	2	1,1%	14	3,5%
83,3	-	-	1	0,4%	1	0,6%	-	-
87,5	4	3,5%	-	-	2	1,1%	4	1,0%
100	80	69,6%	171	74,0%	142	79,3%	322	79,9%
Total	115	100%	231	100%	179	100%	403	100%
Average	85,29		88,59		87,81		89,69	

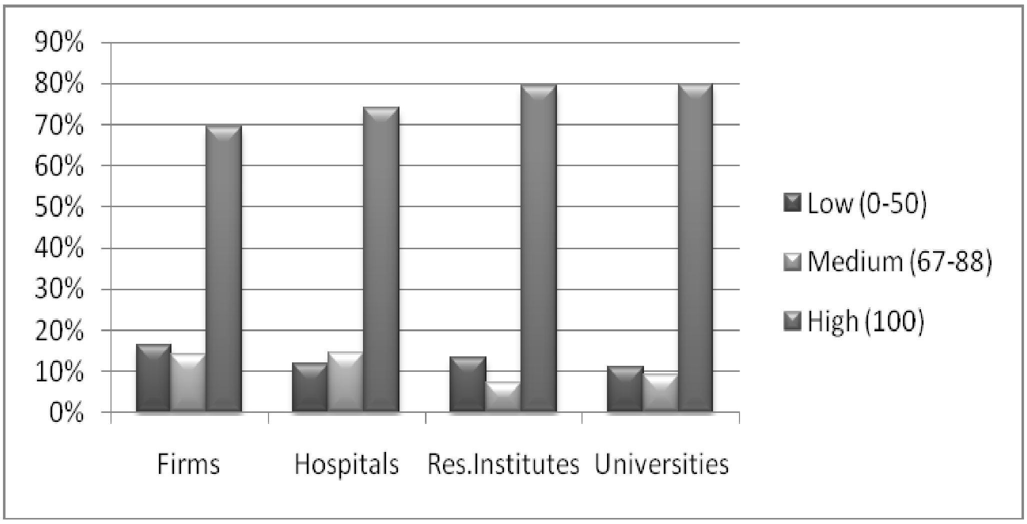

Figure 5. Cognitive distance and kind of collaborators

4.3.Geographic distance

With regard to geographic distance, 75,4% of collaborations have been established with partners located within Italy (table 7). In more detail, 33,6% of collaborations are established with collaborators located in the same Italian region and 12,6% in the same macro region, while 29,2% are established with partners located in rest of Italy. Outside Italy an almost equal percentage of collaborations exists with partners located in the rest of the world outside Europe and with European partners. In particular, 218 collaborations occur with partners located in the United States[2].

Table 7. Geographic distance

Geographic distance	N. of relationships	%	N. of collaborations	%
Same italian region	196	21,1%	749	33,6%
Same Italian macroarea, but outside the region	120	12,9%	281	12,6%
Rest of Italy	281	30,3%	651	29,2%
Total Italy	597	64,3%	1.681	75,4%
Europe	176	19,0%	270	12,1%
Rest of the world	155	16,7%	279	12,5%
Total	928	100,0%	2.230	100,0%

Since almost one half of the collaborations occur in Italy within the same region or the same macro area, local innovation systems seem to have a major influence on the creation of collaborations, in accordance with the literature (Cooke, 1992;1998; Storper and Venables, 2005). However, a non negligible percentage of collaborations occurs with Italian partners located outside the same region or the same macro area. Thus, while it appears that regional embeddedness does not limit the search for a research partner, still the fact that collaborations with Italian partners account for two thirds of the total seems to indicate that geographical distance and cultural proximity are factors affecting the formation of these alliances. Moreover, the importance of regional embeddedness and clusterization of biotech activities appears much more clearly when repeated partnerships are considered. In fact, the average frequency of collaboration rises when the geographic distance of the partners falls(table 8). For example, the share of collaborations within the same macro-area rises from 30% of the total for once only collaborations, to 60% for collaborations repeated from 6 to 10 times, and to 68% for collaborations repeated more than 10 times.

Table 8. Geographic distance and collaborations

Geografic distance	Distribution by number of collaborations							
	1		2 to 5		6 to 10		More than 10	
	N .	%	N .	%	N .	%	N .	%

2 The country ranking second among collaborators is Japan, with 19 collaborations. The rest of collaborations are spread among a vast number of countries of all continents.

A) Same italian region	94	16,5%	184	24,1%	161	47,1%	309	56,0%
B) Same Italian macroarea, but outside the region	75	13,1%	97	12,7%	44	12,9%	65	11,8%
C = A+B	**169**	**29,6%**	**281**	**36,8%**	**205**	**59,9%**	**374**	**67,8%**
Rest of Italy	168	29,4%	265	34,7%	74	21,6%	143	25,9%
Total Italy	**337**	**59,0%**	**546**	**71,6%**	**279**	**81,6%**	**517**	**93,7%**
Europe	125	21,9%	124	16,3%	21	6,1%	-	-
Rest of the world	109	19,1%	93	12,2%	42	12,3%	35	6,3%
Total	**571**	**100%**	**763**	**100%**	**342**	**100%**	**552**	**100%**

