



UNIVERSIDAD VERACRUZANA

Instituto de Investigaciones Biológicas

**“Redes de coautorías del Instituto de Investigaciones
Biológicas (IIB)”**

INFORME TÉCNICO

QUE PRESENTA:

I.Q. Jeanette Chacón Hernández

Dr. Luis Gerardo Abarca Arenas

Dr. Eduardo Morteo Ortiz

Xalapa, Ver.

Agosto, 2024

Resumen

El presente informe se centra en visualizar las colaboraciones entre investigadores del Instituto de Investigaciones Biológicas (IIB) de la Universidad Veracruzana a través de redes de coautoría. Se recopiló información de autores, estandarizándola y organizándola en matrices que representan las interacciones entre ellos y demás colaboradores a lo largo de los años 2011 a 2023. Utilizando los software Excel y R, se crearon matrices de adyacencia para cada tipo de publicación (artículos, libros y capítulos de libros), mientras que con el software Gephi, se elaboraron los grafos. Se calculó la centralidad y la productividad académica, generando gráficos que muestran la dinámica de colaboración y la modularidad de los grupos de investigadores.

Objetivos

General

- Representar de manera gráfica las colaboraciones entre los investigadores activos del Instituto de Investigaciones Biológicas (IIB), destacando tanto las interacciones internas entre ellos como las colaboraciones con autores externos, utilizando herramientas de software especializado

Específicos

- Recopilar, estandarizar y organizar información de autores
- Crear matrices y redes de coautorías
- Calcular centralidad y productividad académica
- Representar de manera gráfica los resultados obtenidos

Introducción

La evaluación de la producción científica se ha vuelto cada vez más relevante tanto para la comunidad académica como para los responsables de formular políticas públicas, debido a su papel crucial en la medición del impacto y la asignación estratégica de recursos hacia la

investigación. Este enfoque se materializa mediante el uso de métodos estadísticos y matemáticos, junto con herramientas como los indicadores bibliométricos. Estos últimos permiten un análisis detallado no solo de los aspectos cuantitativos, como el número de publicaciones, el tipo de revistas y las citas recibidas, sino también de los aspectos cualitativos, incluyendo las colaboraciones entre investigadores (Gutiérrez y Velasco, 2017).

La colaboración entre investigadores, en particular, emerge como un factor crucial en este análisis, revelando las redes informales de conocimiento y estableciendo la amplitud del impacto de la comunidad investigadora (Bordons y Zulueta, 1999).

Las redes de coautoría se presentan como una herramienta para representar visualmente estas colaboraciones. Estas redes no solo identifican quiénes colaboran y con qué frecuencia, sino que también muestran las estructuras de colaboración y patrones de influencia dentro de las comunidades científicas (Tijssen y van Raan, 1994; Abbasi y Altmann, 2010). Al aplicar técnicas de análisis de redes sociales, se pueden explorar diversos aspectos como la centralidad de los actores en la red, la densidad de las conexiones y la modularidad de los grupos, proporcionando una visión sobre la dinámica colaborativa y su impacto en la producción científica (Sanz-Casado, 2000).

Estas redes también facilitan la identificación de áreas de colaboración potencial y la optimización de recursos en la investigación. Permiten a los financiadores y responsables de políticas científicas identificar y apoyar a los grupos más productivos y colaborativos, contribuyendo así a la mejora del impacto y la eficiencia en la investigación (Velden *et al.*, 2010). Además, las redes de coautoría sirven como herramientas estratégicas para visualizar y analizar cómo la colaboración influye en la productividad y el impacto de los grupos de investigación, tanto a nivel local como global (Huamaní y Mayta, 2010).

En ese contexto, este informe se enfoca en explorar y demostrar el uso de las redes de coautoría como indicadores relacionales en la evaluación de la actividad científica del Instituto de Investigaciones Biológicas (IIB) de la Universidad Veracruzana. En la sección de metodología se detallan los métodos y herramientas utilizados para el análisis de estas redes, mostrando su aplicación práctica en la comprensión y mejora de la colaboración

científica y su impacto en las comunidades científicas (Corral-Marfil y Cànoves, 2014; Aguado *et al.*, 2009).

Metodología

Se recopilaron los nombres de todos los autores que colaboraron en artículos con los investigadores del IIB de la Universidad Veracruzana durante el periodo de 2011 a 2023. Se crearon redes por cada año, así como por periodos y totales, tanto para cada tipo de producto (artículos, libros y capítulos de libros) como generales. Los nombres fueron depurados, estandarizados y filtrados debido a las variaciones en el formato de los nombres según los lineamientos de las revistas, con el fin de unificar el formato de la siguiente manera: PrimerApellido-SegundoApellido,Inicial(es)deNombre(s), sin caracteres raros. Los nombres fueron capturados en una hoja de Excel, obteniendo un total de 911 autores.

A cada autor se le asignó un identificador: para los investigadores activos, su identificador fue su propio nombre, mientras que para el resto de los colaboradores se utilizó el formato "ID+número", de acuerdo con el orden alfabético.

Una vez estandarizados los nombres de los autores en un formato uniforme, se procedió a crear bases de datos en hojas de Excel, organizadas por año y tipo de producto (artículos, libros y capítulos de libros) para el periodo 2011-2023. Cada fila representaba un producto y cada columna, un autor diferente. Los archivos fueron guardados en formato “.csv”.

Utilizando un script del programa "R" (ver. 4.3.1, R Core Team, 2020) en “RStudio” (ver. 2024.04.0 + 735, RStudioTeam, 2020) con el paquete "rstudioapi", se crearon las matrices de adyacencia de las colaboraciones: primero, se importaron los datos desde un archivo “.csv”, así como la lista de autores con sus respectivos identificadores. Se generó un vector de autores a partir de los datos importados, eliminando espacios vacíos y duplicados, y ordenándolos alfabéticamente para almacenarlos en un “dataframe”. Posteriormente, se emparejaron los nombres de los autores con sus identificadores correspondientes, creando un nuevo dataframe que combinaba ambos elementos. Dicho script incluye una verificación para detectar y reportar autores que no tienen un identificador asociado. Con esta lista depurada, se construyó una matriz de coautorías, estableciendo enlaces colaborativos entre

autores que han trabajado juntos. Finalmente, los nombres de los autores en la matriz fueron reemplazados por sus respectivos identificadores y la matriz resultante se exportó en formato “.csv”.

Para crear los grafos, se empleó el software libre "Gephi" (ver. 0.10.1, Bastian *et al.*, 2009), donde se importaron las matrices de adyacencia. Se ajustó el tamaño de los nodos de los investigadores activos según el número de productos elaborados con un valor superior (22) e inferior (6), mientras que los nodos de los demás autores se mantuvieron con un tamaño mínimo (1), con el fin aumentar el contraste entre los nodos de interés.

Se calculó la modularidad de cada red y se representaron los módulos con diferentes colores. La modularidad es una característica fundamental de todas las redes, que influye en la heterogeneidad de los contactos entre los nodos (Ravasz *et al.*, 2002; Kashtan y Alon, 2005; Whitehead y Lusseau, 2012). En el contexto de la red de coautorías, la modularidad describe cómo la red se divide en grupos o clusters y el grado de interacción entre estos grupos. Este concepto se puede estimar mediante un coeficiente de modularidad (Q) que varía de 0 a 1. Un valor de Q cercano a 1 indica una red con una estructura fuertemente agrupada, donde las interacciones entre individuos de diferentes clusters son mínimas (Newman, 2006; Marcoux y Lusseau, 2013):

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j)$$

Donde m es el número de vínculos en la red, A_{ij} representa los vínculos ponderados entre el autor i y el autor j , k_i es la suma de todos los vínculos ponderados del autor i (grado), y la función $\delta(C_i, C_j)$ es 1 si i y j están en el mismo grupo y 0 en caso contrario. Las redes resultantes se exportaron en archivos ".png".

Se elaboraron gráficas de puntos con los valores de centralidad (degree, closeness, betweenness y eccentricity), así como el número de productos por cada investigador activo del IIB.

El degree centrality o centralidad de grado ($C'd$) es una medida en los gráficos de redes que se calcula por el número total de enlaces directos que tiene un nodo con otros nodos (Pérez-Beltrán *et al.*, 2015; Zhang y Luo, 2017). La fórmula para calcular $C'd$ es una ecuación que

suma las conexiones directas de un nodo con los demás (X_{ij}), donde n representa el número total de nodos en la red focal (Shaw, 1954; Nieminen, 1974; Bolland, 1988):

$$C'd = \frac{\sum_{j=1}^n X_{ij}}{(n-1)(n-2)} (i \neq j)$$

Por otro lado, closeness centrality o la centralidad de cercanía ($C'c$) mide la capacidad de un autor para conectarse con otros autores a través de pocos pasos en la red (Wasserman y Faust, 1994). Valora a los autores según la distancia, medida en pasos, hacia los demás en la red. Un autor con alta centralidad de cercanía tiene un menor número de pasos necesarios para relacionarse con otros, lo que indica su autonomía y relevancia en términos de conexión rápida y eficiente. Esto representa la facilidad y conveniencia de sus conexiones en la red (Beauchamp, 1965; Wasserman y Faust, 1994; Zhang y Luo, 2017):

$$C'c = \frac{n-1}{\left[\sum_{j=a}^n d(N_i, N_j)\right]} (i \neq j)$$

Donde la función de sumatoria significa el número total de “pasos” desde el autor N hasta los otros autores en la red, siendo n el número total de autores de la red.

