

**MINERÍA DE DATOS DEL RESIDUO TERMINAL-N Y ANÁLISIS DEL TIEMPO
DE VIDA DE LAS ENZIMAS PRESENTES EN PROTEOMAS DE REFERENCIA
(*Escherichia coli*, *Nicotiana tabacum*, *Oryctolagus cuniculus*, *Saccharomyces cerevisiae*
y *virus inmunodeficiencia humana (VIH-1)*)**

DIEGO ÁNDRES ROJAS
-Estudiante de Maestría en Química-

**UNIVERSIDAD DE PAMPLONA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE QUÍMICA
PAMPLONA, NORTE DE SANTANDER, COLOMBIA
2022**

**MINERÍA DE DATOS DEL RESIDUO TERMINAL-N Y ANÁLISIS DEL TIEMPO
DE VIDA DE LAS ENZIMAS PRESENTES EN PROTEOMAS DE REFERENCIA
(*Escherichia Coli*, *Nicotiana tabacum*, *Oryctolagus cuniculus*, *Saccharomyces cerevisiae*
y *virus inmunodeficiencia humana (VIH-1)*)**

DIEGO ÁNDRES ROJAS
-Estudiante de Maestría en Química-

Trabajo de grado para optar el título de Magister en Química

Dr. DANIEL IVÁN BARRERA VALDERRAMA
-director-

Dr. ELISEO AMADO GONZÁLEZ
-director-

UNIVERSIDAD DE PAMPLONA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE QUÍMICA
PAMPLONA, NORTE DE SANTANDER, COLOMBIA
2022

NOTA DE ACEPTACIÓN

Presidente del jurado

Jurado

Jurado

Ciudad y fecha

DEDICATORIA

A mi madre Araminta Rojas
con todo mi amor.

AGRADECIMIENTOS

Un agradecimiento muy especial a mis directores, Dr. Daniel Iván Barrera Valderrama y Dr. Eliseo Amado González por haberme transmitido el apasionamiento por la investigación, por su mentoría, instruidos consejos y por su constante apoyo y aliento durante la elaboración de esta tesis de maestría. Enormemente agradecido por sus asistencias durante este trabajo ya que no hubiera sido posible sin su ayuda y saber hacer.

De igual manera, un agradecimiento a Eugenio Llanos por sus clases magistrales en los lenguajes de programación awk y sed.

TABLA DE CONTENIDO

1	RESUMEN	1
2	INTRODUCCIÓN.....	2
3	JUSTIFICACIÓN.....	3
4	OBJETIVOS.....	4
4.1	Objetivo general.....	4
4.2	Objetivos específicos	4
5	MARCO TEÓRICO Y ESTADO DEL ARTE	5
5.1	Tiempo de vida de las proteínas y regla del Terminal-N.....	5
5.2	Proteomas y Terminales-N de las proteínas.....	12
5.3	Uniprot.....	14
5.3.1	Secuencia de proteínas no redundantes	14
5.4	Clases de enzimas.....	14
6	METODOLOGÍA.....	15
6.1	Equipo.....	15
6.2	Herramientas y software	15
6.2.1	Excel	15
6.2.2	Power Query.....	15
6.2.3	Power BI.....	15

6.2.4	Cytoscape 3.9.1.	15
6.3	Procedimiento	16
6.3.1	Selección de los proteomas de estudio.	16
6.3.2	Descarga de las proteínas de cada proteoma.	16
6.3.3	Clasificación de las proteínas con estatus de revisadas.....	16
6.3.4	Clasificación de las proteínas de acuerdo con el tipo de reacción que catalizan 17	
6.3.5	Clasificación de las enzimas de acuerdo con su residuo Terminal-N	18
6.3.6	Clasificación de cada uno de los residuos terminal-N como estabilizador y desestabilizador.	18
6.3.7	Clasificación de las secuencias de acuerdo con la superfamilia a la que pertenecen.....	18
6.3.8	Clasificación de las enzimas de acuerdo con la ruta metabólica que cataliza	18
7	RESULTADOS Y DISCUSIÓN	19
7.1	Selección de los proteomas de referencia	19
7.2	Proteínas de los proteomas de referencia.....	19
7.3	Clasificación de las proteínas con estatus de <i>revisadas</i>	20
7.4	Clasificación de las proteínas de acuerdo con el tipo de reacción que catalizan... 21	
7.5	Clasificación de las enzimas de acuerdo con su residuo terminal-N.....	31
7.6	Clasificación de las enzimas de acuerdo con la superfamilia a la que pertenecen.41	
7.7	Clasificación de las enzimas de acuerdo con la ruta metabólica que catalizan	59

8	CONCLUSIONES.....	76
9	REFERENCIAS BIBLIOGRÁFICAS	78
10	ANEXOS	85

ÍNDICE DE TABLAS

Tabla 1. Comportamiento estabilizador y desestabilizador de los residuos Terminal-N la β -galactosidasa.	7
Tabla 2. Proteínas con dos actividades enzimáticas para el proteoma <i>Oryctolagus cuniculus</i>	39
Tabla 3. Proteínas con dos actividades enzimáticas para el proteoma <i>Nicotiana tabacum</i> ..	39
Tabla 4. Numero de superfamilias proteicas que tienen en común los proteomas de estudio.	43
Tabla 5. Porcentaje de coincidencia de superfamilias en los proteomas de estudio*.	43
Tabla 6. Superfamilias con mayor población de enzimas en los cuatro proteomas de estudio.	52
Tabla 7. Rutas metabólicas con mayor participación de enzimas.	70

ÍNDICE DE FIGURAS

Figura 1. Ruta del Arg/Terminal-N en células eucariotas.....	8
Figura 2. Ruta del Ac/Terminal-N en células eucariotas.....	9
Figura 3. Ruta del Pro/Terminal-N en células eucariotas.....	10
Figura 4. Ruta del fMet/Terminal-N en células eucariotas y en bacterias..	11
Figura 5. Ruta del Leu/Terminal-N en bacterias..	12
Figura 6. Proteomas de referencia no redundantes.....	13
Figura 7. Conteo de proteínas en cada proteoma.....	20
Figura 9. Conteo de enzimas por cada clase de enzima (EC) para cada uno de los proteomas.	22
Figura 10. Conteo de proteínas con actividad catalítica en cada proteoma.....	23
Figura 11. Conteo de enzimas en cada proteoma.....	23
Figura 12. Visualización de las enzimas que se repiten más de una vez en los cuatro proteomas.....	25
Figura 13. Ampliación de una sección de la Figura 12 del proteoma <i>Saccharomyces cerevisiae</i>	26
Figura 14. Visualización de las proteínas con más de una actividad enzimática en <i>Nicotiana tabacum</i>	28
Figura 15. Visualización de las proteínas con más de una actividad enzimática en <i>Saccharomyces cerevisiae</i>	29
Figura 16. Visualización de las proteínas con más de una actividad enzimática en <i>Saccharomyces cerevisiae</i>	30

Figura 17. Agrupamiento de las enzimas de acuerdo con su terminal-N.....	31
Figura 18. Agrupamiento de las enzimas de acuerdo con su terminal-N en cada proteoma.	32
Figura 19. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma <i>Escherichia coli</i>	33
Figura 20. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma <i>Nicotiana tabacum</i>	34
Figura 21. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma <i>Oryctolagus cuniculus</i>	34
Figura 22. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma <i>Saccharomyces cerevisiae</i>	35
Figura 23. Conteo de enzimas que tienen un terminal-N estabilizador.....	36
Figura 24. Conteo de enzimas que tienen un terminal-N desestabilizador..	37
Figura 25. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador para <i>Oryctolagus cuniculus</i> ..	38
Figura 26. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador para <i>Nicotiana tabacum</i>	40
Figura 27. Número de enzimas que tiene asignada una superfamilia.....	41
Figura 28. Número de enzimas que tiene asignada una superfamilia por cada proteoma..	41
Figura 29. Número de superfamilias proteicas por cada proteoma.....	42
Figura 30. Red de interacción de las superfamilias con los cuatro proteomas.....	44
Figura 31. Superfamilias comunes con los cuatro proteomas.....	45

Figura 32.....	46
Figura 33. Conteo de superfamilias asignadas por cada clase de enzima (EC).....	47
Figura 34. Conteo de enzimas asignadas a cada superfamilia proteica.....	48
Figura 35. Conteo de enzimas asignadas a cada superfamilia en <i>Escherichia coli</i>	50
Figura 36. Conteo de enzimas asignadas a cada superfamilia en <i>Nicotiana tabacum</i>	51
Figura 37. Conteo de enzimas asignadas a cada superfamilia en <i>Oryctolagus cuniculus</i>	52
Figura 38. Conteo de enzimas asignadas a cada superfamilia en <i>Saccharomyces cerevisiae</i>	53
Figura 39. Conteo de enzimas de acuerdo con el terminal-N.....	54
Figura 40. Conteo de enzimas de acuerdo con el terminal-N en <i>Escherichia coli</i>	54
Figura 41. Superfamilias que tienen asignada enzimas con terminal-N estabilizador y desestabilizador.	56
Figura 42. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia.	57
Figura 43. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para <i>Nicotiana tabacum</i>	58
Figura 44. Conteo de enzimas asociadas a rutas metabólica.	59
Figura 45. Conteo de rutas metabólicas.....	60
Figura 46. Red de interacción entre los proteomas y sus diferentes rutas metabólicas.....	61
Figura 47. Rutas metabólicas que coinciden con los cuatro proteomas.	62
Figura 48. Rutas metabólicas únicas para cada proteoma.	62

Figura 49. Conteo de rutas metabólicas por clase de enzima (EC).	64
Figura 50. Conteo de rutas metabólicas por clase de enzima (EC) en cada proteoma.	64
Figura 51. Conteo de enzimas por cada ruta metabólica.....	66
Figura 52. Conteo de enzimas por cada ruta metabólica para <i>Escherichia coli</i>	67
Figura 53. Conteo de enzimas por cada ruta metabólica para <i>Nicotiana tabacum</i>	68
Figura 54. Conteo de enzimas por cada ruta metabólica para <i>Oryctolagus cuniculus</i>	69
Figura 55. Conteo de enzimas por cada ruta metabólica para <i>Saccharomyces cerevisiae</i> ... 70	
Figura 56. Conteo de enzimas que tienen asignado un terminal-N y que participan en alguna ruta metabólica..	71
Figura 57. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador para <i>Escherichia coli</i> que participan en alguna ruta metabólica.	72
Figura 58. Red de interacción de las rutas metabólicas en las que participan enzimas con terminal-N estabilizador y terminal-N desestabilizador.	73
Figura 59. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica sin distinción del proteoma.	74
Figura 60. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en <i>Nicotiana tabacum</i>	75
.....	98

ANEXOS

ÍNDICE DE TABLAS

Tabla 8. Proteomas con las proteínas que tienen más de una actividad enzimática.	104
Tabla 9. Terminal-N presentes en las superfamilias proteicas.	106
Tabla 10. Terminales-N presentes en las rutas metabólicas.	116
Tabla 11. Información referente al virus VIH-1	118
Tabla 12. Tiempos de tiempo de vida de la proteína β -galactosidasa, con los diferentes Terminal-N, en distintos organismos.....	118

ÍNDICE DE FIGURAS

Figura 61. Proteínas con más de una actividad enzimática en <i>Oryctolagus cuniculus</i> . Se presenta 7 clúster, uno por cada clase de enzima (EC).	85
Figura 62. Proteínas con más de una actividad enzimática en <i>Escherichia coli</i> . Se presenta 7 agrupamientos, uno por cada clase de enzima (EC).	86
Figura 63. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador en <i>Escherichia coli</i>	87

Figura 64. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador en <i>Saccharomyces cerevisiae</i>	88
Figura 65. Superfamilias en común en los proteomas <i>Escherichia coli</i> y <i>Saccharomyces cerevisiae</i>	89
Figura 66. Superfamilias en común en los proteomas <i>Nicotiana tabacum</i> y <i>Oryctolagus cuniculus</i>	90
Figura 67. Superfamilias en común en los proteomas <i>Nicotiana tabacum</i> y <i>Saccharomyces cerevisiae</i>	91
Figura 68. Superfamilias que interactúan solamente con los proteomas <i>Escherichia coli</i> y <i>Oryctolagus cuniculus</i>	92
Figura 69. Superfamilias que interactúan solamente con los proteomas <i>Escherichia coli</i> y <i>Nicotiana tabacum</i>	93
Figura 70. Superfamilias que interactúan solamente con los proteomas <i>Saccharomyces cerevisiae</i> y <i>Oryctolagus cuniculus</i>	94
Figura 71. Superfamilias que tienen asignada enzimas con terminal-N estabilizador y desestabilizador en <i>Escherichia coli</i>	95
Figura 72. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma <i>Nicotiana tabacum</i>	96
Figura 73. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma <i>Oryctolagus cuniculus</i>	96
Figura 74. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma <i>Saccharomyces cerevisiae</i>	97
Figura 75. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para <i>Escherichia coli</i>	97

Figura 76. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para <i>Oryctolagus cuniculus</i>	98
Figura 77. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para <i>Saccharomyces cerevisiae</i>	98
Figura 78. Rutas metabólicas que coinciden solo con los proteomas <i>Escherichia coli</i> , <i>Saccharomyces cerevisiae</i> y <i>Oryctolagus cuniculus</i>	99
Figura 79. Rutas metabólicas que coinciden solo con los proteomas <i>Escherichia coli</i> , <i>Saccharomyces cerevisiae</i> y <i>Oryctolagus cuniculus</i>	100
Figura 80. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador para <i>Nicotiana tabacum</i> que participan en alguna ruta metabólica.	100
Figura 81. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador para <i>Oryctolagus cuniculus</i> que participan en alguna ruta metabólica. ..	101
Figura 82. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador para <i>Saccharomyces cerevisiae</i> que participan en alguna ruta metabólica.	101
Figura 83. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en <i>Escherichia coli</i>	102
Figura 84. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en <i>Oryctolagus cuniculus</i>	102
Figura 85. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en <i>Saccharomyces cerevisiae</i>	103

GLOSARIO

Acido Aspártico: Asp, D

Ácido glutámico: Glu, E

Alanina: Ala, A

Arginina: Arg, R

Asparagina: Asn, N

Cisteína: Cys, C

ClpPs: Proteasas caseinolíticas

Complejo Gid4: Es una ligasa RING E3 que consta de seis proteínas diferentes consecutivamente.

Cys: Cisteína, uno de los 20 aminoácidos que usa las células en síntesis de proteínas.

Entry: Identificador asignado por Uniprot a cada proteína.

Espectrometría de masas: Técnica de análisis cualitativo para determinar estructuras orgánicas a partir de iones que se obtiene a partir de la fase gaseosa de la molécula de estudio.

Etiquetado isotópico: Técnica de rastreo atómico.

Fenilalanina: Phe, F

Fmt1: Formiltransferasa, enzima encargada de transferir grupos formilo a un terminal-N.

Genoma: Secuencia completa de instrucciones del ADN que posee un organismo en particular.

Glicina: Gly, G

Glutamina: Gln, Q

Histidina: His, H

Identificador: *Entry* asignado por Uniprot a cada proteína

In silico: Experimento que ocurre o tiene lugar en un entorno computacional.

In vitro: Experimento que ocurre o tiene lugar fuera de un organismo vivo.

In vivo: Experimento que ocurre o tiene lugar dentro de un organismo vivo.

Isoleucina: Ile, I

Leucina: Leu, L

Lisina: Lys, K

Metionina: Met, M

Nr: N-reconocinas, proteínas de reconocimiento proteico para la posterior degradación de una proteína.

Nt-amidasa: Enzimas que catalizan la hidrólisis de una amida del residuo terminal-N.

Nt-desamidación: Eliminación del grupo amino del residuo terminal-n de una proteína.

N α : Nitrógeno que se encuentra enlazado al carbono alfa de un aminoácido.

N^a-acetilación: Reacción química que introduce un grupo acetilo al terminal-N de una proteína.

Prolina: Pro, P

Proteoma: Conjunto completo de proteínas expresadas por un organismo en particular.

Proteosoma: Complejo proteico encargado de realizar la degradación de proteínas.

RING: El dominio RING es un motivo de secuencia rico en cisteína que puede unir dos átomos de zinc.

Serina: Ser, S

Terminal-N: Residuo aminoácido del extremo izquierdo de una proteína o péptido que posee un grupo amino libre.

Tirosina: Tyr, Y

Treonina: Thr, T

Triptófano: Trp, W

Valina: Val, V

1 RESUMEN

La regla del Terminal-N indica que el tiempo de vida de las proteínas está determinado por su residuo aminoacídico Terminal-N. Estudiar esta regla experimentalmente, *in vivo*, ha sido una tarea muy compleja y, por ende, el número de proteínas en las que se ha estudiado es muy pequeño en comparación con el índice de proteínas que pueden ser expresadas por un genoma, el proteoma.

En esta tesis de maestría se analizó el efecto estabilizador o desestabilizador del terminal-N en el tiempo de vida de las enzimas pertenecientes a los proteomas de los organismos *Oryctolagus cuniculus*, *Nicotiana tabacum*, *Saccharomyces cerevisiae*, *Escherichia coli* y del virus de inmunodeficiencia humana (VIH-1). Además, se identificaron las relaciones entre el terminal-N de las enzimas, las rutas metabólicas a las que pertenecen y la superfamilia de plegamiento proteico. Este estudio se enfocó en las enzimas con estatus de *revisada* pertenecientes a los proteomas de estudio y reportadas en la base de datos Uniprot. Las tendencias se identificaron utilizando las herramientas computacionales Power BI, Power Query y la hoja de cálculo de Excel y las relaciones se determinaron por medio del análisis de agrupamientos con el programa Cytoscape.

Los resultados permitieron concluir que las oxidorreductasas (EC:1), transferasa (EC:2) e hidrolasas (EC:3) son las clases de enzimas con mayor número de proteínas con tiempos de vida estables, con mayor asignación de superfamilia y las que participan en un mayor número de rutas metabólicas. Además, al cruzar la información obtenida de cada uno de los proteomas, se encontró que la mayoría de las enzimas de los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae* y *Nicotiana tabacum* pertenecen a la superfamilia proteica SSF52540 (Hidrolasas de trifosfato de nucleósido que contienen bucles P), mientras que la mayoría de las enzimas de del proteoma *Oryctolagus cuniculus* tienden a pertenecer a la superfamilia proteica SSF56112 (similar a la proteína quinasa (*PK-like*)).

2 INTRODUCCIÓN

Diferentes estudios *in vivo* han probado que el tiempo de vida (tiempo necesario para degradar una proteína activa a oligopéptidos de 9 a 11 residuos en el proteosoma o en el sistema de proteasas caseinolíticas (ClpPs) de una proteína, es dependiente de su residuo aminoacídico Terminal-N (Collins & Goldberg, 2017; Dissmeyer et al., 2018; Varshavsky, 2019). El inicio de la degradación de una proteína empieza con el reconocimiento del residuo aminoacídico Terminal-N por una enzima, específica para cada aminoácido terminal-N, que al reconocerlo regula el proceso de marcaje de la proteína; finalmente, la proteína marcada se degrada en el proteosoma (Varshavsky, 2019). El incremento del tiempo de vida de una proteína aumenta el riesgo de daño estructural y por ende una alteración en su función favoreciendo la senescencia celular (Sharpless & Sherr, 2015); por lo tanto, las células han desarrollado varios mecanismos para controlar el tiempo de vida de sus proteínas (Basisty, Holtz, & Schilling, 2019). Los resultados *in vivo* del tiempo de vida de la β -galactosidasa han demostrado que el residuo Terminal-N está muy relacionado con su tiempo de vida, este hallazgo se conoce como la regla del Terminal-N. Esta regla tiene diferentes rutas de degradación: ruta Arg/Terminal-N, ruta Ac/Terminal-N, ruta Pro/Terminal-N y ruta fMet/Terminal-N. El número de proteínas que se han usado para estudiar la regla del Terminal-N es muy pequeño en comparación con el índice de proteínas que pueden ser expresadas por un genoma, el proteoma. Estudiar experimentalmente el Terminal-N ha sido una tarea muy compleja *in vivo* por lo tanto pocos estudios han determinado el tiempo de vida de las proteínas presentes en proteomas (Lange, Huesgen, Nguyen, & Overall, 2014; Li, Yan, & Zhang, 2019; Martínez et al., 2008).

En esta tesis de maestría se analizó la estabilidad del tiempo de vida de las enzimas presentes en los proteomas de referencia del Conejo (*Oryctolagus cuniculus*), el Tabaco (*Nicotiana tabacum*), la levadura de cerveza (*Saccharomyces cerevisiae*), la bacteria Gram (-) *Escherichia coli* y del virus de inmunodeficiencia humana (VIH-1) aplicando la regla del Terminal-N. Usando herramientas de la minería de datos se obtuvieron las secuencias de aminoácidos de cada una de las enzimas pertenecientes a los proteomas en estudio y, posteriormente, se asignó el tiempo de vida a cada enzima usando como referencia los datos

del tiempo de vida dependiente del Terminal-N reportados experimentalmente. El análisis se enfocó en comparar las tendencias encontradas en cada proteoma, especialmente, en las rutas metabólicas a las que pertenecen las enzimas, la superfamilia proteica y el efecto estabilizante en el tiempo de vida determinado por la regla del Terminal-N.

3 JUSTIFICACIÓN

Desde que se encontró la relación entre el residuo Terminal-N de una proteína y su tiempo de vida, cerca de 500 artículos han sido publicados en este tema, no obstante, el número de proteínas y el número de proteomas en las que se ha analizado la regla del Terminal-N es muy pequeño en comparación con el número de proteínas que pueden ser expresadas por un genoma, es decir el proteoma. Experimentalmente, estudiar el Terminal-N, *in vivo*, es una tarea compleja a pesar que se han empleado múltiples enfoques: espectrometrías de masas, etiquetado isotópico del terminal amino del sustrato (TAILS, por sus siglas en inglés *Terminal Amine Isotopic Labeling of Substrates*), proteólisis asistida por láser, entre otros (Lange, Huesgen, Nguyen, & Overall, 2014; Li, Yan, & Zhang, 2019; Martinez et al., 2008). Por lo tanto, para la mayoría de los proteomas se desconoce la información experimental sobre el tiempo de vida de las proteínas dependientes del Terminal-N. De otra parte, el estudio *in silico* de los proteomas muestra que la información disponible en la base de datos Uniprot reporta 17.174 proteomas de referencia. Sin embargo, no se han desarrollado estudios que se enfoquen en determinar el tiempo de vida de las proteínas presentes en cada uno de ellos.

Tomando ventaja de la rapidez y economía de las herramientas computacionales para obtener resultados, en esta tesis de maestría se analizó la estabilidad del tiempo de vida de las enzimas presentes en los proteomas de *Oryctolagus cuniculus*), tabaco (*Nicotiana tabacum*), levadura de cerveza (*Saccharomyces cerevisiae*), la bacteria Gram (-) *Escherichia coli* y del virus de inmunodeficiencia humana (VIH-1) en los cuales ya se ha estudiado la regla del Terminal-N. Los resultados encontrados en cada proteoma de estudio se compararon en busca de

relaciones entre las rutas metabólicas, las superfamilias proteicas y el tiempo de vida determinado por la regla del terminal-N.

4 OBJETIVOS

4.1 Objetivo general

Analizar el tiempo de vida de las enzimas pertenecientes a los proteomas del conejo (*Oryctolagus cuniculus*), el tabaco (*Nicotiana tabacum*), la levadura de cerveza (*Saccharomyces cerevisiae*), la bacteria Gram (-) *Escherichia coli* y del virus de inmunodeficiencia humana (VIH-1) aplicando la regla del Terminal-N.

4.2 Objetivos específicos

Comparar las tendencias de los tiempos de vida de las enzimas pertenecientes a cada uno de los proteomas en estudio: conejo (*Oryctolagus cuniculus*), tabaco (*Nicotiana tabacum*), levadura de cerveza (*Saccharomyces cerevisiae*), la bacteria Gram (-) *Escherichia coli* y del virus de inmunodeficiencia humana (VIH-1).

Identificar las relaciones entre el tiempo de vida de las enzimas pertenecientes a cada uno de los proteomas en estudio con las rutas metabólicas a las que pertenecen y a la superfamilia de plegamiento proteico.

5 MARCO TEÓRICO Y ESTADO DEL ARTE

5.1 Tiempo de vida de las proteínas y regla del Terminal-N

Numerosos estudios *in vivo* han demostrado que las proteínas en la célula pueden presentar afectaciones en su estructura y por ende producir un detrimento o modificación en la eficiencia de su función, facilitando la senescencia celular (Sharpless & Sherr, 2015). Dentro de los factores que pueden desencadenar el daño en las proteínas está el estrés oxidativo (Dissmeyer et al., 2018; Donida et al., 2017), la isomerización o la agregación espontánea (Cheignon et al., 2018), las interacciones con metabolitos tóxicos (Hanson, Henry, Fiehn, & de Crécy-Lagard, 2016) o de factores exógenos como metales pesados (Tamás, Fauvet, Christen, & Goloubinoff, 2018). Por lo tanto, el control de la estabilidad de las proteínas *in vivo* tiene un papel muy importante en la regulación de todos los procesos celulares ya que es necesario disminuir la probabilidad del detrimento de la actividad biológica desde su síntesis hasta su degradación (Alvarez-Castelao & Schuman, 2015; Bozaykut et al., 2014; J. M. Kim et al., 2018; Mccaffrey & Braakman, 2016; Varshavsky, 2008). Múltiples estudios indican que el tiempo de vida de las proteínas está estrechamente relacionado con la regeneración celular, el mecanismo de defensa contra el envejecimiento celular (Basisty, Holtz, & Schilling, 2019; Saez & Vilchez, 2014), la presencia de agentes perjudiciales y patógenos (Klionsky, 2005) y la eliminación de proteínas con problemas de plegamiento (Collins & Goldberg, 2017). El proceso de degradación proteica *in vivo*, se lleva a cabo por mecanismos diferentes, los dos más estudiados son el lisosomal y el proteosomal (Ciechanover, 2005; Kwon & Ciechanover, 2017). Una característica en común en estos dos mecanismos es la unión de conjugados de poliubiquitina, o de proteínas similares a ésta (Striebel, Imkamp, Özcelik, & Weber-Ban, 2014), a proteínas blanco para desencadenar los procesos degradativos o rutas de degradación (Kwon & Ciechanover, 2017) que son activadas por las señales de degradación.

El tiempo de vida de las proteínas depende del control de la activación de la degradación por señales en la secuencia primaria que se conocen como degrones. Los degrones se consideran como sitios de regulación de la abundancia de las proteínas, siendo muy importantes en

diversos procesos celulares ya que permiten la eliminación de las proteínas que no son requeridas y previenen la disfunción biológica en proteínas (Dissmeyer et al., 2018; Mészáros et al., 2017; Varshavsky, 2019). Los degrades más conocidos hasta el momento son los C-degrades y los N-degrades, estos últimos estrechamente relacionados con esta investigación. Los C- y N-degrades pueden incluir motivos de secuencia contiguos y residuos de lisina que sirven como sitios de poliubiquitinación (Varshavsky, 2019). Las proteínas tienen una secuencia perteneciente al Terminal-N que actúa como señal de destino celular (Izard & Kendall, 1994; Martoglio & Dobberstein, 1998; Owji, Nezafat, Negahdaripour, Hajiebrahimi, & Ghasemi, 2018). En algunos casos para completar la maduración de la proteína, una peptidasa elimina la secuencia que actúa como señal de destino, generando así un nuevo residuo Terminal-N que rige el tiempo de vida de la proteína. La determinación del tiempo de vida de una proteína de acuerdo al residuo Terminal-N se conoce como la regla del Terminal-N (Dissmeyer et al., 2018; Gondasb et al., 1989).

La regla del Terminal-N postula que el tiempo de vida de una proteína está determinado por el residuo aminoacídico que corresponde al Terminal-N en la proteína nativa (Bachmair et al., 1986; Gondasb et al., 1989; Tasaki, Sriram, Park, & Kwon, 2012). El inicio de la degradación de la proteína se inicia con el reconocimiento del residuo aminoacídico ubicado en el Terminal-N por una enzima, específica para cada aminoácido, que lo reconoce y regula el proceso de ubiquitinación (Tasaki et al., 2009; Varshavsky, 1997). La secuencia aminoacídica de las enzimas de reconocimiento del Terminal-N difieren en células procariotas o eucariotas (D. A. Dougan, Micevski, & Truscott, 2012; Tasaki et al., 2009; Vittal, Stewart, Brzovic, & Klevit, 2015), en humanos el número de estas enzimas es cercano a 300 y aún no se conoce a fondo su rol en la fisiología celular (Vittal et al., 2015).

Los estudios *in vivo* de la regla del Terminal-N han arrojado resultados cualitativos (ver tabla 1) que demuestran que el residuo Terminal-N sí está relacionado con el tiempo de vida de las proteínas (Ubíeta & Santiago, 1993). Esto sucede debido a que a que la regla del Terminal-N tiene varias rutas que permiten la degradación de la proteína marcada en el proteosoma 26S en eucariotas o en las proteasas caseinolíticas (ClpPs) en procariotas (Collins & Goldberg, 2017; Zeiler et al., 2013). Teniendo en cuenta lo dicho anteriormente y los datos

de la Tabla 1, se puede concluir que la regla del Terminal-N no es universal y que por ende hay rutas de degradación diferentes dependientes del residuo Terminal-N o del tipo de célula.

Tabla 1. Comportamiento estabilizador y desestabilizador de los residuos Terminal-N la β -galactosidasa.

