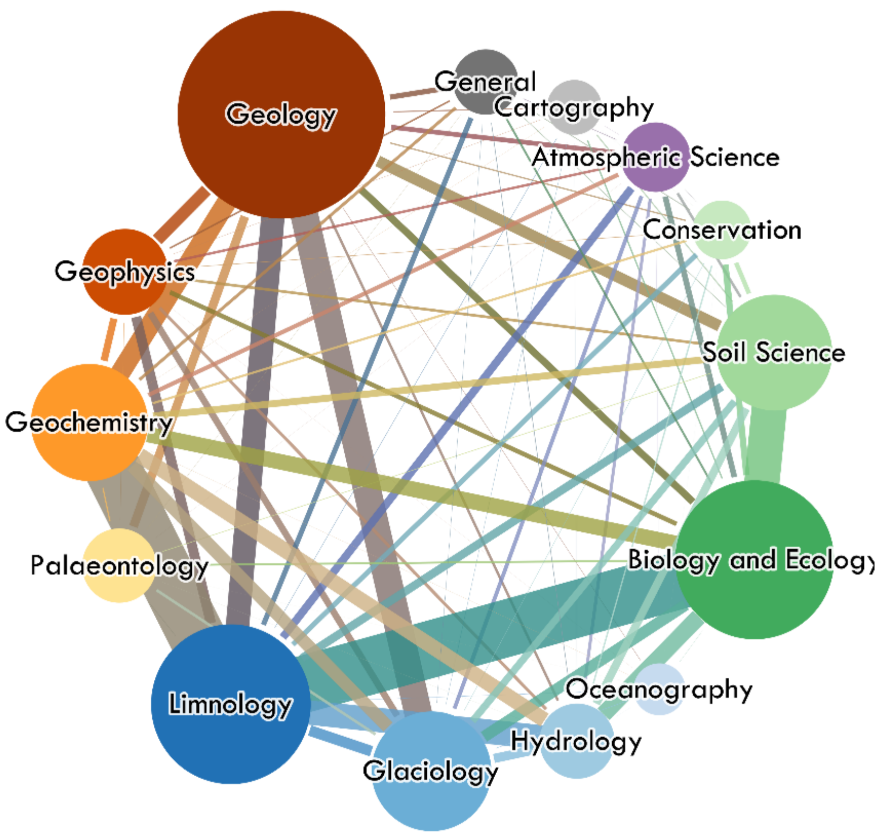
Supplementary material for: Chignell S.M., Howkins A., Gullett P. & Fountain A.G. 2022. Patterns of interdisciplinary collaboration resemble biogeochemical relationships in the McMurdo Dry Valleys, Antarctica: a historical social network analysis of science, 1907–2016. *Polar Research 41*. Correspondence: Stephen Chignell, Institute for Resources, Environment and Sustainability, University of British Columbia, Aquatic Ecosystems Research Lab, 2202 Main Mall, Vancouver, V6T 1Z4 BC, Canada. E-mail: steve.chignell@ubc.ca.

Supplementary Table S1. Publications summarized by format. Minor overlaps in format types are due to different types in reference data sources (i.e., Web of Science, paper bibliography volumes). See Supplementary Data Set S1 for the complete set of publications.

Publication format	Count	Percent of total
Article	2187	79.8
Chapter	114	4.2
Report	94	3.4
Article; proceedings paper	74	2.7
Review	49	1.8
Thesis/dissertation	48	1.8
Proceedings	32	1.2
Proceedings paper	30	1.1
Conference abstract	27	1.0
Article; book chapter	20	0.7
Meeting abstract	19	0.7
Thesis	14	0.5
Book	10	0.4
Note	6	0.2
Letter	5	0.2
Editorial material	4	0.1
Unpublished paper	3	0.1
Review; book chapter	2	0.1
Discussion	1	0.04
Editorial material; book chapter	1	0.04
Correction	1	0.04
Map	1	0.04

Supplementary Table S2. Publications summarized by language.

Language	Count	Percent of total
English	2683	97.8
Japanese	30	1.1
Russian	10	0.4
German	8	0.3
Italian	6	0.2
French	3	0.1
Chinese	2	0.07



Supplementary Fig. S1. Transformed cumulative author–discipline network, showing strength of connections between pairs of academic disciplines as a function of the number of authors who published in both disciplines. Nodes are sized by frequency (number of publications) and arranged in a clockwise circular layout, following the order of academic disciplines in Fig. 2.

Supplementary Fig. S2. Animated version of Fig. 3 (video file available online), showing annual changes in the cumulative co-authorship network of the MDV. Annotations mark key historical moments in the development of the network. Each circle (node) represents an individual scientist, and its size reflects the number of publications that scientist authored. Only scientists with more than 20 publications are labelled. Note that a multi-authored publication is counted as a publication for each listed author. Line (tie) thickness represents the number of collaborations between authors. Colours represent different groups identified, and only those comprising at least 4% of the total network are assigned a colour and label.