Betweenness centrality o centralidad de intermediación ($C'b$), se refiere a la medición de cuánto un autor actúa como intermediario en la comunicación dentro de una red (Pérez-Beltrán *et al.*, 2015). Un autor con alta intermediación conecta diferentes subgrupos, controlando los flujos de información y relaciones (Burt, 2010). Si un autor se encuentra en el camino más corto entre muchos otros autores, se considera que tiene alta intermediación. Esta medida captura la importancia de un autor en la mediación de conexiones, comunicaciones y transacciones dentro de la red (Freeman, 1977; Zhang y Luo, 2017):

$$C'b = \frac{2 \sum_{j < k} \frac{G_{jk}(N_i)}{G_{jk}}}{(n-1)(n-2)}$$

La expresión de sumatoria representa la suma de todas las combinaciones posibles de nodos j y k en la red. El término $G_{jk}(N_i)$ indica el número de caminos más cortos entre los nodos j y k que pasan por el nodo N_i , mostrando cuántas veces el nodo i se encuentra en la ruta

más corta entre esos dos nodos. G_{jk} es el número total de caminos más cortos entre los nodos j y k , sin importar si pasan por N_i o no. n es el número total de nodos en la red.

Como último parámetro, la excentricidad o eccentricity de un nodo es la mayor distancia entre ese nodo y cualquier otro nodo en la red, indicando cuántos pasos se necesitan para conectar el nodo más lejano. El diámetro de la red es la excentricidad máxima, mientras que el radio es la excentricidad mínima (Hage y Harary, 1995) En un grafo conectado, la suma de distancias (s) de un nodo (u) es la suma de las distancias entre u y los demás nodos (Harary, 1959):

$$s = \sum_{j=1}^p d_{ij}$$

La excentricidad ayuda a identificar nodos estratégicos, que pueden alcanzar otros nodos con la menor cantidad de pasos. Los nodos con excentricidad igual al radio se consideran centrales y forman el centro de la red, representando puntos clave de conexión.

Estos parámetros fueron calculados en el programa “Gephi” y se normalizaron en una escala de 0 a 1, donde, independientemente de la magnitud del valor máximo, siempre será representado con 1. Para elaborar dichas gráficas, se utilizó el paquete de “R” “ggplot2”. Las gráficas se guardaron en archivos “.png”.

Finalmente, se creó un mapa de calor con el número de colaboraciones entre los investigadores activos del IIB, utilizando los paquetes de “R” “rstudioapi”, “ggplot2” y “reshape2”.

Resultados

Se realizaron un total de 43 grafos diferentes: 12 correspondientes a capítulos, 4 a libros, 13 a artículos, 13 representando todos los productos por año, y 1 de colaboraciones entre académicos durante el periodo de 2011 a 2023. Estos grafos se encuentran en la sección Anexos. Cabe mencionar que algunos años no cuentan con grafos de ciertos productos, ya sea debido a la falta de producción o porque los investigadores trabajaron de manera individual, lo que impidió la creación de redes.

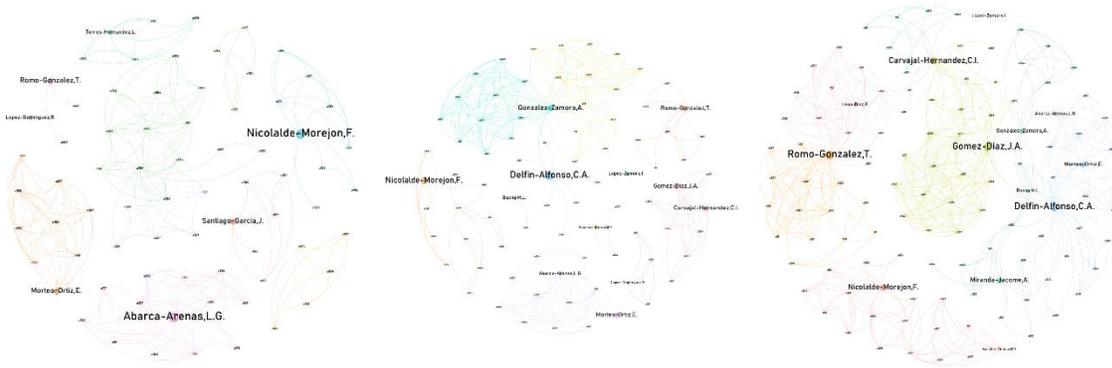


Figura 1: Grafos de coautoría de todos los productos de los investigadores de los años 2011, 2017 y 2023.

Además, se elaboraron 4 gráficas de puntos con los valores de centralidad de cada investigador activo, cada gráfica correspondiente a capítulos, libros, artículos y productos totales producidos durante los años 2011 a 2023, dividido en 3 periodos: 2011-2014, 2015-2018 y 2019-2023.

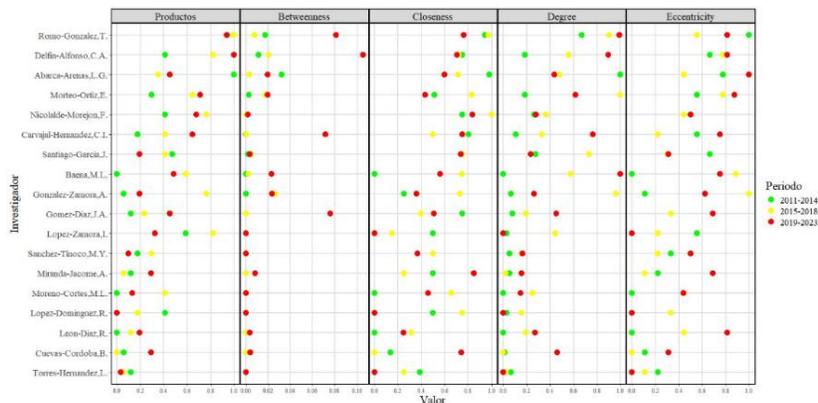


Figura 2: Gráfica de puntos de todos productos de los investigadores.

Todos los archivos generados para la realización de este trabajo se encuentran en una carpeta compartida con el investigador encargado en “Google Drive”.

Discusión

A través de las redes de coautoría generadas para el periodo 2011-2023, se identificaron patrones que muestran la colaboración interna y externa en la producción científica del instituto. Las visualizaciones de las redes mostraron módulos de colaboración, indicativos de la formación de grupos científicos dentro del IIB. Estos módulos no solo reflejan la frecuencia de colaboración entre investigadores, sino también la diversidad de temas de investigación abordados por diferentes grupos dentro del instituto. Esto muestra que la colaboración interdisciplinaria y multidisciplinaria es un componente para la investigación en el IIB, facilitando el intercambio de conocimientos y la generación de nuevos descubrimientos.

Además, el análisis de centralidad mostró la importancia de ciertos investigadores como intermediarios clave dentro de la red de coautoría. Investigadores con alta centralidad actúan como puentes entre diferentes grupos de investigación, contribuyendo a la difusión de conocimientos y la integración de diferentes disciplinas en los trabajos colaborativos.

Conclusión

El análisis de las redes de coautoría en el Instituto de Investigaciones Biológicas muestra las dinámicas colaborativas entre sus investigadores activos, mostrando ser una comunidad científica conectada que ha fortalecido sus lazos de manera interna con el paso de los años.

Fortalecer estas redes es importante para incrementar la productividad académica y el impacto de la investigación en el IIB. Seguir monitoreando las colaboraciones mediante las redes es una buena estrategia para evaluar y mejorar la actividad científica del IIB, promoviendo colaboraciones interinstitucionales.

Fuentes de información

- Abbasi, A., y Altmann, J. (2011). A social network system for analyzing publication activities of researchers. In *On collective intelligence* (pp. 49-61). Springer Berlin Heidelberg.
- Aguado-López, E., Rogel-Salazar, R., Garduño-Oropeza, G., Becerril-García, A., Zúñiga-Roca, M. F., y Velázquez-Álvarez, A. (2009). Patrones de colaboración científica a partir de redes de coautoría. *Convergencia. Revista de Ciencias Sociales*, 16, 225-258.
- Bastian M., S. Heymann & M. Jacomy, 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1), 361-362. URL <https://gephi.org/>.
- Beauchamp, M. A. (1965). An improved index of centrality. *Behavioral science*, 10(2), 161-163.
- Bolland, J. M. (1988). Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social networks*, 10(3), 233-253.
- Bordons, M. (1999). Evaluación de la actividad científica a través de indicadores bibliométricos. *Revista española de cardiología*, 52(10), 790-800.
- Burt, R. S. (2010). *Neighbor networks: Competitive advantage local and personal*. Oxford University Press.
- Corral-Marfil, J. A., y Cànoves, G. (2014). La colaboración científica en turismo: grupos y redes de investigación en Cataluña. *Documents d'anàlisi geogràfica*, 60(1), 57-86.
- Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35-41.
- Gutiérrez, J. K. R., y Velasco, N. Y. G. (2017). Redes de coautoría como herramienta de evaluación de la producción científica de los grupos de investigación. *Revista General de Información y documentación*, 27(2), 279.