In conclusion, the previous results confirm the importance of geographic proximity. After all, 75% of the collaborations occur in Italy in a technological field which is widely, although not uniformly, internationally distributed. However, they do not imply that this is the only factor affecting technological collaboration in biotechnology. In fact, the existence of collaborations located in the USA, Japan, Canada or Australia suggest that a crucial factor is likely to be the distance with respect to the technological frontier of the time. In biotechnology and medical research the frontier is located in the USA (Dosi, Llerena, Sylos Labini, 2006) with other important organizations being located in Canada or Australia. Thus Italian biotechnology firms will opt for local knowledge whenever that is available, but will go anywhere to obtain knowledge which is scarce or unavailable locally. We can then expect a bimodal geographic distribution of collaborations with many located over short geographic distances and few others occurring at much larger distances. Of course the two types of collaboration are not equivalent. The local ones may be aimed at solving recurrent problems which need continuous consultation, as shown by the very high contribution of local alliances to repeated co-publications. On the other hand, the more expensive collaborations with a very distant partner will be used to acquire very scarce but very important knowledge. Similarly, the attractiveness of particular, 'catalyst', institutions (Aygodan and Lyon's, 2004) could explain also the collaborations with Italian universities and research institutes located outside the same region or macro area. Summarizing, direct and continuous interactions and exchanges of complementary and similar knowledge, on which successful of partnerships are based, are easier in geographic proximity, but when locally unavailable knowledge becomes crucial, it does not matter how far the partner is located. In other words, what we suggest is that even though geographic and cultural distance are likely to be barriers to collaboration, representing a "cost", they can be compensated by the benefits arising from collaborating with particularly interesting partners.

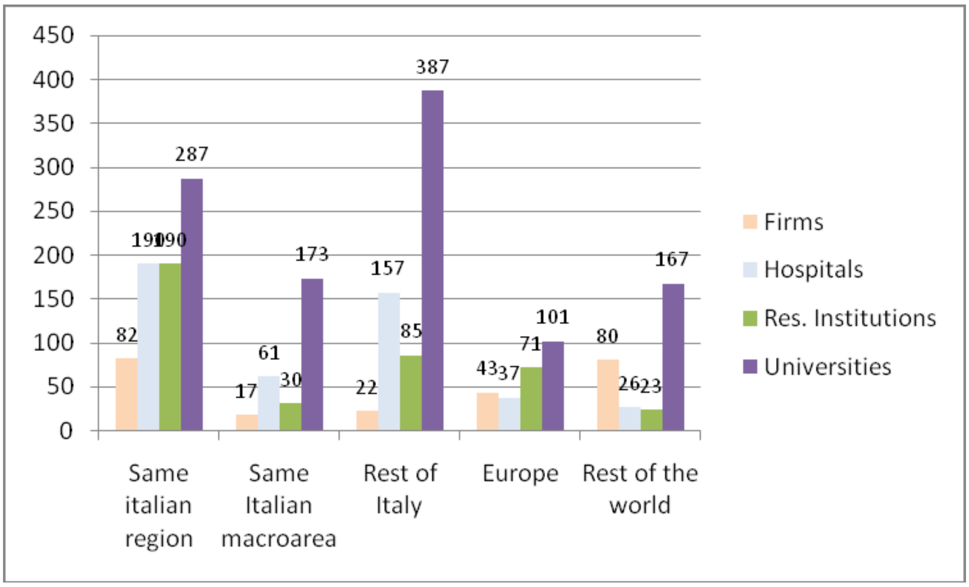

Figure 6. Number of collaborations across various distances subdivided by kind of collaborating institutions

Table 9. The collaborations of the sample of Italian biotechnology firms with different kinds of institutions distributed by geographic distance

	Firms	Research insitutions	Hospitals	Universities	Total
A) Same italian region	33,6%	47,8%	40,3%	25,7%	33,6%
B) Same Italian macroarea, but outside the region	7,0%	7,5%	13,0%	15,5%	12,6%
C = A+B	**40,6%**	**55,3%**	**53,3%**	**41,2%**	**46,2%**
Rest of Italy	9,0%	21,3%	33,3%	34,7%	29,2%
Total Italy	**49,6%**	**76,6%**	**86,6%**	**75,9%**	**75,4%**
Europe	17,6%	17,8%	7,9%	9,1%	11,3%
Rest of the world	32,8%	5,8%	5,5%	15,0%	13,3%
Total	100,0%	100,0%	100,0%	100,0%	100,0%
Total number	244	471	400	1.115	2.230

The non equivalence of the collaborations with different types of partners is confirmed by the different distribution of co-publications by type of partner in different geographical areas. While in any area universities are the most important partner, it is outside of the same region that collaborations with universities are particularly important, especially in the rest of Italy and the rest of the world. The collaborations with research institutions and hospitals are instead concentrated in the same Italian region. Finally, collaborating firms play an especially notable role among the partners from abroad, both from Europe and from the rest of the world (table 9 and fig. 6).

At a first inspection these findings might seem to partially contradict what we could have expected from the literature, which argues that geographic proximity with universities is fundamental in terms of localised knowledge spillovers for biotech firm (Oerlemans, et al, 2000, Vedovello, 1997). A number of factors could affect these apparently divergent results. First, we should remember that, with the exception of Molmed, linked with San Raffaele university, the 32 Italian biotech firms of our sample were not born as university spin-offs.

