Manual de NDM/VNDM:

Programas para la identificación de áreas de endemismo

Szumik, Claudia¹, Dolores Casagranda¹, Sergio Roig Juñent ²

¹Instituto Superior de Entomología - Instituto Miguel Lillo, Miguel Lillo 205, CC 4000, San Miguel de Tucumán, Tucumán, Argentina, cszumik@csnat.unt.edu.ar ² Laboratorio de Entomología, Instituto Argentino de Investigaciones de las Zonas Áridas 9 (IADIZA), CC 507, 5500, Mendoza, Argentina.

AGRADECIMIENTOS:

A Fernando Navarro, Sara Bertelli y Lone Aagesen quienes aplicaron y utilizaron las primerísimas versiones de estos programas, de manera tal que muchas de las mejoras desarrolladas por Goloboff (gracias Pablo!) fueron gracias a sus aportes y críticas. A Juan Manuel Díaz Gómez, Germán San Blas, Andrés Quiroga, Marcela Peralta, Luis Grosso, Vivian Scheinsohn, Cecilia Domínguez y Jonathan Liria quienes valientemente aplican nuestras ideas a sus datos. A Tania Escalante por darnos el impulso necesario para dar a luz este manual. A Florencia Vera Candioti por sus comentarios y su agudeza estética. A Leila Taher por su interés en nuestro trabajo y su enorme colaboración como programadora (por hacernos la vida más fácil!). A Mercedes Lizarralde por su indispensable apoyo. Esta documentación fue realizada con equipo del INSUE e Instituto Argentino de Estudios Filogenéticos. Finalmente: Saludos a todos los que nos conocen!

ÍNDICE

Breve comentario sobre áreas de endemismo	3
Introducción al programa NDM-VNDM	5
Edición de datos	8
Edición de la grilla en VNDM	9
Origen de la grilla, tamaño de las celdas	10
Algunas teclas útiles	10
Cómo rellena la grilla VNDM?	11
Ctrl-s	13
Análisis de endemicidad	14
Características de la búsqueda	15
Exploración de los resultados	16
V <enter> Especies endémicas del área.</enter>	17
Ctrl - c Áreas parcialmente superpuestas	18
Ctrl - u Áreas que incluyen al área analizada	19
Ctrl - w Áreas incluidas dentro del área analizada	19
Areas de consenso	20
Metafiles	22
Acerca de los gráficos	23
Maneras adicionales de editar datos	24
Cómo se citan los programas?	25
Donde está descripto el método?	25
Trabaios que aplican este método?	25

Qué son las áreas de endemismo

Por qué nos importan?

Cuales son los problemas relacionados al tema?

El termino "área de endemismo" es utilizado en Biogeografía para referirnos a un patrón de distribución particular: un área geográfica delimitada por la congruencia en los rangos de distribución de, al menos, dos taxones.

Dado que la distribución de un taxón es producto de factores históricos y actuales, podemos inferir que aquellos taxones que presentan rangos de distribución similares, habrían sido influidos de manera similar por dichos factores (Szumik et al., 2002).

Para poder inferir los procesos que dan origen a las áreas de endemismo, un primer paso imprescindible es su identificación, es decir, el descubrimiento de los patrones de distribución de las especies. Por esto, las áreas de endemismo son consideradas la unidad básica de análisis en Biogeografía.

A pesar de su relevancia, el concepto de área de endemismo ha sido muy discutido, y su comprensión enfrenta diversos problemas, que pueden ser clasificados como: a) problemas semánticos, b) problemas asociados a la ausencia de un marco conceptual claro, y c) problemas analíticos (ver Anderson, 1994).

a) Problemas semánticos:

Estos problemas están relacionados con la gran cantidad (y diversidad) de definiciones referidas a los términos "endémico/a", "endemismo", y "área de endemismo", y a la proliferación de expresiones utilizadas como sinónimos de área de endemismo ("tracks generalizados", "trazos", "elementos bióticos", "centros de endemicidad", etc.). Tanto la diversidad de definiciones propuestas para el término "área de endemismo" como la variedad de palabras empleadas para hacer referencia a este patrón, han ocasionado confusiones semánticas respecto a este tema.

b) Ausencia de un marco conceptual claro:

La dificultad para diferenciar patrones de procesos, se hace evidente al momento de discutir sobre áreas de endemismo. Hay quienes consideran que un patrón de distribución puede denominarse área de endemismo sólo si se reconoce a la vicarianza como proceso formador del mismo, lo que implica el reconocimiento del patrón y del proceso que lo originó, en forma simultánea. Además de la falta de claridad conceptual que supone, esta idea implica obvias dificultades operativas, ya que, en la práctica, la descripción del patrón debe preceder, necesariamente, a cualquier hipótesis explicativa acerca de los procesos que lo formaron.

c) Problemas analíticos:

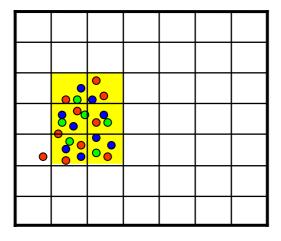
Durante los últimos años se propusieron numerosos protocolos para la delimitación de áreas de endemismo (ver Morrone, 1994; Linder 1998; Hausdorf, 2002; Hausdorf & Hennig, 2003; Harold & Mooi, 1994). Varias de estas propuestas fueron originalmente pensadas para definir otros tipos de patrón, por lo que fueron construidas sobre supuestos y conceptos teóricos que los vuelven inapropiados para la identificación de áreas de endemismo. Además el uso indiscriminado de estos métodos, basados en principios teóricos y procedimientos analíticos diferentes, dificulta más aún la comparación de hipótesis resultantes.