Sistema biológico de estudio	Residuos Terminal-N estabilizadores*	Residuos Terminal-N desestabilizadores*	Método	Referencia
Conejo (<i>Oryctolagus cuniculus</i>)	M, P, V, G	F, L, W, Y, R, K, H, I, N, Q, D, E, C, A, S, T	Unión lineal de aminoácidos a la proteína β -galactosidasa usando muta génesis en el gen que codifica la fusión de Ub en levadura.	(Gonda et al., 1989)
**Tabaco (<i>Nicotiana tabacum</i>)	M, G, V, T, S, A	Q, N, C, E, D, R, K, H, L, I, F, W, Y		(Graciet et al., 2010)
Levadura (<i>Saccharomyces cerevisiae</i>)	M, P, V, G, T, S, A, C	E, D, Q, N, I, H, K, R, Y, W, L, F	Unión lineal de aminoácidos a la proteína β -galactosidasa usando muta génesis en el gen que codifica la fusión de Ub en levadura.	(Bachmair et al., 1986)
**Bacteria (<i>Escherichia coli</i>)	M, V, G, T, S, A, C, E, D, Q, N, I, H	K, R, Y, W, L, F	Unión lineal de aminoácidos a la proteína β -galactosidasa usando muta génesis en el gen que codifica la fusión de Ub en levadura.	(Tobias et al., 1991)
***Virus (VIH-1)	M, A, G	L, R, H, D, E, Q, N, P		(De Groot et al., 1991; Mulder & Muesing, 2000)

*Los tiempos de vida correspondientes a los residuos terminal-N se encuentran en la Tabla 12 en la sección anexos.

** El autor reportó que el resultado de la Prolina es desconocido.

***El autor no reportó de los 9 residuos Terminal-N restantes.

Las rutas de la regla del Terminal-N hasta el momento encontradas son: a) ruta cuando el residuo Terminal-N es una Arginina (Arg/Terminal-N), b) ruta cuando el residuo Terminal-N sufre una acetilación (Ac/Terminal-N), c) ruta cuando el residuo Terminal-N es una Prolina (Pro/Terminal-N), d) ruta cuando el residuo aminoacídico es una Metionina pero sufre una formilación (fMet/Terminal-N) y e) ruta cuando el residuo Terminal-N es modificado por la transferencia de una lisina conllevando a la generación de un nuevo Terminal-N (Leu/Terminal-N). A continuación, se describen brevemente cada una de las rutas de la regla del Terminal-N (Varshavsky, 2019).

a) *Ruta de degradación cuando el residuo Terminal-N es una Arginina (Arg/Terminal-N)*: Es una ruta presente en eucariotas en la cual las N-reconocina (Nr) forman complejos con residuos aminoacídicos no acetilados en el Terminal-N. Los residuos que permiten que se desencadene esta ruta son: Arg, His, Leu, Phe, Tyr, Trp, Ile y Met-únicamente cuando Met está unido a un residuo de aminoácido hidrofóbico voluminoso- (Varshavsky, 2011) (Fig. 1). Adicionalmente los residuos Asn y Gln pueden desaminarse a Asp y Glu respectivamente, luego una Arginina-transferasa añade un residuo de Arg (arginilación) en el Terminal-N y de esta manera pueden ser reconocidos por la Arg-N-reconocina (Arg-Nr). Si el Terminal-N de la proteína es un residuo de Cys puede entrar a esta ruta por arginilación pero es necesaria una oxidación de la Cys (Eldeeb & Fahlman, 2016; Gibbs et al., 2011; Gonda et al., 1989; Hu et al., 2005; J. M. Kim & Hwang, 2014; Varshavsky, 2011).

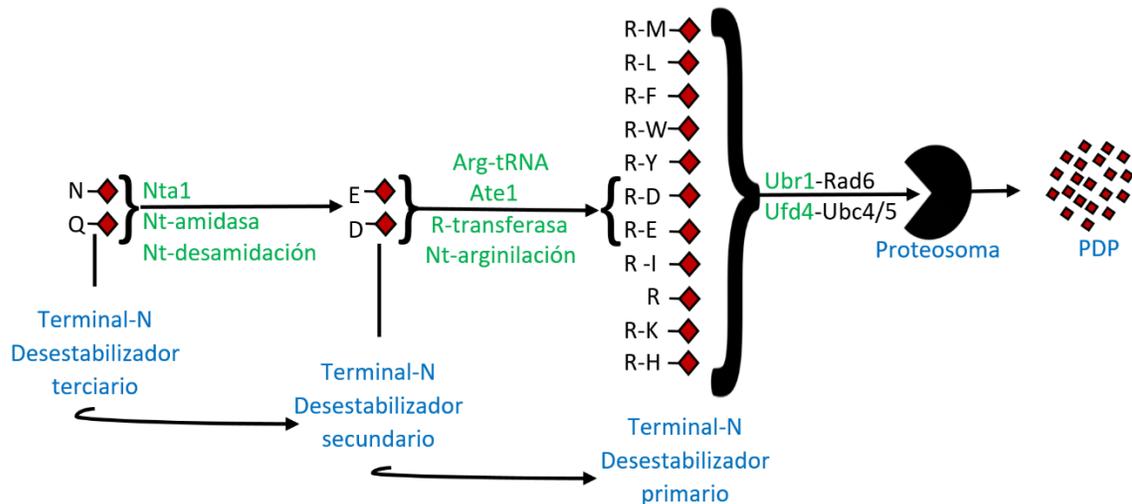


Figura 1. Ruta del Arg/Terminal-N en células eucariotas. Inicialmente los residuos Terminal-N desestabilizadores terciarios son convertidos a Terminal-N desestabilizadores secundarios por las amidasas del terminal-N, Nta1, y por desamidación del terminal-N. Luego los residuos Terminal-N desestabilizadores secundarios son convertidos a Terminal-N primarios por arginilación o R-transferasa en presencia de Arg-tRNA (esta modificación es llevada a cabo por la enzima transferasa Ate1, la cual transfiere un residuo arginina desde Arg-tRNA de la proteína). Los residuos Terminal-N primarios (Arg/Terminal-N) son reconocidos (por Ubr1-Rad6 y Ufd4-Ubc4/5) y dirigidos al proteosoma para su degradación. Los residuos desestabilizadores primarios, son aquellos que proporcionan directamente la señal de degradación; residuos desestabilizadores secundarios, son aquellos que solo mediante la conjugación Arg-tRNA-proteína transferasa a arginina pueden ejercer su acción desestabilizadora; y residuos desestabilizadores terciarios, que son aquellos que sufren una conversión a residuos desestabilizadores secundarios (Ubíeta & Santiago, 1993). PDP (producto de degradación del proteosoma). Figura modificada de (Varshavsky, 2019).

b) *Ruta de degradación cuando el residuo Terminal-N sufre una acetilación (Ac/Terminal-N)*: Es una ruta presente en eucariotas y puede estar presente en archaeas donde las Nr forman complejos con residuos aminoacídicos acetilados en el N α del residuo aminoacídico Terminal-N. Los residuos susceptibles de N α -acetilación en la proteína madura son: Met, Ala, Val, Ser, Thr o Cys. La N α -acetilación es una modificación co-traducciona común (50-80%) en las proteínas de eucariotas. Esta ruta requiere de enzimas especiales como Terminal-N-acetilasa (Nt-acetilasa) y Met-aminopeptidasas para que se pueda dar el reconocimiento por parte de la Ac/Nr (Eldeeb & Fahlman, 2016; H.-K. Kim et al., 2014; Martinez et al., 2008; Nguyen, Mun, Lee, & Hwang, 2018; Varshavsky, 2011) (Fig. 2).

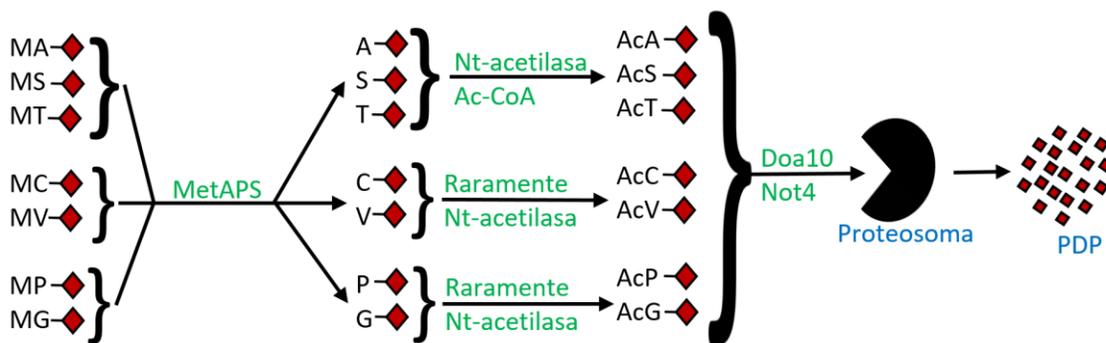


Figura 2. Ruta del Ac/Terminal-N en células eucariotas. Inicialmente las aminopeptidasas MetAPS eliminan el residuo Terminal-N Met (M) si está unido a los residuos A, S, T, C, V, P, o G. Luego hay una acetilación de los nuevos residuos Terminal-N. Los residuos mayormente acetilados son los A, S y T, los menos acetilados son los C, V, P, y G. Seguidamente, las enzimas de reconocimiento de Terminales-N acetilados, Doa10 y Not4, se unen específicamente al residuo Terminal-N acetilado para la degradación por el proteosoma. Figura modificada de (Varshavsky, 2019).

c) *Ruta de degradación cuando el residuo Terminal-N es una Prolina (Pro/Terminal-N)*: Es una ruta reportada en *Sacharomyces cerevisiae*, donde la N-reconocina forman complejos con el residuo aminoacídico Pro en el Terminal-N. Para que se active esta ruta de degradación es necesario que haya un estrés por baja concentración de glucosa y una alta concentración de etanol. La célula responde a estos cambios en el medio extracelular mediante la modificación de los Terminal-N de algunas de sus proteínas cambiando el residuo normal por el residuo Pro/Terminal-N, hasta el momento, en plantas o en mamíferos no ha sido reportada esta ruta (Chen, Wu, Wadas, Oh, & Varshavsky, 2017; Dong et al., 2018; David A. Dougan & Varshavsky, 2018) (Fig. 3).

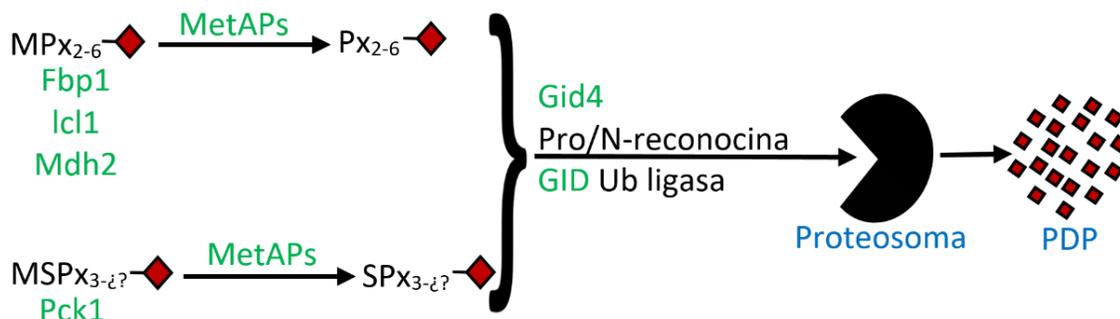


Figura 3. Ruta del Pro/Terminal-N en células eucariotas. MPX_{2-6} es el sustrato compuesto por; M, que es el código para la metionina; P, el código para la prolina; y X_{2-6} son los residuos siguientes (hasta el sexto residuo) de la cadena polipeptídica que son permitidos para la enzima Gid4 se pueda unir eficientemente al sustrato. Los residuos permitidos para cada posición son; en la posición 2 son permitidos los residuos G, A, S, T, C, D, N, Y y H; en la posición 3 son permitidos los residuos A, V, L, I, K y R; en la posición 4 son permitido los residuos V, C, P, L, I y W; en la posición 5 son permitidos los residuos G, A, S, T, C, N, Q, M, L, I, F, Y, W, H, K y R; y en la posición 6 son permitidos todos los residuos excepto D y E (Chen et al., 2017). Esta ruta inicia con la eliminación del Met/Terminal-N de las proteínas MPX_{2-6} (enzimas gluconeogénicos Fbp1, Icl1, Mdh2) y $MSPX_{3-i?}$ (enzima gluconeogénica Pck1) por el enzima Met-aminopeptidasa (MetAPs). Posteriormente, una subunidad de la enzima Gid4 se une al residuo Pro/Terminal-N del sustrato que es degradado por el proteosoma S26. La misma ruta también reconoce sustratos en los que el residuo Terminal-N Pro está en segunda posición. Figura modificada de (Varshavsky, 2019).

d) Ruta de degradación cuando el residuo aminoacídico es una Metionina, y sufre una formilación (*fMet/Terminal-N*). Para esta ruta, en bacterias y en eucariotas (mitocondrias y cloroplastos), antes de la traducción de la proteína, las formiltransferasas (Fmt1) en presencia de 10-formiltetrahydrofolato (10-fTHF) formilan las Met de los Met-tARN-iniciadores (Met-tARNi) generando *fMet-tARNi*. Solo en las células eucariotas la retención citosólica de Fmt1 y el aumento resultante en los niveles de proteínas citosólicas con Terminal-N formilado, en células nutricionalmente estresadas, requieren una proteína quinasa (Gcn2) (Dissmeyer, 2019). Además, las proteínas citosólicas formiladas en el Terminal-N están dirigidas a la degradación selectiva por la ligasa Psh1-Ubc3, que actúa como la *fMet/N-reconocina* que, conjuntamente, con las ubiquitin ligasas Psh1 y E3 se dirigen a los residuos *fMet* Terminal-N de las proteínas citosólicas eucariotas (como Cse4, Pgd1 y Rps22a) para la degradación medida por la poliubiquilación (J.-M. Kim et al., 2018). Mientras que en bacterias el grupo formilo de *fMet* Terminal-N se elimina por transferencia mediante la enzima péptido deformilasa (PDF) (Piatkov, Vu, Hwang, & Varshavsky, 2015) (Fig. 4).

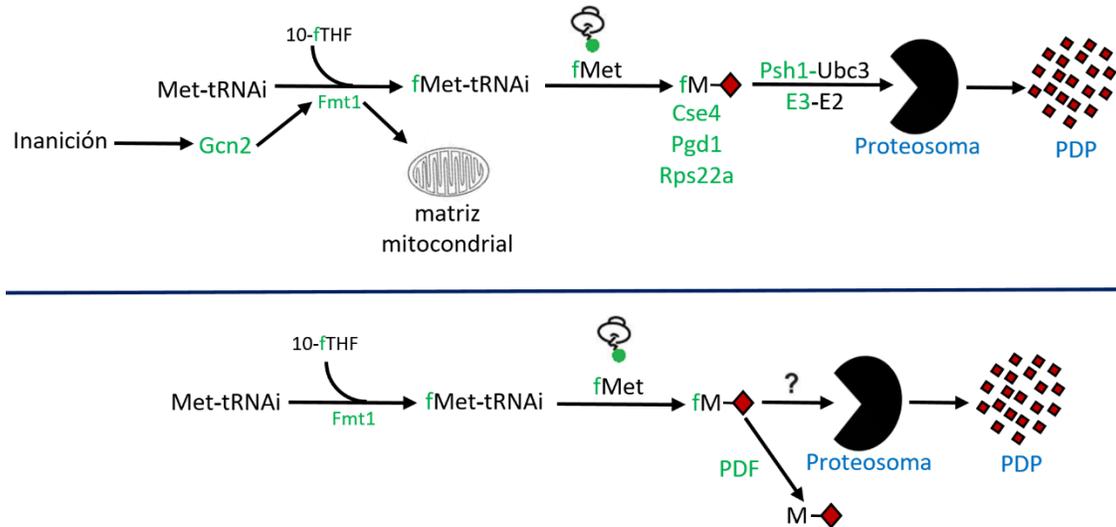


Figura 4. Ruta del fMet/Terminal-N en células eucariotas (parte superior) y en bacterias (parte inferior). Tanto en eucariotas como en bacterias las proteínas no maduras son formiladas por la enzima formiltransferasas (Fmt), usando 10-fTHF como co-sustrato, para posteriormente traducir la proteína formilada. A) En inanición las células, en el citosol, usan Gcn2 para aumentar el número de proteínas formiladas en su Terminal-N Met. El sustrato Met-tRNAi es formilado por la enzima Fmt1 a Fmet-TRNAi, que luego los ribosomas lo traducen en proteínas formiladas (Cse4, pgd1, Rps22a). Los residuos Terminal-N formilados de las proteínas son reconocidos por las enzimas Psh1-Ubc3 y E3-E2 para su degradación por el proteosoma. B) En bacterias, el sustrato que no es degradado por el proteosoma es des-formilado por la peptidasa PDF. Figura modificada de (Varshavsky, 2019).

e) Ruta cuando el residuo Terminal-N es modificado por la transferencia de una lisina conllevando a la generación de un nuevo Terminal-N. (Leu/Terminal-N). Esta ruta está conformada por el proteosoma ClpAP, dependiente de ATP, que se encarga de degradar los sustratos proteicos con Leu/Terminal-N, -Phe, -Trp o Tyr que van unidos a la enzima Leu/Nr ClpS (Schmidt, Zahn, Bukau, & Mogk, 2009). En la ruta también participa la Aat, una L/F-transferasa que emplea Leu-tRNA o Phe-tRNA como co-sustrato para conjugar en gran medida Leu (y ocasionalmente Phe) al extremo Terminal-N de proteínas que llevan Lys/Terminal-N o Arg/Terminal-N; y Bpt, una L-transferasa que emplea Leu-tRNA para conjugar Leu con Asp/Terminal-N, -Glu y (posiblemente) -Cys oxidada (Varshavsky, 2011) (Fig. 5).

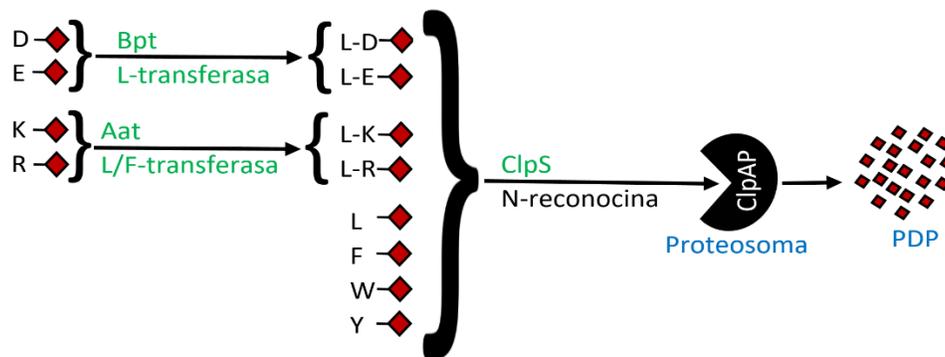


Figura 5. Ruta del Leu/Terminal-N en bacterias. Al Terminal-N (D, E, K y R) de un sustrato proteico, por medio de una L-transferasa y L/F-transferasa, se le adiciona un residuo L, seguidamente, el sustrato con el nuevo residuo Terminal-N Leucina es reconocido por la enzima ClpS y degradado por el proteosoma ClpAP. Figura modificada de (Varshavsky, 2019).

5.2 Proteomas y Terminales-N de las proteínas

En cuanto al estudio *in silico* de proteomas, la información disponible es numerosa, de hecho, la base de datos Uniprot, a la fecha 05-2022, reporta 459.108 proteomas. Sin embargo, algunos proteomas han sido seleccionados de forma manual o algorítmica como proteomas de referencia de acuerdo con el interés médico, biotecnológico, filogenético histórico o académico y están distribuidos en proteomas de eucariotas, procariotas, archaeas y de virus, en total se encuentran reportados 21.787 proteomas de referencia. A manera de ejemplo se grafican el conteo de proteínas en proteomas de referencia no redundantes de bacterias reportado en la base de datos Uniprot (Figura 6). Este trabajo de investigación se enfocó en estudiar el tiempo de vida de las enzimas de proteomas a los cuales se les haya estudiado la regla de Terminal-N ya sea de forma cuantitativa o cualitativa (Tablas 1 y 2).

Experimentalmente, estudiar el Terminal-N *in vivo* ha sido una tarea muy compleja a pesar que se han utilizado espectrómetros de masas de última generación (Lange et al., 2014; Li et al., 2019; Martinez et al., 2008). Los estudios se han enfocado en estudiar los Terminales-N de proteínas humanas encontrándose que el número de Terminal-N es mayor al de las proteínas analizadas debido a modificaciones postraduccionales como la acetilación.

Además, se ha reportado información sobre la modificación del tiempo de vida de las proteínas y sobre las poblaciones de los Terminal-N de las proteínas como consecuencia de la proteólisis selectiva (Lange et al., 2014; Li et al., 2019; Martinez et al., 2008). Los resultados *in vivo* e *in vitro* muestran que se deben aunar esfuerzos para mejorar la comprensión de la regulación de la activación de las rutas del Terminal-N, desarrollar protocolos experimentales reproducibles que permitan tamizar los sustratos en cada una de las rutas de la regla del Terminal-N, determinar las bases moleculares de la interacción de las distintas N-reconocina con sus blancos moleculares para así desencadenar la degradación proteica y determinar el significado fisiológico de la N-acetilación del Terminal-N. Adicionalmente, analizar las poblaciones de terminarles-N de las proteínas y cruzar esta información con su función biológica y su estructura tridimensional, lo cual es el propósito de esta investigación.

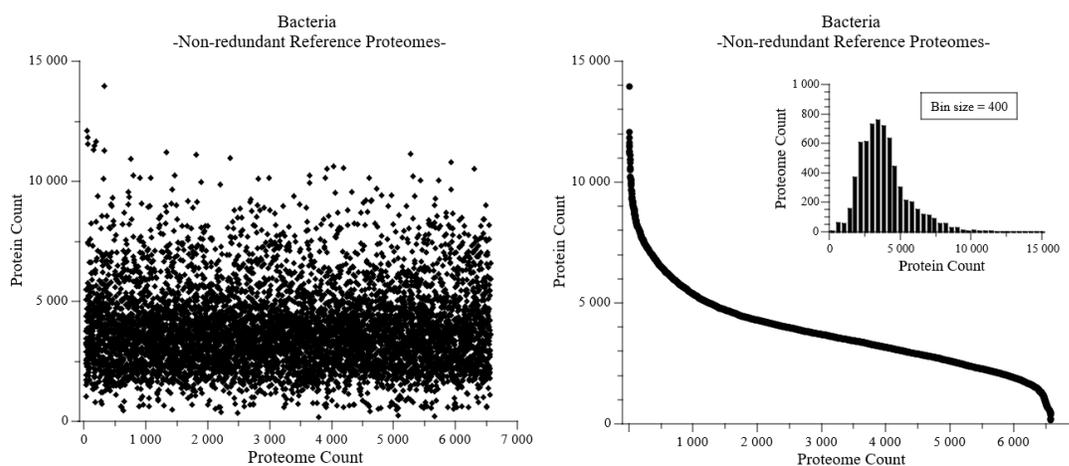


Figura 6. Proteomas de referencia no redundantes. A la izquierda se muestra el conteo de proteínas en cada proteoma de referencia no redundante de bacterias. A la derecha se muestra de forma ordenada y descendente el conteo de proteínas en los mismos proteomas. El recuadro muestra el histograma de proteínas en cada proteoma usando un *bin* de 400 proteínas. Los datos fueron tomados de la base de datos Uniprot.

5.3 Uniprot

Uniprot es una base de datos gratuita que proporciona secuencias (no redundantes) y notaciones funcionales de excelente calidad de las proteínas expresadas por proteomas. En esta base de datos, ciertos proteomas fueron seleccionados de forma manual y de forma algorítmica (o como proteomas de referencia) de utilidad para la investigación científica. Además, en esta base de datos las proteínas se encuentran catalogadas de acuerdo con el tipo de reacción que catalizan, ruta metabólica, superfamilia, etc (Bateman et al., 2021; Greer et al., 2022).

5.3.1 Secuencia de proteínas no redundantes

La secuencia de proteínas no redundantes son las secuencias que no contienen información análoga con otra secuencia dentro de la misma base de datos (Bateman et al., 2021; Greer et al., 2022).

5.4 Clases de enzimas.

Las enzimas son biocatalizadores necesarios y responsables de que las reacciones biológicas en los proteomas sean eficientes. De acuerdo con los tipos de las reacciones catalizadas, la Comisión de Enzimas (EC) agrupo las enzimas en 7 clases, oxidorreductasas (EC:1): Catalizan reacciones de oxidorreducción; transferasas (EC:2): Catalizan la transferencia de grupos catalizados; hidrolasas (EC:3): Catalizan el rompimiento hidrolítico de enlaces químicos; liasas (EC:4): Catalizan la escisión por medio de procesos diferentes a la oxidación y la hidrolisis; isomerasas (EC:5): Catalizan los cambio geométricos o estructurales de las biomoléculas, ligasas (EC:6): Catalizan la unión de dos moléculas mediante la formación de un nuevo enlace; y translocasas (EC:7): Cataliza el paso de iones y moléculas a través de las membranas (Balaji, 2021; Porto de Souza Vandenberghe et al., 2020; Tao et al., 2020).

6 METODOLOGÍA

6.1 Equipo

Procesador Intel Core i5-7200U CPU @ 2.50GHz, Gráficos Intel HD Graphics 620 y Memoria RAM de 8 GB.

6.2 Herramientas y software

6.2.1 Excel

Excel es una herramienta especializada para el tratamiento de datos para su posterior análisis.

6.2.2 Power Query

Power Query es una extensión de Excel que permite la transformación de datos y, además, permite la obtención de datos de diferentes fuentes, tanto internas como externas (*Download Microsoft Power Query Para Excel from Official Microsoft Download Center*, n.d.).

6.2.3 Power BI

Es una herramienta computacional que permite el análisis (por medio de graficas) de datos ya transformados en Excel / Power Query (*Power BI Desktop: Informes Interactivos / Microsoft Power BI*, n.d.).

6.2.4 Cytoscape 3.9.1.

Cytoscape es un programa computacional de libre acceso que permite hacer el análisis y visualización de redes complejas de proteomas, superfamilias proteínas, rutas metabólicas y otros datos (Otasek et al., 2019).

6.3 Procedimiento

6.3.1 Selección de los proteomas de estudio.

Los proteomas de referencia no redundantes seleccionados fueron aquellos pertenecientes a las especies del mismo orden taxonómico a los cuales se le ha determinado experimentalmente la regla del Terminal-N y que están contenidos en la Tabla 1.

6.3.2 Descarga de las proteínas de cada proteoma.

Para cada uno de los proteomas de referencia (*Escherichia coli*, *Nicotiana tabacum*, *Oryctolagus cuniculus*, *Saccharomyces cerevisiae* y *virus inmunodeficiencia humana (VIH-1)*) se descargaron todas las proteínas como se describe a continuación:

- a. Se ingresó a al buscador de internet.
- b. En el buscador se escribió el *string*, por ejemplo, para el proteoma *Oryctolagus cuniculus*: *organism:"Oryctolagus cuniculus (Rabbit) [9986]" AND proteome:up000001811*.
- c. Se ingresó a la interfaz de búsquedas avanzadas y se filtró la información de interés (*Entry, Status, Entry name, Protein names, Organism, Length, Sequence, EC number, Pathway, Chain y Signal peptide y superfamilia*)
- d. Por último, se descargaron las proteínas en formato Excel.

Los pasos anteriores se repitieron para de cada uno de los proteomas.

6.3.3 Clasificación de las proteínas con estatus de revisadas

Para cada uno de los proteomas se filtraron las proteínas con estatus de *revisadas* utilizando la hoja de cálculo Excel. Para lo cual:

- a. Primero se dio formato de tabla a los datos.
- b. Seguidamente se seleccionó la columna de *status*.
- c. Se seleccionaron los filtros, luego datos.
- d. Se seleccionó la flecha del encabezado de la columna *status*.
- e. Se dejó solo seleccionada la columna de reviewed.

Este procedimiento se repitió para cada uno de los proteomas.

6.3.4 Clasificación de las proteínas de acuerdo con el tipo de reacción que catalizan

Para cada uno de los proteomas se filtraron las proteínas con actividad catalítica utilizando la hoja de cálculo Excel. Para lo cual:

- a. Primero se dio formato de tabla a los datos.
- b. Seguidamente se seleccionó la columna de EC number.
- c. En inicio, se seleccionó ordenar y filtros > filtros.
- d. Se seleccionó la flecha del encabezado de la columna *status* y se dejó solo seleccionada la casilla con números EC.
- e. Luego se dio clic en datos, posteriormente en tabla o rango.
- f. Luego, en Power Query, se seleccionó transformar datos > extraer > texto antes del limitador.
- g. En la tabla se escribió como delimitador un punto (.), quedando de esta forma solo el primer número de *EC number* que indica la clase de enzima de la proteína.

Este procedimiento se repitió para cada uno de los proteomas.

Para visualizar las enzimas con más de una actividad catalítica en cada uno de los proteomas se construyó una red de interacción enzimática a partir de las proteínas con actividad catalítica contenidas en cada uno de los proteomas de estudio (Otasek et al., 2019).

6.3.5 Clasificación de las enzimas de acuerdo con su residuo Terminal-N

Utilizando la función EXTRAER de la hoja de cálculo Excel se identificó el Terminal-N para las proteínas de cada uno de los proteomas de estudio, como se indica a continuación:

```
=EXTRAE([@Sequence];[@Chain];1
```

Donde *Sequence* es la columna de la secuencia de residuos aminoácidos, de cada proteína, representados con el símbolo de una letra. *Chain* es la columna que me indica la posición de el residuo aminoácido terminal-N y 1 es la cantidad de caracteres que se extrae.

6.3.6 Clasificación de cada uno de los residuos terminal-N como estabilizador y desestabilizador.

Posteriormente, utilizando la herramienta buscar y emplazar de la hoja de cálculo Excel, se clasificó cada Terminal-N como estabilizador y desestabilizador, según la Tabla 1.