- Hage, P., y Harary, F. (1995). Eccentricity and centrality in networks. *Social networks*, 17(1), 57-63.
- Harary, F. (1959). Status and contrastatus. *Sociometry*, 22(1), 23-43.
- Huamaní, C., y Mayta-Tristán, P. (2010). Producción científica peruana en medicina y redes de colaboración, análisis del Science Citation Index 2000-2009. *Revista Peruana de Medicina Experimental y Salud Pública*, 27, 315-325.
- Kashtan, N., y Alon, U. (2005). Spontaneous evolution of modularity and network motifs. *Proceedings of the National Academy of Sciences*, 102(39), 13773-13778.
- Marcoux, M., y Lusseau, D. (2013). Network modularity promotes cooperation. *Journal of theoretical biology*, 324, 103-108.
- Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23), 8577-8582.
- Nieminen, J. (1974). On the centrality in a graph. *Scandinavian journal of psychology*, 15(1), 332-336.
- Pérez-Beltrán, J. E., Valerio-Ureña, G., y Rodríguez-Aceves, L. (2015). Análisis de redes sociales para el estudio de la producción intelectual en grupos de investigación. *Perfiles educativos*, 37(150), 124-142.
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., y Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks. *science*, 297(5586), 1551-1555.
- RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.
- Sanz-Casado, E. (2000). Proyecto docente para la provisión de una plaza de Catedrático de Universidad sobre Bibliometría. *Madrid: Universidad Carlos III*.
- Shaw, M. E. (1954). Group structure and the behavior of individuals in small groups. *The Journal of psychology*, 38(1), 139-149.

- Tijssen, R. J. y Van Raan, A. F. J. (1994). Mapping changes in Science and Technology. Bibliometric co-occurrence analysis of the R&D literature. *Evaluation Review*, 1 (18), 98-115.
- Velden, T., Haque, A. U., y Lagoze, C. (2010). A new approach to analyzing patterns of collaboration in co-authorship networks: mesoscopic analysis and interpretation. *Scientometrics*, 85(1), 219-242.
- Wasserman, S., y Faust, K. (1994). Social network analysis: Methods and applications. Cambridge University Press.
- Whitehead, H., y Lusseau, D. (2012). Animal social networks as substrate for cultural behavioural diversity. *Journal of theoretical biology*, 294, 19-28.
- Zhang, J., y Luo, Y. (2017). Degree centrality, betweenness centrality, and closeness centrality in social network. En *2017 2nd international conference on modelling, simulation and applied mathematics (MSAM2017)* (pp. 300-303). Atlantis press.

Anexos

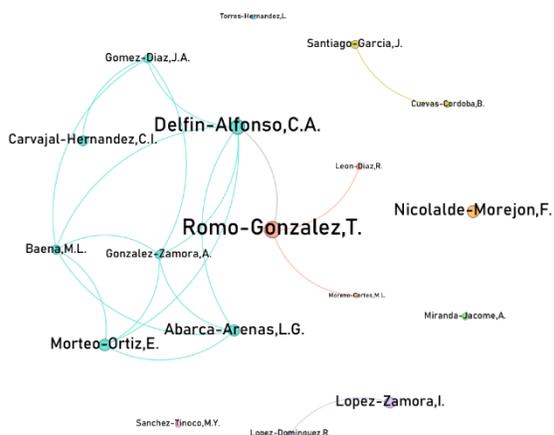


Figura A1: Colaboraciones entre investigadores desde 2011 a 2023.

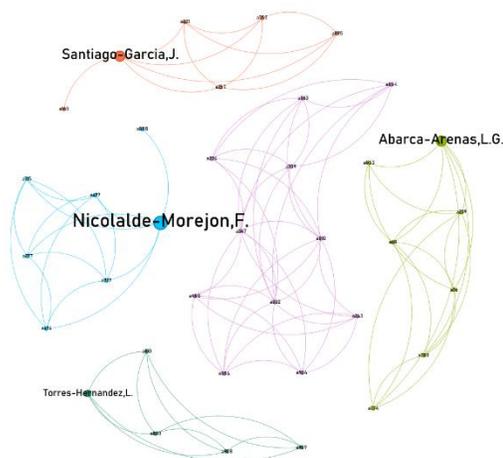


Figura A2: Artículos de 2011.

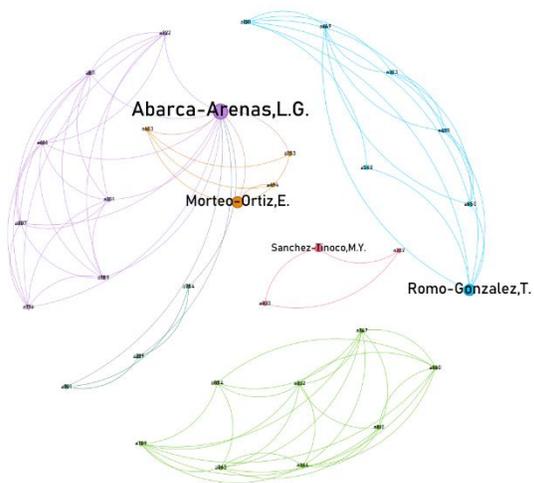


Figura A3: Artículos de 2012.

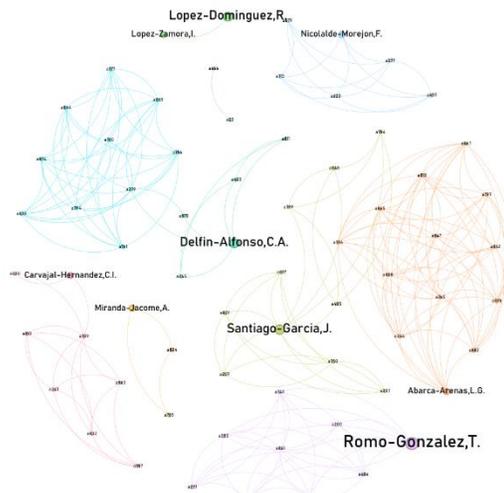


Figura A4: Artículos de 2013.

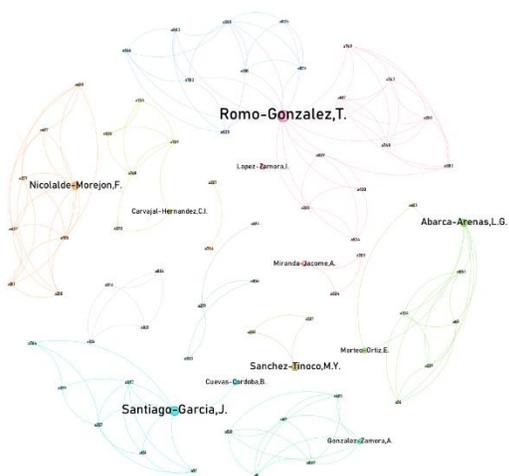


Figura A5: Artículos de 2014.

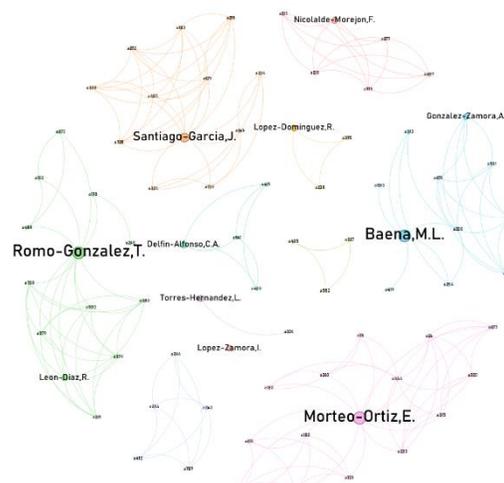


Figura A6: Artículos de 2015.