A continuación se describen los programas NDM - VNDM que aplican la propuesta de Szumik et al. (2002) y Szumik & Goloboff (2004). A diferencia de propuestas previas, este método, al basarse explícitamente en el concepto de áreas de endemismo según Platnick (1991), incluye el componente espacial y aplica un criterio de optimización durante la evaluación de las hipótesis y no después de la obtención de estas (Szumik, et al., 2002). Para una descripción completa del método ver los trabajos originales.

INTRODUCCIÓN AL PROGRAMA NDM-VNDM.

Los algoritmos desarrollados en los programas de computadora NDM/VNDM ver. 2.5 (Goloboff, 2005) implementan el método para la identificación de áreas de endemismo propuesto por Szumik et al. (2002) y Szumik & Goloboff (2004). El criterio de optimalidad que aplica el programa evalúa patrones de distribución sobre la base del concepto de áreas de endemismo. Dado que la distribución de un taxón es producto de factores históricos y actuales; si diferentes taxones responden de igual manera a esos factores debería haber concordancia en los rangos de distribución de dichos taxones (Szumik et al., 2002). El criterio evalúa mediante un índice de endemicidad cuántos y cuán endémicos son los taxones para un área dada. Aquellas áreas mejor apoyadas por los datos serán seleccionadas como áreas de endemismo. A semejanza de protocolos anteriores (ver por ejemplo Morrone, 1994) el método está obligado al uso de una grilla de ausencia/presencia de taxones para una región dada; si bien también es posible trabajar con datos georeferenciados (ver más abajo en edición de datos).

El índice de endemicidad que proponen Szumik & Goloboff (2004) es simple:



Dada esta grilla, un grupo de celdas (por ejemplo las amarillas) tendrá un valor de endemicidad que dependerá de cuán ajustadas están las distribuciones de los taxones "azul", "rojo" y "verde". El taxón azul tendrá un valor máximo dado que se encuentra en cada una de las celdas amarillas y está ausente en el resto de la grilla. El taxón verde tendrá un valor de endemicidad menor dado que está ausente en una de las celdas amarillas. Por último, el taxón rojo tendrá un valor menor aún dado

que, si bien está presente en cada una de las celdas amarillas, también está presente en una celda adyacente a dicha área. El valor de endemicidad del área amarilla será la suma de los índices de cada taxón endémico que posee.

La formula que aplica esta idea es:

$$p + (i \times Fi) + (a \times Fa)$$
 $IE_x =$
 $t + (o \times 1/Fo) + (d \times 1/Fd) + (n \times 1/Fn)$

donde:

p: número de celdas del área donde el taxón X está presente.

i: número de celdas del área donde el taxón X está inferido (cuando satisface la regla de homogeneidad, ver Szumik & Goloboff, 2004).

a: número de celdas del área donde el taxón X está asumido (determinado por el usuario).

t: número total de celdas que tiene el área.

o: número de celdas adyacentes al área donde el taxón X está presente.

d: número de celdas adyacentes al área donde el taxón X está asumido.

n: número de celdas no-adyacentes al área donde el taxón X está asumido.

F_i factor para presencias inferidas dentro del área (default 0.50)

F_a factor para presencias asumidas dentro del área (default 0.75)

F_o factor para presencias observadas fuera del área (default 0.50)

F_d factor para presencias asumidas adyacentes al área (default 2.00)

F_n factor para presencias asumidas no-adyacentes al área (default 0.50).

Estos factores hacen que cada uno de los términos de la formula sea más o menos influyente; todos ellos pueden ser modificados por el usuario, de manera que es posible tratar las presencias fuera del área, o ausencias dentro, en forma desde muy benévola hasta muy estricta (Szumik & Goloboff, 2004).

Cuando una celda satisface la regla de homogeneidad y ha sido registrada por el usuario como asumida, se utiliza el factor con mayor valor.

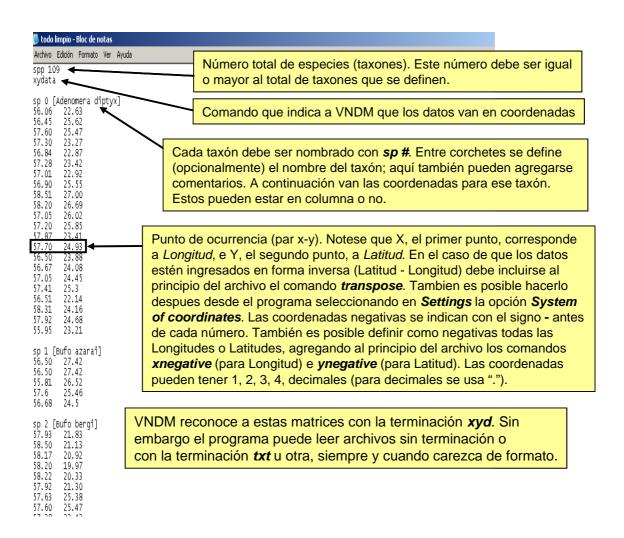
De manera que cuanto más especies se consideren como endémicas, y mientras mayor sea su *grado de endemicidad*, el grupo de celdas estará mejor apoyado como "área de endemismo" (Szumik & Goloboff, 2004).