6.3.7 Clasificación de las secuencias de acuerdo con la superfamilia a la que pertenecen.

Para cada una de las enzimas descargadas (de todos los proteomas no redundantes seleccionados) se filtró y graficó solo aquellas que tienen asignada una superfamilia.

6.3.8 Clasificación de las enzimas de acuerdo con la ruta metabólica que cataliza

Para cada una de las enzimas descargadas (de todos los proteomas no redundantes seleccionados) se filtró y graficó solo aquellas que participan en una ruta metabólica.

7 RESULTADOS Y DISCUSIÓN

7.1 Selección de los proteomas de referencia

El estudio de las tendencias se realizó solo con los proteomas de los organismos *Escherichia coli*, *Nicotiana tabacum*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus*. De los cinco proteomas seleccionados originalmente se descartó para el análisis de resultados el proteoma del virus de inmunodeficiencia humana (VIH-1) por tener solo una enzima con estatus de revisada en la base de datos Uniprot. No obstante, los resultados para esta enzima se encuentran en la sección de anexos (ver Tabla 11).

7.2 Proteínas de los proteomas de referencia.

En la Figura 7 se muestra el número de proteínas presentes en los proteomas en estudio. En esta gráfica se observa que el proteoma con mayor número de proteínas reportadas (en Uniprot) es del proteoma *Nicotiana tabacum*, con 73606 proteínas, mientras que el proteoma con menor número de proteínas reportadas es del *Saccharomyces cerevisiae*, con 6050 proteínas. La sumatoria de las proteínas reportadas de cada uno de los proteomas da un total de 130885.

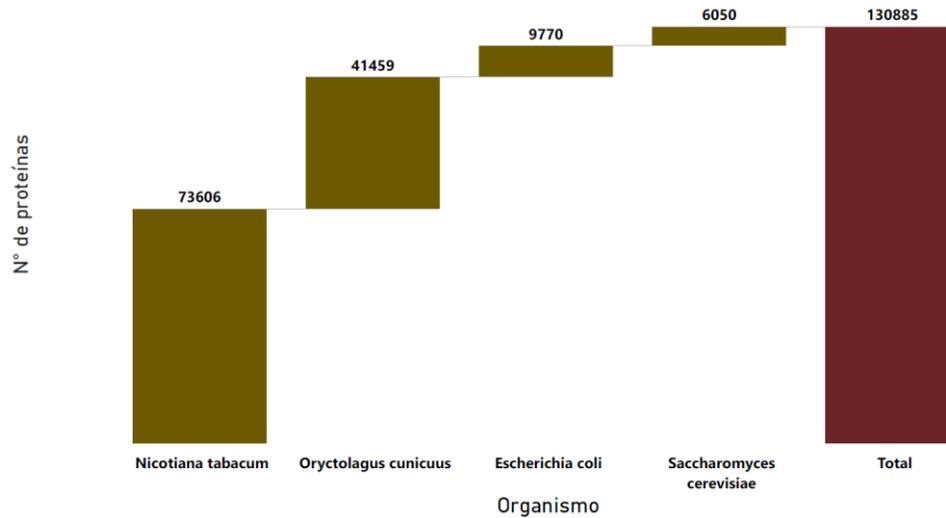


Figura 7. Conteo de proteínas en cada proteoma. En la columna color verde zapote se representa las cantidades de proteínas de cada proteoma y en la columna color vino tinto se representa el total de las proteínas.

7.3 Clasificación de las proteínas con estatus de *revisadas*.

De las proteínas obtenidas se seleccionaron con estatus de *revisadas*. Se encontró que el proteoma con mayor número de proteínas *revisadas* pertenece a *Escherichia coli* (6423 proteínas), mientras que el proteoma *Saccharomyces cerevisiae* (392 proteínas) tiene el menor número de proteínas *revisadas*. Adicionalmente, se encontró que el número total de proteínas *revisada* es de 13758 (Figura 8).

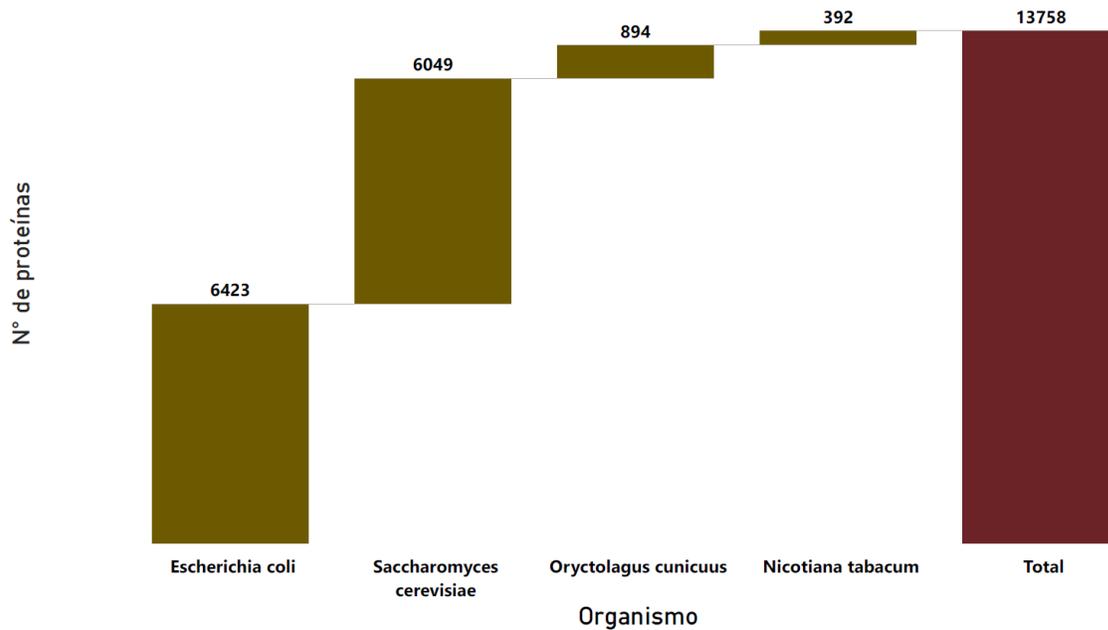


Figura 8. Conteo de proteínas con estatus de revisadas en cada proteoma. En la columna color verde zapote se muestra el número de proteínas con estatus de revisadas para cada uno de los proteomas y en la columna color vino tinto se muestra el total de proteínas con estatus de revisadas.

7.4 Clasificación de las proteínas de acuerdo con el tipo de reacción que catalizan.

De las proteínas con estatus de *revisadas* se seleccionaron solo aquellas clasificadas con las actividades enzimáticas de oxidoreductasas (EC:1), transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En la Figura 9 se muestra el conteo de enzimas agrupadas de acuerdo con el tipo de reacción que catalizan.

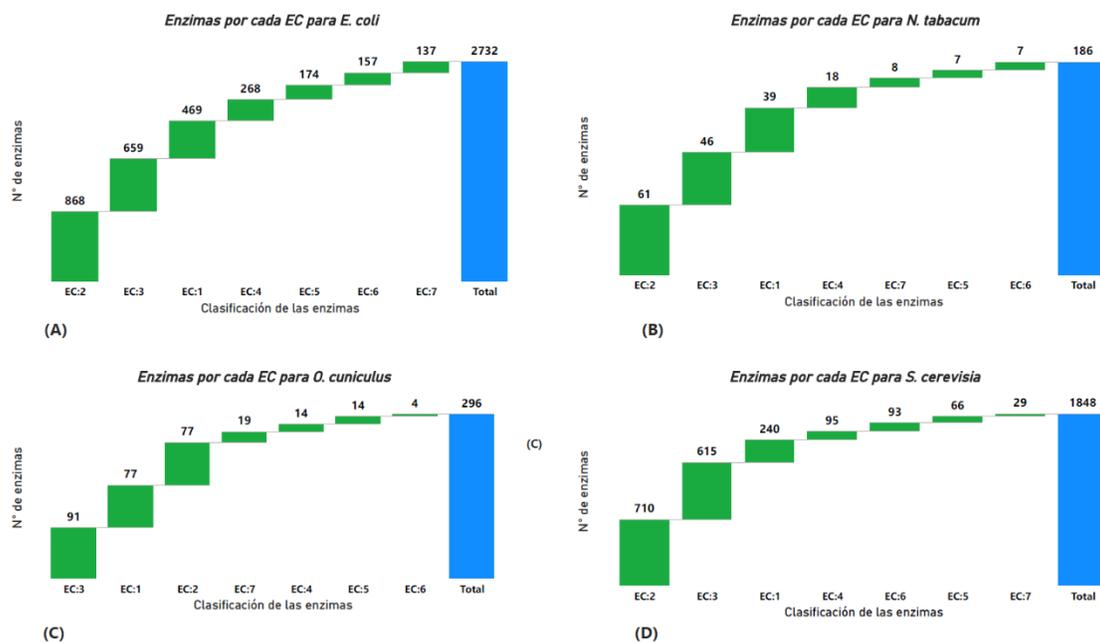


Figura 9. Conteo de enzimas por cada clase de enzima (EC) para cada uno de los proteomas. Oxidorreductasas (EC:1), Transferasas (EC:2), Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En esta figura se muestra las enzimas agrupadas por clase de enzima para cada proteoma (columna color verde) y la sumatoria, por proteoma, de las enzimas de cada actividad enzimática (columna azul). En (A) están representadas las enzimas del proteoma *Escherichia coli*, en (B) están representadas las enzimas del proteoma *Nicotiana tabacum*, en (C) están representadas las enzimas del proteoma *Oryctolagus cuniculus* y en (D) están representadas las enzimas del proteoma *Saccharomyces cerevisiae*.

En cada uno de los cuatro proteomas, las enzimas clasificadas como oxidorreductasas (EC:1), transferasas (EC:2) e hidrolasas (EC:3) tienden a ser las más abundantes. En los proteomas *Escherichia coli*, *Nicotiana tabacum* y *Saccharomyces cerevisiae*, las enzimas más abundantes son las transferasas (EC:2), seguido de las hidrolasas (EC:3) y las oxidorreductasas (EC:1). En *Oryctolagus cuniculus*, las hidrolasas (EC:3) son las más abundantes, seguido de las oxidorreductasas (EC:1) y transferasas (EC:2). Adicionalmente, se observa que el proteoma con menor número de enzimas es el *Nicotiana tabacum*, con 186 enzimas, seguido de los proteomas *Oryctolagus cuniculus*, *Saccharomyces cerevisiae* y *Escherichia coli* con 296, 1848 y 2732 enzimas, respectivamente (Figura 9, Figura 10).

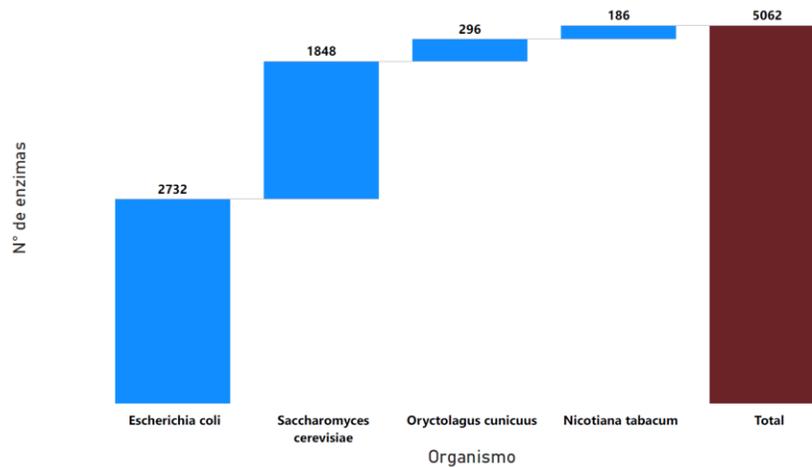


Figura 10. Conteo de proteínas con actividad catalítica en cada proteoma. En las columnas verde zapote se encuentra la sumatoria del conteo de enzimas por actividad catalítica y en la columna color rojo se representa el conteo del total de enzimas.

El conteo total de enzimas por cada proteoma es la sumatoria del número de enzimas agrupadas por cada clase de enzima (EC) (Figura 9). Eso quiere decir que si una enzima pertenece a más de un proteoma y/o que tenga más de una actividad catalítica se va a contar más de una vez. Al obtener los resultados evitando el doble conteo de la entrada (*Entry*) en todos los proteomas el número de enzimas disminuyó (Figura 10 y Figura 11), confirmando que efectivamente hay enzimas que se repiten.

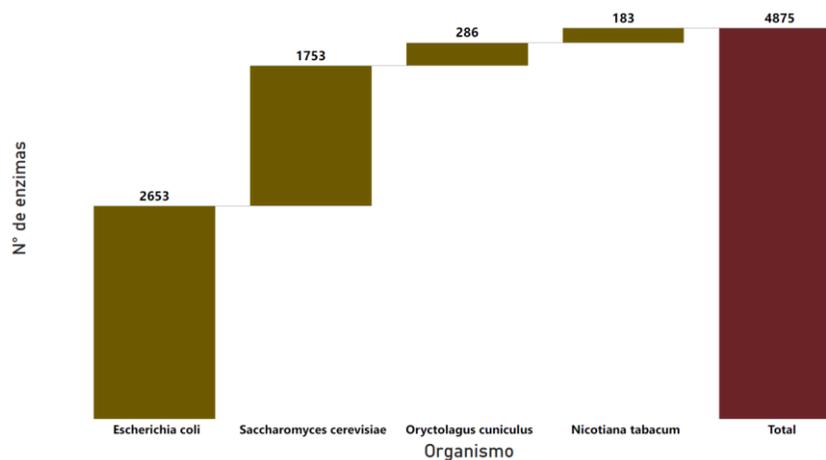


Figura 11. Conteo de enzimas en cada proteoma. En las columnas verde zapote esta las enzimas pro cada proteoma, y en la columna vino-tinto esta la sumatoria de las enzimas de cada proteoma.

Para visualizar si las enzimas se repiten en más de un proteoma, se construyó una red de interacción enzimática con las proteínas que tienen alguna actividad catalítica y que están contenidas en cada uno de los proteomas de estudio (Figura 12). Al graficar el conjunto de datos, con el objetivo de observar el comportamiento de la red de interacción de las enzimas, se muestra la formación de 4 agrupaciones, uno por cada proteoma (*Escherichia coli*, *Saccharomyces cerevisiae*, *Oryctolagus cuniculus* y *Nicotiana tabacum*). Cada una de las cuatro agrupaciones está formada por un nodo central color rojo (que representa a los proteomas de estudio) al cual están unidos varios nodos de color naranja (que representan las enzimas de dicho proteoma). Como resultado se encontró un total de 4879 nodos: 2654 nodos para la agrupación *Escherichia coli*; 1754 nodos para la agrupación *Saccharomyces cerevisiae*; 287 nodos para la agrupación *Oryctolagus cuniculus* y 184 nodos para la agrupación *Nicotiana tabacum*. Al restar los nodos color rojo al total de nodos de cada proteoma los resultados coinciden con los datos ilustrados en la Figura 11.

Para la misma gráfica, el número de aristas (líneas que une a los nodos color naranja con el nodo color rojo) dio como resultado un total de 5062. Ahora, si cada enzima (nodos color naranja) está conectado solo una vez, directa y únicamente, con un solo nodo rojo, el número de aristas debería de dar un total de 4875 aristas, o sea, debería de coincidir con el número de nodos que representa a las enzimas. Por lo tanto, un resultado de 5062 aristas nos indica que algunas enzimas se repiten. Igualmente, se observa que los nodos que representan a las enzimas de cada agrupación están conectados a un solo nodo rojo indicando que las enzimas de cada uno de los proteomas tienden a ser únicas para cada proteoma.

Para identificar las enzimas que se repiten más de una vez se discriminaron los nodos que tienen más de una arista conectada con el nodo central encontrándose que las enzimas pueden tener de 1 a 4 aristas (Figura 12).

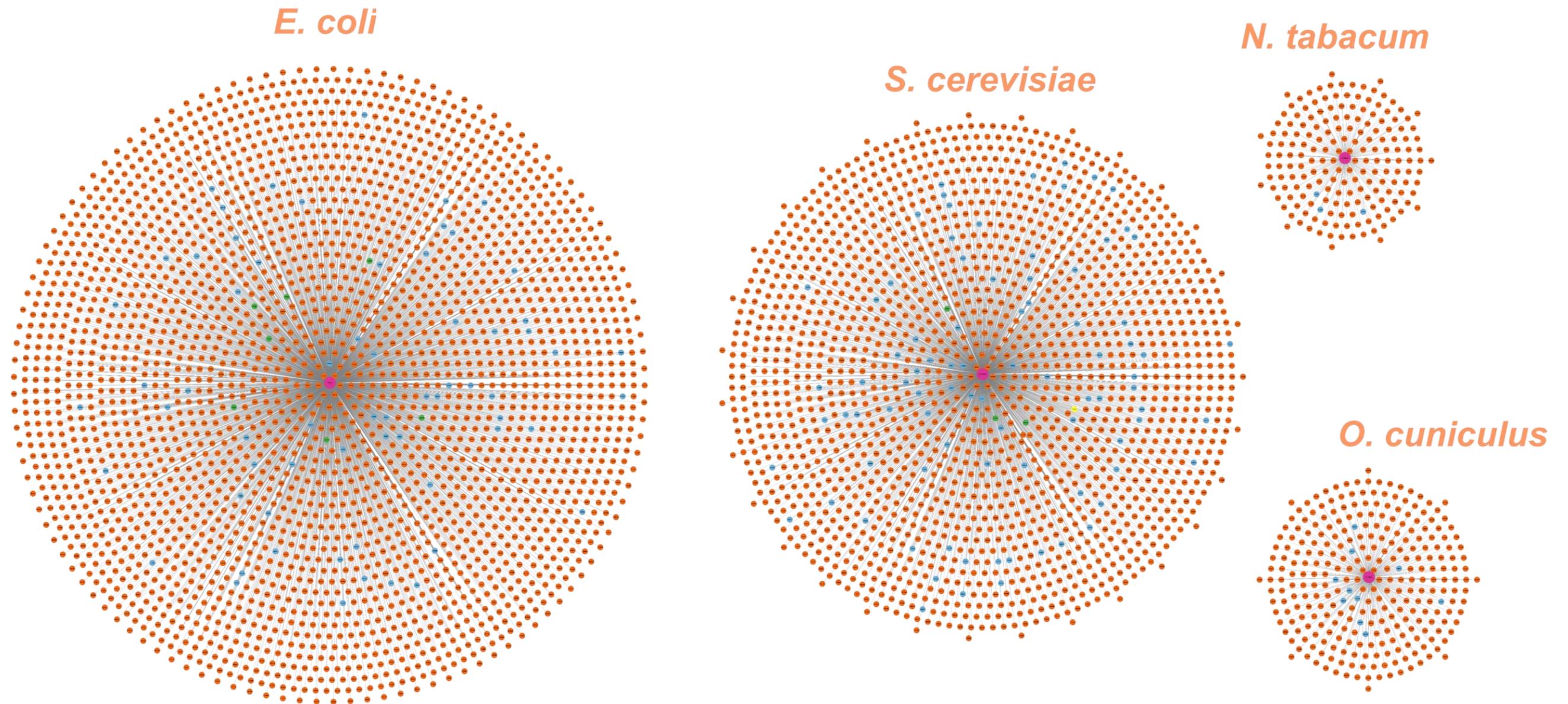


Figura 12. Visualización de las enzimas que se repiten más de una vez en los cuatro proteomas. En esta figura se observan cuatro clústeres, cada grupo (o clúster) está formado por un nodo central color rojo que representa al proteoma (*Escherichia coli*, *Saccharomyces cerevisiae*, *Oryctolagus cuniculus* y *Nicotiana tabacum*) y del cual se desprenden los nodos color naranja, azul, verde y amarillo, los cuales representan cada una de las enzimas. Los nodos de color naranja están unidos al nodo central por medio de una sola arista, los nodos de color azul están unidos al nodo central por medio de dos aristas, los nodos de color verde están unidos al nodo central por medio de tres aristas y el nodo de color amarillo está unido al nodo central por medio de cuatro aristas. Cada arista representa las veces que se repite la enzima.

La Figura 13 es la ampliación de una sección de la Figura 12, más exactamente de la agrupación del proteoma *Saccharomyces cerevisiae*, en la región que rodea al nodo amarillo y que a la vez coincide vecinalmente con los nodos de color naranja, verde y azul. En esta grafica se visualiza perfectamente como las aristas (o líneas) se desprenden de cada nodo (o enzima).

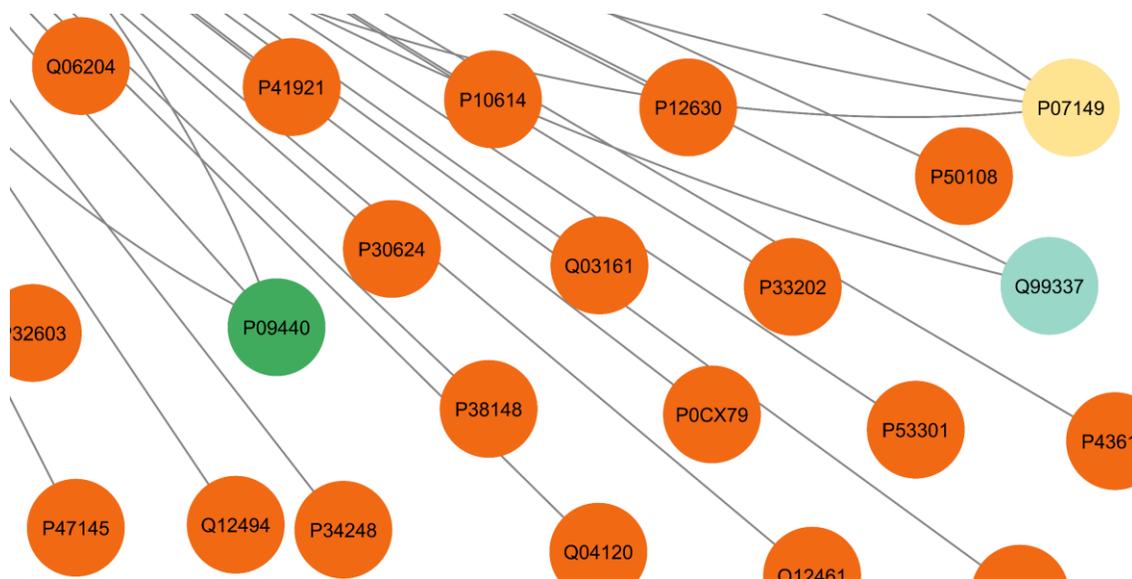


Figura 13. Ampliación de una sección de la Figura 12 del proteoma *Saccharomyces cerevisiae*, en la región que rodea al nodo amarillo y que a la vez coincide vecinalmente con nodos de color naranja, verde y azul. En esta grafica se visualiza con mayor detalle la forma y propiedades de cada nodo. Obsérvese que de los nodos de color naranja se depende solo una arista o línea, del nodo color azul se desprenden dos aristas, del nodo color verde se desprenden tres aristas y del nodo color amarillo se despende cuatro aristas. Gráfico elaborado con Cytoscape.

Para todos los agrupamientos, la sumatoria de las enzimas (o nodos) que se repiten más de una vez (o que tienen más de una arista) representan un total 175 nodos. De estos 175 nodos el proteoma *Escherichia coli* tiene 72 nodos, 7 con 3 aristas y el resto con 2 aristas; el proteoma *Saccharomyces cerevisiae* tiene 90 nodos, 1 con 4 aristas, 3 con 3 aristas y el resto con 2 aristas; el proteoma *Oryctolagus cuniculus* tiene 10 nodos, todos con dos aristas; y el proteoma *Nicotiana tabacum* tiene 3 nodos, todos con dos aristas (Figura 12, en anexos la Tabla 8).

Partiendo de que las enzimas se repiten dentro del proteoma al que pertenece, se realizaron redes de interacción para visualizar cual es el comportamiento de la red. Se elaboraron cuatro redes de interacción (Figura 14, Figura 15 y en anexos la Figura 61 y Figura 62), una por cada proteoma, para lo cual, se utilizaron las enzimas agrupadas de acuerdo con el tipo de reacción que catalizan.

En estos gráficos se muestra claramente que hay enzimas que tienen más de una arista, debido a que pertenece a más de una clase de enzima (EC) del mismo proteoma. En el proteoma *Nicotiana tabacum* se muestran tres proteínas con nodos azules que tienen dos actividades enzimáticas. Por ejemplo, el identificador (*Entry*) P10978 presenta las actividades transferasa (EC:2) e hidrolasa (EC:3), el identificador Q42942 presenta las actividades transferasa (EC:2) y liasa (EC:4) e identificador Q42965 presenta las actividades hidrolasa (EC:3) y liasa (EC:4).

Asimismo, los resultados muestran que las enzimas del agrupamiento para las translocasas (EC:7), de los grafos de los proteomas *Oryctolagus cuniculus*, *Nicotiana tabacum* y *Escherichia coli*, tienden a no interactuar con alguno de los otros 5 grupos de enzimas, indicando que estas proteínas solo se clasifican en esa única actividad enzimática. (Figura 14, en anexos la Figura 61 y Figura 62).

La Figura 16 es una versión de la Figura 15 donde se muestra solo las proteínas que tienen más de una clase de enzima. En esta gráfica se visualiza perfectamente el identificador de entrada de cada enzima y las clases de enzimas en las que se clasifica. Los resultados muestran que en el proteoma *Saccharomyces cerevisiae* (Figura 16), en comparación con los otros tres proteomas (Figura 14, Figura 15, en anexos la Figura 61 y Figura 62) es el único que posee una única enzima (identificador P07149, nodo amarillo) que se clasifica en cuatro clases de enzimas (oxidorreductasas, hidrolasas, transferasas y liasas).

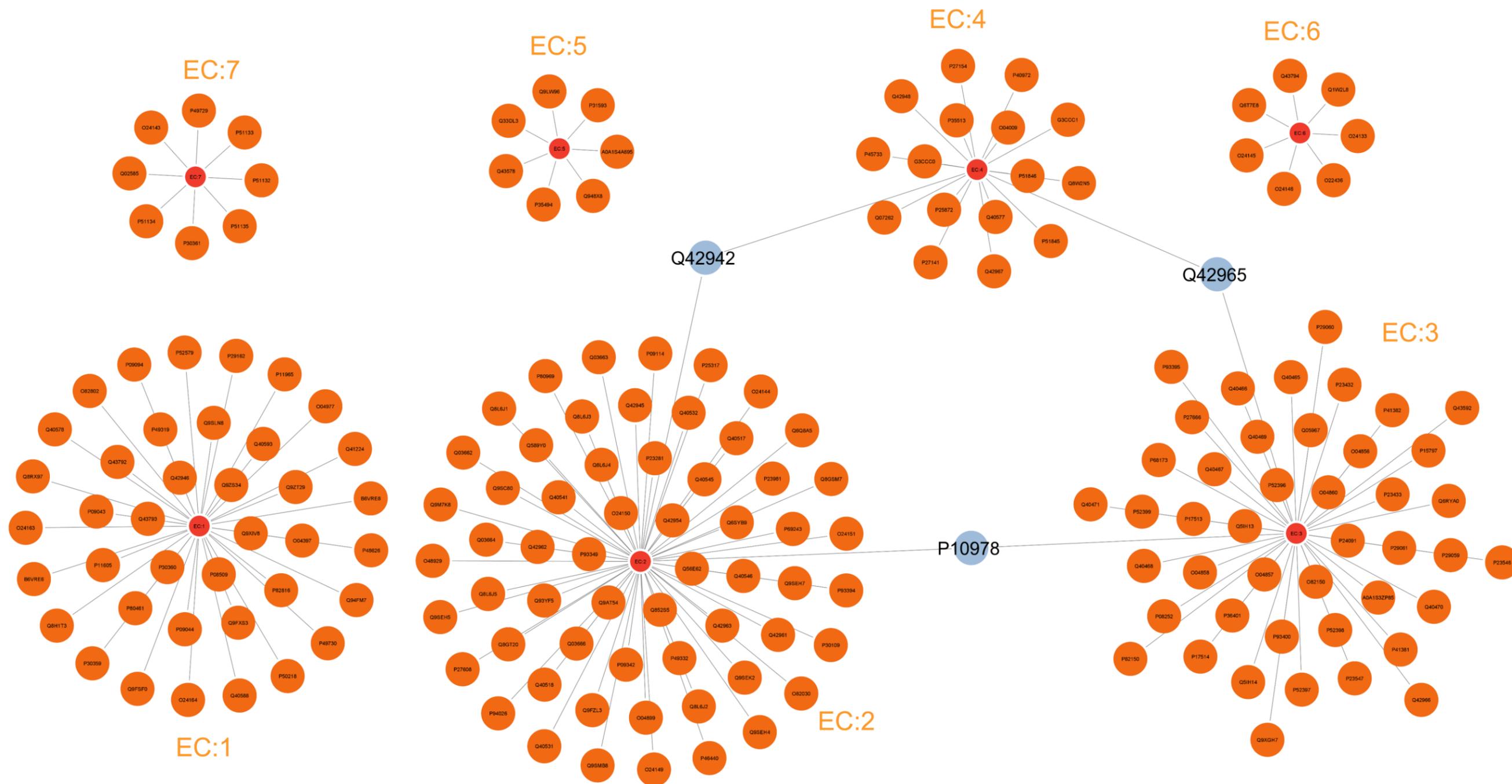


Figura 14. Visualización de las proteínas con más de una actividad enzimática en *Nicotiana tabacum*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). Para el proteoma *Nicotiana Tabacum* se presenta 7 clúster, uno por cada clase o actividad enzimática (EC). Los nodos color rojo son los nodos centrales que representa la clasificación de las enzimas, los nodos color azul son enzimas que tienen dos aristas y los nodos color naranja son las enzimas que tienen una sola arista. La fuente del texto de los identificadores de entrada (*Entry*) para las proteínas con más de una clasificación enzimática se colocaron en mayo tamaño para una mejor visualización.