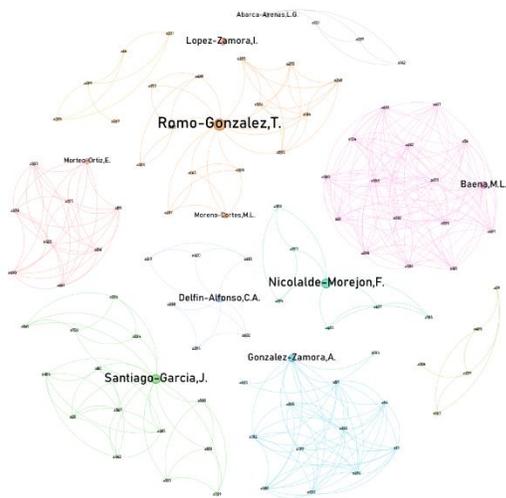


Figura A7: Artículos de 2016.

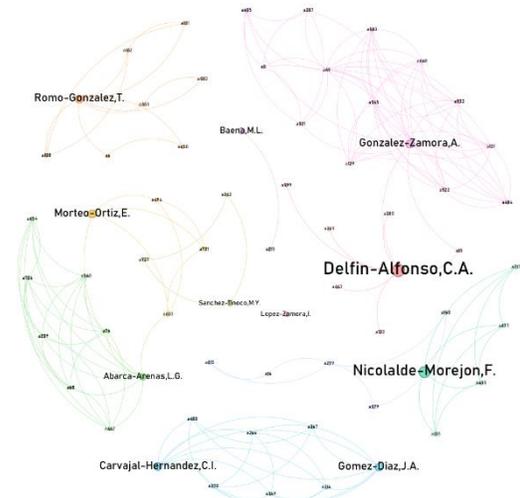


Figura A8: Artículos de 2017.

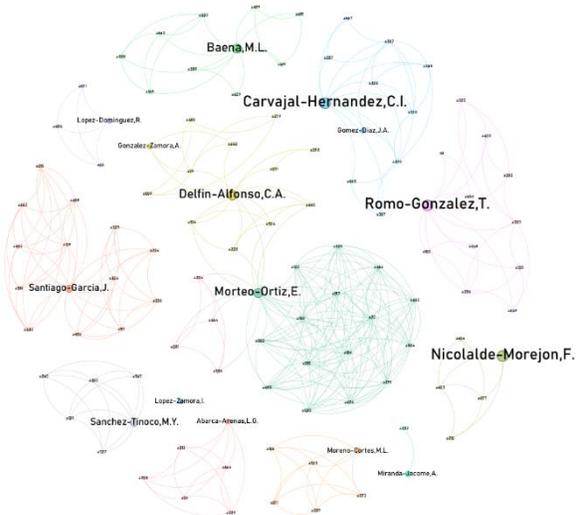


Figura A9: Artículos de 2018.

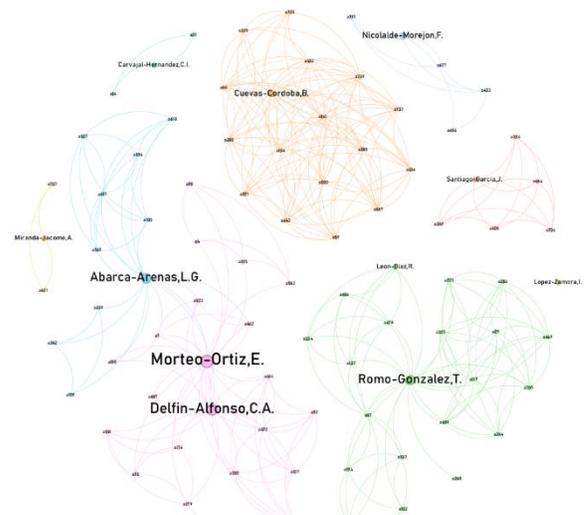


Figura A10: Artículos de 2019.

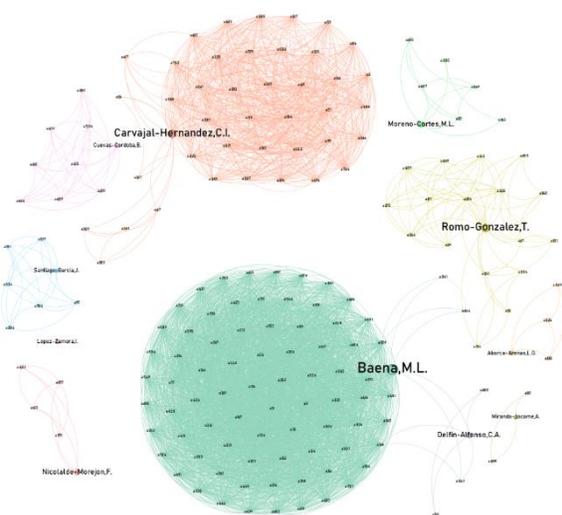


Figura A11: Artículos de 2020.

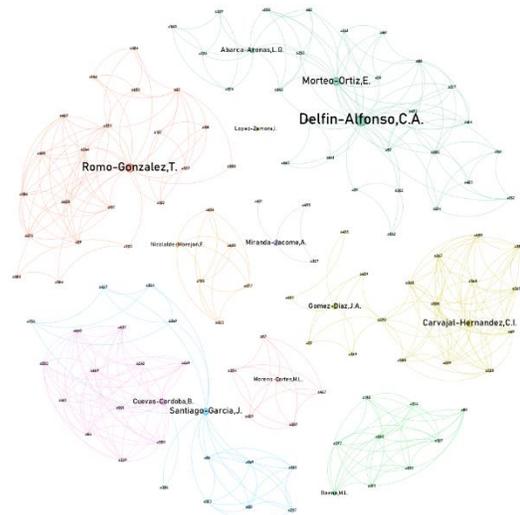


Figura A12: Artículos de 2021.

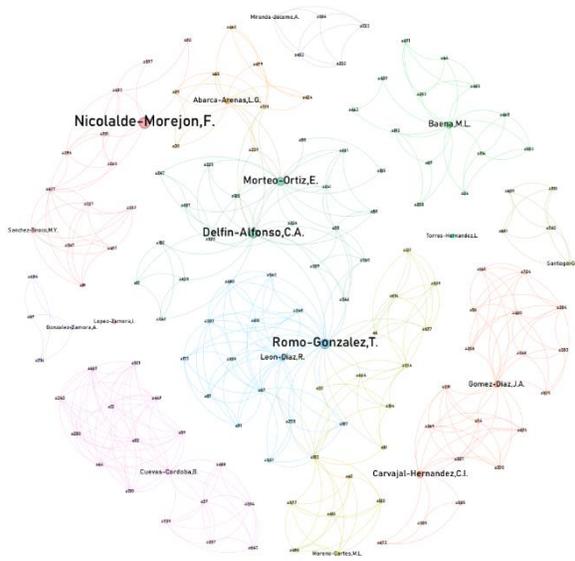


Figura A13: Artículos de 2022.

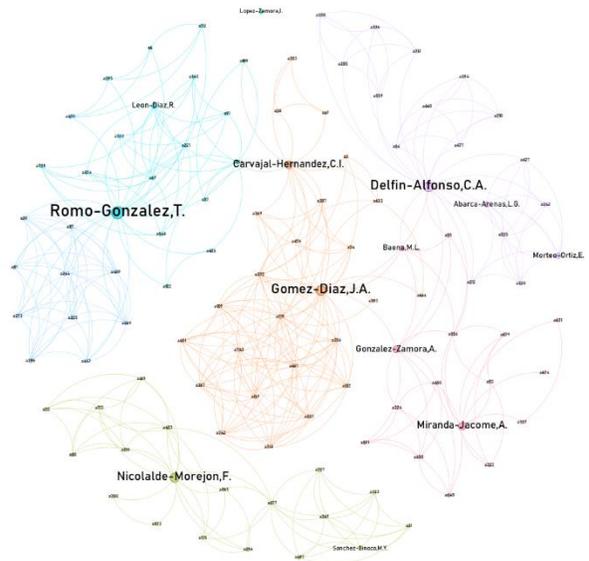


Figura A14: Artículos de 2023.

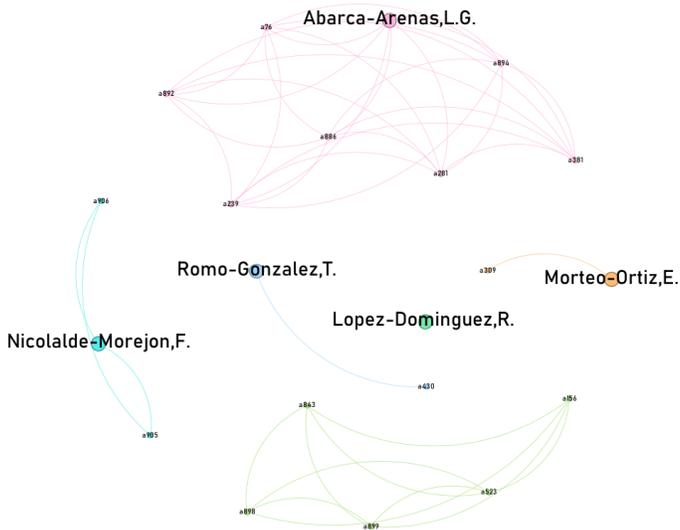


Figura A15: Capítulos de 2011.