Los algoritmos desarrollados por Goloboff (financiado por CONICET y GBIF) están agrupados en dos programas: NDM que realiza las búsquedas de áreas y VNDM que es el visor del anterior. NDM realiza soluciones heurísticas y permite definir un buen número de variables (e.g. modificar las constantes de la formula, dar un tope mínimo de endemicidad, evaluar y retener subóptimos, etc.) y trabaja con líneas de comandos tipo DOS; sin embargo, con el sistema de ventanas desarrollado en VNDM no hay necesidad alguna de invocar a NDM directamente. El programa VNDM edita grillas, envía líneas de comandos de búsqueda a NDM, analiza los resultados obtenidos por NDM, genera output files, metafiles, y exporta resultados a otros programas como Global-Mapper. Debe por lo tanto tenerse en cuenta que ambos programas DEBEN estar en la misma carpeta o en el path. Debe aclararse que tanto NDM como VNDM son programas gratuitos y de código abierto.

EDICIÓN DE DATOS.

NDM (el buscador) opera con matrices binarias de presencia/ausencia de especies por celda. VNDM (el visor) permite trabajar con las coordenadas geográficas de los taxones, convirtiéndolas fácilmente a datos de presencia/ausencia en una grilla. VNDM permite múltiples formas de generar matrices, lo cual dependerá de la información original que se posee:

1)- Coordenadas geográficas. Los datos en forma de coordenadas geográficas (latitud/longitud) nos permiten explotar al máximo estos programas. VNDM permite generar en tiempo mínimo una gran variedad de grillas dado que es posible definir: origen de la grilla, tamaño de las celdas, etc. (ver mas abajo) además de facilitarnos la visualización y el análisis de los resultados. En un archivo de texto sin formato (notepad, editor de DOS, etc.) se incluyen los puntos de ocurrencia para cada taxón (coordenadas geográficas x-y), este archivo debe incluir:



En este tipo de matrices VNDM permite poner de fondo los límites de un mapa. Esto se logra incluyendo al final del archivo: map

> line 66.217 21.767 66.183 21.783 66.167 21.800 66.117 21.817 66.067 21.833 66.050 21.850 66.050 21.867 66.050 21.883

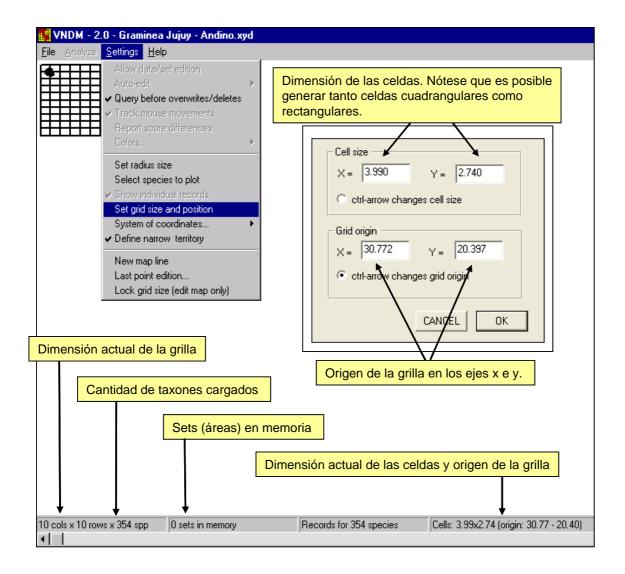
Cada línea (máximo 300 líneas) del mapa puede tener hasta 5000 referencias planas. Nótese que si se termina de definir la península de Yucatán y se pasa a la costa occidental de México debe definirse una nueva línea, sino ese salto aparecerá dibujado también en el mapa. También es posible editar las líneas en modo gráfico y salvar posteriormente las coordenadas resultantes.

Edición de la grilla en VNDM. El archivo xyd (el input file) puede ser leído arrastrandolo hasta el ícono de VNDM, o cliqueando sobre el ícono de VNDM y seleccionando la ruta donde se encuentra el archivo xyd (o, si ya se ha establecido la asociación entre la extensión xyd y el programa VNDM, haciendo doble clic en el archivo mismo):

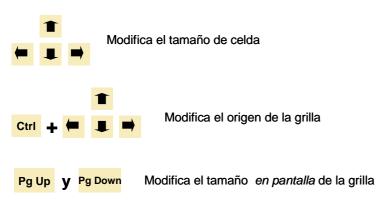


Una vez seleccionado el archivo, VNDM automáticamente asigna un origen en los ejes x e y; además, genera una grilla de 10 celdas x 10 celdas. Está claro que el origen y tamaño de celdas adecuados dependerán de los datos y del usuario.

Origen de la grilla, tamaño de las celdas.