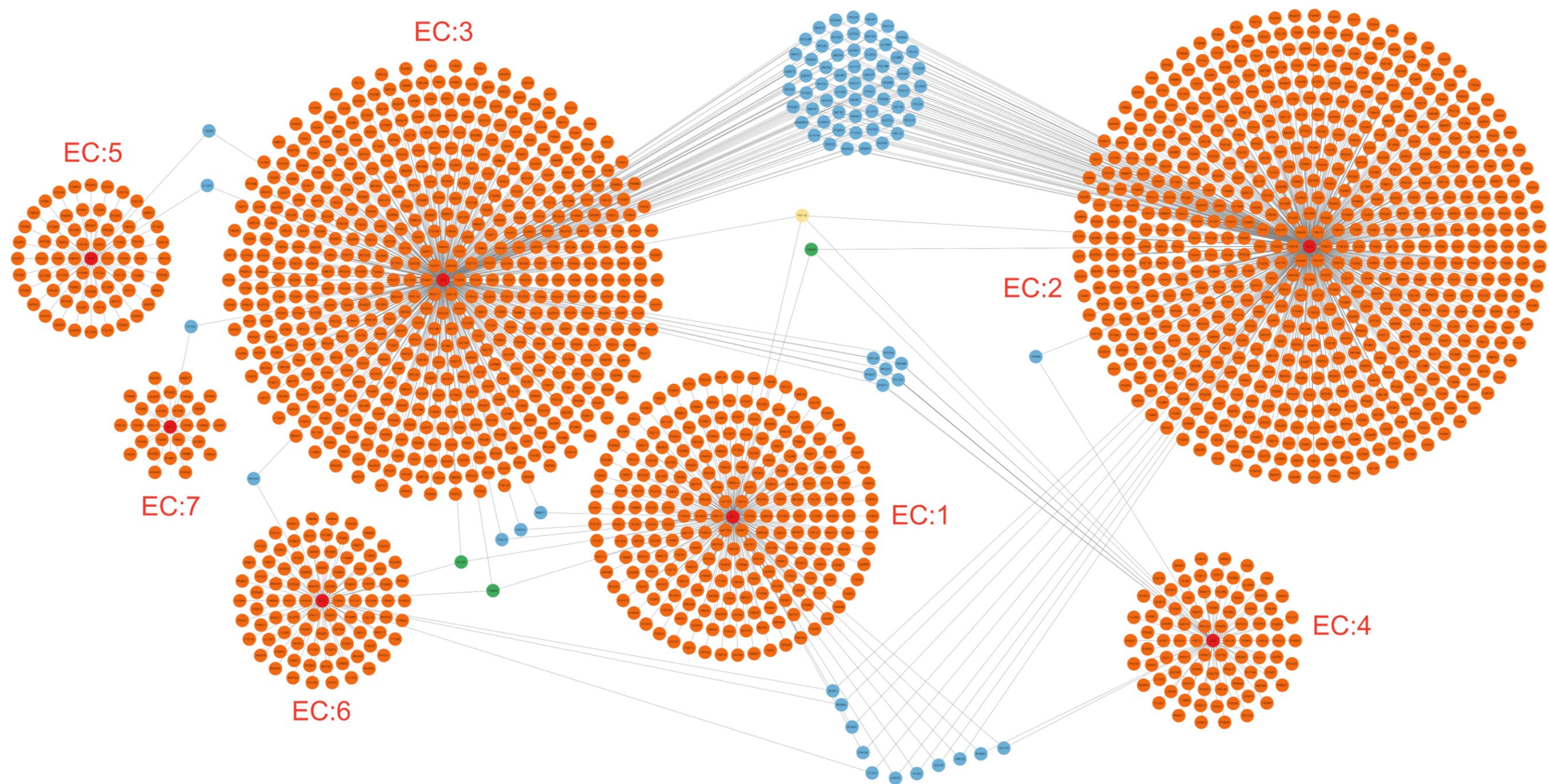


Figura 15. Visualización de las proteínas con más de una actividad enzimática en *Saccharomyces cerevisiae*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En la gráfica se observa la formación de 7 clúster, un clúster por cada clase de enzima (EC) para el proteoma *Saccharomyces Cerevisiae*. El nodo color amarillo es una enzima con 4 aristas, los nodos color verde son enzimas que tienen 3 aristas, los nodos color azules tienen 2 aristas y los nodos color naranja son enzimas que tienen 1 arista.

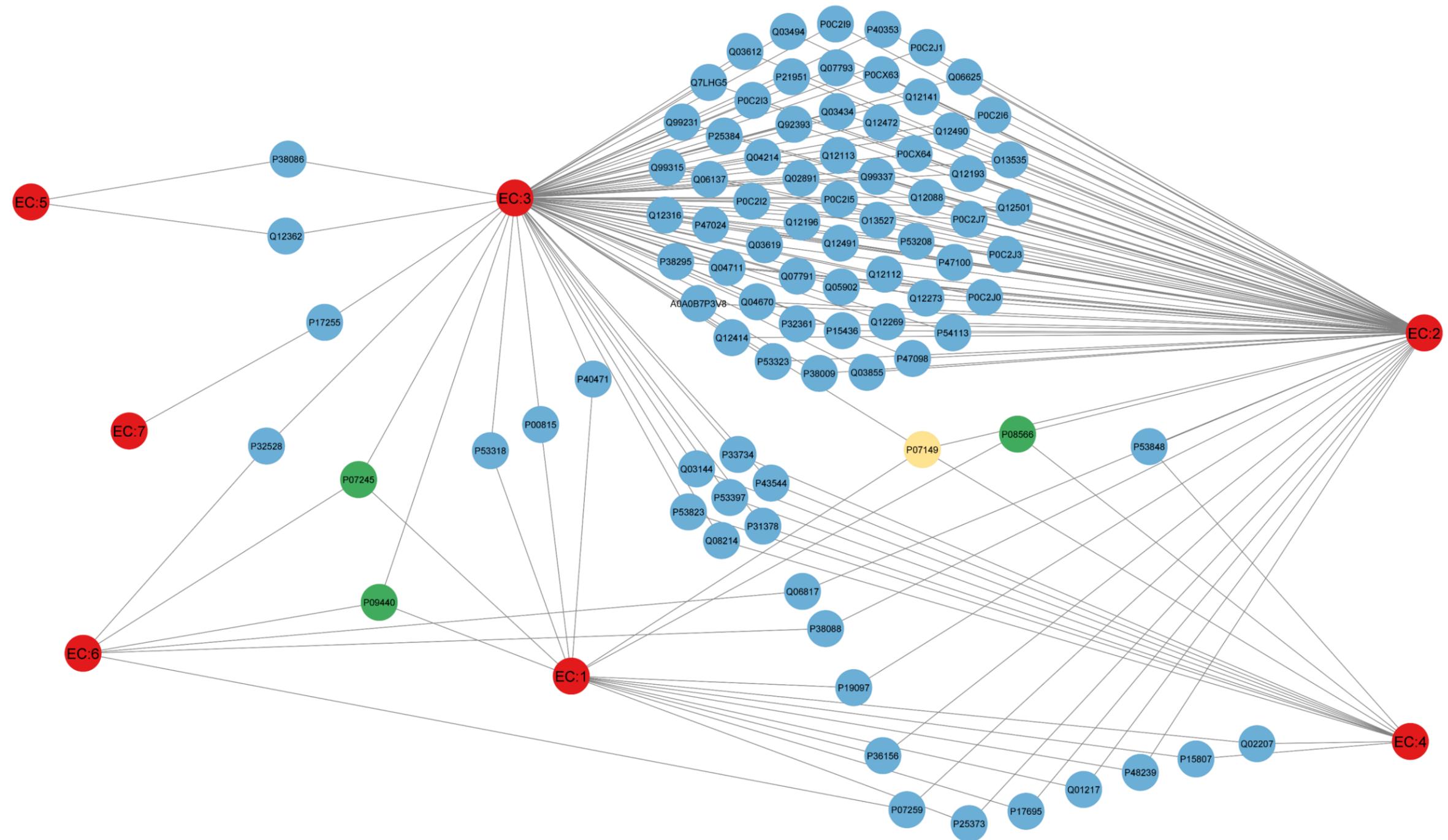


Figura 16. Visualización de las proteínas con más de una actividad enzimática en *Saccharomyces cerevisiae*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En la gráfica se observa solo las proteínas que se clasifican en más de una clase de enzima (EC) para el proteoma *Saccharomyces Cerevisiae*. El nodo color amarillo es una enzima con 4 aristas, los nodos color verde son enzimas que tienen 3 aristas y los nodos color azules tienen 2 arista

7.5 Clasificación de las enzimas de acuerdo con su residuo terminal-N

Después de descartar la secuencia de aminoácidos del péptido señal, dejando solo la secuencia de aminoácidos de la proteína madura, las enzimas se agruparon de acuerdo con su residuo terminal-N y los resultados se muestran en la Figura 17. En esta gráfica se observa que, de las 4875 enzimas presentes en los cuatro proteomas, la metionina (M) es el terminal-N que agrupa mayor cantidad de enzimas, con 3744 enzimas y, de forma contraria, los residuos terminal-N R y W son los que agrupa menos enzimas, con 5 enzimas cada uno.

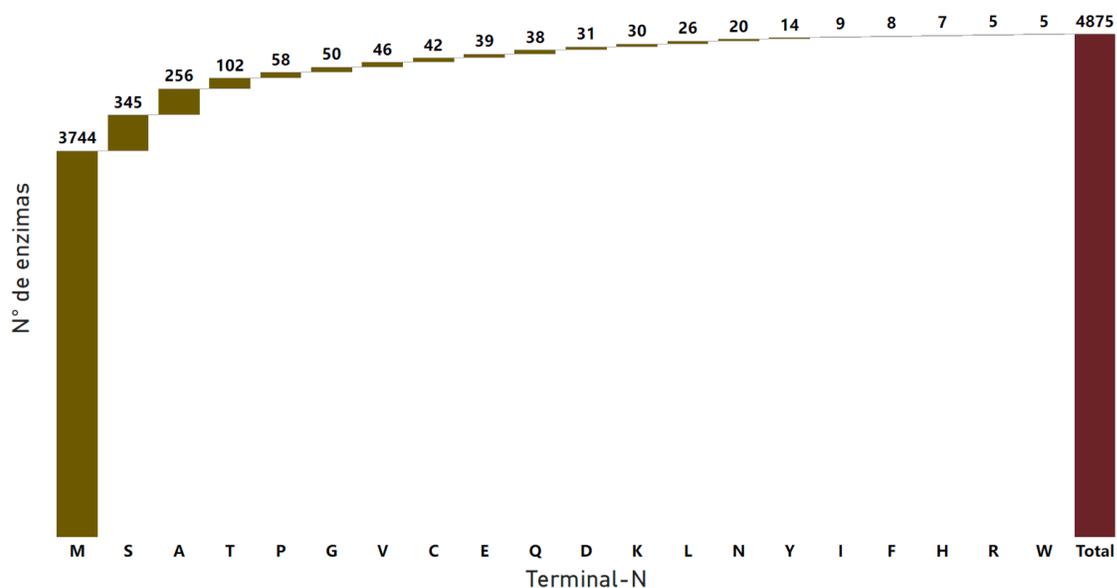


Figura 17. Agrupamiento de las enzimas de acuerdo con su terminal-N. Conteo de enzimas agrupadas por residuo terminal-N sin discriminar al tipo de proteoma que pertenezca.

Del 100% de (las 4875) enzimas aproximadamente el 77% tienen como terminal-N un residuo aminoácido metionina (M) y aproximadamente el 91% de las (4875) enzimas se agrupa solo en 4 residuos terminal-N (M, S, A y T). Para identificar si el agrupamiento de las enzimas por proteoma individual sigue las mismas tendencias de la Figura 17 se realizó la Figura 18.

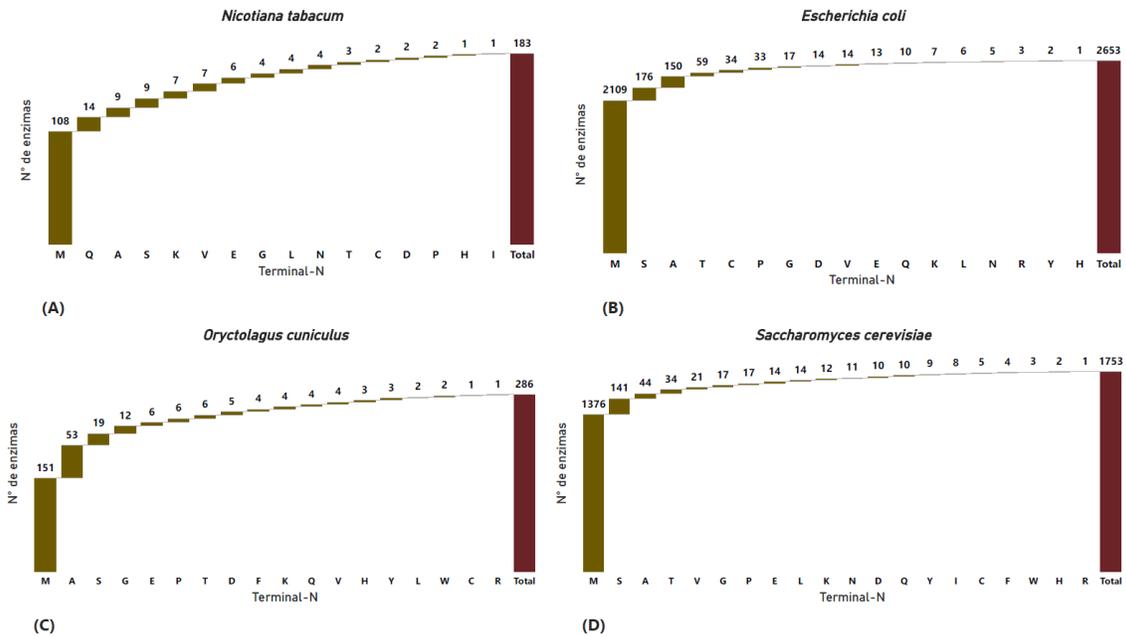


Figura 18. Agrupamiento de las enzimas de acuerdo con su terminal-N en cada proteoma. En (A) enzimas agrupadas por su residuo terminal-N en el proteoma *Nicotiana tabacum*. En (B) enzimas agrupadas por su residuo terminal-N en el proteoma *Escherichia coli*. En (C) enzimas agrupadas por su residuo terminal-N en el proteoma *Oryctolagus cuniculus*. En (D) enzimas agrupadas por su residuo terminal-N en el proteoma *Saccharomyces cerevisiae*.

De los resultados de la Figura 18, se encontró que la mayoría de las enzimas de cada uno de los proteomas tienden a tener como terminal-N al residuo aminoácido metionina, 59% en *Nicotiana tabacum*, 79% en *Escherichia coli*, 53% en *Oryctolagus cuniculus* y 78% en *Saccharomyces cerevisiae*, lo cual coincide con la tendencia que se analizó para la Figura 17.

Otra tendencia encontrada es que dentro de los cuatro terminal-N (M, Q, A y S para *Nicotiana tabacum*; M, S, A y T para *Escherichia coli*; M, A, S y G para *Oryctolagus cuniculus*; y M, S A y T para *Saccharomyces cerevisiae*) donde se agrupan la mayoría de las enzimas en cada uno los cuatro proteomas se encuentran los residuos aminoácidos M, A y S. Las sumatoria del conteo de enzimas que se agrupan en estos cuatro terminal-N para los proteomas *Nicotiana tabacum*, *Escherichia coli*, *Oryctolagus cuniculus* y *Saccharomyces cerevisiae* corresponden aproximadamente al 77%, 94%, 92 y 91% respectivamente (Figura 18).

Para identificar otras tendencias en cada uno de los proteomas se graficaron los datos agrupando las enzimas de acuerdo con el residuo aminoácido terminal-N estabilizador y

desestabilizador (de acuerdo con la Tabla 1) que poseen las enzimas en estudio (Figura 19 a la Figura 22).

Los resultados muestran que, para el proteoma *Escherichia coli* (Figura 19), aproximadamente el 0,7% de las proteínas con actividad catalítica están agrupadas como enzimas con residuos aminoácidos terminal-N desestabilizadores y aproximadamente el 98% se agrupan como enzimas con residuos aminoácidos terminal-N estabilizadores. De aquí en adelante las enzimas que tienen un residuo terminal-N desestabilizador también serán llamadas enzimas con tiempos de vida inestables y a las enzimas que tienen residuos terminal-N estabilizador también serán llamadas enzimas con tiempos de vida estables.

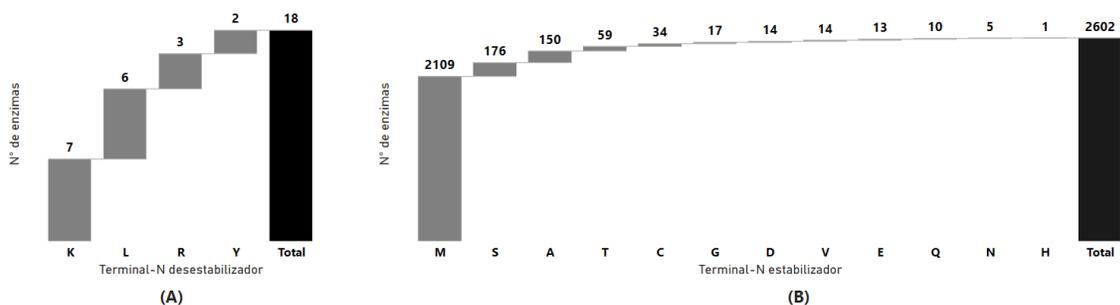


Figura 19. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma *Escherichia coli*. En (A) número de enzimas con el terminal-N desestabilizador y en (B) número de enzimas con el terminal-N estabilizador.

En la Figura 20 del proteoma *Nicotiana tabacum*, aproximadamente el 22% de las proteínas con actividad catalítica tienen tiempos de vida inestables, debido a que tienen residuos aminoácidos terminal-N desestabilizadores, y aproximadamente el 77% tienen tiempos de vida estable debido a que se agrupan como enzimas con residuos aminoácidos terminal-N estabilizadores.

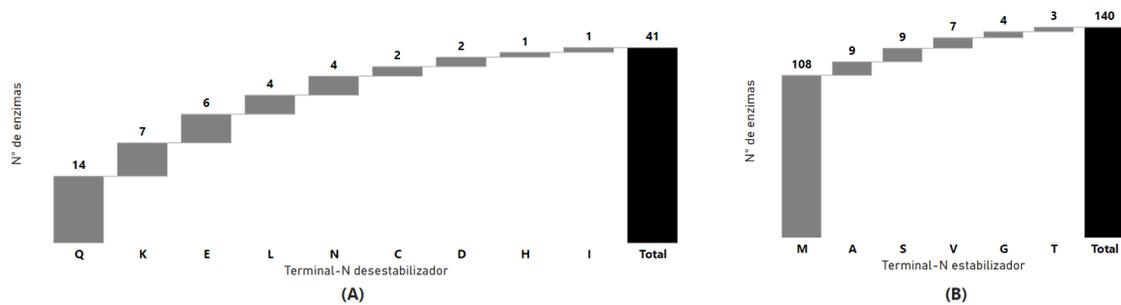


Figura 20. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma *Nicotiana tabacum*. En (A) está representado el número de enzimas que tienen un terminal-N desestabilizador y en (B) está representado el número de enzimas que tienen un terminal-N estabilizador.

En la Figura 21, del proteoma *Oryctolagus cuniculus*, aproximadamente el 40% de las proteínas con actividad catalítica están agrupadas como enzimas con residuos aminoácidos terminal-N desestabilizadores y aproximadamente el 60 % se agrupan como enzimas con residuos aminoácidos terminal-N estabilizadores.

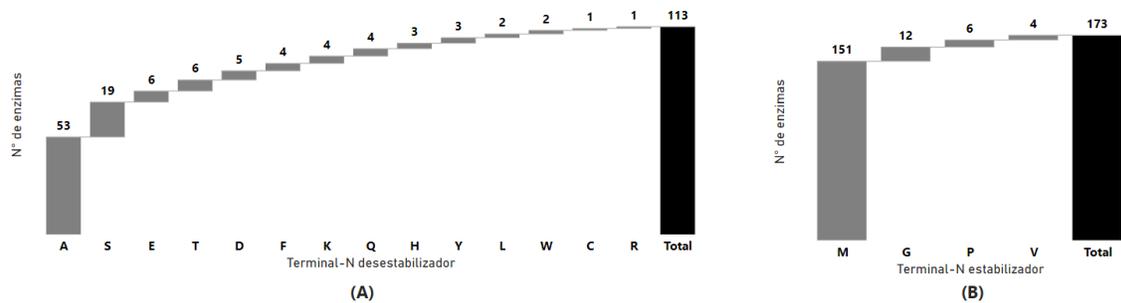


Figura 21. Agrupamiento de las enzimas por su terminal-N estabilizador y desestabilizador para el proteoma *Oryctolagus cuniculus*. En (A) está representado el número de enzimas que tienen un terminal-N desestabilizador y en (B) está representado el número de enzimas que tienen un terminal-N estabilizador.

En la Figura 22, del proteoma *Saccharomyces cerevisiae*, aproximadamente el 6 % de las proteínas con actividad catalítica están agrupadas como enzimas con residuos aminoácidos terminal-N desestabilizadores y aproximadamente el 94% se agrupan como enzimas con residuos aminoácidos terminal-N estabilizadores.

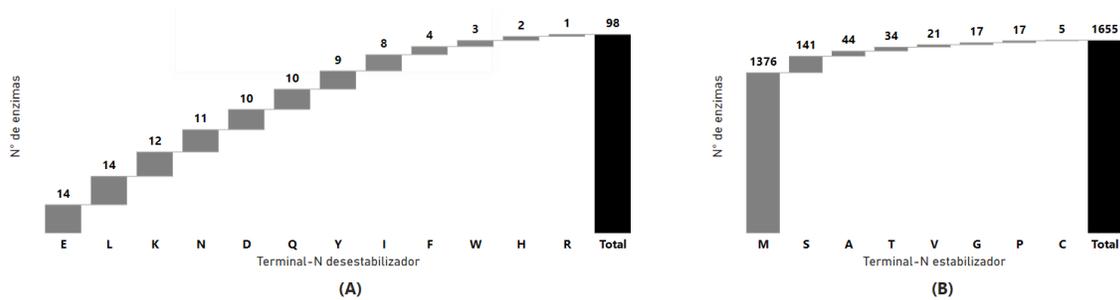


Figura 22. Agrupamiento de las enzimas por su terminal-N estabilizador y destabilizador para el proteoma *Saccharomyces cerevisiae*. En (A) está representado el número de enzimas que tienen un terminal-N destabilizador y en (B) está representado el número de enzimas que tienen un terminal-N estabilizador.

Los resultados muestran que el residuo aminoácido terminal-N estabilizador metionina es predominante en las enzimas de los cuatro proteomas. Por lo tanto, se puede decir que la mayoría de las proteínas con función catalítica tienen tiempos de vida estables. En estas gráficas se observa que hay residuos terminal-N que son estabilizadores en un proteoma y en otro son destabilizadores, por ejemplo, el residuo aminoácido alanina (A) se encuentra como terminal-N estabilizador en los proteomas *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Escherichia coli*, mientras que en el proteoma *Oryctolagus cuniculus* se encuentra como residuo aminoácidos termina-N destabilizador (Tabla 1), además, en el *Oryctolagus cuniculus* es el residuo terminal-N destabilizador (Alanina, A) que está como terminal-N en la mayoría (aproximadamente el 45%) de las proteínas catalíticas que se encuentran agrupadas en el conjunto de enzimas con terminal-N destabilizador.

Adicionalmente, para entender la distribución de las enzimas respecto a su residuo terminal-N y su clasificación EC, se graficó el número de enzimas con residuo terminal-N estabilizador o destabilizador (Figura 23, Figura 24).

Los resultados de la Figura 23 evidencian que la clase transferasa (EC:2) presenta mayoritariamente un residuo aminoácido terminal-N estabilizador, excepto en el proteoma *Oryctolagus cuniculus* donde predomina la oxidorreductasa. En general, se observa que el residuo terminal-N estabilizadores son frecuentes en las clases oxidorreductasas (EC:1), transferasas (EC:2) e hidrolasas (EC:3).

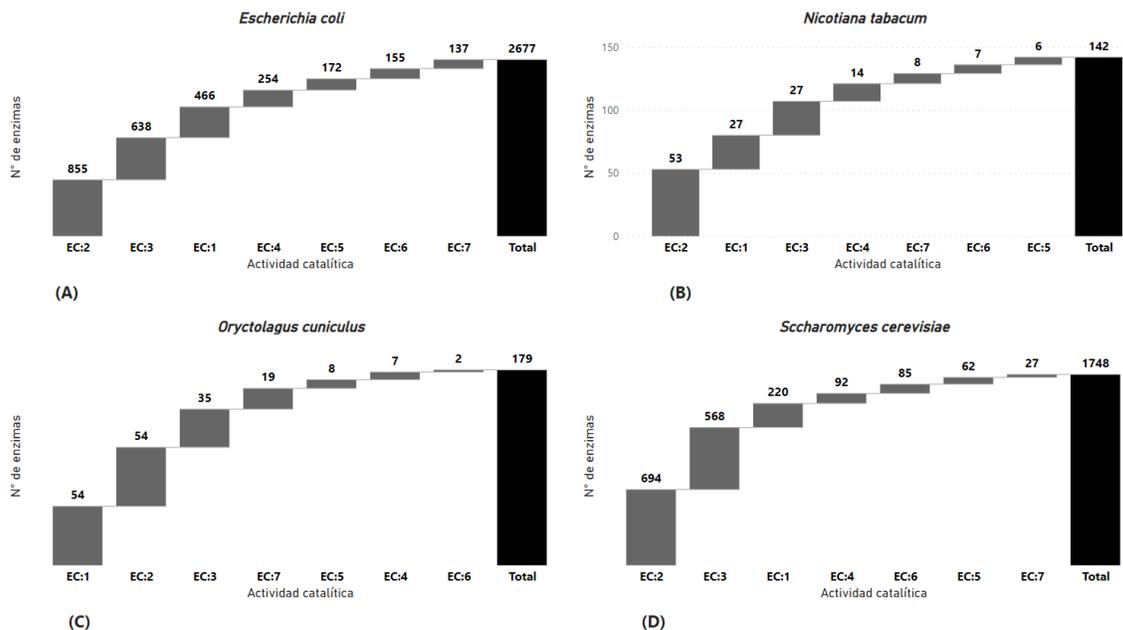


Figura 23. Conteo de enzimas que tienen un terminal-N estabilizador. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). Número de enzimas con residuo aminoácido terminal-N estabilizador por cada clase de enzima para cada proteoma. (A) *Escherichia coli*, (B) *Nicotiana tabacum*, (C) *Oryctolagus cuniculus* y (D) *Saccharomyces cerevisiae*.

Por otro lado, en la Figura 24 se muestra que las hidrolasas (EC:3) son las que tienen más enzimas con un residuo aminoácido terminal-N desestabilizador, particularmente los proteomas *Oryctolagus cuniculus* y *Saccharomyces cerevisiae*. Asimismo, se observa que la clase de enzima translocasa (EC:7) de los proteomas *Escherichia coli*, *Oryctolagus cuniculus* y *Nicotiana tabacum* carecen de enzimas con residuos aminoácidos terminal-N desestabilizadores, excepto el *Saccharomyces cerevisiae* que tiene dos enzimas con terminal-N desestabilizador.

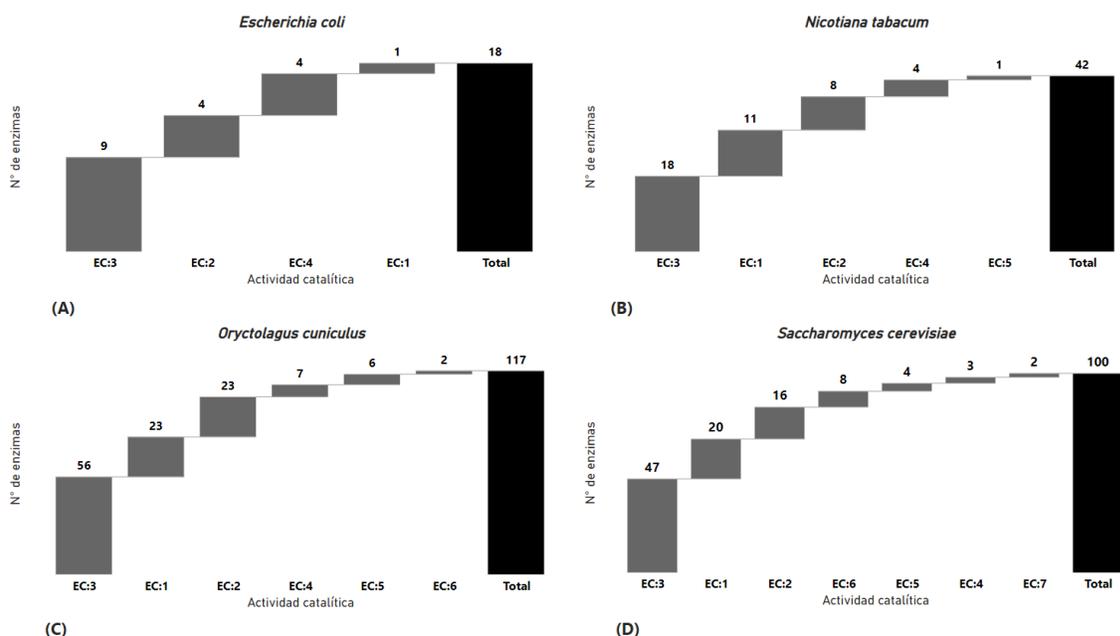


Figura 24. Conteo de enzimas que tienen un terminal-N desestabilizador. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). Número de enzimas con residuo aminoácido terminal-N estabilizador por cada clase de enzima para cada proteoma. (A) *Escherichia coli*, (B) *Nicotiana tabacum*, (C) *Oryctolagus cuniculus* y (D) *Saccharomyces cerevisiae*.