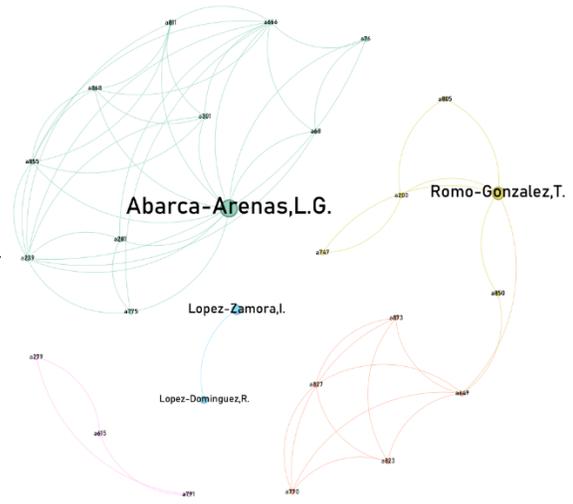


Figura A16: Capítulos de 2012.

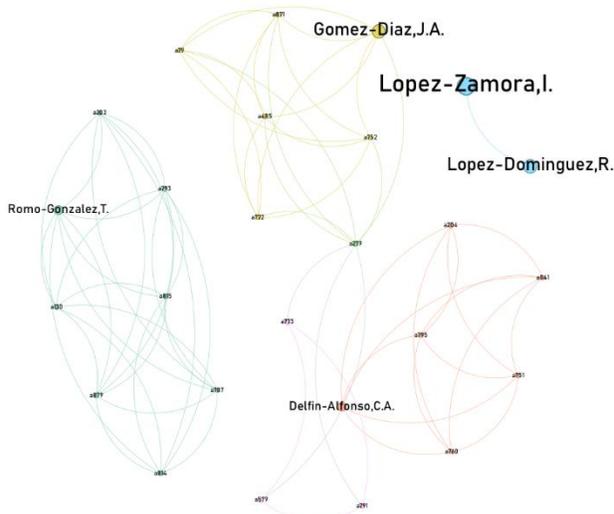


Figura A17: Capítulos de 2013.

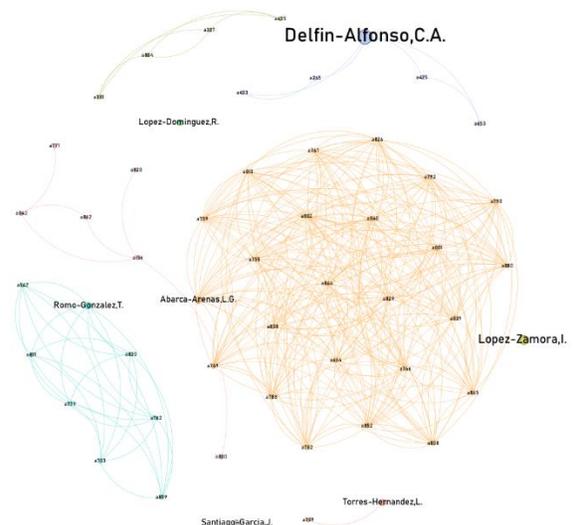


Figura A18: Capítulos de 2014.

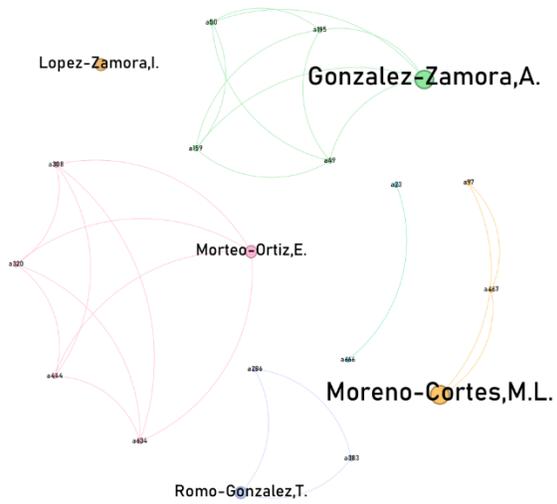


Figura A19: Capítulos de 2015.

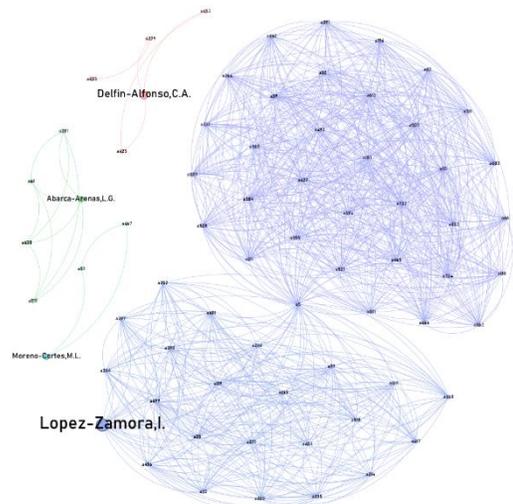


Figura A20: Capítulos de 2016.

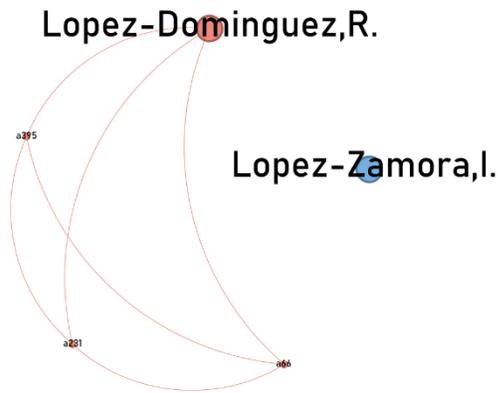


Figura A21: Capítulos de 2017.

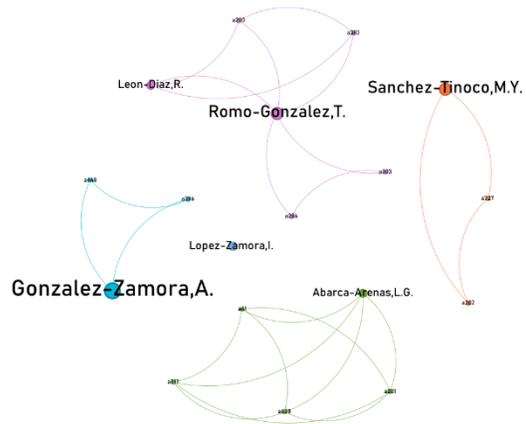


Figura A22: Capítulos de 2018.

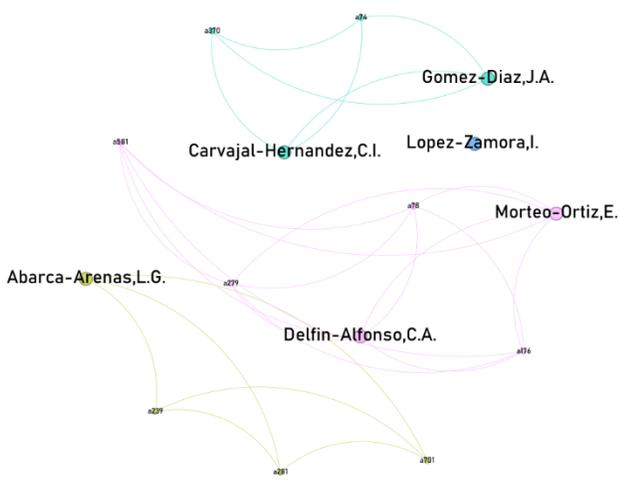


Figura A23: Capítulos de 2019.

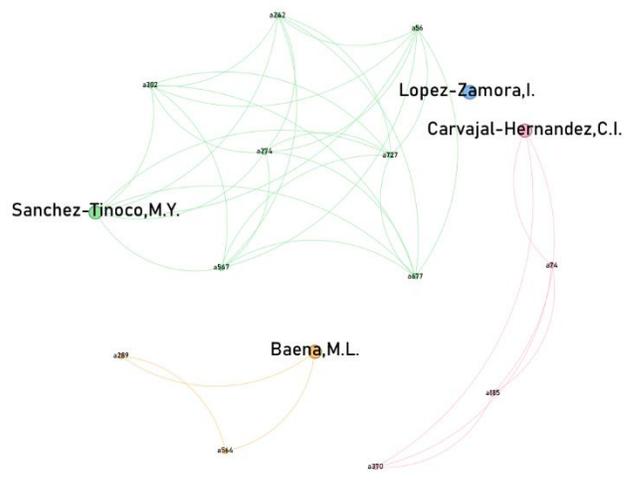


Figura A24: Capítulos de 2020.