Algunas teclas útiles:

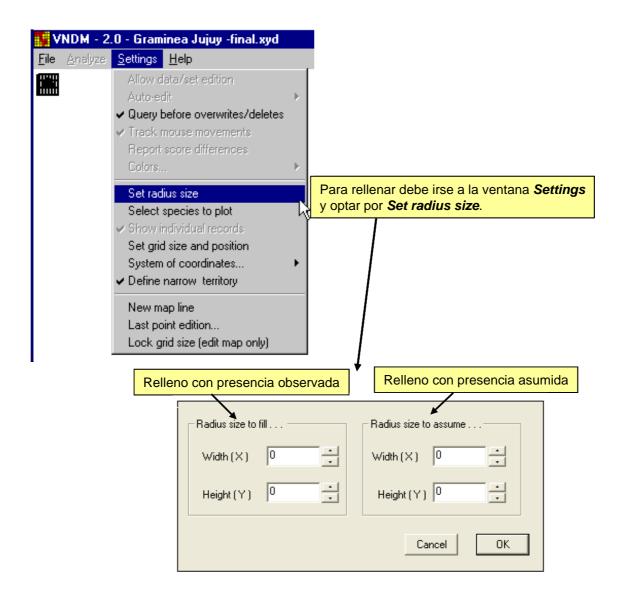


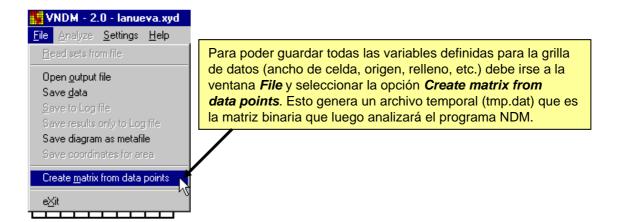
Instituto Argentino de Estudios Filogenéticos, Año V, Vol. (3)

Cómo rellena la grilla VNDM?

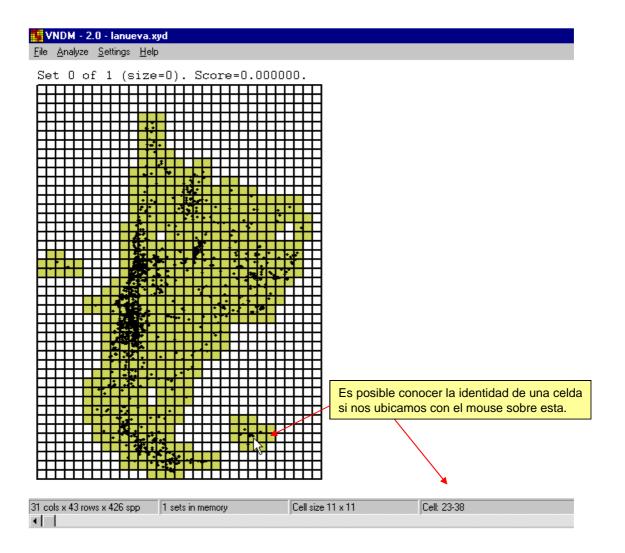
Los puntos que caen sobre el limite de dos celdas son considerados *siempre* como presencia observada para ambas celdas; lo mismo vale para los vértices, en ese caso las cuatro celdas tendrán presencia observada.

VNDM también permite hacer un relleno especial en el caso de aquellos puntos que caen cerca del límite de una celda. Debe especificarse el radio de relleno que va de 0 a 100 (el default es 0) y el tipo de relleno: presencia observada (de igual valor que la celda que posee el punto) o presencia asumida (altamente probable que este presente, y vale menos que la observada). Este relleno puede también hacerse en forma manual, taxón por taxón, utilizando el mouse (ver más abajo).

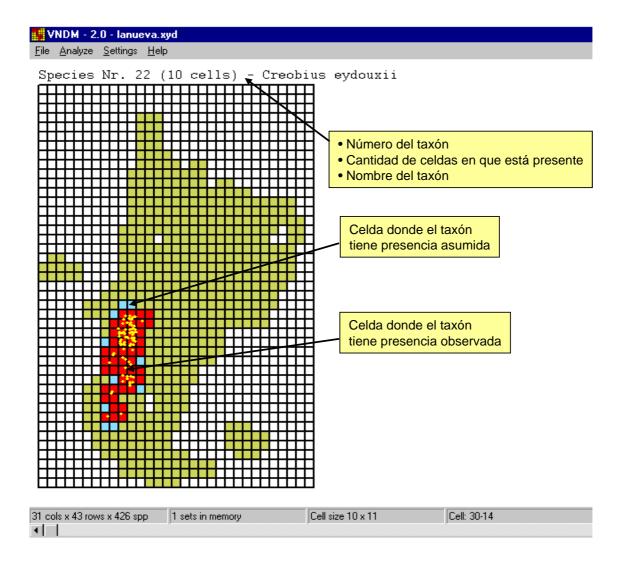




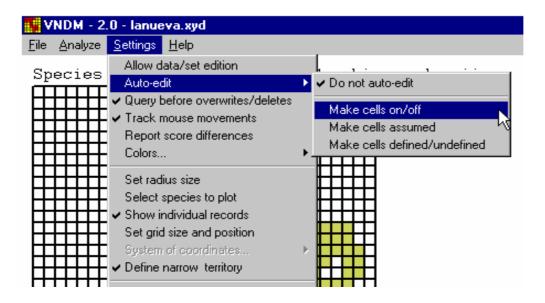
Una vez seleccionada esta opción, VNDM ya nos muestra la grilla con la totalidad de puntos de referencia:



Con las teclas *Ctrl-s* se puede visualizar las distribuciones de los taxones:

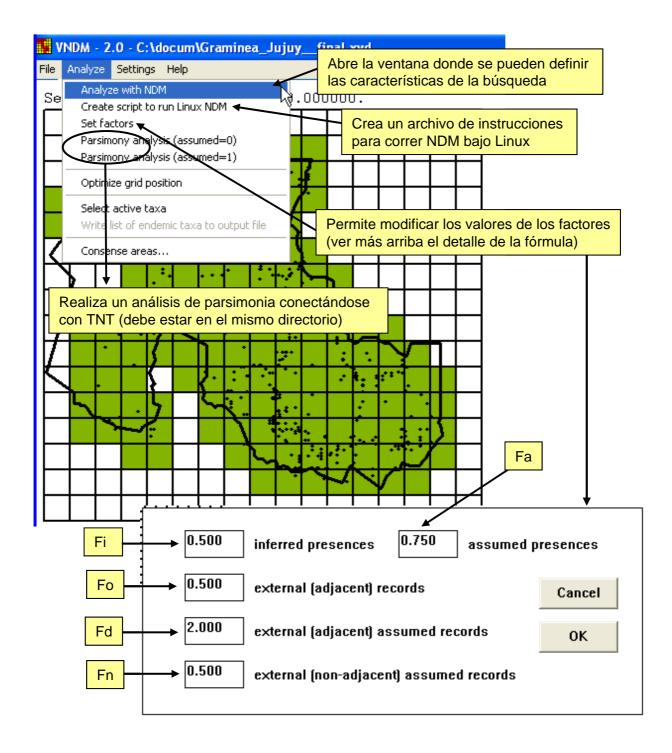


Aquí es posible eliminar o reincorporar las presencias observadas y asumidas en forma manual con las teclas del mouse (*Settings/AutoEdit*, etc.). Es claro que una vez finalizada la edición manual debe guardarse con la opción *File/SaveData*.

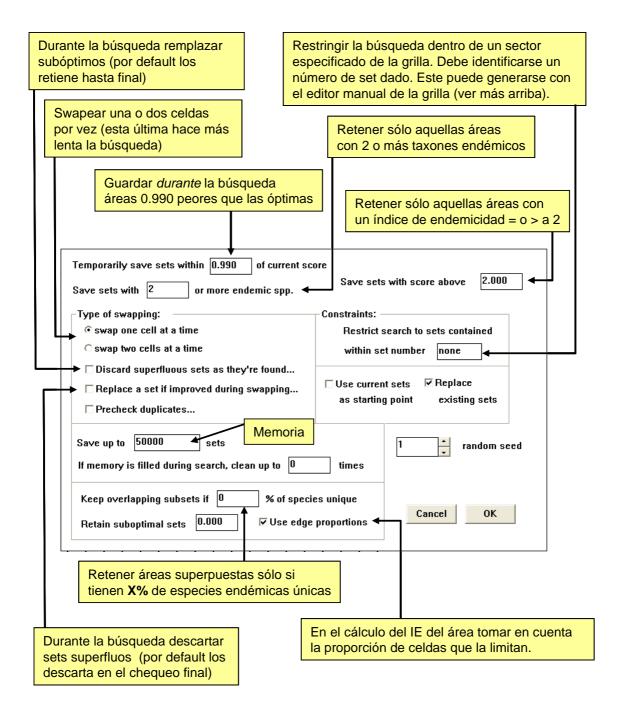


ANÁLISIS DE ENDEMICIDAD.

Si bien las características de la búsqueda pueden definirse desde VNDM, el análisis de endemicidad lo realiza NDM, por tanto ambos programas deben estar en el mismo directorio para poder comunicarse.

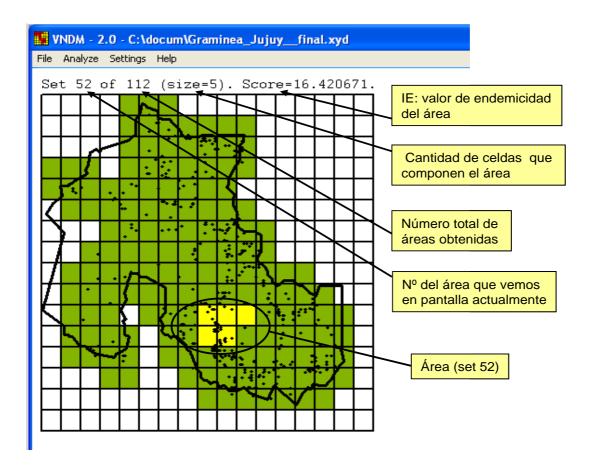


Ventana de características de la búsqueda (*Analyze with NDM*): Las búsquedas que realiza NDM son heurísticas; aquí se comentan brevemente las variables más comúnmente utilizadas. Para la descripción de las búsquedas heurísticas ver Szumik & Goloboff (2004).



EXPLORACIÓN DE LAS ÁREAS RESULTANTES.