Para determinar si la presencia del terminal-N estabilizador depende de la clase de enzimas se graficaron las redes de interacción utilizando como variable los terminales-N estabilizadores y desestabilizadores (Figura 25, Figura 26, en anexos Figura 63, Figura 64). Adicionalmente, de los resultados obtenidos se analizó si hay una relación entre el número de residuo aminoácidos de la secuencia de la enzima madura con la clase de enzima y con el terminal-N.

Los resultados en la Figura 25 muestran 10 nodos que tienen doble arista (o proteínas que se clasifica en dos clases de enzimas), de estas 10 enzimas, el 60 % tiene como terminal-N un residuo aminoácido estabilizador y el 40 % un residuo aminoácido terminal-N desestabilizador (Tabla 2).

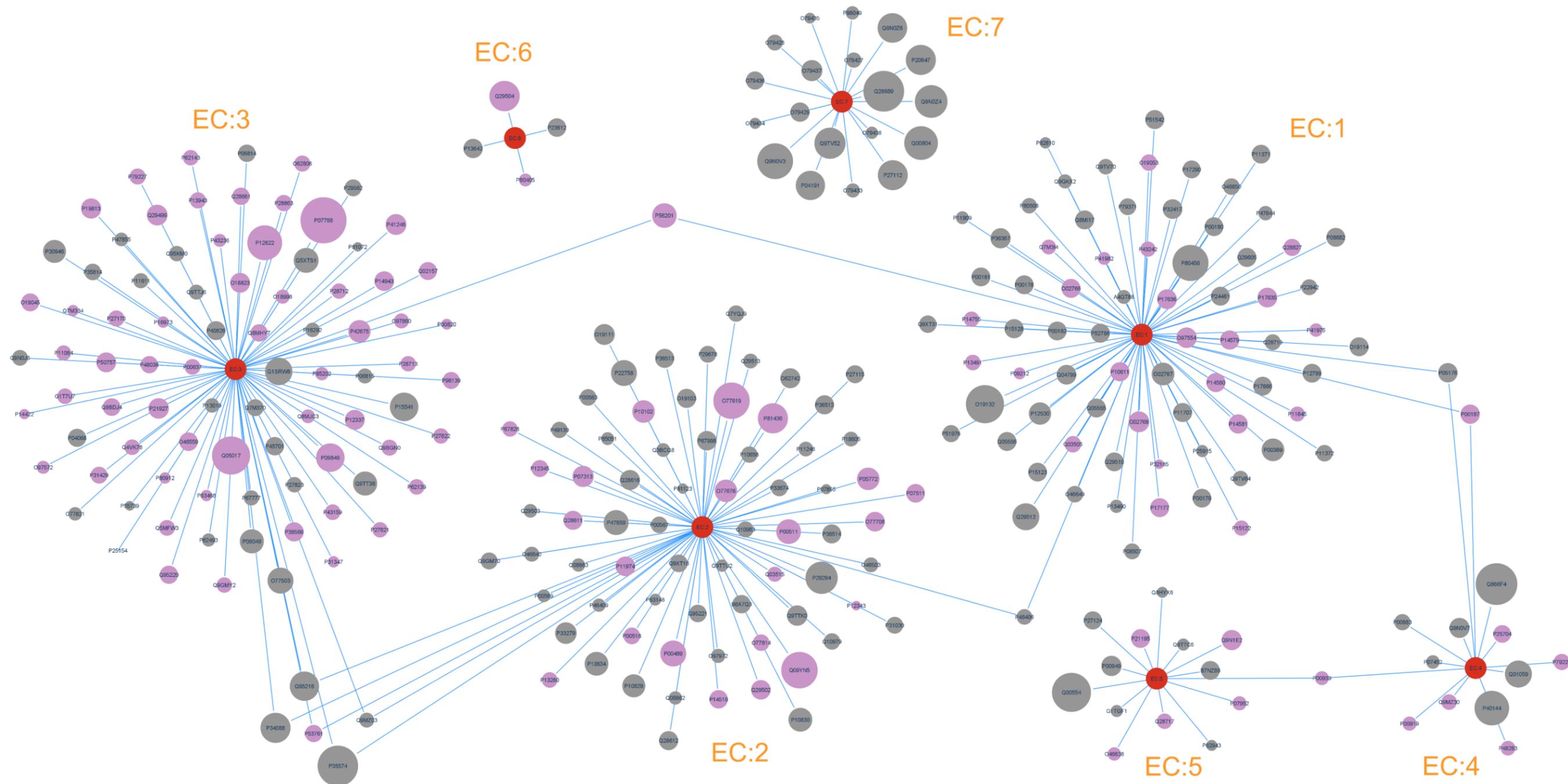


Figura 25. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N destabilizador para *Oryctolagus cuniculus*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En esta gráfica, se representa los enzimas con un Terminal-N destabilizador (nodos color rosa) y las enzimas con un Terminal-N estabilizador (nodos color gris). El tamaño del nodo es directamente proporcional al número de residuos aminoácidos de la secuencia de la enzima madura.

Tabla 2. Proteínas con dos actividades enzimáticas para el proteoma *Oryctolagus cuniculus*.

Enzima (<i>Entry</i>)	Terminal-N	N° de aminoácidos	N° de aristas	Clase de enzima
P00939	Estabilizador	247	2	Liasa; Isomerasa
Q9MZ03	Estabilizador	297	2	Transferasa; Hidrolasa
P46406	Estabilizador	332	2	Oxidoreductasa; Transferasa
P53761	Desestabilizador	415	2	Transferasa; Hidrolasa
P00187	Desestabilizador	514	2	Oxidoreductasa; Liasa
P05176	Estabilizador	517	2	Oxidoreductasa; Liasa
P56201	Desestabilizador	772	2	Oxidoreductasa; Hidrolasa
Q95216	Estabilizador	1004	2	Transferasa; Hidrolasa
P34088	Estabilizador	1041	2	Transferasa; Hidrolasa
P35574	Estabilizador	1554	2	Transferasa; Hidrolasa

Los resultados de la Figura 26 exponen que solo hay tres proteínas que presentan dos actividades catalíticas (EC). Dos tienen un residuo aminoácido terminal-N estabilizador y la restante un residuo aminoácido terminal-N desestabilizador (Tabla 3).

Tabla 3. Proteínas con dos actividades enzimáticas para el proteoma *Nicotiana tabacum*.

Enzima (<i>Entry</i>)	Terminal-N	N° de aminoácidos	N° de aristas	Clase de enzima
Q42965	Estabilizador	348	2	Hidrolasa; Liasa
Q42942	Desestabilizador	460	2	Transferasa; Liasa
P10978	Estabilizador	1327	2	Transferasa; Hidrolasa

En estas gráficas (Figura 25, Figura 26, en anexos Figura 63, Figura 64) se evidencia que el 100 % de las enzimas, de los proteomas *Escherichia coli*, *Nicotiana tabacum* y *Oryctolagus cuniculus*, que se clasifican como translocasas (EC:7) tienden a tener solo residuos terminal-N estabilizadores, excepto en el proteoma *Saccharomyces cerevisiae* ya que el 7 % de sus enzimas translocasas tienen un residuo terminal-N desestabilizador.

Por otro lado, se muestra que el tamaño del nodo es directamente proporcional con el número de residuos aminoácidos de la secuencia de la enzima madura. Los resultados muestran claramente que no hay una relación, o tendencia, entre el número de residuos aminoácidos de la secuencia de la enzima con el terminal-N y con el tipo de reacciones que catalizan (Figura 25, Figura 26, en anexos Figura 63, Figura 64).

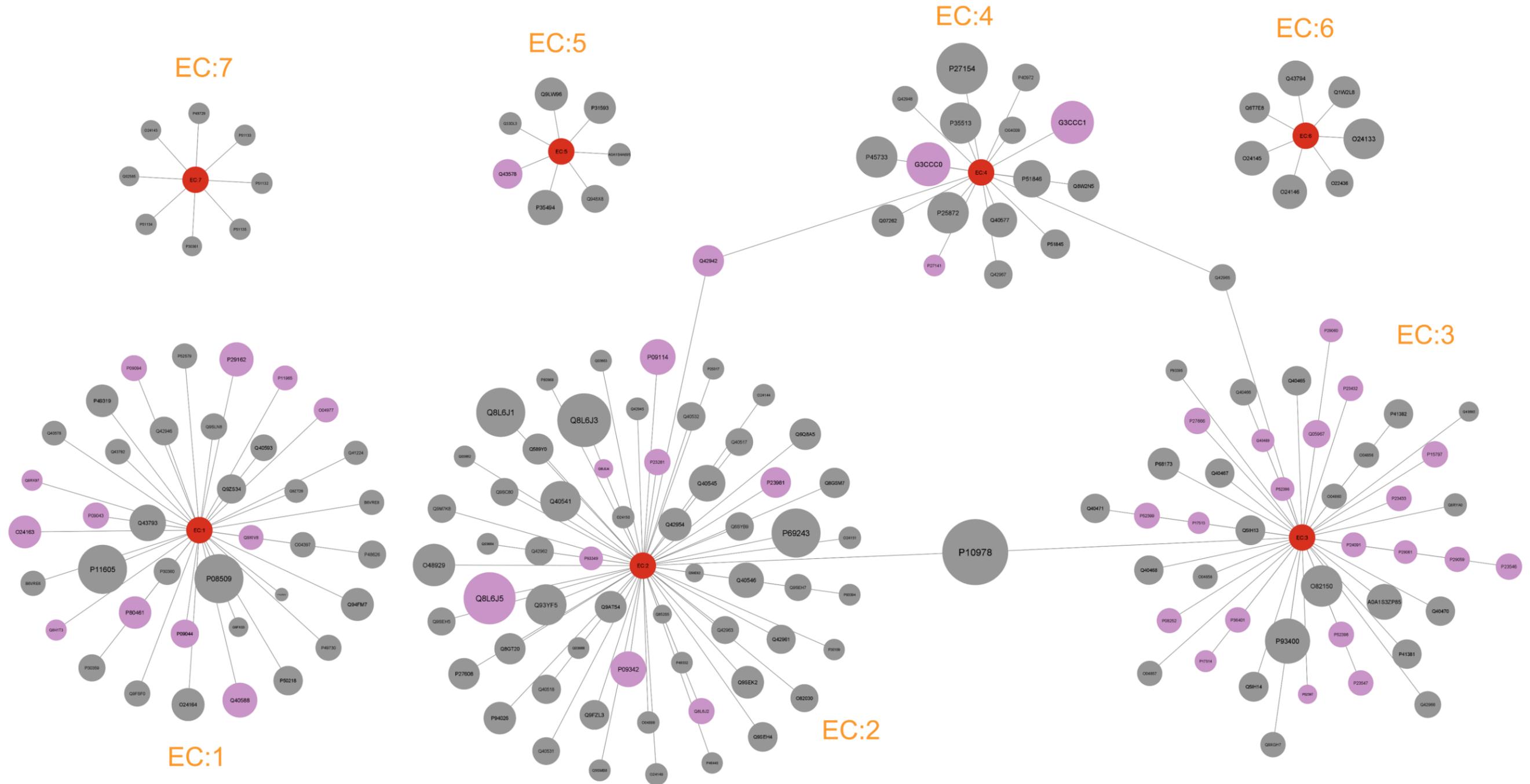


Figura 26. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N destabilizador para *Nicotiana tabacum*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerazas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En esta gráfica, se representa los enzimas con un Terminal-N destabilizador (nodos color rosa) y las enzimas con un Terminal-N estabilizador (nodos color gris). El tamaño del nodo es directamente proporcional al número de residuos aminoácidos de la secuencia de la enzima madura.

7.6 Clasificación de las enzimas de acuerdo con la superfamilia a la que pertenecen.

Se filtraron las enzimas con información sobre la superfamilia en la base de datos Uniprot y se encontró que 3950 enzimas tienen asignada al menos una superfamilia proteica (Figura 27).

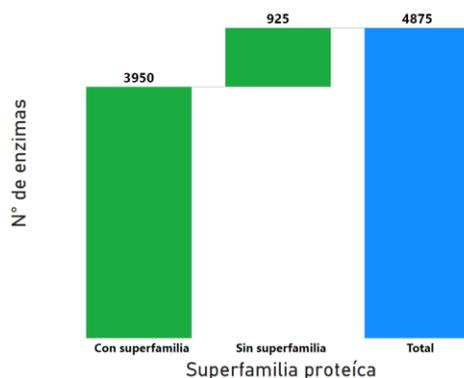


Figura 27. Número de enzimas que tiene asignada una superfamilia. Representación del conjunto de enzimas que tiene asignada al menos una superfamilia y las enzimas que no tiene asignada una superfamilia.

De las 3950 enzimas distribuidas en 541 superfamilias, 2136 son del proteoma *Escherichia coli*, 1421 del proteoma *Saccharomyces cerevisiae*, 235 del proteoma *Oryctolagus cuniculus* y 158 del proteoma *Nicotiana tabacum* (Figura 28).

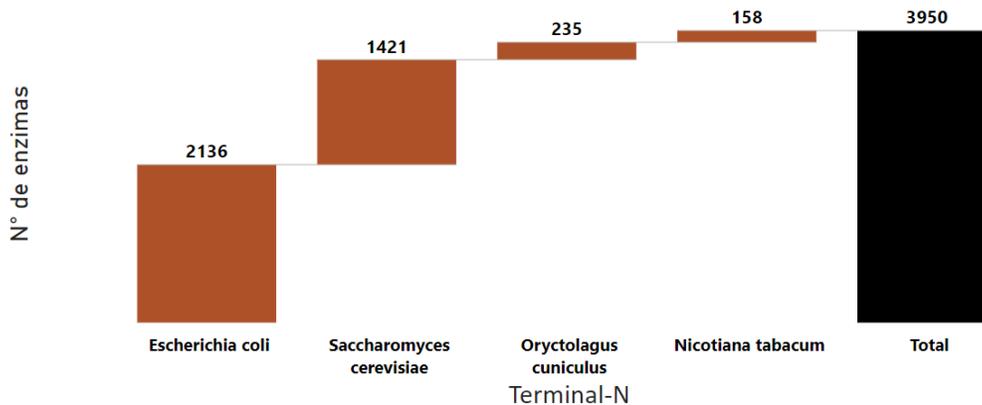


Figura 28. Número de enzimas que tiene asignada una superfamilia por cada proteoma.

De las 541 superfamilias, en las que se distribuyen las enzimas de los cuatro proteomas, (Figura 41) se encontró que 429 superfamilias proteicas están presentes en *Escherichia coli*, 351 en *Saccharomyces cerevisiae*, 119 en *Oryctolagus cuniculus* y 70 en *Nicotiana tabacum* (Figura 28, Figura 29).

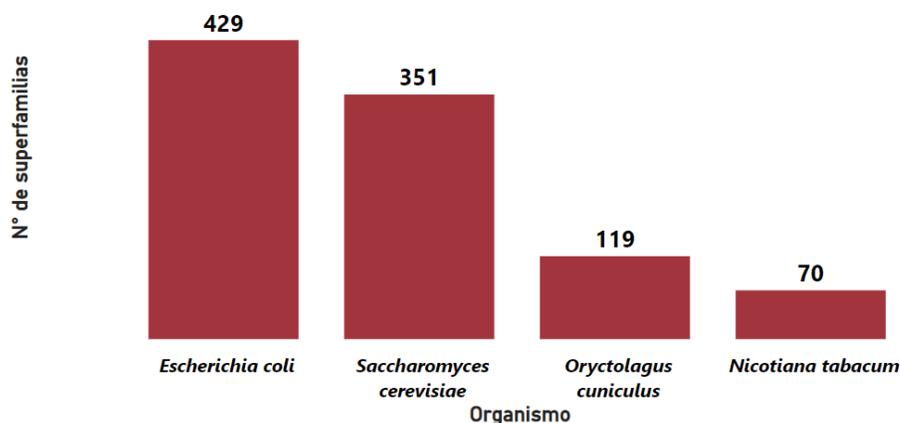


Figura 29. Número de superfamilias proteicas por cada proteoma. Conteo de las superfamilias en las que están asignadas las diferentes enzimas de cada proteoma.

La sumatoria del conteo de superfamilias de la Figura 29 de cada proteoma es de 899 (obviamente mayor a las 541 superfamilias), lo que indica que hay superfamilias en común entre los cuatro proteomas. Para identificar cuáles son las superfamilias que coinciden en más de un proteoma se construyó una red de interacción con el conjunto de superfamilias por proteoma.

Como resultado se muestra la Figura 30, donde se visualiza cuatro agrupamientos, uno por cada proteoma de estudio. Los resultados revelan que hay enzimas de los diferentes proteomas que tienen en común una misma superfamilia proteica. Además, se observa que los dos proteomas que más superfamilias tienen en común son los proteomas *Escherichia coli* y *Saccharomyces cerevisiae*, ya que coinciden en 262 superfamilias, lo cual se puede visualizar mejor en las gráficas de los anexos (Figura 65 a la Figura 70). Siguiendo el mismo

análisis para los demás proteomas, en la Tabla 4, se observa que los proteomas *Nicotiana tabacum* y *Oryctolagus cuniculus* son los que menor número de superfamilias tienen en común, con 29 superfamilias proteicas.

Tabla 4. Numero de superfamilias proteicas que tienen en común los proteomas de estudio.

Proteomas	<i>Escherichia coli</i>	<i>Saccharomyces cerevisiae</i>	<i>Oryctolagus cuniculus</i>
<i>Escherichia coli</i>	---	---	---
<i>Saccharomyces cerevisiae</i>	262	--	--
<i>Oryctolagus cuniculus</i>	79	88	--
<i>Nicotiana tabacum</i>	60	62	29

A pesar de que entre los proteomas *Escherichia coli* y *Saccharomyces cerevisiae* tienen el mayor número de superfamilias en común (Figura 29, Tabla 4), no tienen el mayor porcentaje de coincidencia (Tabla 5). Teniendo en cuenta que la cantidad de superfamilias que se agrupan en el proteoma *Escherichia coli* y *Saccharomyces cerevisiae* son de 429 y 351 respectivamente se concluye que el proteoma *Escherichia coli* tiene una coincidencia de aproximadamente del 74% sobre el proteoma *Saccharomyces cerevisiae* y que el proteoma *Saccharomyces cerevisiae* tiene un 61% de coincidencia sobre el proteoma *Escherichia coli*. Sin embargo, se observa que el mayor porcentaje de coincidencia se da entre los proteomas *Saccharomyces cerevisiae* y *Nicotiana tabacum*, donde el proteoma *Saccharomyces cerevisiae* tiene un 87% de coincidencia sobre el proteoma *Nicotiana tabacum*, y el menor porcentaje de coincidencia es del 13% del proteoma *Nicotiana tabacum* sobre el proteoma *Escherichia coli*.

Tabla 5. Porcentaje de coincidencia de superfamilias en los proteomas de estudio*.

Proteomas	<i>Escherichia coli</i>	<i>Saccharomyces cerevisiae</i>	<i>Oryctolagus cuniculus</i>	<i>Nicotiana tabacum</i>
<i>Escherichia coli</i>	-----	74%	66%	86%
<i>Saccharomyces cerevisiae</i>	61%	-----	73%	87%
<i>Oryctolagus cuniculus</i>	18%	25%	-----	41%
<i>Nicotiana tabacum</i>	13%	18%	24%	-----

*En la parte superior de las líneas puntuadas que divide la tabla se muestra el porcentaje de coincidencia del proteoma con mayor número de superfamilias sobre el proteoma de menor número de superfamilias, y en la parte inferior el porcentaje de coincidencia del proteoma con menor número de superfamilias sobre el proteoma de mayor número de superfamilias.

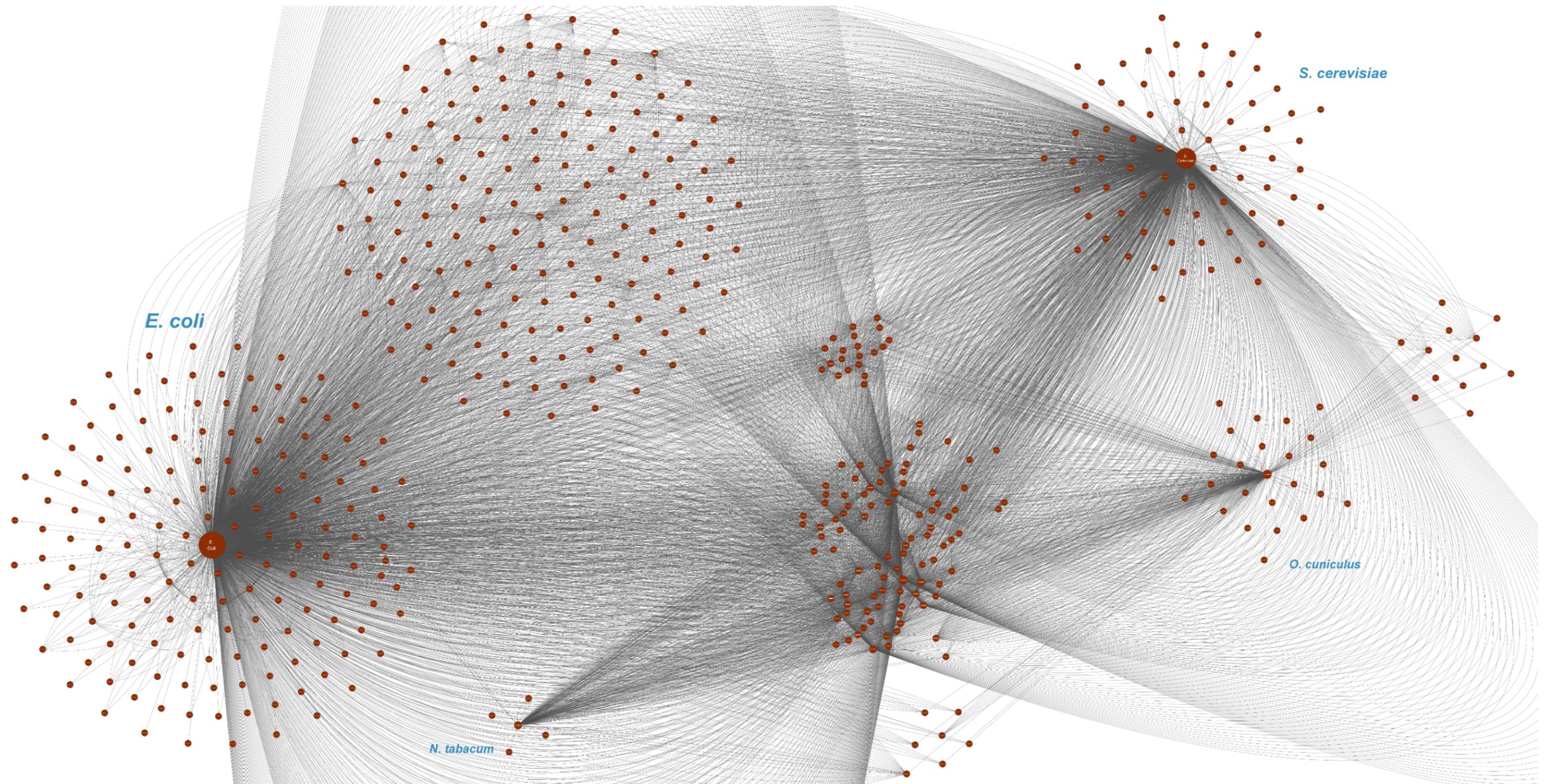


Figura 30. Red de interacción de las superfamilias con los cuatro proteomas. Se visualiza como interaccionan las superfamilias (o se conectan) con los proteomas de los diferentes proteomas.

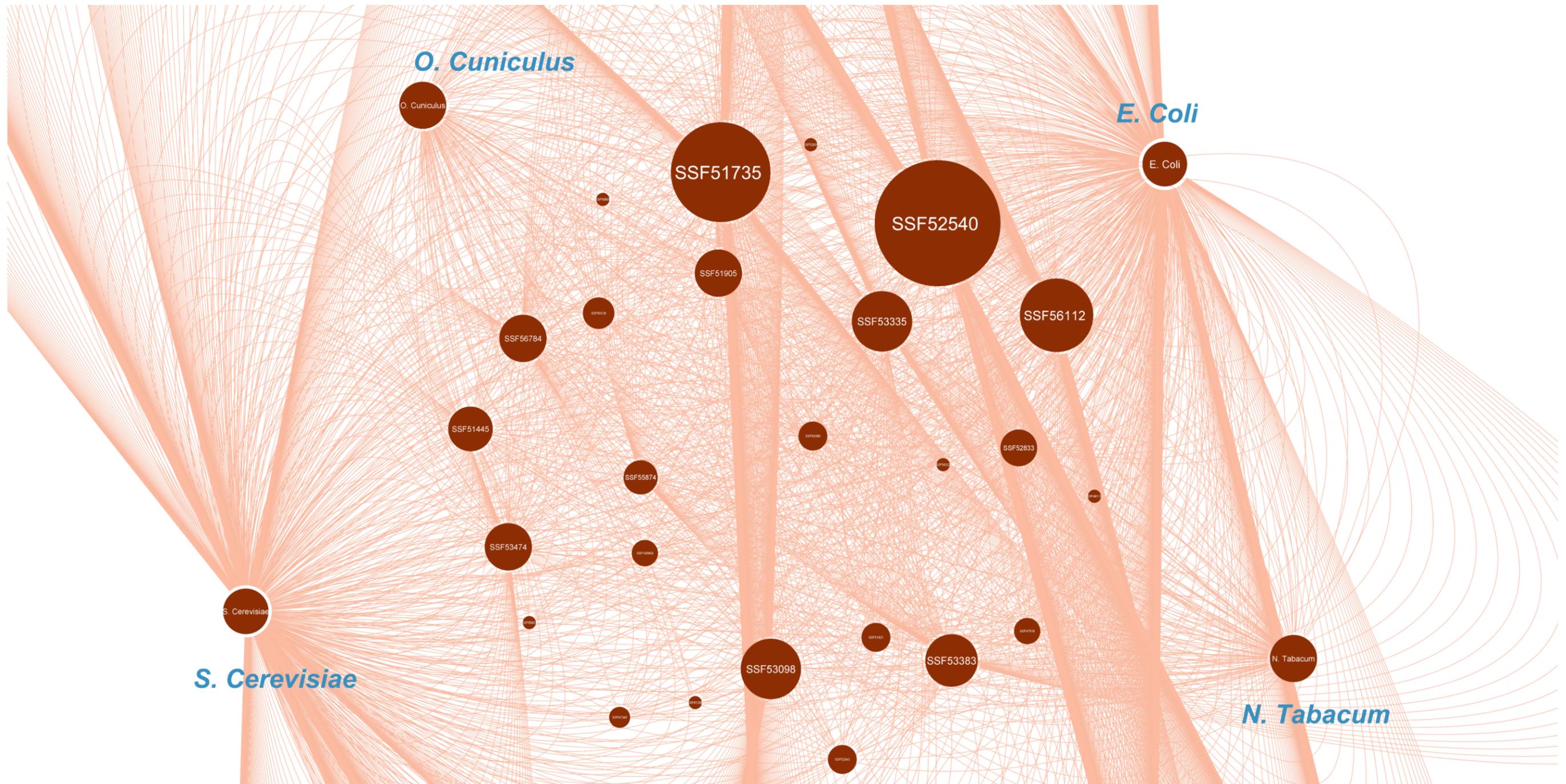


Figura 31. Superfamilias comunes con los cuatro proteomas. El tamaño del nodo de las superfamilias es directamente proporcional con el número de enzimas asignas a esa superfamilia.

Continuando con el mismo análisis, se graficaron las superfamilias comunes en los cuatro proteomas. La Figura 31 muestra como resultado 25 superfamilias en común entre los cuatro proteomas. Indicando que cerca del 95% de las superfamilias proteicas no se comparten o no son comunes entre los cuatro proteomas.

Ahora, para entender la distribución de las superfamilias en las actividades enzimáticas en los proteomas se realizó el conteo de las superfamilias presentes en cada clasificación EC. En la Figura 32 se encontró que la clase de enzimas con mayor número de superfamilia son las transferasas (EC:2), seguida de las hidrolasas (EC:3) y oxidorreductasa (EC:1) con 225, 224 y 120 superfamilias respectivamente. Para evidenciar si cada proteoma sigue esta misma tendencia, se realizó la gráfica de las agrupaciones de las superfamilias por cada clase de enzima, esta vez con distinción por proteoma (Figura 33).

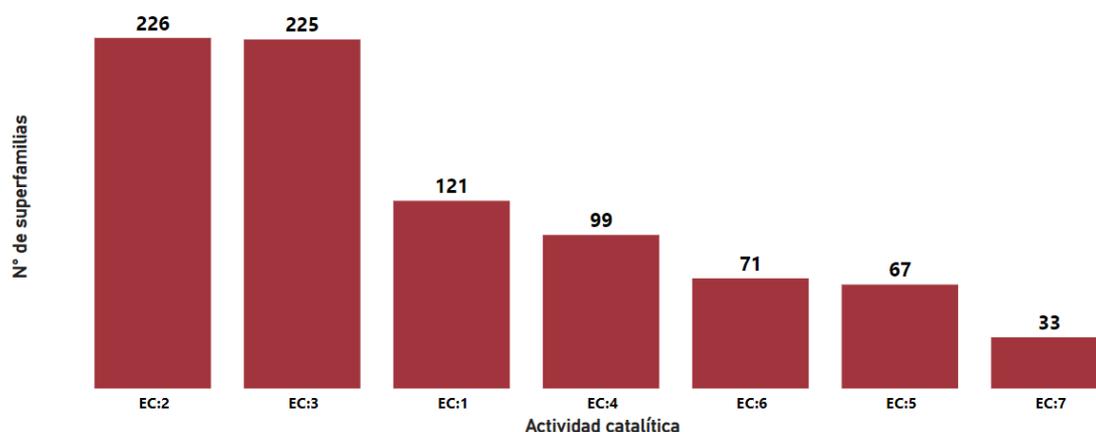


Figura 32. Conteo de superfamilias agrupadas por cada clase de enzima (EC). Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En la gráfica se observa la agrupación de superfamilias por clase de enzima sin distinción del proteoma.