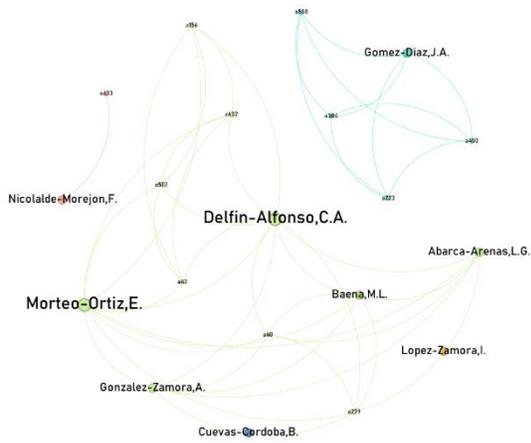


Figura A25: Capítulos de 2022.

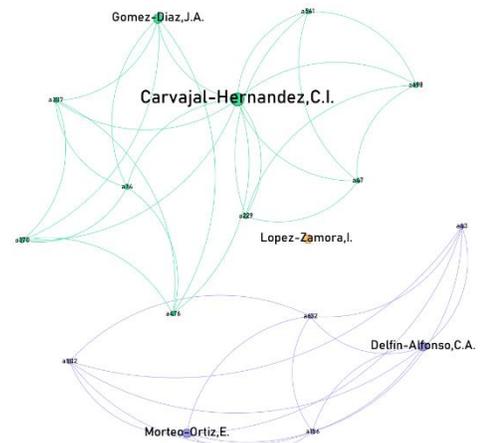


Figura A26: Capítulos de 2023.

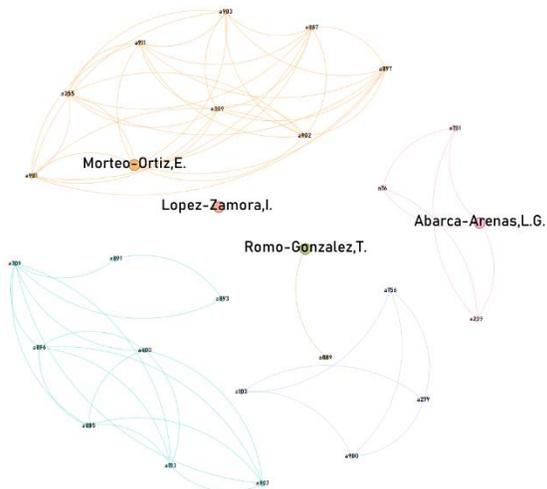


Figura A27: Libros de 2011.

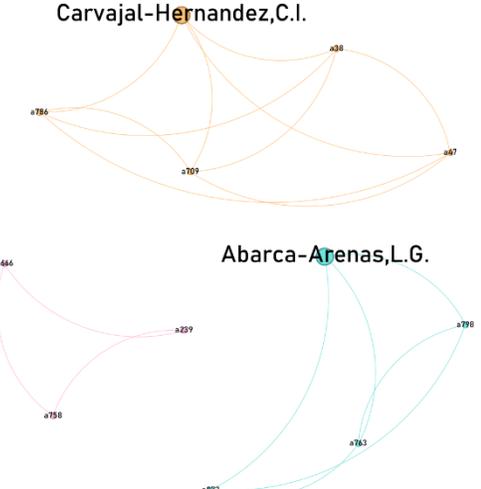


Figura A28: Libros de 2012.

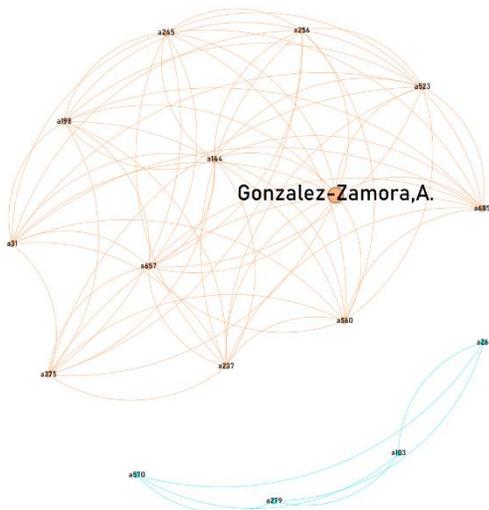


Figura A29: Libros de 2017.

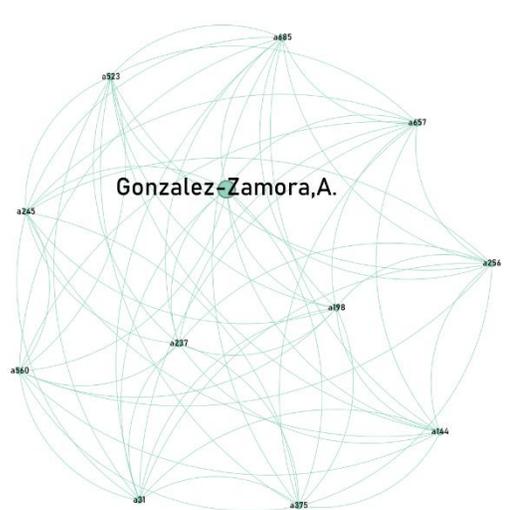


Figura A30: Libros de 2021.

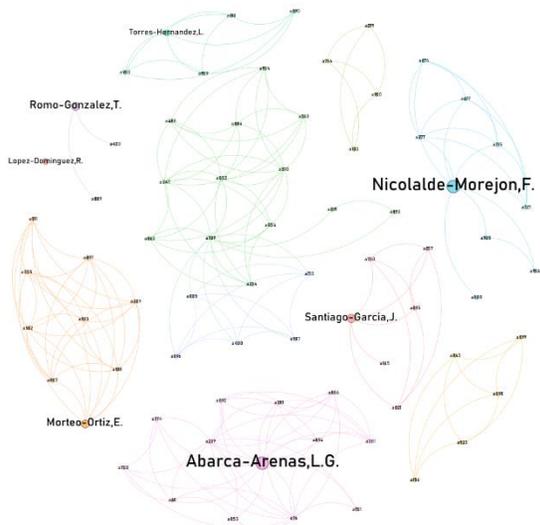


Figura A31: Producción total de 2011.

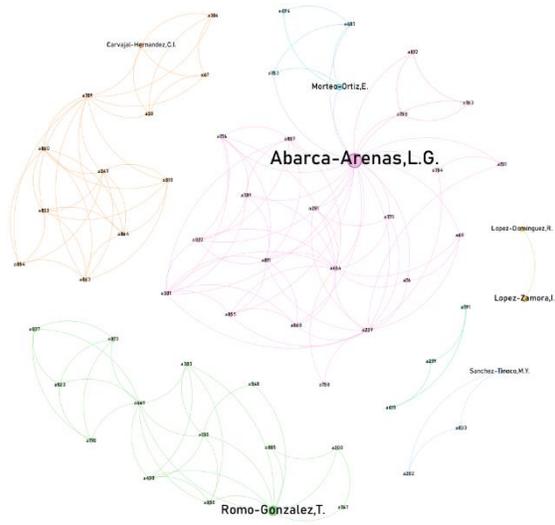


Figura A32: Producción total de 2012.

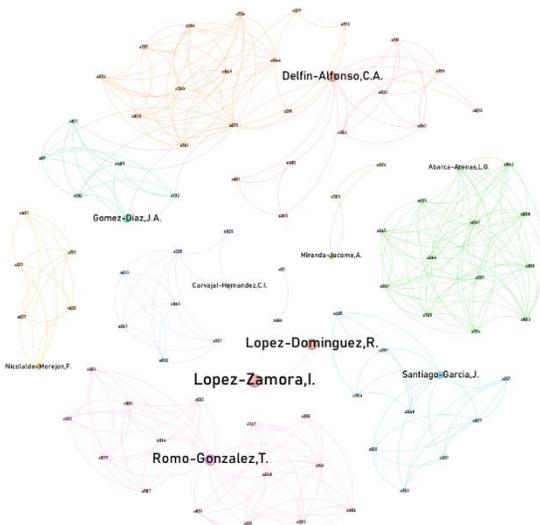


Figura A33: Producción total de 2013.

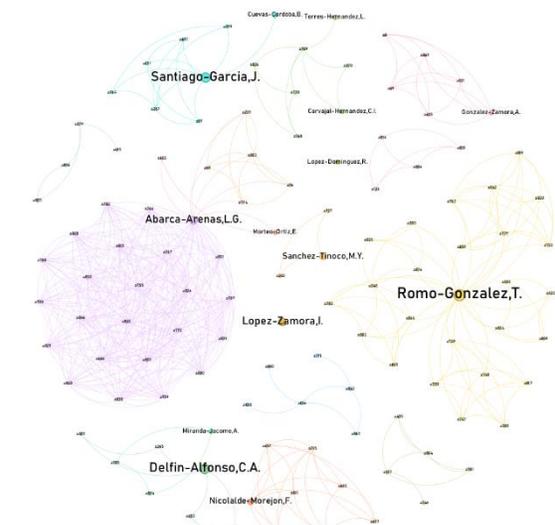


Figura A34: Producción total de 2014.