En la primera vista de los resultados podemos obtener información básica acerca de cada una de las áreas obtenidas (con <enter> y <backspace> se avanza o retrocede):



Además de estos datos primarios, podemos extraer una buena cantidad de información a través de una serie de teclas:

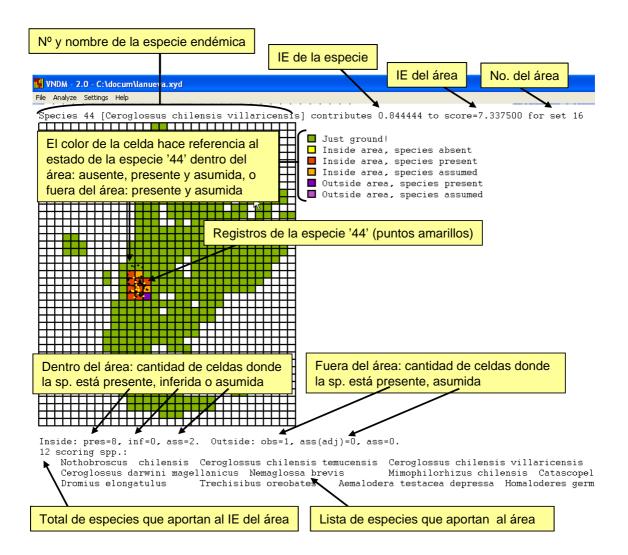
v <enter> Indica cuántas y cuáles taxones endémicos aportan al IE del área, así como
 Los registros que posee y el valor con que contribuyen.
 v - muestra las spp que NO dan score.

Ctrl + c Compara áreas que tienen celdas en común (parcialmente superpuestas), e informa diferencias en composición de especies e IEs de las áreas.

Ctrl + u Indica si el área analizada está incluida en otras, y si este anidamiento es sólo espacial, o si involucra la composición de especies endémicas.

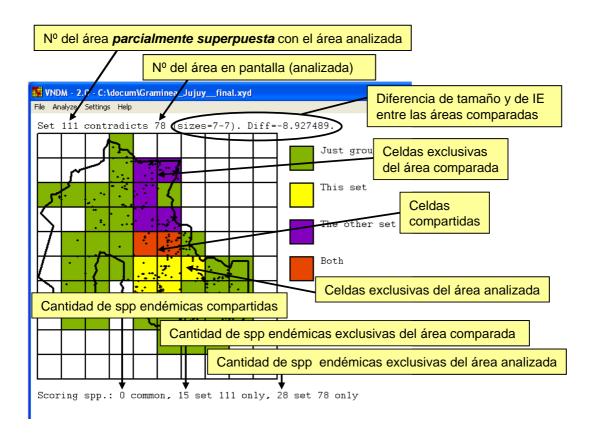
- Ctrl + w Indica si el área analizada incluye otras áreas.
- Ctrl + f Muestra el total de especies presentes en el área analizada sean endémicas o no.
- Ctrl + 1 Muestra todas aquellas áreas de endemismo compuestas por una única celda.
- *n* Marca aquellas celdas de la grilla de estudio que carecen de datos.
- **d** Elimina áreas duplicadas; **z** muestra áreas duplicadas.
- Ctrl + m Permite seleccionar áreas =, > o < a un valor dado de ie, tamaño del área, etc.
- Ctrl + i Invierte la selección.
- Ctrl + delete Elimina la selección.

V <*enter*> Especies endémicas del área.

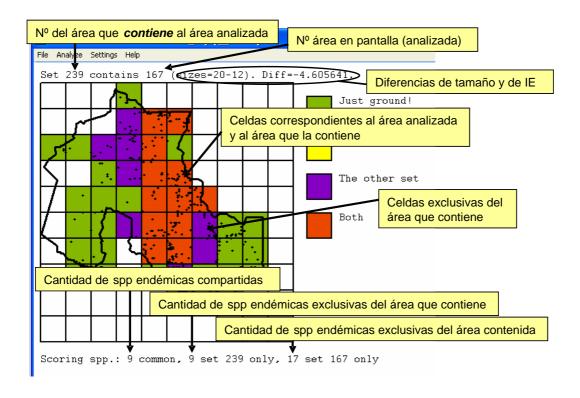


Otras teclas útiles. F1 y F2: modifican el grosor de las líneas de la grilla y de los registros; <enter> y <backspace>: se avanza o retrocede en la lista de spp endémicas del área; Settings/ReportScoreDifferences: permite evaluar con el mouse las diferencias de score al agregar o eliminar una celda; Shf-m o F12: guarda el área en pantalla en un metafile.

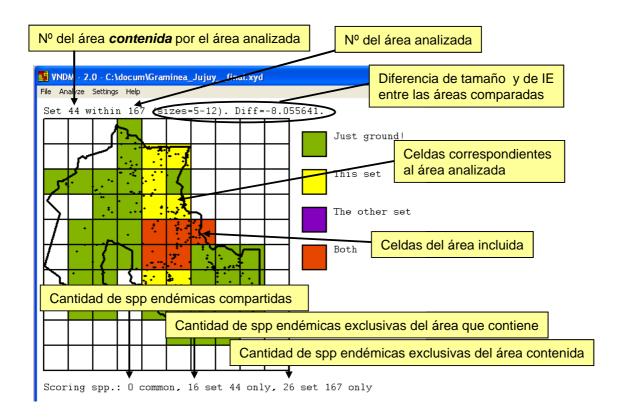
Ctrl - c Áreas parcialmente superpuestas con el área en vista actual.



Ctrl - u Áreas que incluyen al área analizada.



Ctrl - w Áreas incluidas dentro del área analizada.