En la Figura 33, gráfica (A), del proteoma *Escherichia coli* se encontró que las enzimas agrupadas como transferasas (EC:2) son las que mayor número de superfamilias tienen asignadas, con 174, seguida por las hidrolasas (EC:3) y oxidorreductasas (EC:1) con 149 y 92 superfamilias respectivamente. Para el proteoma *Nicotiana tabacum*, gráfica (B), la clase

de enzima que agrupa mayor número de superfamilias son las transferasas (EC:2), seguida por las oxidorreductasas (EC:1), hidrolasas (EC:3) y liasas (EC:4) con 28, 20, 16 y 16 superfamilias respectivamente. En la gráfica (C) de la misma figura, del proteoma *Oryctolagus cuniculus*, se halló que las enzimas agrupadas como hidrolasas (EC:3) son las que tienen asignado un mayor número de superfamilias, seguida de las oxidorreductasas (EC:1) y transferasas (EC:2), con 41, 35 y 32 superfamilias, respectivamente. Por último, para el proteoma *Saccharomyces cerevisiae*, gráfica (D), se evidencia que el mayor número de superfamilia se agrupa en la clase enzimática de las hidrolasas (EC:3), seguidas por las transferasas (EC:2) y oxidorreductasas (EC:1) con 149, 132 y 66 superfamilias respectivamente.

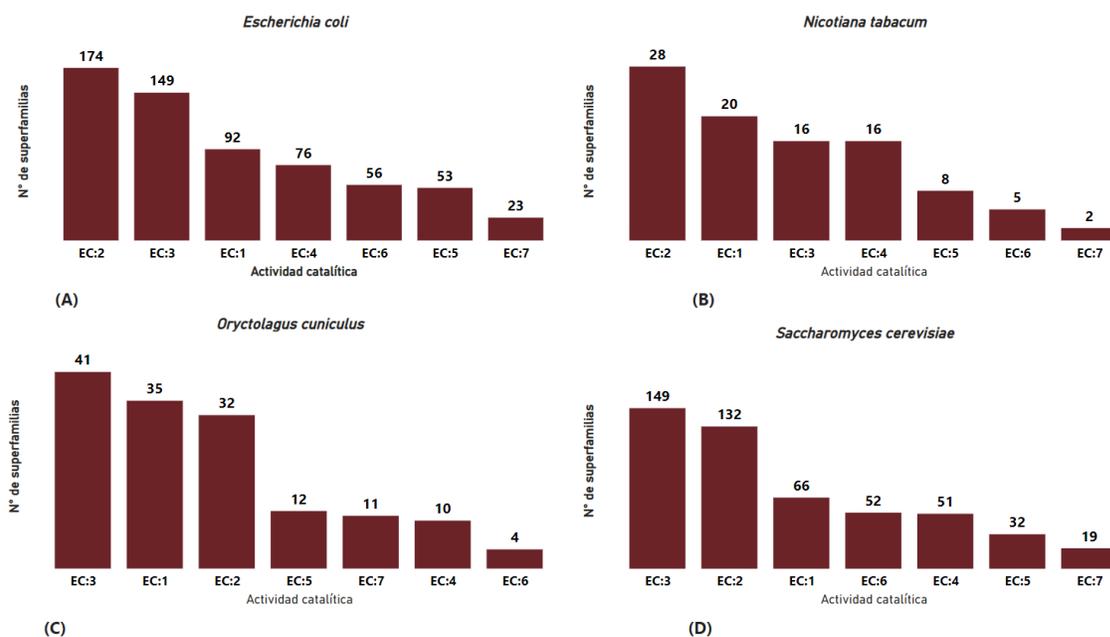


Figura 33. Conteo de superfamilias asignadas por cada clase de enzima (EC). En la gráfica se observa la agrupación de superfamilias por clase de enzima con distinción del proteoma. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7).

De lo escrito anteriormente se analiza que los datos de la gráfica (A) (Figura 33), del proteoma *Escherichia coli*, sigue exactamente la tendencia descrita para la Figura 32, donde se muestra que las clases de enzimas con mayor número de superfamilia son las transferasas, seguida de la hidrolasas y oxidorreductasa. Sin embargo, los datos de los demás proteomas

no siguen dicha tendencia, pero es evidente que coinciden en las tres clases de enzimas (EC) con mayor número de superfamilia son las clases de enzimas oxidorreductasas (EC:1), transferasas (EC:2) e hidrolasas (EC:3).

Hasta aquí, en esta sección, se ha realizado un análisis general del comportamiento de las superfamilias proteicas. A continuación, se presenta el conteo de enzimas asignadas a cada una de las superfamilias con la finalidad de saber cuál es la superfamilia que tiene mayor número de enzimas asignadas. Los resultados obtenidos se representan por medio de gráficas tipo anillo (Figura 34).

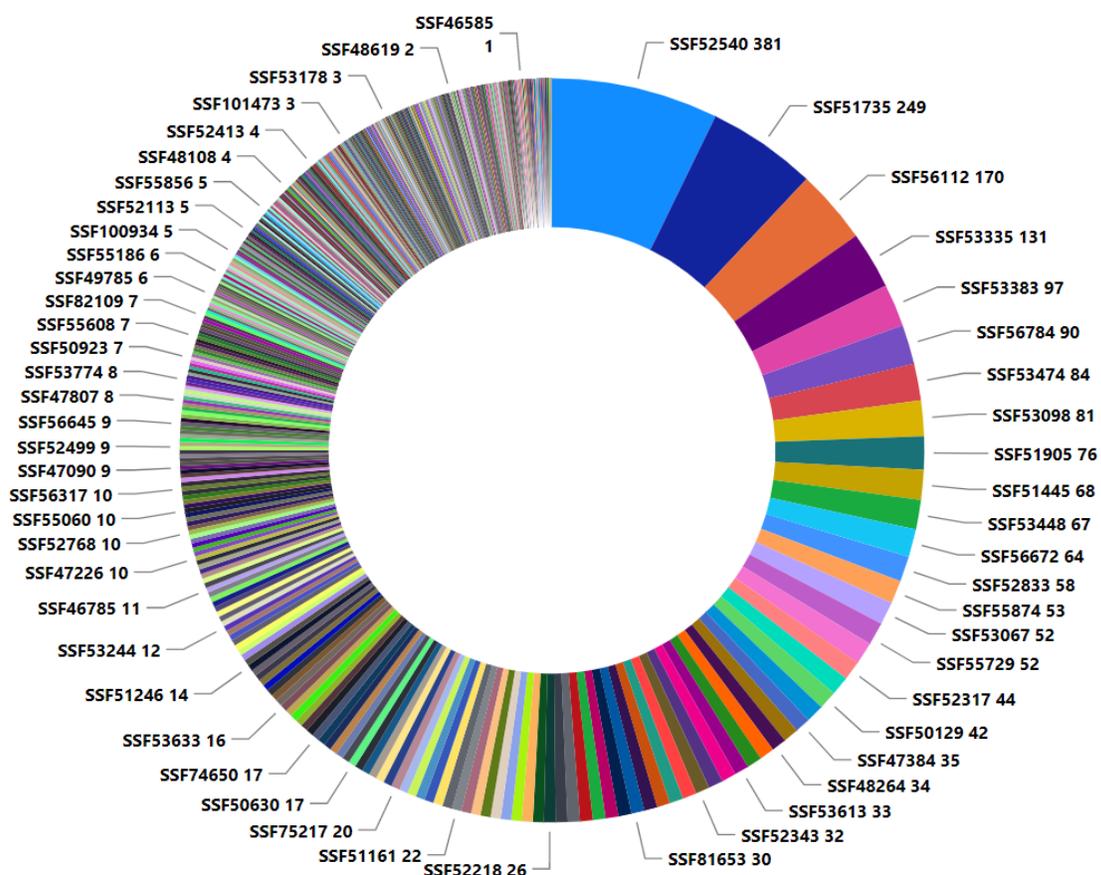


Figura 34. Conteo de enzimas asignadas a cada superfamilia proteica. En la gráfica se muestra el número de enzimas por cada superfamilia.

De la Figura 34, se puede establecer que las superfamilias con mayor número de enzimas asignadas son: SSF52540 (P-loop que contiene nucleósidos trifosfato hidrolasas), con 381 enzimas, seguidas de la SSF51735 (con 294 enzimas), la SSF56112 (con 170 enzimas) y dentro de la superfamilia con una única enzima se encuentra la SSF81496 (Subunidad del citocromo c1 del complejo del citocromo bc1 (ubiquinol-citocromo c reductasa), anclaje transmembrana), SSF81648, SSF82215. Para identificar si estas tendencias aplican individualmente para cada proteoma, se realizaron cuatro gráficas tipo anillo una por cada proteoma (Figura 35 a la Figura 38).

Los resultados obtenidos en el proteoma *Escherichia coli* (Figura 35) demuestran que la superfamilia que tiene mayor número de enzimas asignadas es SSF52540 (P-loop que contiene nucleósidos trifosfato hidrolasas), con 198 enzimas, seguida de SSF51735 (Dominios de plegado de Rossmann de unión a NAD(P), SSF53335. Mientras que dentro las superfamilias con una única enzima se encuentran la SSF48173 (Dominio de enlace FAD criptocromo/fotoliasa), SSF102220 (Subunidad psi de la ADN polimerasa III), SSF103515 (Autotransportador).

Por otra parte, en el proteoma *Nicotiana tabacum* (Figura 36), se revela que la superfamilia que tiene mayor número de enzimas asignadas es SSF51445, SSF51735, SSF52540 ((Trans)glucosidasas, Dominios de plegado de Rossmann de unión a NAD(P), P-loop que contiene nucleósidos trifosfato hidrolasas, respectivamente), con 13 enzimas, seguida de SSF53335 (Metiltransferasas dependientes de S-adenosil-L-metionina), con 10 enzimas, SSF52833 (similar a la Tiorredoxina), con 9 enzimas, etc. Mientras que dentro las superfamilias con una única enzima se encuentran la SSF100950 (NagB/RpiA/CoA transferasa-like), SSF53056 (beta-anhidrasa carbónica, cabina), SSF56634 (Catalasa dependiente del hemo).

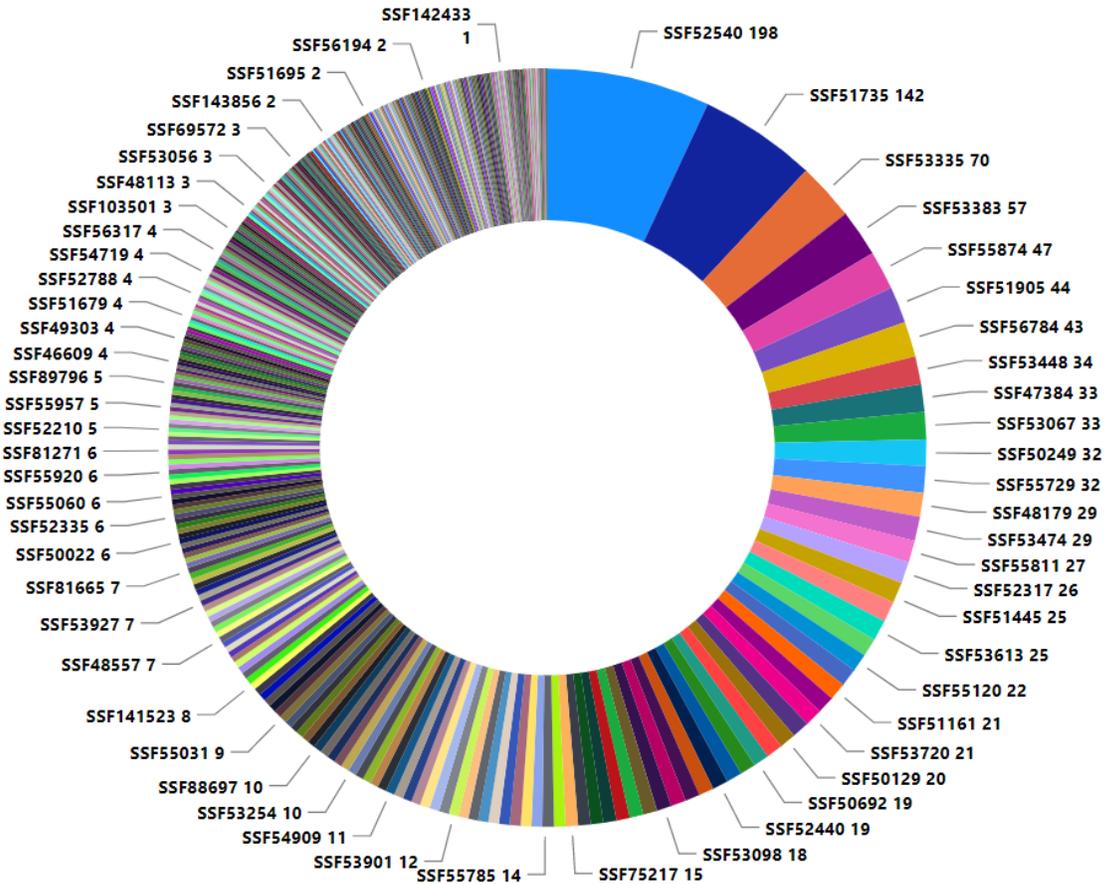


Figura 35. Conteo de enzimas asignadas a cada superfamilia en *Escherichia coli*. En la gráfica se muestra el número de enzimas por cada superfamilia.

Además, en el proteoma *Oryctolagus cuniculus* (Figura 37) se contempla que la superfamilia que tiene mayor número de enzimas asignadas es SSF48264 (Citocromo P450), con 29 enzimas, seguida de SSF56112 (Proteína similar a la quinasa (pk-like)), con 25 enzimas, SSF51735 (Dominios de plegado de Rossmann de unión a NAD(P)), con 17 enzimas. Mientras que dentro las superfamilias con una única enzima se encuentran la SSF101690 (Dominio PAZ), SSF48726 (Inmunoglobulina), SSF56542 (Dominio de unión al sustrato de la HMG-CoA reductasa).

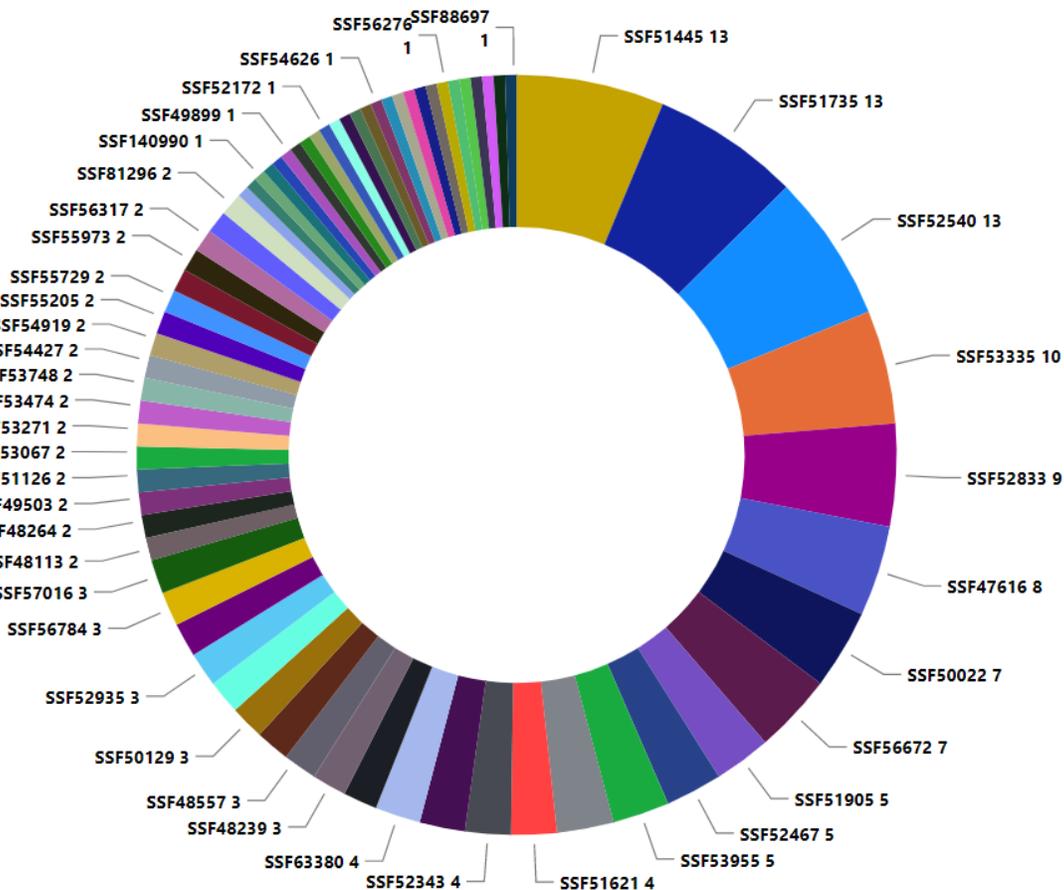


Figura 36. Conteo de enzimas asignadas a cada superfamilia en *Nicotiana tabacum*. En la gráfica se muestra el número de enzimas por cada superfamilia.

Asimismo, para el proteoma *Saccharomyces cerevisiae* (Figura 38) la superfamilia que tiene mayor número de enzimas asignadas es SSF52540 (P-loop que contiene nucleósidos trifosfato hidrolasas), con 156 enzimas, seguida de SSF56112 (Proteína similar a la quinasa (pk-like)), con 128 enzimas, SSF51735 (Dominios de plegado de Rossmann de unión a NAD(P)), con 77 enzimas. Mientras que dentro las superfamilias con una única enzima se encuentran la SSF101353 (Dominio putativo de unión a anticodones de la alanil-ARNt sintetasa (AlaRS)), SSF47384 (Dominio homodimérico de la señal transductora de histidina quinasa), SSF52121 (Lumazina sintasa).

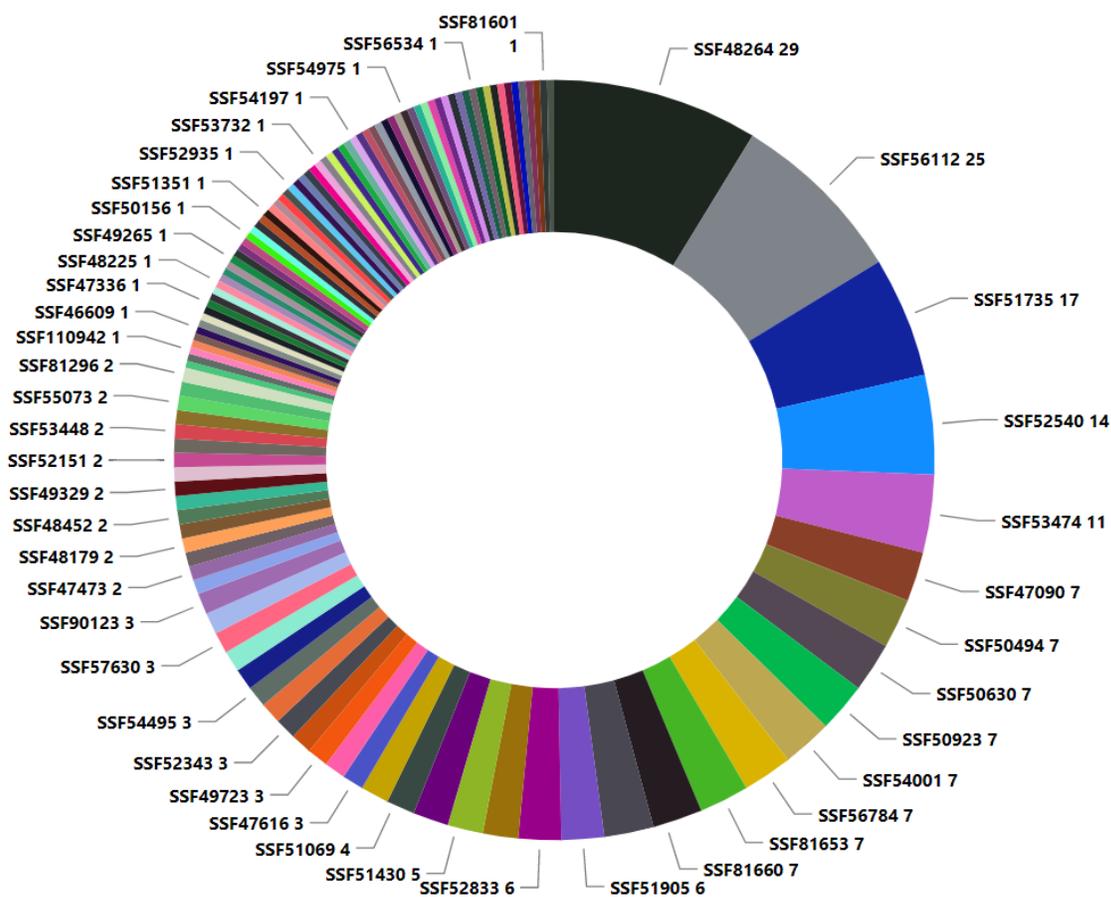


Figura 37. Conteo de enzimas asignadas a cada superfamilia en *Oryctolagus cuniculus*. En la gráfica se muestra el número de enzimas por cada superfamilia.

Las tres superfamilias grandes son SSF52540 (P-loop que contiene nucleósidos trifosfato hidrolasas), SSF1735 (Dominios de plegado de Rossmann de unión a NAD(P)) y SSF56112 (Proteína similar a la quinasa (*PK-like*)), (Figura 35 a la Figura 38 y Tabla 6).

Tabla 6. Superfamilias con mayor población de enzimas en los cuatro proteomas de estudio.

superfamilia	<i>Escherichia coli</i>	<i>Saccharomyces cerevisiae</i>	<i>Oryctolagus cuniculus</i>	<i>Nicotiana tabacum</i>	Total
SSF52540	198	156	14	13	381
SSF51735	142	77	17	13	249
SSF56112	12	128	25	5	170

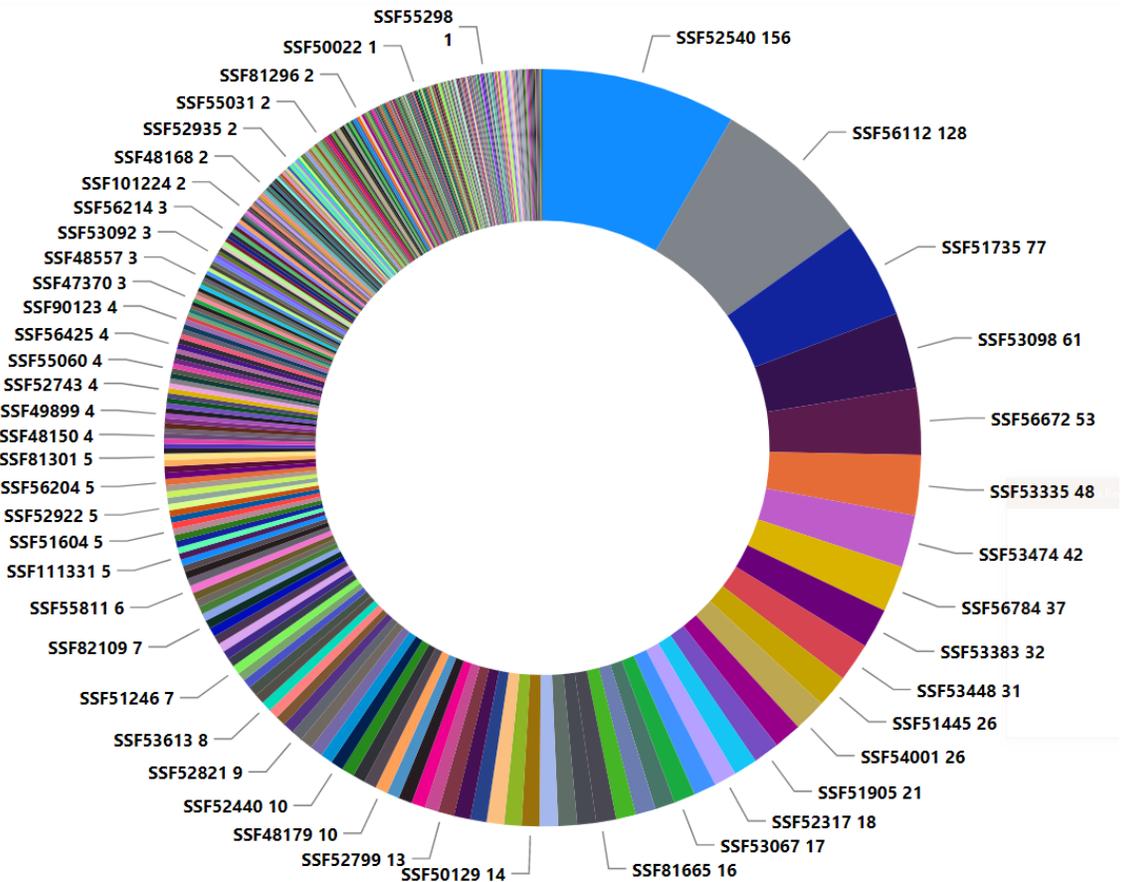


Figura 38. Conteo de enzimas asignadas a cada superfamilia en *Saccharomyces cerevisiae*. En la gráfica se muestra el número de enzimas por cada superfamilia.

Para estudiar la distribución de las enzimas con terminal-N estabilizador y desestabilizador en cada superfamilia, se filtraron aquellas enzimas que tiene asignado un residuo aminoácido terminal-N (estabilizador o desestabilizador) y luego se contabilizaron las enzimas con tiempos de vida estable y con tiempos de vida inestable. Los resultados se representaron por medio de gráficas de barras (Figura 39).

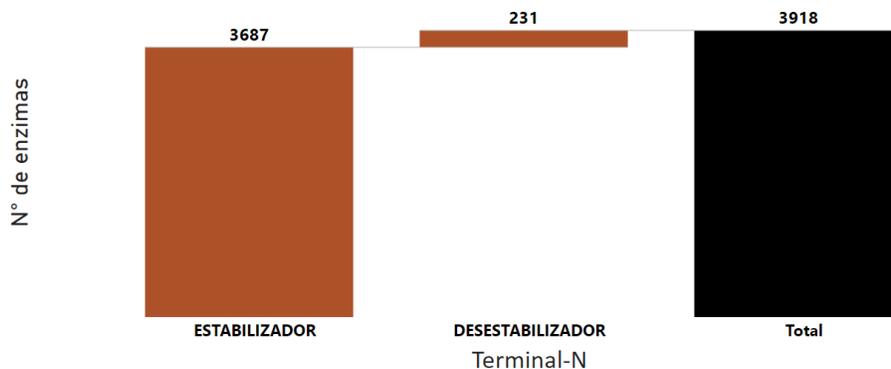


Figura 39. Conteo de enzimas de acuerdo con el terminal-N. De las 3950 enzimas que tiene asignada una superfamilia, solo 3918 está caracterizada como una enzima con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 3918 enzimas, 3687 se clasifican como enzimas con tiempos de vida estable (debido a que tienen un terminal-N estabilizador) y 231 como enzimas con tiempos de vida inestable (debido a que tienen un terminal-N desestabilizador).

De los resultados obtenidos, se observó que de las 3950 enzimas reportadas (en Uniprot) que tiene asignada alguna superfamilia (Figura 27), solo 3918 están caracterizadas, ya sea, como una enzima con un tiempo de vida estable (3687 enzimas) o como enzimas con un tiempo de vida inestable (231 enzimas) (Figura 39), dependiendo del residuo aminoácido terminal-N que posee (según la Tabla 1).

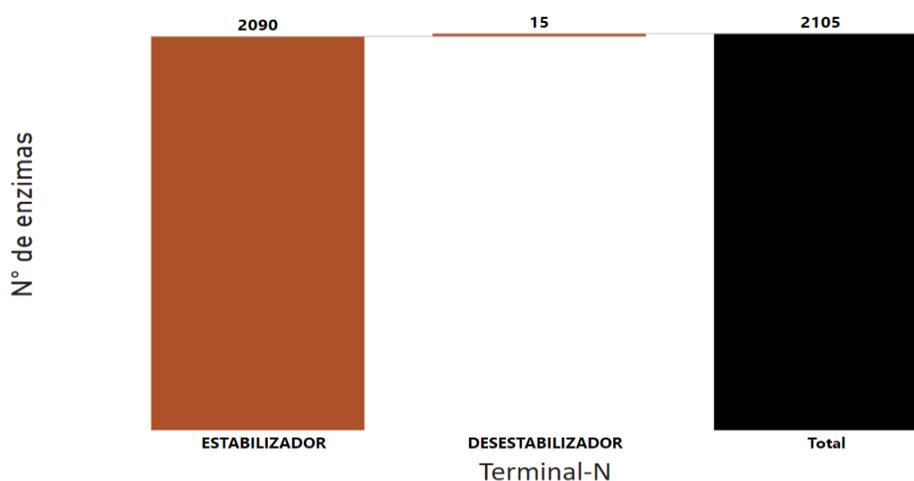


Figura 40. Conteo de enzimas de acuerdo con el terminal-N en *Escherichia coli*. De las 2136 enzimas que tiene asignada una superfamilia, solo 2105 están caracterizadas como enzimas con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 2105 enzimas, 2090 se clasifican como enzimas con tiempos de vida estable y 15 como enzimas con tiempos de vida inestable.