Thus, to establish contacts with local rather than other Italian universities would not have been the obvious first choice for them. The important role of local research institutes can probably partly explain the relatively limited use of local universities in co-publications. Amongst the types of co-publishing partners research institutes are the closest to universities in terms of objectives and procedures. In other words, although not identical, they are reasonably close substitutes. Thus, if the local research system contains a high percentage of research institutes it is not surprising that they account for a significant share of co-publications. One should not forget that national research systems differ considerably with respect to the relative importance of universities and research institutes: for example, research institutes are less important in the USA and in the UK than in France or Germany. Thus, the important role of local research institutes in co-publications is either due to the local organisation of research or to their intrinsic characteristics (e.g. pattern of specialisation, more applied research then universities etc).

With regard to collaborations with firms, the very high frequency of co-publications with firms in the rest of the world and in Europe indicates that the firms of our sample, when looking for partners with an objective closer to exploitation than exploration, have to resort to foreign firms, for the simple reason that in Italy domestic big pharmaceutical firms are simply non existent

This finding might be due to the situation of biotechnology in Italy. As already pointed out, modern biotechnology was created in the USA where the frontier of knowledge is still located. During the 1970s and 1980s DBFs were an almost uniquely USA phenomenon. In Europe the number of DBFs only started rising substantially during the 1990s. In Italy the rapid growth in the number of DBFs only started growing in the 2000s (Blossom Associati, 2008). Thus, if in biotechnology the most advanced European countries were latecomers, Italy was a late latecomer. This is important because the strategies required to enter an industry are likely to vary according to the period of its life-cycle. Thus, in the USA, during the 1970s and the 1980s when modern biotechnology was in its infancy, a very high percentage of DBFs were founded by scientific entrepreneurs (Zucker et al, 1998; Audtretsch and Stephan, 1996, Oliver, 2004), while this does not seem to be the case at all in Italy, where most entries started taking place in the 2000s when at least some subsets of biotechnology were already maturing.

With respect to this point we have to bear in mind that in biotechnology we can identify two generations linked to recombinant DNA and monoclonal antibodies and to genomics respectively (Saviotti, Catherine, 2008). Within the first generation R&D collaborations had virtually disappeared by the end of the 1990s, while marketing agreements continued. If we assume that the end of R&D collaborations in the first generation implies the onset of maturation of this subset of biotechnology, it follows that only the second generation linked to genomics still represented the frontier by the beginning of the 2000s. This situation would open possibilities for different types of collaboration in the two generations. In particular we might expect a late latecomer like Italy to opt for incremental innovations in the more maturing subset of biotechnology, that is in the first generation, together with local partners and to participate in innovations near the technological frontier of the time with advanced foreign partners.

In order to test the proposition that the collaborations with non Italian, and in particular with Rest of the World, co-publishing partners are different from those with Italian partners we constructed a list of the journals in which such co-publications appeared and of the themes describing the co-publications. We found no differences when comparing the lists of co-publications of the firms of our sample with other firms, research institutes, hospitals or universities without taking into account their geographical location. On the other hand, when we compared the distribution of the co-publications of the firms of our sample with Italian partners to those with non-Italian partners (external co-publications) by means of the key words associated with each publication we found that they differed considerably. Some examples of these differences are shown in Table 10. If we bear in mind that the percentages of each keyword in the total of Italian and of external co-publications represent the pattern of specialisation of each set of co-publications, we can see that the Italian and external co-publications specialise in different subsets of knowledge. For example, the most common keyword associated with all Italian co-publications, haematology, accounts for 16.0 percent of all Italian co-publications but only for 3.8 percent of the external ones. On the other hand, developmental biology accounts for 4.0 percent of external co-publications but for 0.45 percent of the Italian ones and behavioural sciences which account for 1.1 percent of external co-publications, are absent in the Italian ones. Thus, Italian biotechnology firms use co-publications with distant partners to acquire types of knowledge different from those which they can obtain by co-publishing with close by partners.

Table 10. Pattern of specialisation of the co-publications of the firms of our sample with Italian and with non-Italian partners, as detected by the keywords associated with the articles

Field of Knowledge (Keyword)	Percentage of Italian co-publications	Percentage of non-Italian co-publications
Haematology	16.0	3.8
Biochemistry and molecular biology	12.6	25.0
Neurosciences	4.8	7.7
Developmental biology	0.45	4.0
Behavioural sciences	0	1.1

Only the examples showing the greatest differences are displayed in this table. The complete list of co-publications is shown in tables A.1 and A.2 in the appendix.