Supplementary Table S3. Selected statistics for top 20 authors with highest frequency in the cumulative co-authorship network. Following standard SNA procedures, we normalized network scores to facilitate interpretation and comparison.

Author	Frequency	Degree	Weighted degree	Eigenvector centrality	Betweenness centrality	No. of triangles
Torii T.	97	77	240	0.0462	0.031495	198
Lyons W.B.	91	198	451	0.276433	0.105531	1244
Priscu J.C.	84	143	292	0.164517	0.038255	667
McKnight D.M.	74	143	336	0.16808	0.026247	741
McKay C.P.	64	163	312	0.214468	0.117331	960
Fountain A.G.	62	105	294	0.153739	0.053088	548
Doran P.T.	58	145	322	0.19878	0.041383	783
Virginia R.A.	57	77	247	0.134984	0.010597	478
Parker B.C.	55	43	151	0.024932	0.028574	108
Welch K.A.	55	107	245	0.144891	0.007488	534
Wall D.H.	54	113	290	0.18598	0.030182	876
Denton G.H.	53	64	144	0.076769	0.030924	204
Matsumoto G.I.	47	45	125	0.022139	0.026254	111
Bockheim J.G.	47	64	95	0.077203	0.040579	307
Marchant D.R.	47	72	153	0.093202	0.028773	246
Friedmann E.I.	47	81	124	0.146793	0.053898	569
Barrett J.E.	44	77	230	0.127741	0.014514	457
Wharton R.A.J.	44	59	160	0.084574	0.017505	250
McGinnis L.D.	43	27	65	0.027502	0.010702	44
Webb P.N.	43	19	44	0.054773	0.010544	18

Network diameter

The network diameter is the path length (number of ties) between the two most distant nodes in the network. Network diameter is considered one of the most robust measures of network topology, as it provides an indication of how well information travels from person to person (Sivakumar et al. 2018). The diameter for the MDV cumulative network is 14 and increased only slightly throughout the study period. The mean path length between two nodes is 4.5 and remained fairly stable through the entire study period.

Giant component

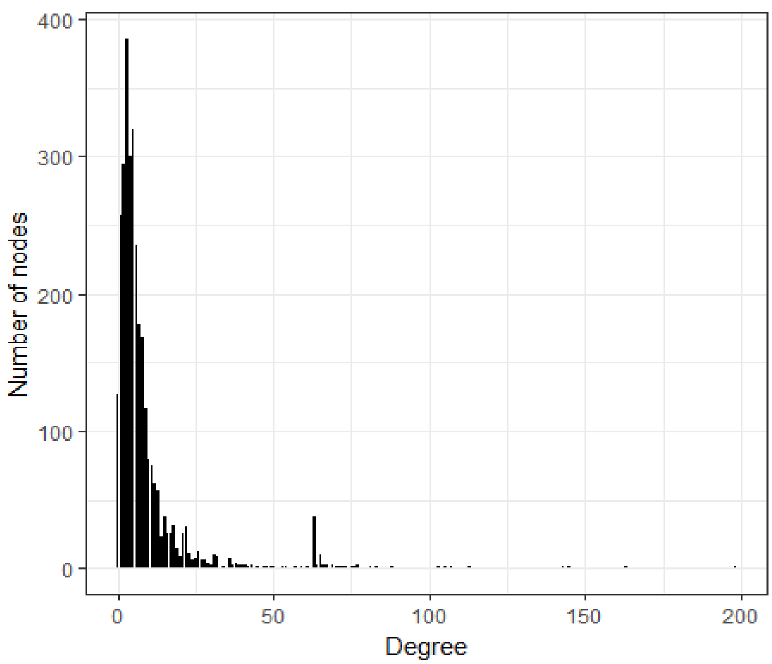
The giant component represents a large group of individuals connected by paths of intermediate acquaintances (Newman 2001). The giant component of the cumulative MDV network contains 76% of the nodes and 91% of the ties. This indicates that the vast majority of collaborations have taken place within a series of large and well-connected groups, while nearly a quarter of MDV researchers have worked as academic 'lone wolves'. Comparing the three periods, we see that the giant component did not contain a majority of the network during the early period (42%) and increased slightly during the middle period (58%). In the recent period, the giant component is considerably larger (81%), which indicates a significant densification of collaborative ties during this period. These recent shifts in the giant component may relate to broader changes in policy structures as well as the culture of scientific publishing (i.e., including more people who work on a project as co-authors), and represent potential lines of future research.

Modularity

Modularity algorithms partition the network into communities such that nodes within a given community have denser relationships to each other than to nodes outside the community. A high modularity score thus indicates complex internal structure. Modularity began low during the International Geophysical Year of 1957–58 but increased rapidly during the early period before stabilizing at very high values for the remainder of the study period. The modularity of the cumulative co-authorship network was high (0.81) (Table 1), with seven main communities, each comprising at least 4% of the total network (Fig. 3). The largest group mostly comprises the MDV LTER, a US National Science Foundation-funded project comprised of research scientists and students from multiple universities and academic disciplines (MCM LTER 2018). The second-largest group includes glacial geologists and palaeoclimatologists, primarily with connections to the University of Maine and universities in NZ. The third largest, includes the Japanese geochemistry research program and members of the DVDP. The next group includes limnologists and astrobiologists affiliated with NASA and Virginia Polytechnic Institute and State University. This is followed by a group of scientists connected via American astrobiologist E. Imre Friedmann and distributed throughout the centre of the network. The sixth group is comprised of abiotic soil scientists from NZ, the US and the UK, and the seventh group are primarily microbial ecologists affiliated with the NZ Antarctic programme.