De las 3687 enzimas que están caracterizadas con tiempos de vida estable, 2090 son del proteoma del organismo *Escherichia coli*, 116 son del proteoma del organismo *Nicotiana tabacum*, 142 son del proteoma del organismo *Oryctolagus cuniculus* y 1339 son del proteoma del organismo *Saccharomyces cerevisiae*. Asimismo, de las 231 enzimas que están caracterizadas con tiempos de vida inestable, 15 corresponden al proteoma del organismo *Escherichia coli*, 41 corresponden al proteoma del organismo *Nicotiana tabacum*, 93 corresponden al proteoma del organismo *Oryctolagus cuniculus* y 82 corresponden al proteoma del organismo *Saccharomyces cerevisiae* (Figura 40, en anexos de la Figura 72 a la Figura 74).

Ahora, para identificar qué cantidad superfamilia tiene asignado enzimas de tiempos de vida estable y de enzimas de tiempo de vida inestable se realizó una red de interacción (Figura 41). De los resultados obtenidos, se ve cuáles son las superfamilias que tienen asignada solo enzimas con algún residuo aminoácidos terminal-N estabilizador, cuales las que tienen asignado solo enzimas con algún residuo aminoácidos terminal-N desestabilizadores y cuales superfamilias tiene asignado enzimas con las dos clases de residuos aminoácidos terminal-N (estabilizadores y desestabilizadores).

En la Figura 41 se puede distinguir que hay una mayor cantidad de superfamilias que tienen solamente enzimas con residuos aminoácidos terminal-N estabilizador (nodos color morado) que superfamilias con enzimas con residuos aminoácidos terminal-N desestabilizador (nodos color azul). Mas específicamente, de las 541 superfamilias, 12 agrupan enzimas con tiempo de vida inestable, 107 agrupan enzimas con tiempos de vida estable y con tiempos de vida inestable y, por último, 422 agrupan enzimas con tiempo de vida estable (en anexos Tabla 9).

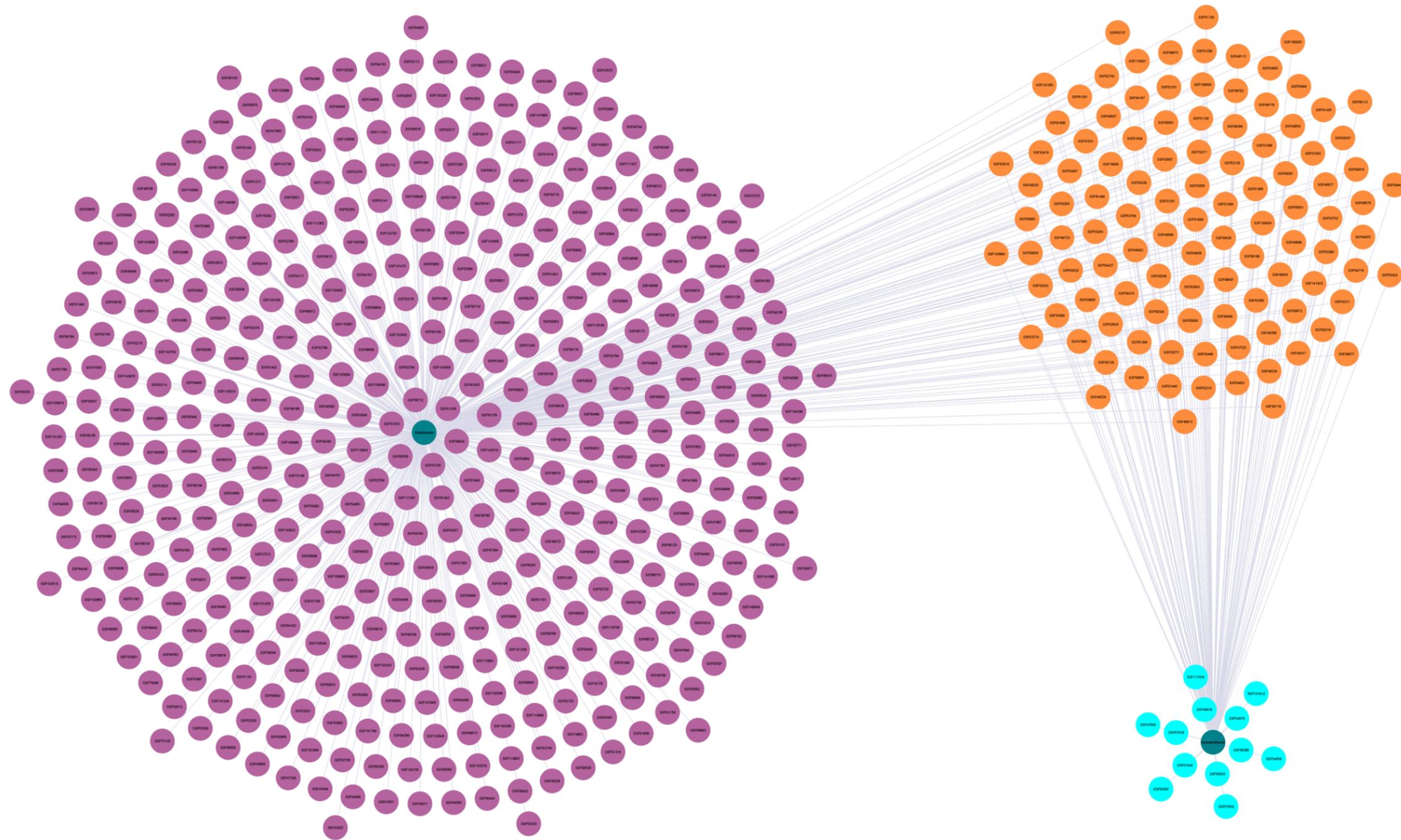


Figura 41. Superfamilias que tienen asignada enzimas con terminal-N estabilizador y desestabilizador. En la gráfica muestra el agrupamiento de las superfamilias que tienen enzimas con terminal-N estabilizador y terminal-N desestabilizador. Los nodos color naranja representa las superfamilias que tienen enzimas con terminal-N estabilizador y desestabilizador, los nodos color morado representa las superfamilias con solo enzimas que tienen un terminal-N estabilizador y Los nodos color azul representa las superfamilias con solo enzimas que tienen un terminal-N desestabilizador (Tabla 9).

Para ver el conteo de enzimas con tiempos de vida estable e inestable asignado a cada una de las superfamilias se muestran cinco figuras, una figura donde se muestra un conteo sin hacer distinción del proteoma y una figura para cada proteoma (Figura 42, Figura 43, en anexos de la Figura 75 a la Figura 77).

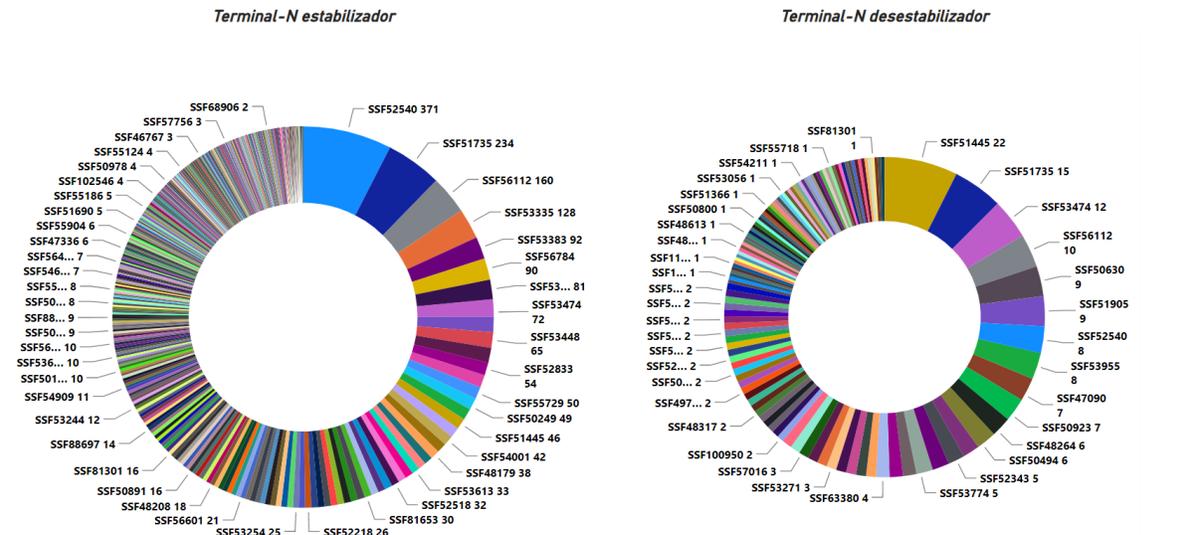


Figura 42. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia. A la izquierda se ilustra el número de enzimas, por cada superfamilia, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada superfamilia, con un terminal-N desestabilizador.

En la Figura 42 se encuentran ilustrado el conteo de las enzimas con terminal-N estabilizador (ilustración de la izquierda) y terminal-N desestabilizador (ilustración de la derecha) para cada superfamilia. En esta gráfica se muestra que las tres superfamilias que tiene asignada un mayor número de enzimas con tiempo de vida estable son las SSF52540, SSF51735 y SSF56112 con 371, 234 y 160 enzimas, respectivamente. Mientras que las tres superfamilias con las enzimas con tiempos de vida inestables son las SSF51445, SSF51735 y SSF53474 con 22, 15 y 12 enzimas, respectivamente.

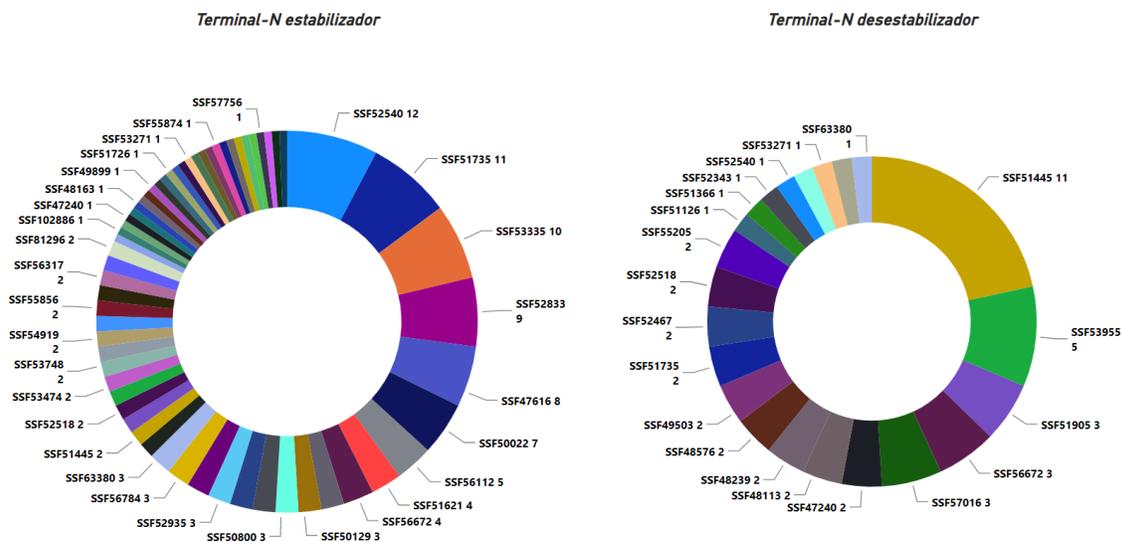


Figura 43. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para *Nicotiana tabacum*. A la izquierda se ilustra el número de enzimas, por cada superfamilia, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada superfamilia, con un terminal-N desestabilizador.

Ahora bien, analizando los datos individualmente por cada uno de los cuatro proteomas se observa que el proteoma *Escherichia coli* (en anexos Figura 75) tiene la superfamilia (SSF52540) con mayor número asignado de enzimas con tiempos de vida estable, 196 enzimas; seguido del proteoma *Saccharomyces cerevisiae*, (en anexos Figura 77) para la misma superfamilia (SSF52540) con 151 enzimas; en el proteoma *Oryctolagus cuniculus* (en anexos Figura 76), con la superfamilia SSF48264 con 23 enzimas; y el proteoma *Nicotiana tabacum* (Figura 43) con 12 enzimas en la superfamilia SSF52540. Mientras que la superfamilia con mayor número asignado de enzimas con tiempo de vida inestable es la SSF51445 (con 11 enzimas) en el proteoma *Nicotiana tabacum*, seguido de la SSF56112, SSF51445 y SSF48317 en los proteomas *Oryctolagus cuniculus*, *Saccharomyces cerevisiae* y *Escherichia coli*, respectivamente (Figura 43, en anexos de la Figura 75 a la Figura 77).

7.7 Clasificación de las enzimas de acuerdo con la ruta metabólica que catalizan

En la base de datos Uniprot, de las 4875 enzimas con estatus de revisadas, de los cuatro proteomas de estudio, se reporta 1730 enzimas que participan en al menos una ruta metabólica (Figura 44). De los cuatro proteomas, *Escherichia coli* reporta un mayor número de enzimas que participa en rutas metabólicas, 1092, seguido por *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus* con 507, 73 y 58 respectivamente.

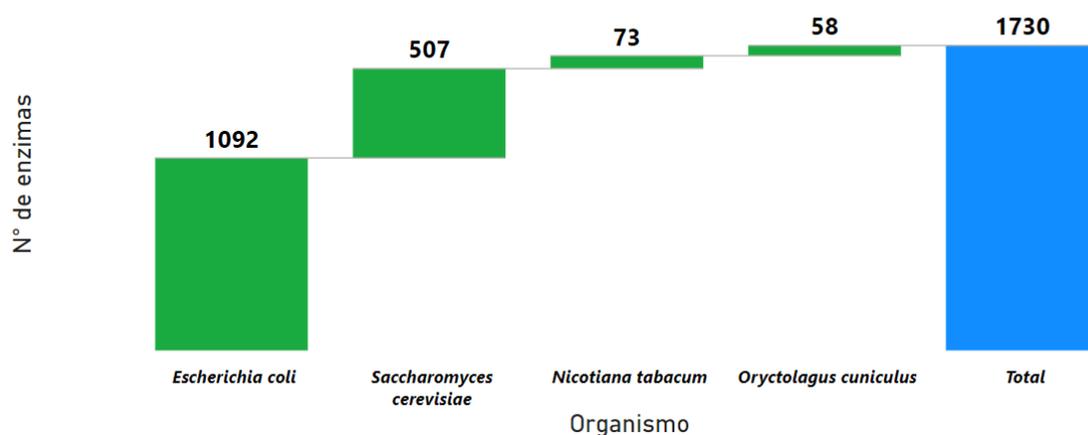


Figura 44. Conteo de enzimas asociadas a rutas metabólicas. Representación de la cantidad de enzimas, por proteomas, a las cuales se le reporta al menos una ruta metabólica en la base de datos de Uniprot.

El número de rutas metabólicas en las cuales participan las 1730 enzimas, de los cuatro proteomas, es de 70. En la Figura 45 se muestra que *Escherichia coli* tiene el mayor número de rutas metabólicas (53 de las 70), seguido por *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus* con 42, 28 y 22 respectivamente. Al realizar la sumatoria de las rutas metabólicas de cada uno de los proteomas se obtiene un total de 145, esto indica que hay rutas metabólicas que están asociadas a más de un proteoma.

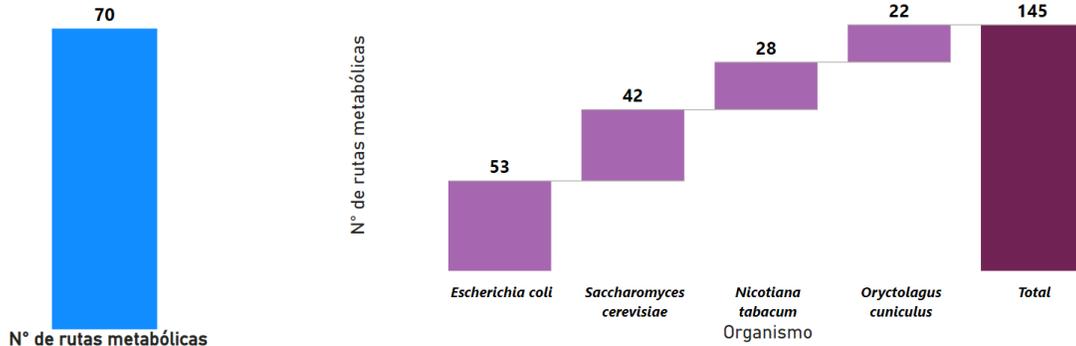


Figura 45. Conteo de rutas metabólicas. A la izquierda se presenta el conteo de rutas metabólicas, sin distinción del proteoma, en los cuales se agrupan las 1730 enzimas. A la derecha se presenta el conteo de rutas metabólicas reportadas para cada proteoma.

Para identificar las rutas metabólicas que son comunes en los proteomas (*Escherichia coli*, *Nicotiana tabacum*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus*) se realizó una red de interacción. De estas interacciones se contempla que hay rutas metabólicas presentes en los cuatro proteomas, en tres, en dos y rutas metabólicas que están presentes en un proteoma (de la Figura 46 a la Figura 48, en anexos Figura 78, Figura 79).

De estas gráficas se observa que, de las 70 rutas metabólicas, 10 se encuentran presente en los cuatro proteomas (Figura 47). Esta es una tendencia muy clara que indica que cada uno de estos proteomas tiene enzimas especializadas en catalizar reacciones químicas que tienen en el mismo objetivo final, por ejemplo, el metabolismo energético, el cual consiste en producir energía (como objetivo final) a partir de un sustrato para ser utilizada por la célula del proteoma (Wouters & Baarends, 2006).

Los nodos naranjas (o rutas metabólicas) tienen una sola arista debido a que son rutas metabólicas que están presentes solo en un proteoma (Figura 48). Se encontró que el proteoma *Escherichia coli* tiene mayor número de vías metabólicas únicas (o propias), con 17, seguido de *Nicotiana tabacum*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus* con 8, 6 y 2 respectivamente, para un total de 33 rutas metabólicas. Como ejemplo, se observa que la vía metabólica de la fotosíntesis es una vía metabólica única para el proteoma *Nicotiana tabacum*.

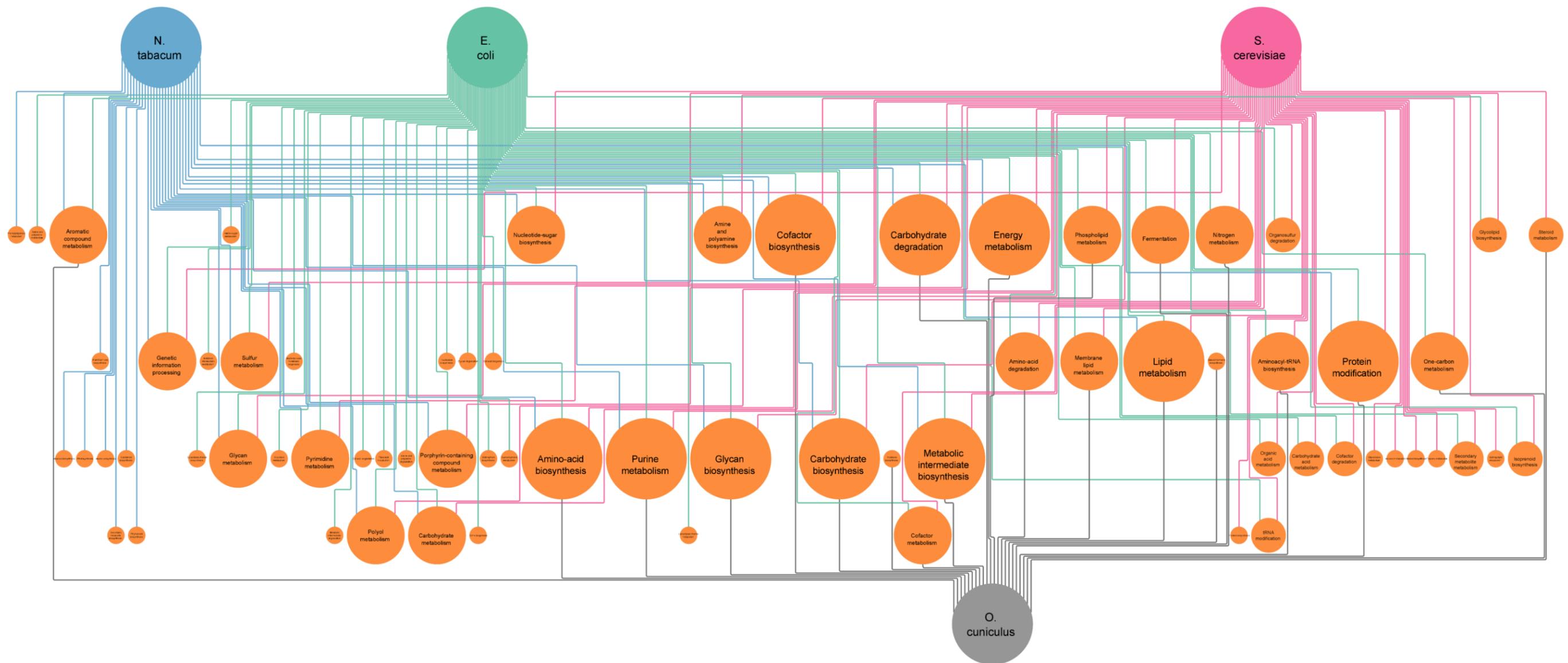


Figura 46. Red de interacción entre los proteomas y sus diferentes rutas metabólicas. Los nodos color naranja representa las rutas metabólicas. El tamaño del nodo naranja es directamente proporcional con el número de proteomas que coincide. Los nodos color gris, rosa, verde y azul representa los proteomas. Las aristas, que coincide con el color de los nodos del proteoma, indica la interacción de la vía metabólica con dicho nodo.

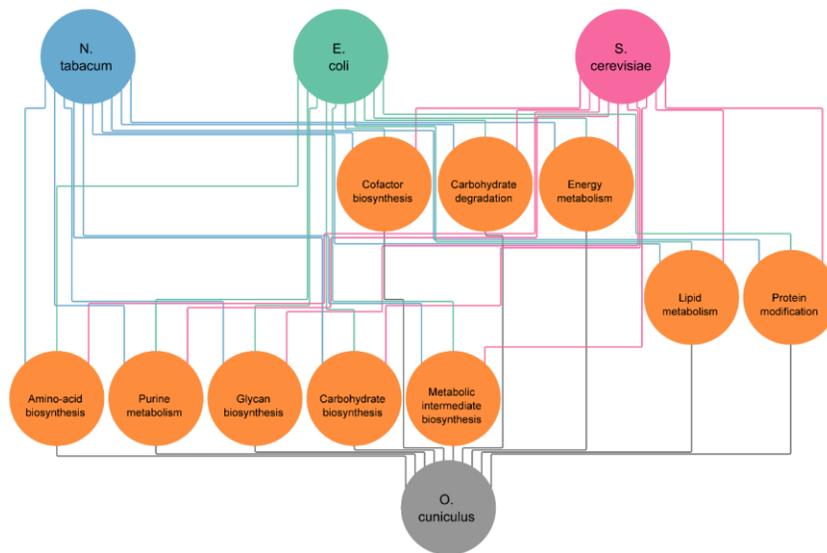


Figura 47. Rutas metabólicas que coinciden con los cuatro proteomas. De la Figura 46 se filtraron las rutas metabólicas que coinciden en los cuatro proteomas.



Figura 48. Rutas metabólicas únicas para cada proteoma.

En la Figura 78 (en anexos) están representadas aquellas rutas metabólicas que están presentes solo en dos proteomas. Se encontró que el proteoma *Saccharomyces cerevisiae* tiene 9 rutas metabólicas, de las cuales, una coincide con el proteoma *Oryctolagus cuniculus* y 8 con el proteoma *Escherichia coli*. De las 8 rutas metabólicas que tienen en común el proteoma *Escherichia coli* y *Saccharomyces cerevisiae* está la degradación de organosulfurados, biosíntesis de isoprenoides, metabolismo de ácidos orgánicos, entre otras.

Por último, en la Figura 79 (en anexos) se muestra aquellas rutas metabólicas que tienden a estar presentes en tres proteomas. Dentro de las 18 rutas metabólicas representadas en la gráfica, como ejemplo está la ruta metabólica del metabolismo de lípidos de membrana, la degradación de aminoácidos, el metabolismo de fosfolípidos que tiende a estar presente solo en los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus*, pero no en el proteoma del *Nicotiana tabacum*. También se observa que las rutas metabólicas del metabolismo de pirimidinas, metabolismo de carbohidratos, metabolismo de sulfuros están presentes solo en los proteomas *Escherichia coli*, *Saccharomyces cerevisiae* y *Nicotiana tabacum*, pero no en el proteoma *Oryctolagus cuniculus*.

Con el propósito de entender la distribución de las rutas metabólicas por clase enzimática (EC) se realizaron agrupaciones de acuerdo con al tipo de reacción que catalizan las enzimas que participan en cada una de las rutas metabólicas. Los resultados obtenidos por estas agrupaciones se representaron por medio de gráficos de barras (Figura 49).

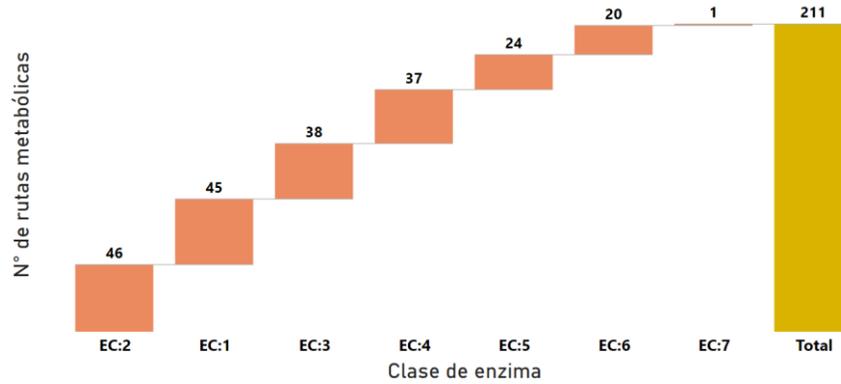


Figura 49. Conteo de rutas metabólicas por clase de enzima (EC). Agrupación de las rutas metabólicas por clase de enzima sin distinción del proteoma a la que pertenece. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7).

En la Figura 49 muestra que la clase de enzima con mayor número de rutas metabólicas son las transferasas (EC:2), seguido de las oxidoreductasas (EC:1) e hidrolasas (EC:3) con 46, 45 y 38 rutas metabólicas, respectivamente (Figura 49). Para evidenciar si cada proteoma sigue esta misma tendencia se realizaron gráficas con las agrupaciones de las rutas metabólicas por cada clase de enzima (Figura 50).

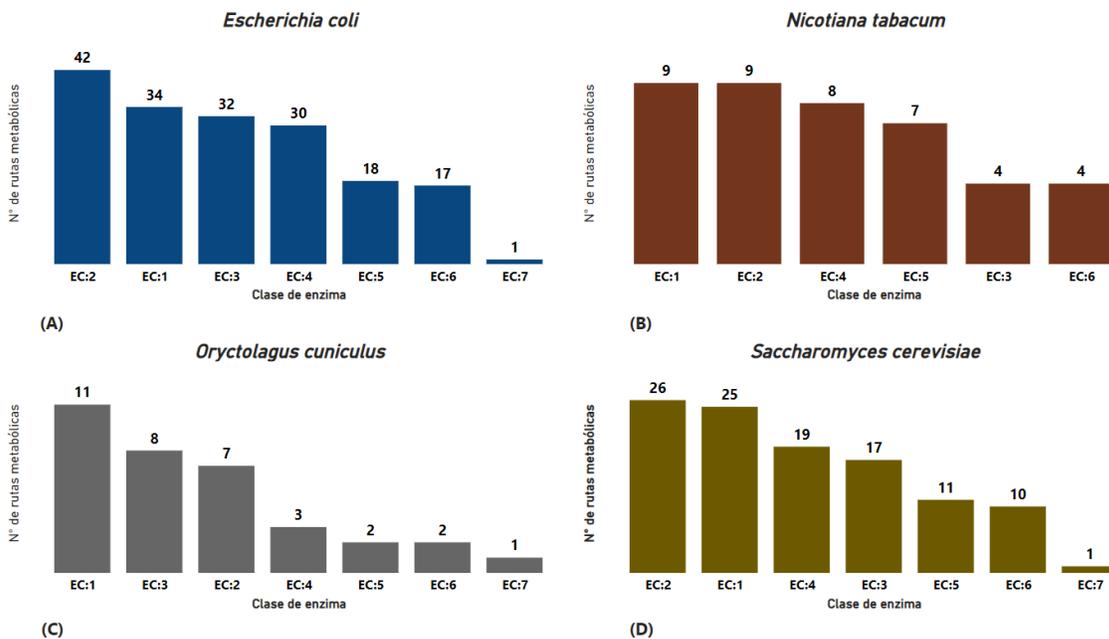


Figura 50. Conteo de rutas metabólicas por clase de enzima (EC) en cada proteoma. (A) *Escherichia coli*, (B) *Nicotiana tabacum*, (C) *Oryctolagus cuniculus* y (D) *Saccharomyces Cerevisiae*.