Table 11. The relationship between geographic and cognitive distance for the firms of our sample and their collaborators

Geographic location of	N.Partners	Cognitive Distance		
		0-50	67-88	100

Universities				
Italy	236	9%	9%	82%
Europe	80	14%	6%	80%
RoW	87	15%	9%	76%
Firms				
Italy	51	20%	8%	73%
Europe	33	18%	9%	73%
RoW	31	10%	10%	81%
Research Institutes				
Italy	122	11%	8%	80%
Europe	36	11%	11%	78%
RoW	21	5%	10%	86%
Hospitals				
Italy	188	11%	19%	71%
Europe	27	15%	11%	74%
RoW	16	19%	6%	75%

Table 11 shows that there is no direct relationship between geographic and cognitive distance. Such lack of a direct relationship could have been expected from our previous discussion. As we previously pointed out, the choice of foreign, or, more generally, a geographically distant partner is likely to be dictated by the proximity of this partner to the technological frontier of the time in the desired sub-field of knowledge. The distance from the technological frontier of the time needs to be clearly distinguished from the cognitive distance that we measure. To understand the difference between the two we can imagine to represent the different sub-fields of knowledge in which we are interested on an horizontal axis ranking them in order of growing dissimilarity (or of growing cognitive distance) and of representing on a vertical axis the technological capabilities of each country or organisation in each sub-field of knowledge on a scale ranging from zero (0) to 100, where 100 would be the frontier. First, we cannot expect any direct relationship between cognitive distance (CD) and distance from the technological frontier (DF). We can only expect firms and research organisations in an imitating country to choose at least some of their partners in the country and in the organisations which are as close as possible to the technological frontier of the time in the desired sub-field of knowledge. The lack of a direct relationship between geographic and cognitive distance follows from the lack of a corresponding relationship between cognitive distance and distance from the technological frontier of the time.

5. Summary and conclusions

In this paper we studied the influence of cognitive and of geographical distance on the collaborations of Italian biotechnology firms. The paper has an important methodological component in that there is no agreed procedure for the calculation of cognitive distances. Furthermore, while measures of similarity exist which could be used if we had data on the patents and on the technological classes of the firms of our sample (see for example the measure of similarity developed by Jaffe, (1986), very few of the firms of our sample have a number of patents sufficient to use these measures. While this might seem a very serious limitation, we have tried to overcome it, because the case of biotechnology firms having no or very few patents is by no means rare. Thus our paper (i) develops an original measure of cognitive distance and (ii) it does so without using data on patents. In this paper we study the co-publications of a sample of 32 Italian biotechnology firms with partners located in different geographical areas and classified by type of organisation. Amongst the co-publishing partners of our firms we distinguish universities, research institutes, hospitals and firms. Such co-publishing partners are located all over the world although their distribution is not uniform. Using various sources of data we have been able to assign to each of our firms and of their co-publishing partners some field of knowledge on the basis of which we have been able to construct a competence vector for each co-publishing partner. This vector gives information about the presence or absence of the fields of knowledge of a given range in the knowledge base of each of the firms and organisations studied. We considered a large number of measures of similarity and of distance used mostly by ecologists to measure the similarity of animal species, since their objective is identical to ours, once we substitute biological traits with fields of knowledge. Amongst the various measures of distance available we chose one called percentage remoteness (PR), because it was the most appropriate for a data set (i) containing information about the presence or absence of given technologies and (ii) in which the fields of knowledge vectors of each firm and organisation studied contain many zeros (absences) and few ones (presences). The results of our calculations show that most co-publications have a high cognitive distance, the average for the whole set being 88.4 out of a maximum of 100. In general we can expect firms collaborating to acquire new knowledge to choose partners with a knowledge base different from theirs and closer to the target knowledge they want to acquire. However, while the advantage of the collaboration for the learning firm may be expected to increase with the cognitive distance, collaboration costs may be expected to rise in the same direction. Thus, one should expect the observed cognitive distances to reflect a trade off between advantages and costs of collaboration (Nooteboom, 2000). According to this argument, it may seem that the cognitive distances we measure are too high. However, any measure of cognitive distance depends on the level of aggregation at which fields of knowledge are defined. When the fields of knowledge are defined at a very low level of aggregation, firms and organizations can specialise in a set of fields different from that of any of their partners while sharing with them a wide range of knowledge. In other words, any measure of cognitive distance will always be 'local' in the sense of measuring distances as a percentage of the maximum possible within a narrow range of knowledge. Absolute measures of cognitive distance could only be calculated for a set including fields of knowledge at all possible levels of aggregation. We conclude that the cognitive distances we observe are large because the organizations we study are highly specialized within a narrow range of knowledge, which allows them to collaborate with partners sharing a lot of background knowledge but having competencies different from theirs.

The cognitive distances we observe vary, although not a lot, with the type of co-publishing partner. Collaborations with partners with the highest cognitive distances are particularly frequent with universities and research institutes, followed by hospitals and firms in the order. Thus, high cognitive distances seem to be more frequent the more the collaborating organisations are exploration oriented and to fall the more they are exploitation oriented. The distribution of co-publications by geographical area shows that about two thirds of the co-publications are with Italian partners and almost one half with partners from the same region or macro-area within Italy. Amongst the co-publications with non Italian partners, those with the rest of the world are slightly more numerous than the ones with European partners. Perhaps the most interesting results of our analysis have been obtained by studying jointly the distributions by kind of partner and geographical distances. The distribution of co-publications by type of partner varies considerably in different geographical areas. While in any area universities are the most important partner, hospitals and research institutes are especially important in the same Italian region, universities in the other Italian regions and abroad, and firms in the rest of the world. These differences are likely to reflect the organization of the Italian research system, where also research institutes are an important actor in applied research. However, we interpreted these variations as arising from the different roles which can be played by collaborations with close by or with very distant partners. The geographical distribution of competencies in biotechnology is by no means uniform. For example, biotechnology firms, universities and research institutes in the USA can be expected to be much closer to the technological frontier than Italian ones. The distance of potential collaborators from the technological frontier of the time is likely to be a very important factor affecting the choice of partners by biotechnology firms. Such choice is likely to be dictated by the balance between costs and benefits of the collaboration. Collaborations with partners located in very distant geographical areas but very close to the technological frontier will in general be more expensive but impossible to replace with local collaborations, given the scarcity of the knowledge required. On the other hand, local collaborations can be used to acquire or improve knowledge required in the everyday practice of research and development. Thus, in general we cannot expect alliances with local or with far away partners to be substitutable but to play a systematically different role. To test this idea we compared the co-publications of the firms of our sample with Italian and with non Italian partners. To do this we calculated the fraction of co-publications corresponding to the keywords associated with each publication in the two groups. The distribution of co-publications by keyword, which represents the relative patterns of specialisation of the two groups, is considerably different for Italian and non Italian collaborations (Table 10). Thus, Italian biotechnology firms use alliances with Italian and non Italian partners to look for different types of knowledge.