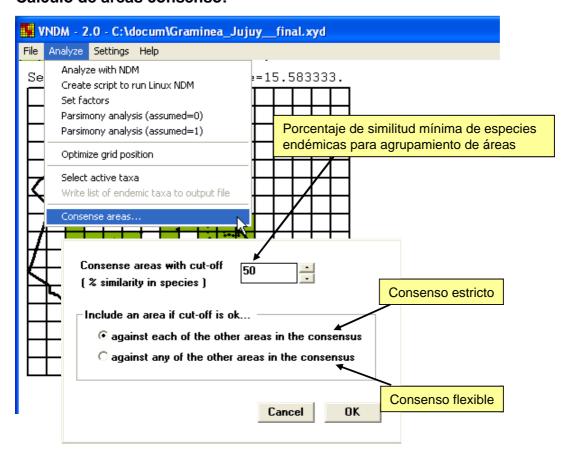


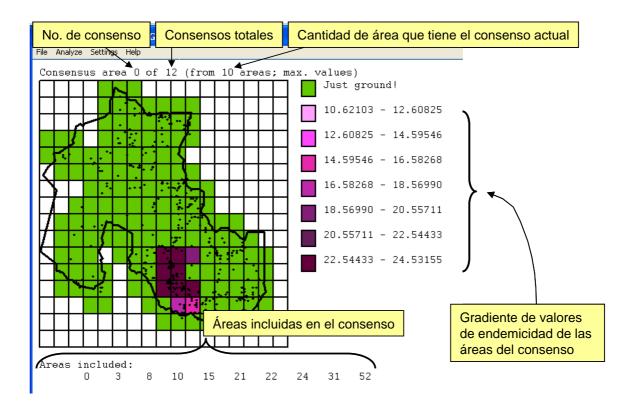
ÁREAS CONSENSO.

Como resultado del criterio de búsqueda aplicado por NDM/VNDM, muchas veces se obtienen áreas que difieren levemente en la presencia de algunas celdas o especies endémicas. En vez de descartar estos resultados, otra posibilidad es sintetizar toda esa información contenida en todas esas áreas "semejantes" en un área consenso. Las áreas de consenso (Goloboff & Szumik, en prep.) resumen la información contenida en aquellas áreas individuales que comparten un porcentaje dado de especies endémicas, facilitandonos en gran medida la comparación y evaluación de los resultados.

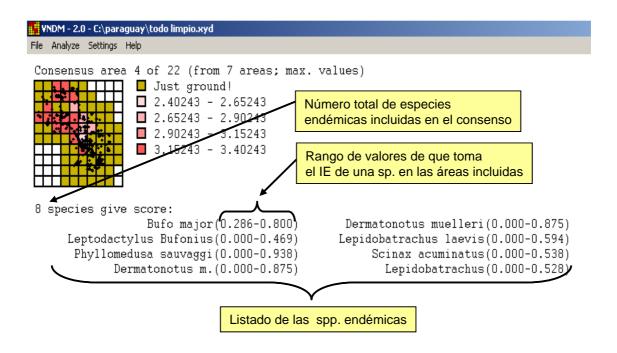
Los consensos pueden clasificarse, según el criterio aplicado en **Consenso estricto** (un área individual formará parte del consenso siempre y cuando comparta un porcentaje dado de especies endémicas con *todas* las áreas que componen dicho consenso) y **Consenso flexible** (un área individual será incluida en el consenso mientras comparta el porcentaje dado de especies endémicas con *alguna* de las áreas que componen dicho consenso).

Cálculo de áreas consenso:

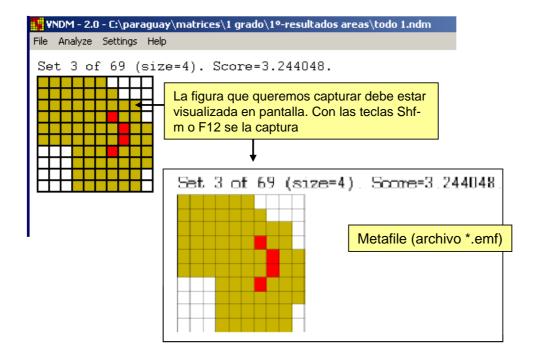




V Muestra cuales son las especies que aportan valor al IE del consenso, y el rango de valores que toma cada una de ellas en las áreas individuales incluidas en el consenso.

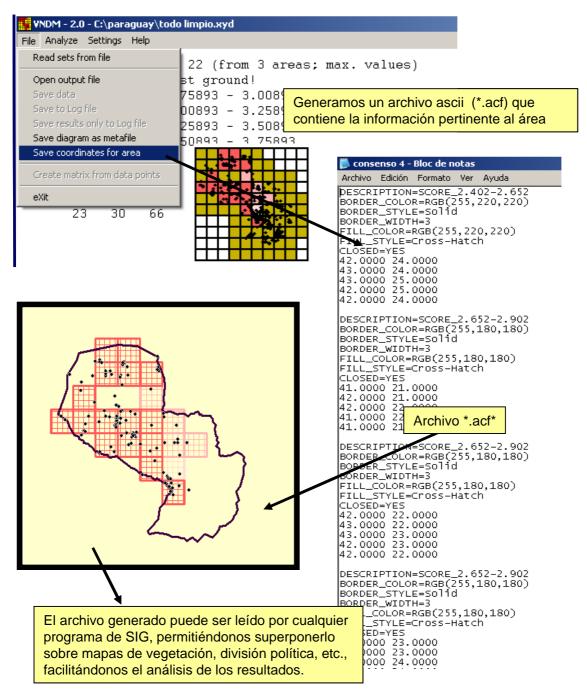


METAFILES. Otra manera de capturar gráficos desde NDM/VNDM es mediante la generación de metafiles; éstos son archivos de imagen (*.emf) que capturan la figura en pantalla, y posteriormente pueden ser modificados desde cualquier programa de edición de imágenes (ej: paint, photoshop, etc.)