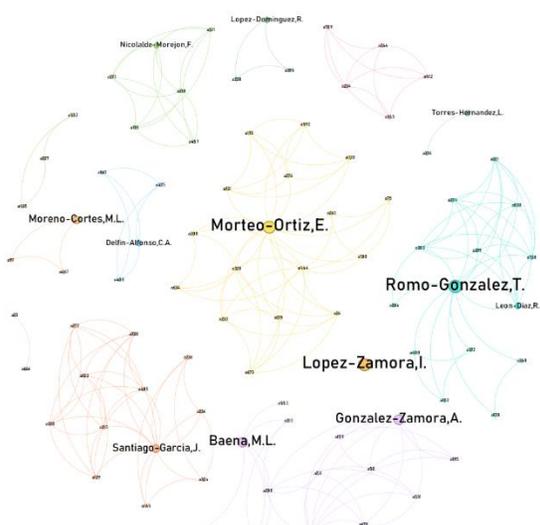


Figura A35: Producción total de 2015.

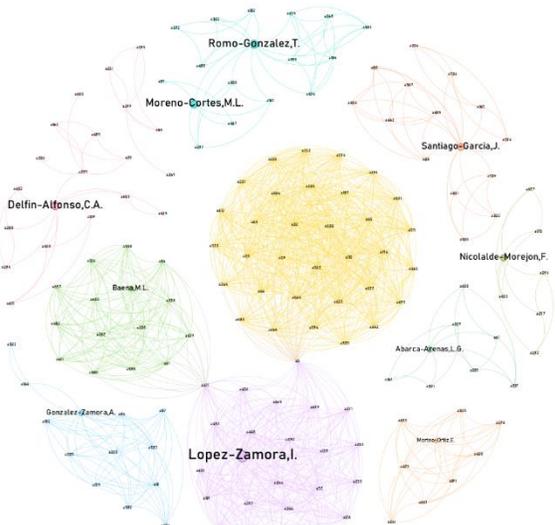


Figura A36: Producción total de 2016.

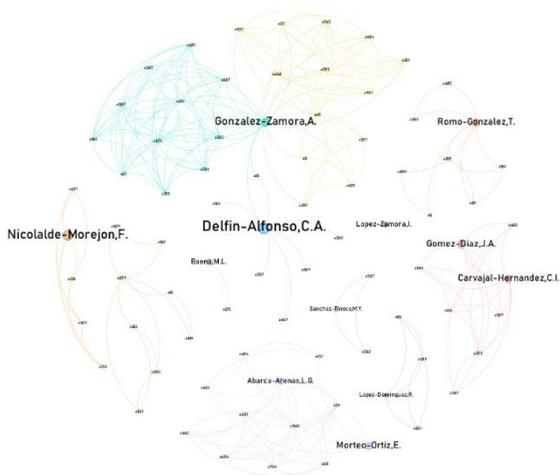


Figura A37: Producción total de 2017.

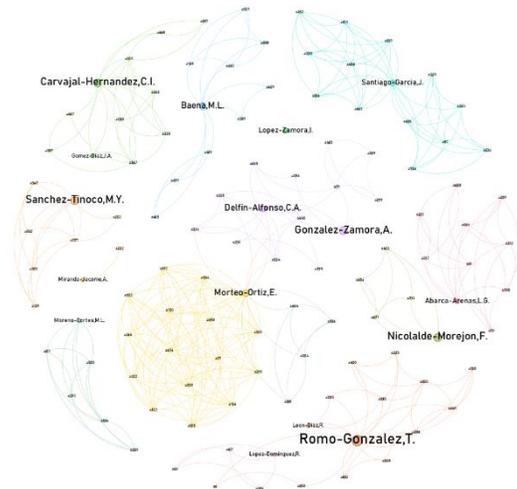


Figura A38: Producción total de 2018.

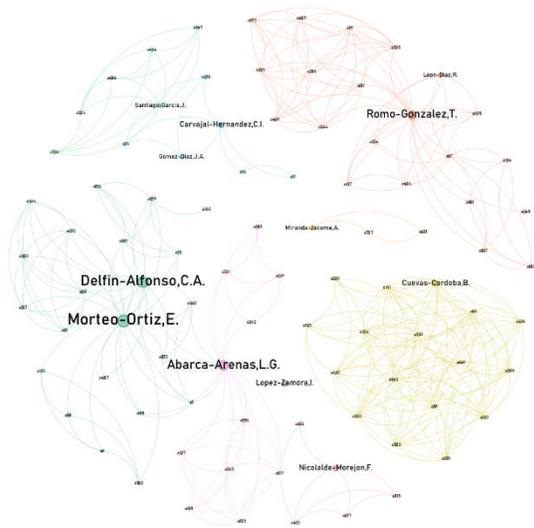


Figura A39: Producción total de 2019.

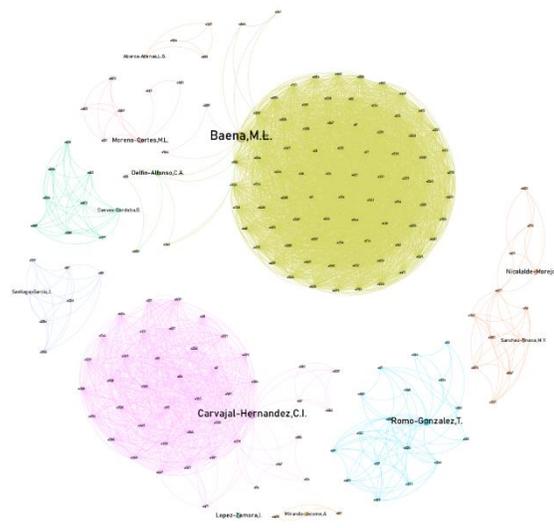


Figura A40: Producción total de 2020.

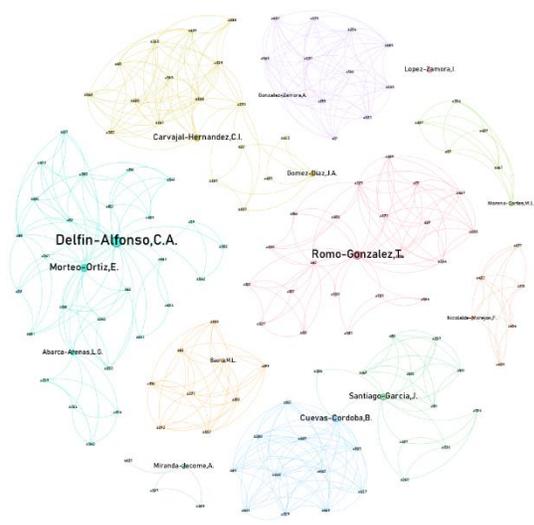


Figura A41: Producción total de 2021.

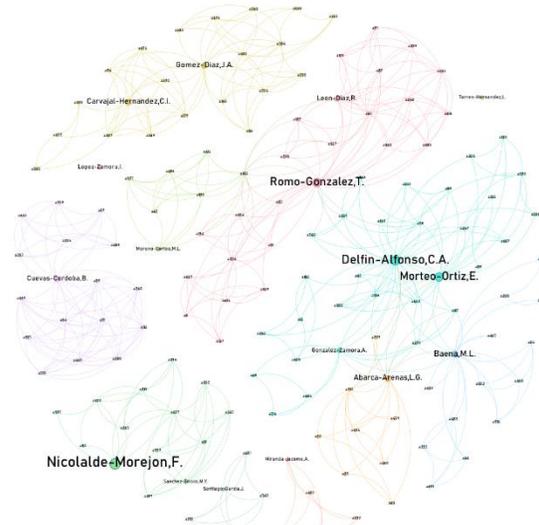


Figura A42: Producción total de 2021.

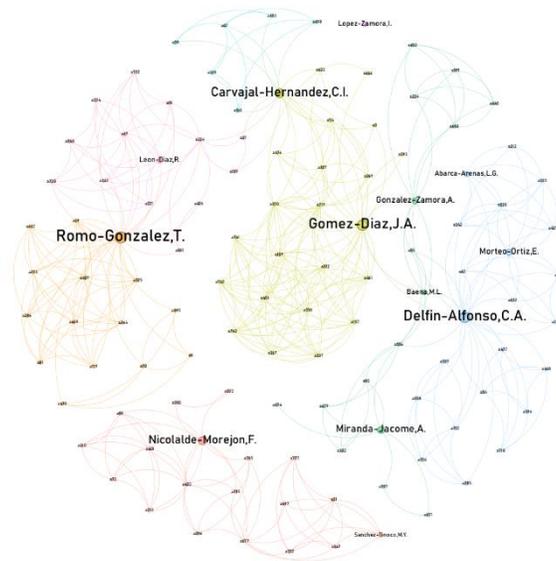


Figura A43: Producción total de 2023.