The previous results show that in order to understand technological alliances in biotechnology (but equally in other high technology sectors) we have to take into account not only cognitive and geographical distances but also the distance from the technological frontier of the time. As a consequence the distribution of the alliances of the biotechnology firms based in a given country will depend on the distance of the country with respect to the technological frontier. Firms based in a country far behind the technological frontier of the time in biotechnology will either need to position themselves in already maturing subsets of biotechnology or to have a fraction of their alliances in countries on or near the technological frontier.

In an IB literature framework, the working of the studied INs recalls the hypothesis of the knowledge seeking internationalisation (Cantwell, 1989), relying on collaborations to reach the needed knowledge on a global scale, selecting countries (innovation systems) and partners on the basis of knowledge needs and trading off cognitive distances. Our findings also support the hypotheses of firm internationalisation driven by the need to access to global scientific reservoirs, recently highlighted by Kafouros, Buckley and Clegg (2009).

The study of the co-publications of a sample of Italian biotechnology firms which we described in this paper gives some interesting results but raises a number of issues for further investigation. First, since the measure of cognitive distances we proposed in this paper is not the only possible one, other measures should be tested and compared to the one we used. Second, the results obtained for co-publications should be compared to those obtained for different types of technological collaborations, for example those aimed at the joint creation of a new drug. Furthermore, the previous results suggest that technological collaborations can evolve during the life cycle of the technology considered as it diffuses from the originating country to imitating countries behind the technological frontier. Thus, the mechanisms of technological alliances in biotechnology should be compared for different countries and in different periods of the life cycle of the technology.

Acknowledgments

We are grateful to Roger Coronini for his help in identifying and classifying the co-publications of our Italian biotechnology firms.

The authors would like to acknowledge the financial support of the Italian Ministry for Education, Universities and Research (FIRB, Project RISC - RBNE039XKA: "Research and entrepreneurship in the knowledge-based economy: the effects on the competitiveness of Italy in the European Union"

References

Acs, Z., L. Anselin, and A. Varga. 2002. Patents and innovation counts as measures of regional production of new knowledge. *Research Policy* 31: 1069-1085.

Ahuja, G. 2000. Collaboration networks, structural holes and innovation: a longitudinal study. *Administrative Science Quarterly* 45, no.3: 425-455.

Amin, A., and N. Thrift. 1993. Globalization, institutional thickness and local prospects. *Revue d'économie régionale et urbaine* 3: 413-418.

Arundel, A. and A. Geuna. 2001. Does proximity matter for knowledge transfer from public institutes and universities to firms?. *SPRU, University of Sussex*.

Asheim, B., and A. Isaksen. 1997. Location, agglomeration and innovation: towards regional innovation systems in Norway. *European Planning Studies* 5, no.3: 299-330.

Audretsch, D. and Stephan P. 1996. Company-scientists locational links: the case of biotechnology. *American Economic Review* 86: 641-652.

Aydogan, N. and T. P. Lyon. 2004. Spatial proximity and complementarities in the trading of tacit knowledge. *International Journal of Industrial organisation* 22: 1115-1135.

Balconi, M., A. Pozzali, and R. Viale. 2007. The "codification debate" revisited: a conceptual framework to analyze the role of tacit knowledge in economics. *Industrial and Corporate Change* 16, no. 5: 823–849.

Bellet, M., G. Colletis and Y. Lung. 1993. Introduction au numéro spécial 'Economie de proximités'. *Revue d'Economie Régionale et Urbaine* 3 : 357-61.

Blossom Associati. 2008. *Biotecnologie in Italia 2008. Analisi strategica e finanziaria*, Blossom Associati, Milano.

Boschma, R. 2004. Does geographical proximity favour innovation? *The 4th congress on Proximity Economics.* Marseilles, France.

Breschi, S. 1999. Economie di agglomerazione e concentrazione geografica delle attività innovative: analisi del caso italiano. *Working Paper LIUC.*

Breschi, S., F. Lissoni and F. Malerba. 2003. *Knowledge networks from patent citations? Methodological issues and preliminary results*. DRUID Summer Conference, Copenhagen.

Brown, J. S., and P. Duguid**.** 1991. Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science* 2, no.1: 40–57.