Degree

The degree of a node is defined simply as the number of connections it has. The weighted degree takes into account the weight of the connected ties. In this study, the mean degree is the mean total number of people with whom a scientist collaborated within a given time period (Newman 2001). Figure 7b shows a positive linear trend for mean weighted degree throughout the entire period, indicating that collaboration steadily increased regardless of the changes in the number of authors. Plotting the degree distribution as a histogram is useful for identifying the type of network (Barabási & Albert 1999; Sivakumar et al. 2018). The degree distribution of the MDV network is highly right-skewed (Supplementary Fig. S3), indicating that the network is made up of many authors with low or no connectivity and a few highly connected authors with a degree far above the mean (Newman 2003).



Supplementary Fig. S3. Degree distribution of the cumulative MDV co-authorship network, showing strong right skew. The nodes with degree scores of 63-65 are primarily driven by a single paper with 64 authors (Wilson et al. 2012).

Clustering

Clustering is a local measure of connection density. In a co-authorship network, a node's clustering coefficient represents the probability that two of its collaborators also wrote a paper together (Barabási et al. 2002). Clustering is related to the number of triangles, which form when three nodes are connected to each other (Newman 2004). The mean clustering coefficient was low in the early years of MDV science and rapidly increased in the DVDP period as the network grew, stabilizing in about 1970 (Fig. 7b). The mean clustering coefficient for the cumulative co-authorship network was high (0.85). In other words, two MDV authors have an 85% or greater probability of collaborating with each other if both have collaborated with a third scientist. The combination of a high clustering coefficient and low value for mean shortest path length indicates that the MDV co-authorship network is a 'small world' network (Newman et al. 2011).

References

Barabási A.-L. & Albert R. 1999. Emergence of scaling in random networks. *Science 286*, 509–512, doi: 10.1126/science.286.5439.509.

Barabási A.-L., Jeong H., Néda Z., Ravasz E., Schubert A. & Vicsek T. 2002. Evolution of the social network of scientific collaborations. *Physica A 311*, 590–614, doi: 10.1016/S0378-4371(02)00736-7.

MCM LTER 2018. McMurdo Dry Valleys personnel. *McMurdo Dry Valleys LTER*. Accessed on the internet at https://www.mcmlter.org/mcmurdo-lter-personnel on 2 April 2018.

Newman M., Barabási A.-L. & Watts D.J. 2011. *The structure and dynamics of networks*. Princeton University Press.

Newman M.E.J. 2001. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America 98*, 404–409, doi: 10.1073/pnas.98.2.404.

Newman M.E.J. 2003. The structure and function of complex networks. *SIAM Review 45*, 167–256, doi: 10.1137/S003614450342480.

Newman M.E.J. 2004. Who is the best connected scientist? A study of scientific coauthorship networks. In E. Ben-Naim (et al.): *Complex networks*. Pp. 337–370. Berlin: Springer.

Sivakumar B., Puente C.E. & Maskey M.L. 2018. Complex networks and hydrologic applications. In A. Tsonis (ed.): *Advances in nonlinear geosciences*. Pp. 565–586. Cham, Switzerland: Springer.

Wilson G.S., Levy R.H., Naish T.R., Powell R.D., Florindo F., Ohneiser C., Sagnotti L., Winter D.M., Cody R., Henrys S., Ross J., Krissek L., Niessen F., Pompillio M., Scherer R., Alloway B.V., Barrett P.J., Brachfeld S., Browne G., Carter L., Cowan E., Crampton J., DeConto R.M., Dunbar G., Dunbar N., Dunbar R., von Eynatten H., Gebhardt C., Giorgetti G., Graham I., Hannah M., Hansaraj D., Harwood D.M., Hinnov L., Jarrard R.D., Joseph L., Kominz M., Kuhn G., Kyle P., Läufer A., McIntosh W.C., McKay R., Maffioli P., Magens D., Millan C., Monien D., Morin R., Paulsen T., Persico D., Pollard D., Raine J.I., Riesselman C., Sandroni

S., Schmitt D., Sjunneskog C., Strong C.P., Talarico F., Taviani M., Villa G., Vogel S., Wilch T., Williams T., Wilson T.J. & Wise S. 2012. Neogene tectonic and climatic evolution of the western Ross Sea, Antarctica—chronology of events from the AND-1B drill hole. *Global and Planetary Change 96–97*, 189–203, doi: 10.1016/j.gloplacha.2012.05.019.