ACERCA DE LOS GRÁFICOS:

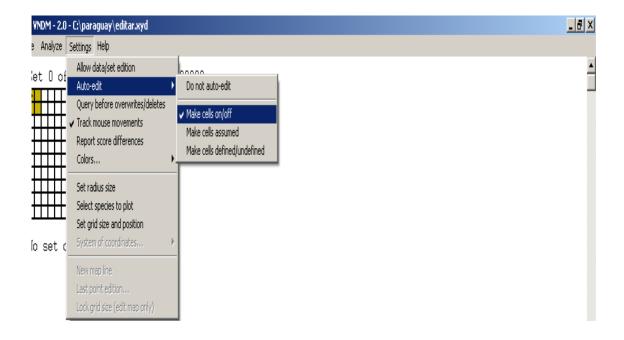
Los gráficos de las áreas de endemismo, así como de las áreas de consenso, pueden capturarse desde NDM/VNDM en formato ascii, para luego ser leídos por cualquier programa de SIG. Además de los gráficos, el archivo ascii generado guarda los datos relacionados a las áreas, así como también los datos de las especies contenidas en éstas. La edición y mapeo de las figuras se facilita en gran medida, ya que la información espacial (coordenadas geográficas) está incluidas en éstos archivos.



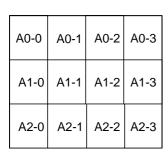
MANERAS ADICIONALES DE EDITAR DATOS.

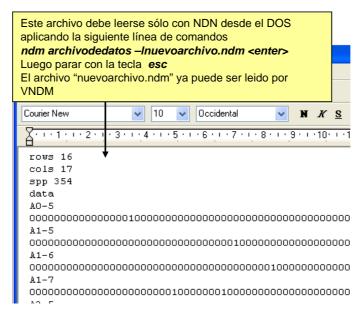
Como dijimos mas arriba, además de los mencionados, existen otras formas de edición de grillas (matrices):

Edición manual desde VNDM. A veces la información que se posee está en forma de mapas, etc. En estos casos esta información puede editarse de manera "manual": primero se genera un archivo con datos ficticios (según se explicó al principio), y luego se rellenan las distribuciones de cada taxón desde VNDM en forma manual. Estas matrices NO se "guardan" automáticamente, de manera que deben ser salvarse como archivo de datos de VNDM con la opción save to logfile o como matriz binaria con la opción save data.



Edición de una matriz binaria. En el caso de que los datos originales se encuentren en formato de matriz binaria, donde las especies son las columnas y las celdas de la grilla las filas. Las celdas para ser identificadas espacialmente siguen la siguiente nomenclatura:





Cómo se citan los programas?

Goloboff, P. 2005. NDM/VNDM ver. 2.5. Programs for identification of areas of endemism. Programs and documentation available at www.zmuc.dk/public/phylogeny/endemism

Donde está descripto el método?

- Szumik C., F. Cuezzo, P. Goloboff & A. Chalup. 2002. An optimality criterion to determine areas of endemism. *Syst. Biol.* 51: 806-816.
- Szumik C. & P. Goloboff. 2004. Areas of endemism. An improved optimality criterion. *Syst. Biol.* 53:968-977.

Trabajos que aplican este método?

- Aagesen, L., C. Szumik, F. Zuloaga, & O. Morrone. 2005. Áreas endémicas y comunidades de Gramíneas en la provincia de Jujuy, noroeste de Argentina. XXX Jornadas Argentinas de Botánica: 117.
- Bertelli S., C. Szumik, P. Goloboff, A. Navarro, A. Townsend Peterson & J. Cracraft. 2005. Identification of Areas of Endemism in Mexico. *AOU*: 417.
- Díaz Gómez, J.M. 2005. Los *Liolaemus* de la Puna Argentina: identificación de áreas de endemismo con NDM y PAE. *VI Congreso Argentino de Herpetología*: 16.
- Domínguez M.C., S. Roig-Juñent, J.J. Tassin, F. Ocampo & G. 2006. Flores. Areas of endemism of the Patagonian steppe: an approach based on insect distributional patterns using endemicity analysis. *J. Biogeogr.* 33:1527-1537.
- Casagranda D. & E. Lavilla. 2005. Anuros de Nordeste Argentino: Identificación de Patrones biogeográficos. *Il Congreso Brasilero de Herpetología.*
- Navarro F., F. Cuezzo, M. Lizarralde de Grosso & C. Szumik. 2004. Áreas de endemismo en las yungas argentinas: un análisis de endemicidad aplicado a insectos. *V Reunión Argentina de Cladística y Biogeografía*: p. 46.
- Roig-Juñent S., C. Szumik & D. Casagranda. 2004. Determinación de las áreas de endemismo de América del Sur Austral: un ejemplo aplicando el criterio de optimalidad. *V Reunión Argentina de Cladística y Biogeografía*: 60.
- Szumik C. & S. Roig-Juñent. 2005. Criterio de optimalidad para áreas de endemismo: El caso de América del Sur Austral. [en "Regionalización biogeográfica en Iberoamérica y tópicos afines" Morrone & Llorente Bousquets (eds). UNAM] pags. 495-508.
- Brusquetti F., Casagranda D., 2005. Anuros de Paraguay: Análisis de endemicidad y riqueza de especies. *VI Congreso Argentino de Herpetología: 39.*