Cantwell, J. 1989. *Technological Innovation and Multinational Corporations.* Basil Blackwell, Oxford, U.K.

Chiesa V. 2003. *La bioindustria.* Milano: Etas.

Cohendet, P., and P. Llerena. 1997. Learning, technical change and public policy: how to create and exploit diversity. In *System of innovation. Technologies, institutions and organizations*, ed. C. Edquist: 223-234, London: Francis Pinter,

Cooke, P. 1992. Regional Innovation Systems: competitive regulation in the Europe, *Geoforum* 23: 365-382.

Cooke, P. 1998. Introductions: origins of the concept. In *Regional Innovation Systems*, ed. H. Braczly, P. Cooke and M. Heidenreich: 2-25, London, UCL Press.

Cooke, P. 2002. Biotechnology clusters as regional, sectoral innovation systems. *International Regional Science Review* 25, no. 1: 8-37.

Dosi G., P. Llerena and M. Sylos Labini. 2006. The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called European Paradox. *Research Policy* 35, no.10: 1450-1464

Pielou E.C. 1984. *The Interpretation of Ecological Data*, New York: Wiley.

Feldman, M. P. 2000. Location and innovation: the new economic geography of innovation, spillovers and agglomeration. In *Oxford handbook of economic geography*, ed. G. Clark, M. P. Feldman and M. Gertler: 373-394. Oxford: Oxford University Press.

Fontana, R., and L. Nesta. 2007. Product Innovation and Survival in a High-Tech Industry. CESPRI Working Papers 208, Università Bocconi, Milano.

Freeman, C. 1991. Network of innovators: a synthesis of research issues. *Research Policy* 20: 499-514.

Gallaud, D., and Rherrad, I. 2002. Knowledge production and patterns of proximity. Paper presented at the DRUID Summer Conference on '*Industrial Dynamics of the New and Old Economy - Who is embracing whom?'*, Copenhagen/Elsinore 6-8 June 2002.

Gertler, M.S., D.A. Wolfe, and D. Garkut. 2000. No place like home? The embeddedness of innovation in a regional economy. *Review of International Political Economy* 7: 1-31.

Gilsing, V., and B. Nooteboom. 2005. Density and strength of ties in innovation networks: an analysis of multimedia and biotechnology. *European Management review* 2: 179-197.

Gilsing, V., B. Nooteboom, W. Vanhaverbeke, G. Duysters, and A.van den Oord. 2008. Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy* 37: 1717–1731.

Haas, S. 1995. *Economies externes technologiques, apprentissage et rendement d'agglomération.* In *Coordination Economique et Apprentissage des Firmes* : 180-225, ed. N. Lazaric and J. M. Mouvier, pp. 180-225. Paris: Economica.

Hakansson H. 1987. *Industrial Technological Development: a Network Approach*, London: Croomhelm.

Hakansson H. 1990. Technological Collaboration in Industrial Network. *European Journal of Marketing*, September: 113-123.

Harrison, B. 1992. Industrial districts: old wines in new bottles. *Regional Studies* 26: 469-483.

Hofstede, G. 1980. Culture's consequences: International differences in work-related values, Beverly Hills, Sage Publications.

Hornell E., J.E.Vahlne and F.Wiedersheim-Paul. 1973. Export och utlandsetableringard, Uppsala, Almqvist & Wicksell.

Jaffe, A. B. 1989. Real effects of academic research. *American Economic Review* 79, no. 5: 957-970.

Jaffe, A. B., M. L. Trajtenberg and R. Henderson. 1993. Geographic localisation of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 63: 577-598.

Jaffe, A.B. 1986. Technological opportunity and spillovers of R&D: evidence from firms' patents, profits, and market value. *American Economic Review* 76, no. 5: 984–1001.

Kafouros M.I., Buckley P.J and Clegg J. 2009. The role of international diversification and the global reservoirs of scientific knowledge in explaining performance outcomes, AIB Conference Proceedings, San Diego.

Kirat, T., and Y. Lung. 1999. Innovation and proximity. Territories as loci of collective learning processes. *European Urban and Regional Studies* 6, no. 1: 27-38.

Lambooy, J. G. 2003. The role of intermediate structures and regional context for the evolution of knowledge networks and structural change, *Papers on Economics and Evolution,* Jena: Max Planck Institute, no. 0309.

Nesta, L., and P.P. Saviotti. 2005. Coherence of the knowledge base and the firm's innovative performance: evidence from the U.S Pharmaceutical industry. Journal of Industrial Economics 53, no.1: 23-142.

Nesta, L., and P.P. Saviotti. 2006. Firm knowledge and market value in biotechnology. Industrial and Corporate Change 15, no.4: 625-652.

Lundvall, B. A. 1992. *National systems of innovation*. London: Francis Pinter.

Lundvall, B.-A. 1988. Innovation as an Interactive Process: From User-Producer Interaction to the National System of Innovation. In *Technical Change and Economic Theory,* ed. G. Dosi, C. Freeman, R. Nelson, G. Silverberg & L. Soete: 349-369. London: Pinter.

Malerba, F., and L. Orsenigo . 2001. Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnoogy: a hystory friendly model. *Industrial and corporate change* 11, no.4: 667-703

March, J. G. 1991. Exploration and exploitation in organisational learning. *Organization Science* 2 , no.1: 71-86.