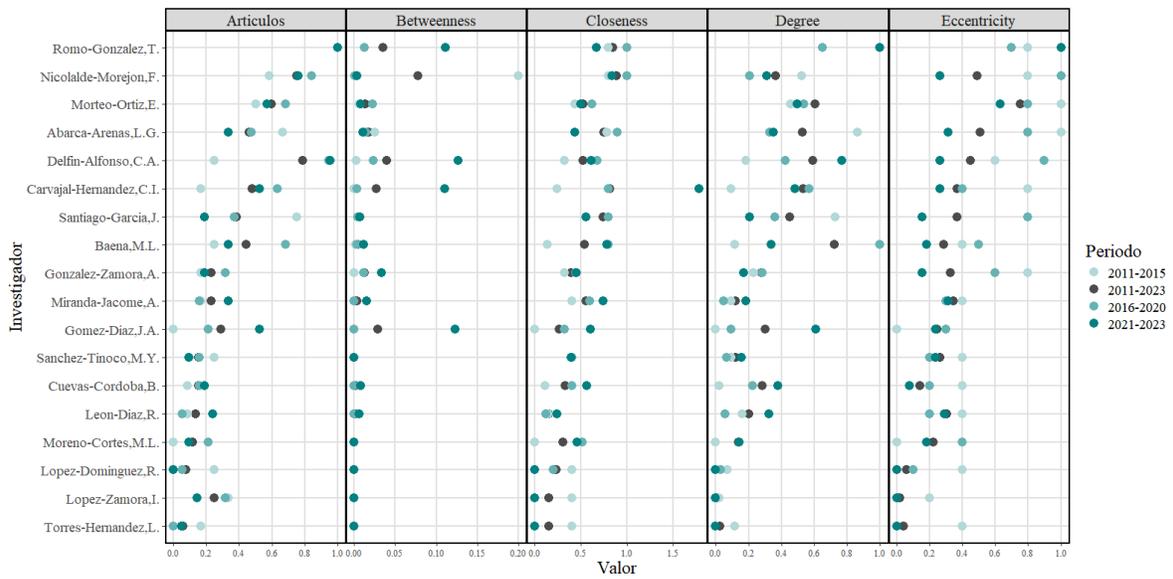


Figura A44: Centralidades de artículos de 2011 a 2023.

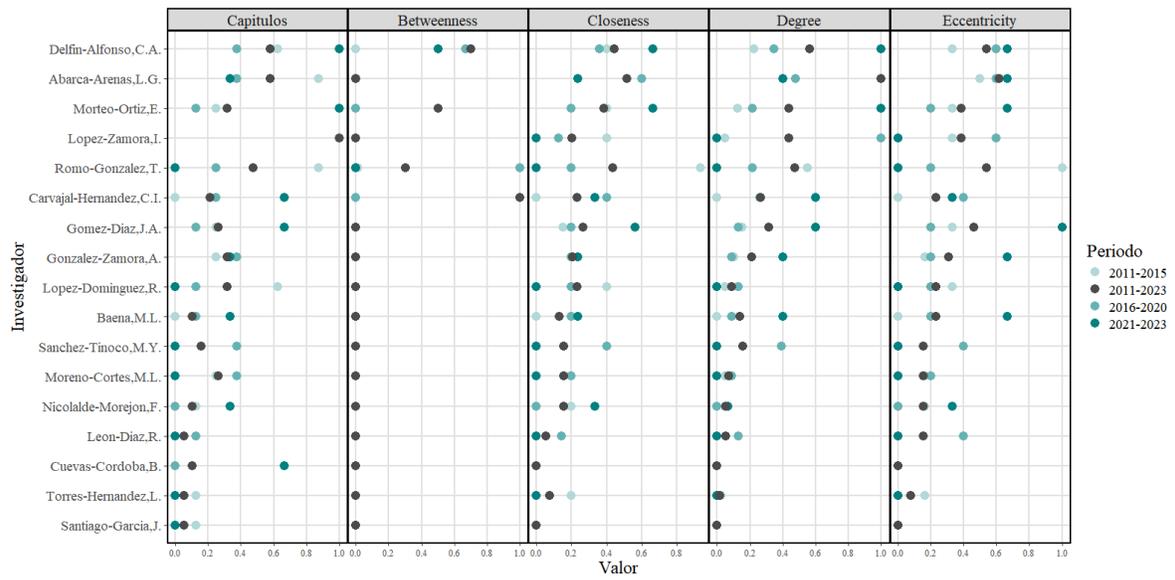


Figura A45: Centralidades de capítulos de 2011 a 2023.

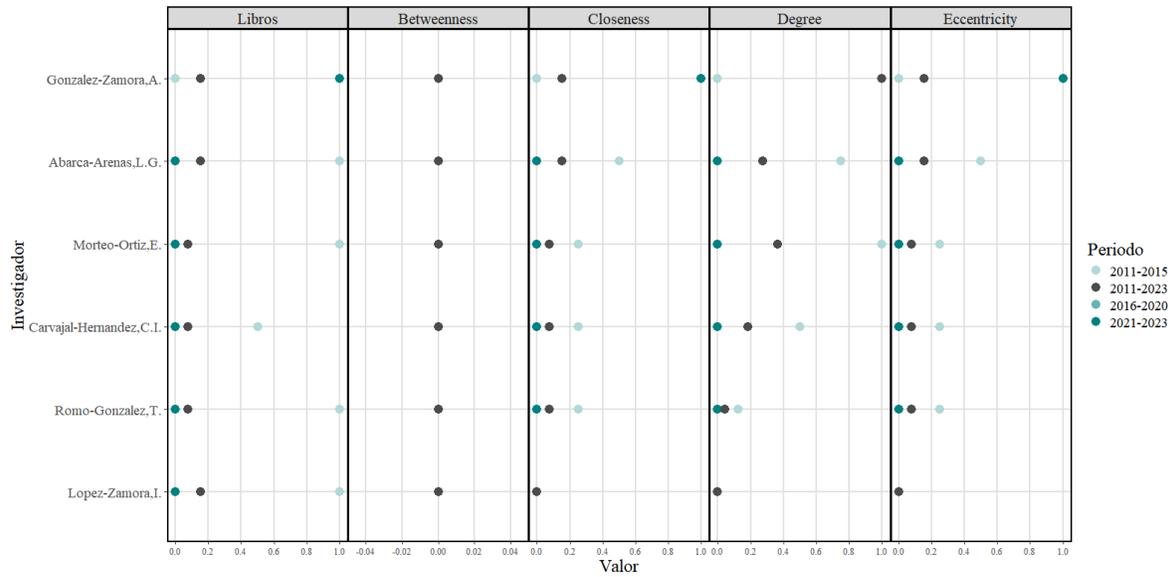


Figura A46: Centralidades de libros de 2011 a 2023.

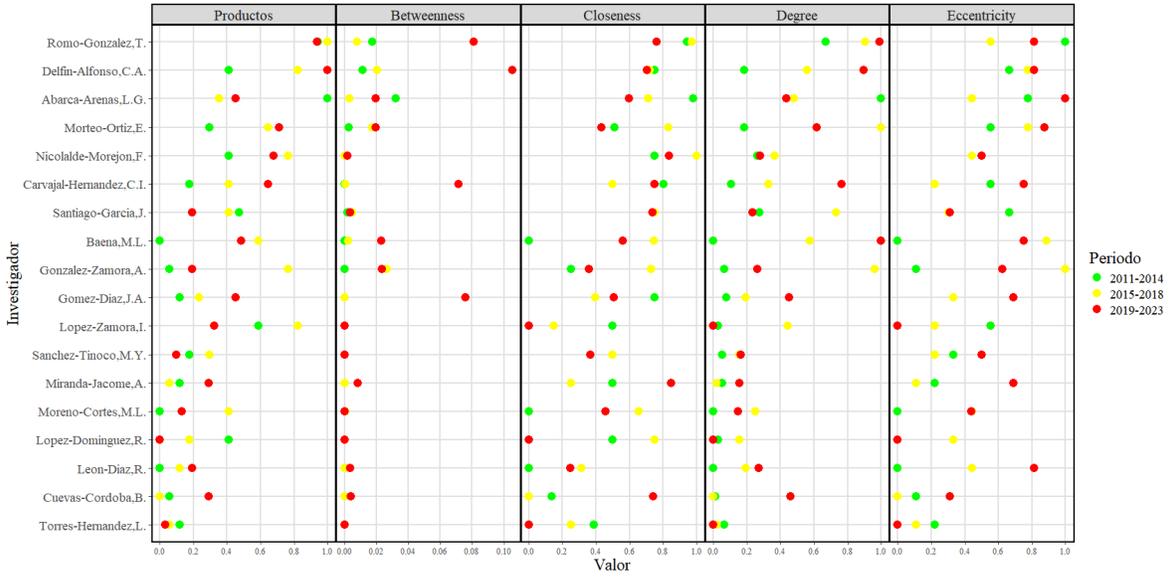


Figura A47: Centralidades de producción total de 2011 a 2023.