Mowery, D.C., and B.N. Sampat. 2001. Patenting and licensing university inventions: lessons from the history of the research corporations. *Industrial and Corporate Change* 10, no.2: 317-355.

Mowery, D.C., J.E. Oxley, and B.S. Silverman. 1996. Strategic Alliances and Inter-firm Knowledge Transfer. *Strategic Management Journal* 17: 77-91.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation, *Organization Science* 5, no. 1: 14-37.

Nonaka, I., and H. Takeuki. 1995. *The Knowledge-Creating Company*, New York and Oxford, UK: Oxford University Press.

Nooteboom, B., 2000. *Learning and innovation in organizations and economies*, Oxford: Oxford University Press.

Oerlemans, L.A.G., M.T.H. Meeus, F.M.W. Boekema. 2000. *On the spatial embeddedness of innovation networks: an exploration of the proximity effect.* Paper presented at the 40th European Congress of the European Regional Science Association, Barcelona

Oliver A. 2004. Biotechnology entrepreneurial scientists and their collaborations. *Research Policy* 33: 583-597.

Pennings, J. M., and F. Harianto. 1992. Technological networking and innovation implementation. *Organization Science* 3, no. 3: 356-382.

Pisano, G.P. 2007. *Science Business.* Boston: Harvard Business School Press.

Powell, W.W., W.K. Koput and L. Smith-Doerr. 1996. Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly* 41: 116-145.

Powell, W. W. 1990. Neither market nor hierarchy: network forms of organisations. *Research in Organisational behaviour* 12: 295-336.

Powell, W. W., and P. Brantley. 1992. Competitive cooperation in biotechnology: learning through networks? In *Networks and organisations: structure, form and Action*, ed. R. Eccles. Boston: Harvard Business School Press.

Pyka, A., and G. Kuppers. 2001. *Innovation networks.* London: Edward Edgar.

Pyka, A., and P. Saviotti. 2005. The evolution of R&D networking in the biotech industries. *International Journal of Entrepreneur and Innovation Management* 5: 49-68.

Rothermael, F.T., D. L. Deeds. 2004. Exploration and Exploitation Alliances in Biotechnology: A System of New Product Development. *Strategic Management Journal* 25, no. 6: 201-221.

Saviotti, P.P., and Catherine D. 2008. Innovation networks in biotechnology. In *Handbook of Bioentrepreneurship,* ed. H. Patzeltand T. Brenner: 53-82. Heidelberg: Springer Verlag.

Shan, W., G. Walker, and B. Kogut, B. 1994. Interfirm cooperation and startup innovation in the biotechnology industry. *Strategic Management Journal* 15, no.5: 387-94.

Smith K. 1995. Interactions in knowledge systems: foundations, policy implications and empirical methods. *STI Review* 16: 69-102.

Storper, M. 1997. *The regional world: territorial development in a global economy, perspectives on economic change*. New York: Guilford Press.

Storper, M., and A. Venables. 2005. Buzz: face to face contact and the Urban Economy, in *Clusters, Networks and innovation*, ed. S. Breschi and F. Malerba: 319-342. New York: Oxford Univeristy Press.

Stuart, T., S. Ozdemir, and W. Ding. 2007. Vertical alliance networks: The case of university–biotechnology– pharmaceutical alliance chains. *Research Policy,* 36: 477–498.

Tijssen, R.J.W. 2001. Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. *Research Policy* 30, 35-54.

Torre A. 1993. Proximité geographique et dynamiques industrielles, *Revue d'economie Regionale et Urbaine* 3: 431-448.

Vedovello, C. 1997. Science parks and university-industry interaction: geographical proximity between the agents as a driving force, *Technovation* 17, no.9: 491-502.

Wuyts, S., M.G. Colombo, S. Dutta, and B. Nooteboom. 2006. Empirical tests of optimal cognitive distance, *Journal of Economic Behaviour and Organisation* 58: 277-302.

Zucker L.G., M.R. Darby and M.B. Brewer. 1998). Intellectual human capital and the birth of US biotechnology enterprises. *American Economic Review* 88: 290-306.

APPENDIX

Table A1. Journals and subjects of the co-publications of the sample of Italian biotechnology firms with Italian partners. Only the co-publications with a frequency higher than 10 are shown.

Freq	Journal	Subject	Country
108	Blood	Hematology	Italy
101	Haematologica	Hematology	Italy
34	Clinical chemistry and laboratory medicine	Medical Laboratory Technology	Italy
32	Proceedings of the national academy of sciences of the united states of america	Multidisciplinary Sciences	Italy
31	Journal of immunology	Immunology	Italy
27	Journal of biological chemistry	Biochemistry & Molecular Biology	Italy
25	Clinical cancer research	Oncology	Italy
25	Leukemia	Oncology; Hematology	Italy
24	Clinical chemistry	Medical Laboratory Technology	Italy
22	Molecular therapy	Biotechnology & Applied Microbiology; Genetics & Heredity; Medicine, Research & Experimental	Italy
19	Transplantation proceedings	Immunology; Surgery; Transplantation	Italy
19	Transplantation	Immunology; Surgery; Transplantation	Italy
18	Haematologica-the hematology journal	Hematology	Italy
17	Cancer	Oncology	Italy
16	American journal of human genetics	Genetics & Heredity	Italy
15	Molecular cancer therapeutics	Oncology	Italy
15	Transplant international	Surgery; Transplantation	Italy
15	Diabetes nutrition & metabolism	Endocrinology & Metabolism; Nutrition & Dietetics	Italy
14	Circulation	Cardiac & Cardiovascular Systems; Hematology; Peripheral Vascular Disease	Italy
13	Biochemistry	Biochemistry & Molecular Biology	Italy
13	Gene	Genetics & Heredity	Italy
13	Hemoglobin	Biochemistry & Molecular Biology; Hematology	Italy
12	Journal of the american society of nephrology	Urology & Nephrology	Italy
12	Febs letters	Biochemistry & Molecular Biology;	Italy

		Biophysics; Cell Biology	
12	Nephrology dialysis transplantation	Transplantation; Urology & Nephrology	Italy
12	Protein science	Biochemistry & Molecular Biology	Italy
11	International journal of artificial organs	Engineering, Biomedical; Transplantation	Italy
11	Clinical therapeutics	Pharmacology & Pharmacy	Italy
11	European urology	Urology & Nephrology	Italy
11	Biological chemistry	Biochemistry & Molecular Biology	Italy
11	Journal of neurology	Clinical Neurology	Italy
11	Annals of the rheumatic diseases	Rheumatology	Italy
11	Neuroscience letters	Neurosciences	Italy
11	Journal of clinical investigation	Medicine, Research & Experimental	Italy
11	Biochemical journal	Biochemistry & Molecular Biology	Italy
11	Science	Multidisciplinary Sciences	Italy
11	Journal of leukocyte biology	Cell Biology; Hematology; Immunology	Italy
11	Human mutation	Genetics & Heredity	Italy
10	Cancer research	Oncology	Italy
10	Journal of neuroimmunology	Immunology; Neurosciences	Italy
10	Journal of medicinal chemistry	Chemistry, Medicinal	Italy
10	Journal of cellular physiology	Cell Biology; Physiology	Italy
10	Faseb journal	Biochemistry & Molecular Biology; Biology; Cell Biology	Italy
10	Annals of tropical medicine and parasitology	Public, Environmental & Occupational Health; Parasitology; Tropical Medicine	Italy

Table A2. Journals and subjects of the co-publications of the sample of Italian biotechnology firms with foreign partners. Only the co-publications with a frequency higher than 3 are shown.

Freq	Journal	Subject	Country
19	Journal of biological chemistry	Biochemistry & Molecular Biology	USA
13	Cell	Biochemistry & Molecular Biology; Cell Biology	USA
11	Blood	Hematology	USA
9	Aging cell	Cell Biology; Geriatrics & Gerontology	USA
8	Cell	Biochemistry & Molecular Biology; Cell Biology	Germany
8	Molecular therapy	Biotechnology & Applied Microbiology; Genetics & Heredity; Medicine, Research & Experimental	USA
8	Development	Developmental Biology	USA
7	Journal of medicinal chemistry	Chemistry, Medicinal	USA
7	Gene therapy	Biochemistry & Molecular Biology; Biotechnology & Applied Microbiology; Genetics & Heredity; Medicine, Research & Experimental	USA
7	Molecular and cellular biology	Biochemistry & Molecular Biology; Cell Biology	USA
6	Microbes and infection	Immunology; Microbiology; Virology	Brazil
6	Development	Developmental Biology	Germany
6	Proceedings of the national academy of sciences of the united states of america	Multidisciplinary Sciences	USA
6	Journal of immunology	Immunology	Germany
6	Journal of immunology	Immunology	USA
5	Nature medicine	Biochemistry & Molecular Biology; Cell Biology; Medicine, Research & Experimental	USA
4	Cancer research	Oncology	Australia
4	Proceedings of the national academy of sciences of the united states of america	Multidisciplinary Sciences	England
4	British journal of pharmacology	Pharmacology & Pharmacy	Sweden
4	Behavioural brain research	Behavioral Sciences; Neurosciences	Canada

4	Gene	Genetics & Heredity	USA
4	Journal of intellectual disability research	Education, Special; Genetics & Heredity; Clinical Neurology; Psychiatry; Rehabilitation	USA
4	Journal of neural transmission	Clinical Neurology; Neurosciences	USA
4	Experimental hematology	Hematology; Medicine, Research & Experimental	Germany
4	Cell	Biochemistry & Molecular Biology; Cell Biology	Netherlands
3	Movement disorders	Clinical Neurology	USA
3	Matrix biology	Biochemistry & Molecular Biology; Cell Biology	USA
3	Infection and immunity	Immunology; Infectious Diseases	France
3	Development	Developmental Biology	England
3	Cancer	Oncology	USA
3	Journal of computer-aided molecular design	Biochemistry & Molecular Biology; Biophysics; Computer Science, Interdisciplinary Applications	USA
3	Nature medicine	Biochemistry & Molecular Biology; Cell Biology; Medicine, Research & Experimental	Germany
3	European journal of cancer care	Oncology; Health Care Sciences & Services; Rehabilitation	USA
3	Blood	Hematology	England
3	Molecular therapy	Biotechnology & Applied Microbiology; Genetics & Heredity; Medicine, Research & Experimental	Germany