En la Figura 50, ilustración (A), del proteoma *Escherichia coli* se halló que las enzimas agrupadas como transferasas (EC:2) son las que participan en un mayor número de rutas metabólicas, con 42, seguida por las oxidorreductasas (EC:1) e hidrolasas (EC:3) con 34 y 32 rutas metabólicas, respectivamente. Siguiendo con el proteoma *Nicotiana tabacum*, de la ilustración (B), se identifica que la clase enzimática que participan en un mayor número de rutas metabólicas son las oxidorreductasas (EC:1), seguida por las transferasas (EC:2), liasas (EC:4) e isomerasas (EC:5) con 9, 9, 8 y 7 rutas metabólicas, respectivamente. Continuando con la ilustración (C) del proteoma *Oryctolagus cuniculus*, se encontró que las enzimas agrupadas como oxidorreductasas (EC:1) son las que participan en un mayor número de rutas metabólicas, seguida de las hidrolasas (EC:3) y transferasas (EC:2), con 11, 8 y 7 rutas metabólicas, respectivamente. Por último, para el proteoma *Saccharomyces cerevisiae*, ilustración (D), se muestra que las enzimas que participan en un mayor número de superfamilia son las transferasas (EC:2), seguidas por las oxidorreductasas (EC:1) y liasas (EC:4) con 26, 25 y 19 rutas metabólicas, respectivamente.

De lo escrito anteriormente se analiza que los datos de la ilustración (A) (Figura 50), del proteoma *Escherichia coli*, sigue exactamente la misma tendencia descrita para la Figura 49. Sin embargo, los datos de los demás proteomas no siguen dicha tendencia, pero es evidente que dentro las tres clases de enzimas (EC) que participan en un mayor número de rutas metabólicas, en los cuatro proteomas, se encuentran la oxidorreductasa (EC:1) y la transferasa (EC:2). Por otro lado, se puede apreciar que para las enzimas translocasas (EC:7) del proteoma *Nicotiana tabacum* no reporta participación en alguna ruta metabólica, mientras que las en los otros tres proteomas reportan participación en una sola ruta metabólica.

Hasta el momento en esta sección, se ha realizado un análisis general de las rutas metabólicas en los cuatro proteomas y, además, se realizó el conteo de las rutas metabólicas en las que participan las diferentes clases de enzimas. Lo que sigue es el análisis de la cantidad de proteínas que actúan como catalizadores en cada una de las rutas metabólicas. Para lo cual, se realizó un conteo en cada una de las rutas metabólicas con la finalidad de saber cuál es la

ruta metabólica que tiene mayor participación de enzimas. Los resultados obtenidos se representan por medio de gráficas tipo anillo (Figura 51).

Se encontraron 70 rutas metabólicas y en éstas pueden participar desde 1 hasta 217 enzimas. En la Figura 51 se muestra que la ruta metabólica con mayor número de participación de enzimas es la biosíntesis de aminoácidos (*Amino-acid biosynthesis*), con 217 enzimas, seguida de la biosíntesis de cofactores (*Cofactor biosynthesis*, con 190 enzimas), la degradación de carbohidratos (*Carbohydrate degradation*, con 118 enzimas). y entre las rutas metabólicas donde participan una, dos y tres enzimas se encuentra la degradación de organosulfurados, metabolismo de flavonoides (*Flavonoid metabolism*) y, entre otras, biosíntesis de nucleósidos (*Nucleoside biosynthesis*), respectivamente. Ahora, para identificar si estas tendencias aplican individualmente para cada proteoma, se realizaron cuatro gráficas tipo anillo (de la Figura 52 a la Figura 55), una por cada proteoma.

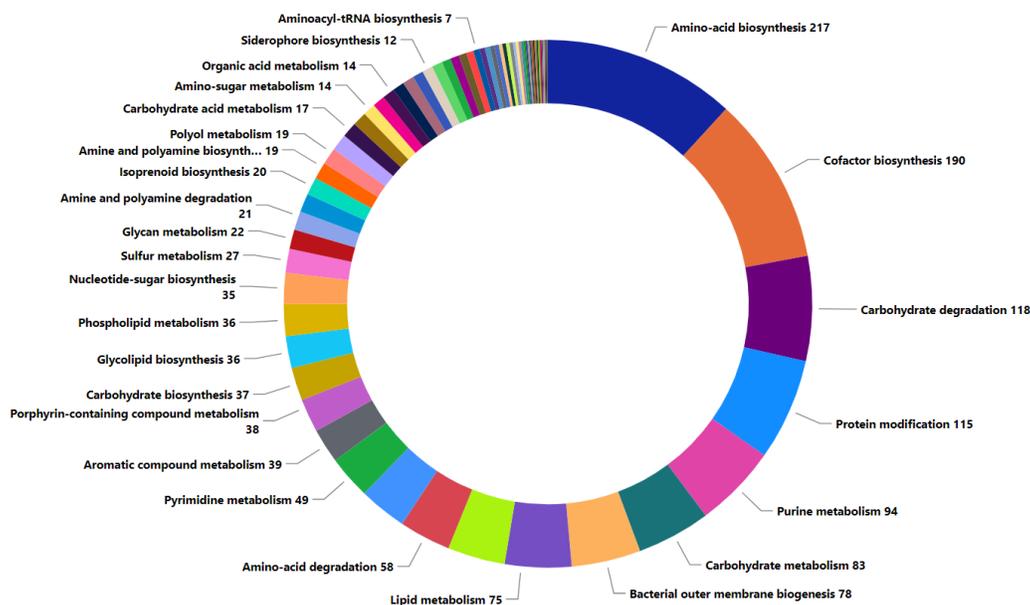


Figura 51. Conteo de enzimas por cada ruta metabólica. En la gráfica se muestra el número de enzimas que participan en cada una de las rutas metabólicas sin discriminar las enzimas por proteoma.

Los resultados encontrados, para el proteoma *Escherichia coli* (Figura 52), muestran que de las rutas metabólicas que tiene mayor participación de enzimas esta la biosíntesis de aminoácidos (*Amino-acid biosynthesis*), con 128 enzimas, seguida de la biosíntesis de cofactores (*Cofactor biosynthesis*, con 127 enzimas), biogénesis de la membrana interna bacteriana (*Bacterial outer membrane biogenesis*, con 78 enzimas). Mientras que entre las rutas metabólicas en donde solo participan una, dos y tres enzimas se encuentra la degradación de organosulfurados, metabolismo de flavonoides (*Flavonoid metabolism*) y biosíntesis de nucleósidos (*Nucleoside biosynthesis*), respectivamente.

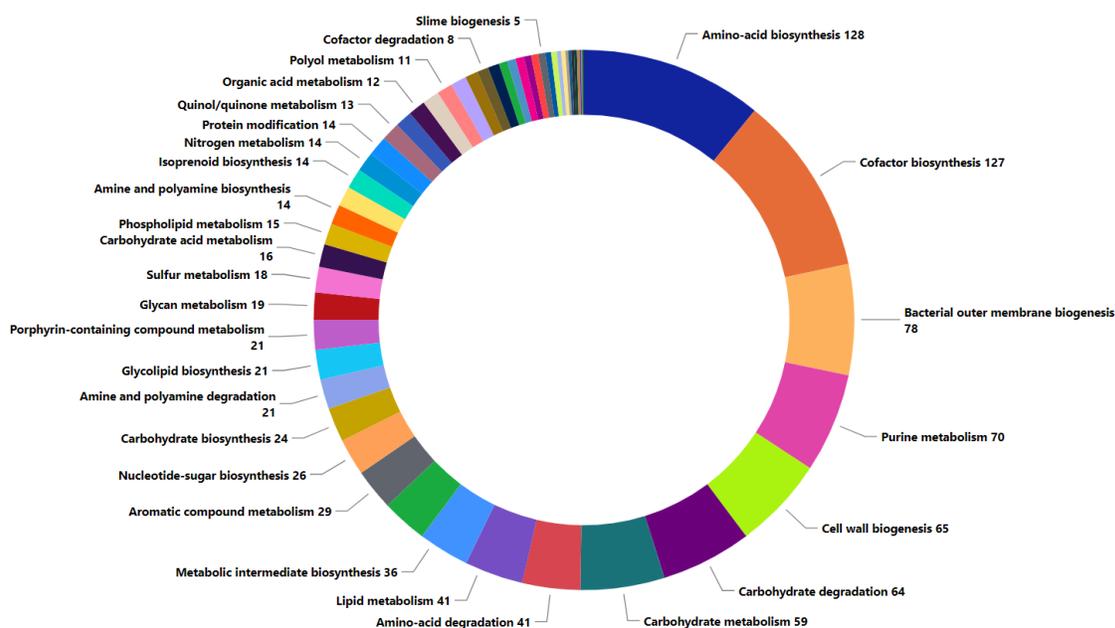


Figura 52. Conteo de enzimas por cada ruta metabólica para *Escherichia coli*. En la gráfica se muestra el número de enzimas que participan en cada una de las rutas metabólicas.

Para el proteoma *Nicotiana tabacum* (Figura 53) se muestra que de las rutas metabólicas que tiene mayor participación de enzimas está el metabolismo de compuestos aromáticos (*Aromatic compound metabolism*, con 9 enzimas), seguida de la degradación de carbohidratos (*Carbohydrate degradation*, con 9 enzimas), biosíntesis de aminoácidos (*Amino-acid biosynthesis*, con 8 enzimas). Mientras que entre las rutas metabólicas en donde solo participan una, dos y tres enzimas se encuentra el metabolismo de sulfuros (*Sulfur*

metabolism), biosíntesis de alcanos (*Alkane biosynthesis*) y biosíntesis de carbohidratos (*Carbohydrate biosynthesis*), respectivamente.

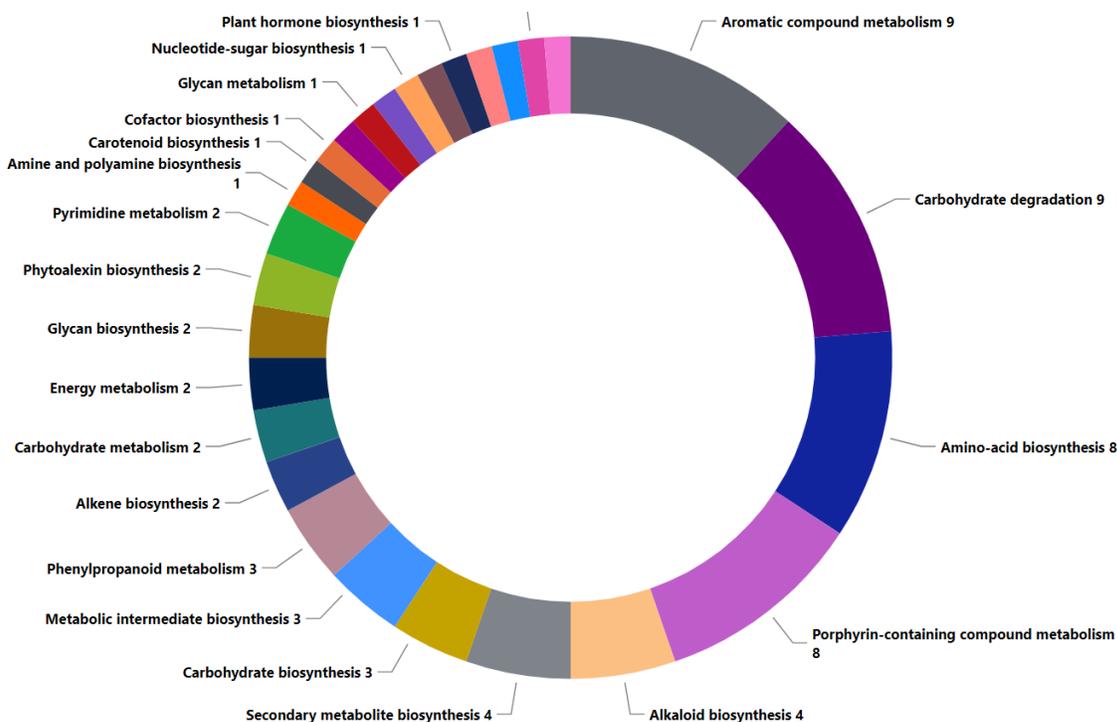


Figura 53. Conteo de enzimas por cada ruta metabólica para *Nicotiana tabacum*. En la gráfica se muestra el número de enzimas que participan en cada una de las rutas metabólicas.

Para el proteoma *Oryctolagus cuniculus* (Figura 54) se halló que de las rutas metabólicas que tiene mayor participación de enzimas está el metabolismo de lípidos (*Lipid metabolism*, con 13 enzimas), seguida de la degradación de carbohidratos (*Carbohydrate degradation*, con 10 enzimas), modificación de proteínas (*Protein modification*, con 10 enzimas). Mientras que entre las rutas metabólicas en donde participan una, dos y tres enzimas se encuentra la biosíntesis de hormonas esteroideas (*Steroid hormone biosynthesis*), metabolismo de nitrógeno (*Nitrogen metabolism*) y biosíntesis de carbohidratos (*Carbohydrate biosynthesis*), respectivamente.

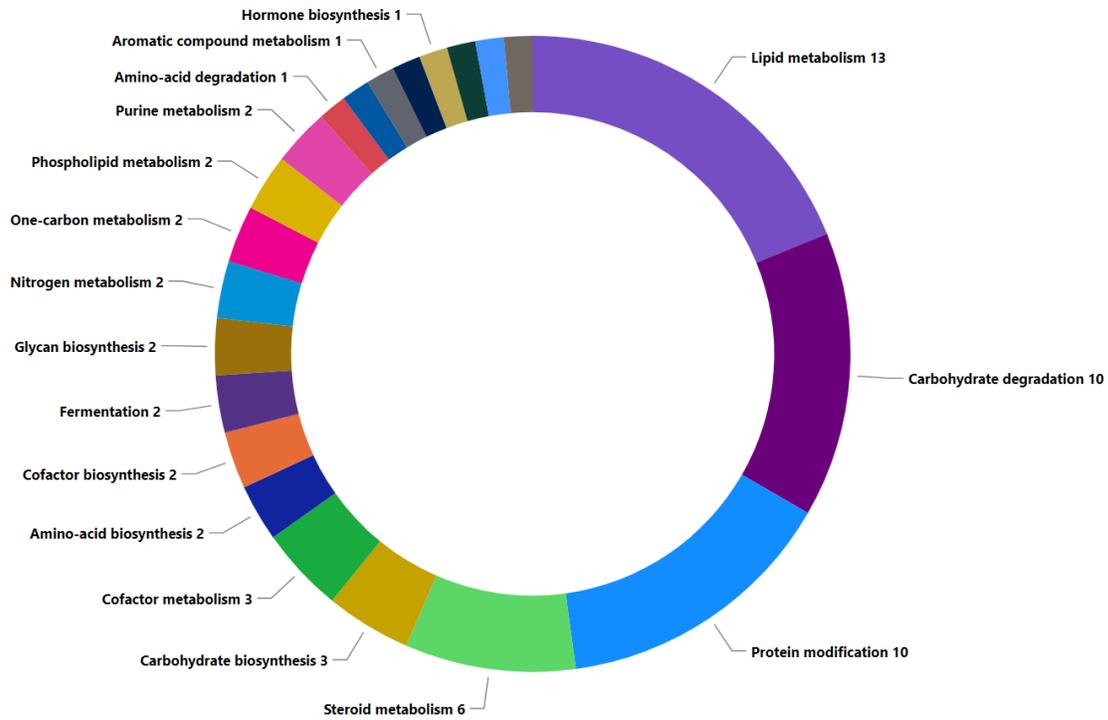


Figura 54. Conteo de enzimas por cada ruta metabólica para *Oryctolagus cuniculus*. En la gráfica se muestra el número de enzimas que participan en cada una de las rutas metabólicas.

Para el proteoma *Saccharomyces cerevisiae* (Figura 55) se encontró que de las rutas metabólicas que tiene mayor participación de enzimas está la modificación proteínas (*Protein modification*, con 90 enzimas), seguida de la biosíntesis de aminoácidos (*Amino-acid biosynthesis*, con 79 enzimas), biosíntesis de cofactores (*Cofactor biosynthesis*, con 60 enzimas). Mientras que entre las rutas metabólicas en donde participan una, dos y tres enzimas se encuentra el metabolismo de esfingolípidos (*Sphingolipid metabolism*), biosíntesis de proteínas (*Protein biosynthesis*) y biosíntesis de aminas y poliaminas (*Amine and polyamine biosynthesis*), respectivamente.

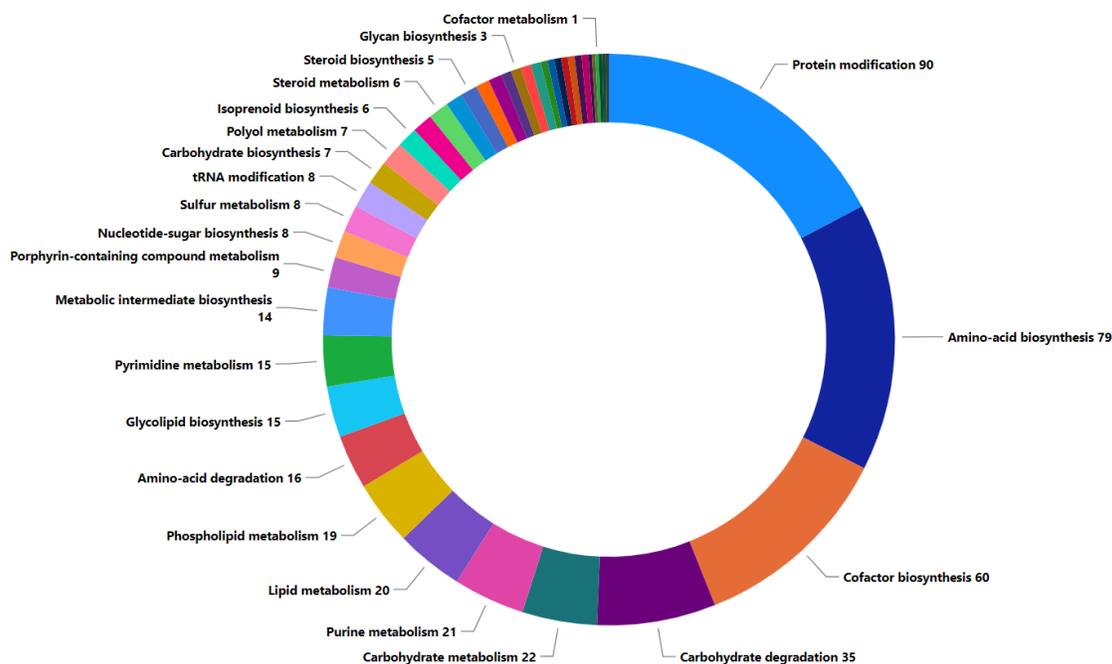


Figura 55. Conteo de enzimas por cada ruta metabólica para *Saccharomyces cerevisiae*. En la gráfica se muestra el número de enzimas que participan en cada una de las rutas metabólicas.

No todos los proteomas siguen las tendencias ilustradas en la Figura 51 (ver Tablas 7 y 8 y Figuras 52 a la 55).

Tabla 7. Rutas metabólicas con mayor participación de enzimas.

Rutas metabólicas	<i>Escherichia coli</i>	<i>Nicotiana tabacum</i>	<i>Oryctolagus cuniculus</i>	<i>Saccharomyces cerevisiae</i>	Total
Biosíntesis de aminoácidos	128	8	2	79	217
Biosíntesis de cofactores	127	1	2	60	190
Degradación de carbohidratos	64	9	10	35	118

Para ver la distribución de las enzimas con terminal-N en cada ruta metabólica, se filtraron aquellas que tienen asignado un aminoácido terminal-N estabilizador y desestabilizador, según la Tabla 1. Luego, se contabilizaron las enzimas con tiempos de vida estable y con tiempos de vida inestable. Los resultados se representan por medio de graficas de barras (Figura 56).

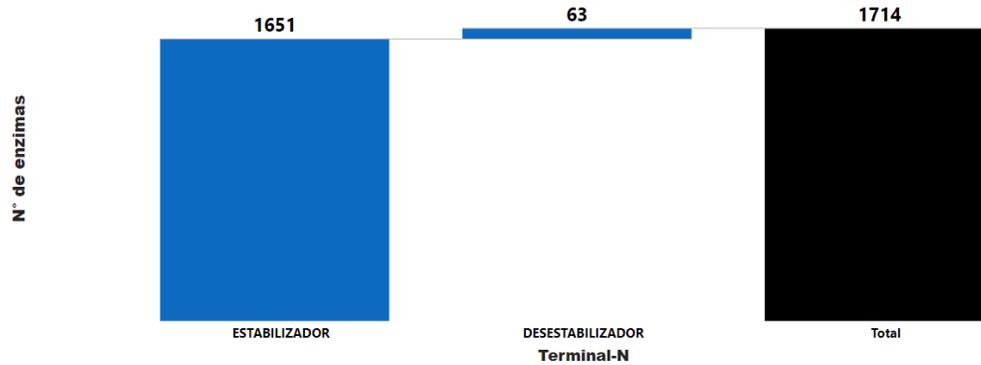


Figura 56. Conteo de enzimas que tienen asignado un terminal-N y que participan en alguna ruta metabólica. De las 1730 enzimas que participan en alguna ruta metabólica, solo 1714 están caracterizadas como enzima con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 1714 enzimas, 1651 se clasifican como enzimas con tiempos de vida estable y 63 como enzimas con tiempos de vida inestable.

De los resultados se muestra que de las 1730 enzimas que participan en alguna ruta metabólica (Figura 44), solo 1714 están caracterizadas, ya sea, como una enzima con un tiempo de vida estable (1651 enzimas) o como enzimas con un tiempo de vida inestable (63 enzimas), dependiendo del residuo aminoácido terminal-N que posea (Figura 56).

Además, de las 1651 enzimas que están caracterizadas con tiempos de vida estable, 1071 son de *Escherichia coli* (Figura 57), 60 son de *Nicotiana tabacum* (en anexos Figura 80), 35 son de *Oryctolagus cuniculus* (en anexos Figura 81) y 485 son de *Saccharomyces cerevisiae* (en anexos Figura 82).

Asimismo, de las 63 enzimas que están caracterizadas con tiempos de vida inestable, 6 corresponden a *Escherichia coli* (Figura 57), 12 corresponden a *Nicotiana tabacum* (en anexos Figura 80), 23 corresponden a *Oryctolagus cuniculus* (en anexos Figura 81) y 22 corresponden a *Saccharomyces cerevisiae* (en anexos Figura 82).

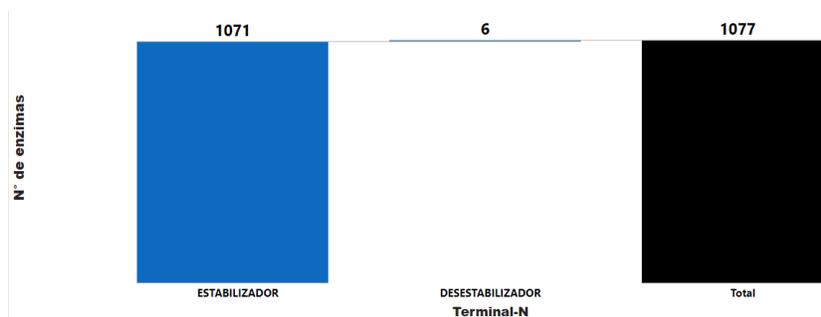


Figura 57. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador para *Escherichia coli* que participan en alguna ruta metabólica. De las 1077 enzimas que están caracterizadas como enzimas con tiempo de vida estable o inestable, 1071 se clasifican como enzimas con tiempos de vida estable y 6 como enzimas con tiempos de vida inestable.

Para identificar en qué cantidad de rutas metabólicas participan las enzimas de tiempos de vida estable y las enzimas de tiempo de vida inestable se realizó una red de interacción (Figura 58). Se encontró que hay mayor cantidad de rutas metabólicas en las que participan solamente enzimas con residuos aminoácidos terminal-N estabilizador (nodos color morado) que rutas metabólicas en las que participan solo enzimas que tienen residuos aminoácidos terminal-N desestabilizador (nodos color azul).

De las 70 rutas metabólicas (Figura 58) en 2 participan solo enzimas que tienen residuos aminoácidos terminal-N desestabilizadores, en 22 de las rutas metabólicas participan aquellas enzimas con residuos aminoácidos terminal-N desestabilizadores y estabilizadores y, por último, en 46 de las rutas metabólicas participan aquellas solamente enzimas que tienen residuos aminoácidos terminal-N estabilizadores (en anexos Tabla 10).

Se realizaron cinco gráficas para ver el conteo de enzimas con tiempos de vida estable e inestable que participan en cada una de las diferentes rutas metabólicas. De las cinco gráficas, en una se muestra un conteo sin hacer distinción del proteoma, y en las otras cuatro se muestra el conteo para cada proteoma. Cada gráfica tiene dos ilustraciones, una ilustración para las rutas metabólicas donde participan las enzimas de tiempos de vida estable y otra ilustración para las rutas metabólicas donde participan las enzimas de tiempos de vida inestable (Figura 59, Figura 60, en anexos la Figura 83 a la Figura 85).

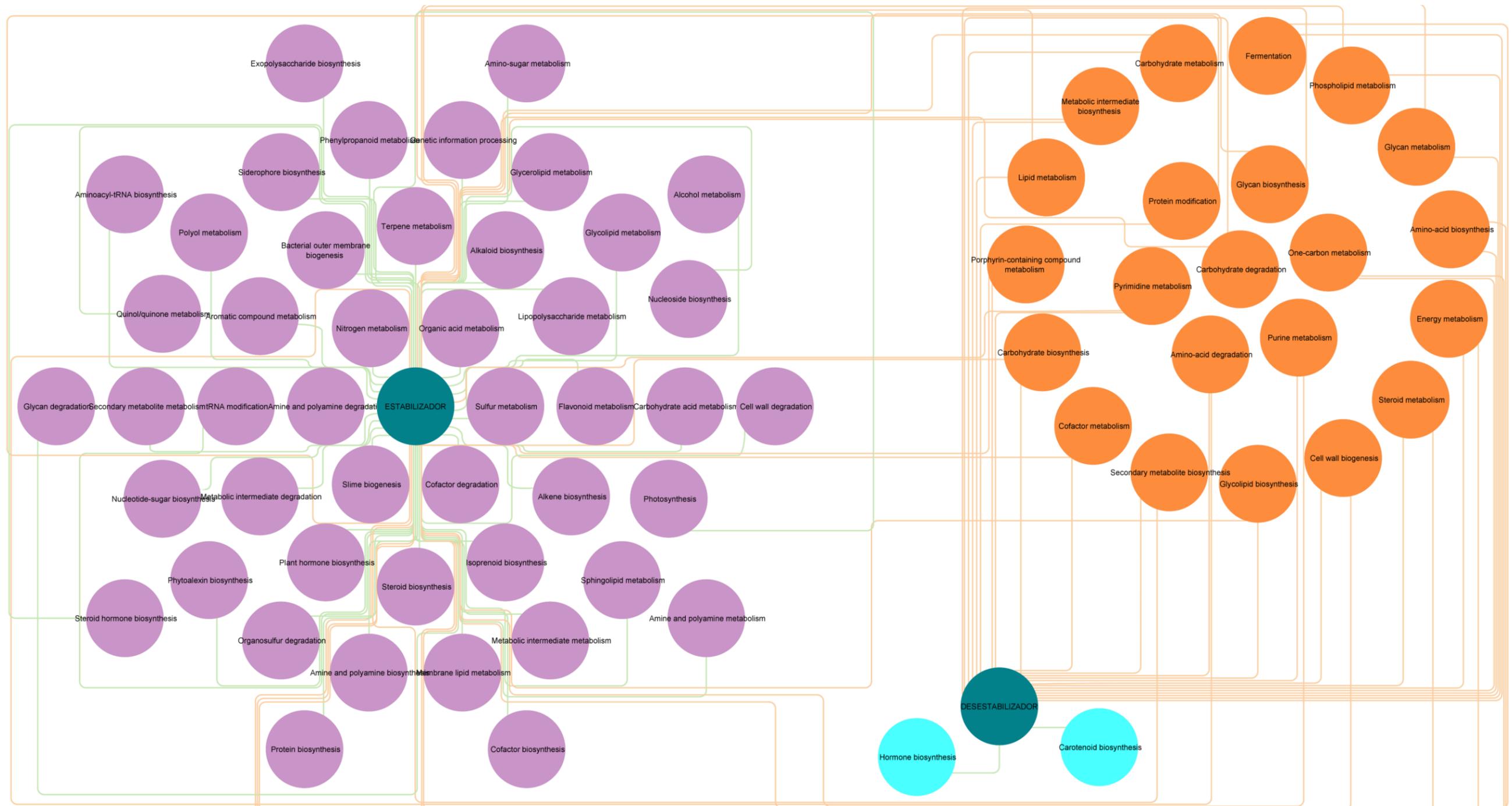


Figura 58. Red de interacción de las rutas metabólicas en las que participan enzimas con terminal-N estabilizador y terminal-N desestabilizador. Los nodos color naranja representa las rutas metabólicas en las que participan enzimas con terminal-N estabilizador y desestabilizador, los nodos color morado representa las rutas metabólicas en las que participan solo enzimas que tienen un terminal-N estabilizador y Los nodos color azul representa las rutas metabólicas en las que participan solo enzimas que tienen un terminal-N desestabilizador.

En la Figura 59 se encuentran ilustrado el conteo de las enzimas con terminal-N estabilizador (ilustración de la izquierda) y terminal-N desestabilizador (ilustración de la derecha) para cada ruta metabólica (sin distinción del proteoma). Teniendo en cuenta que el mayor número de enzimas tiende a tener un residuo terminal-N estabilizador, se espera que en la mayoría de las rutas metabólicas participen las enzimas con tiempos de vida estables. En esta gráfica se encontró que las tres rutas metabólicas donde participa un mayor número de enzimas con tiempo de vida estable son las biosíntesis de aminoácidos, biosíntesis de cofactores y degradación de carbohidratos con 201, 187 y 110 enzimas, respectivamente. Mientras que las tres rutas metabólicas donde participa el mayor número de enzimas con tiempos de vida inestables son las biosíntesis de aminoácidos, degradación de carbohidratos y metabolismo de lípidos con 10, 8 y 8 enzimas, respectivamente.

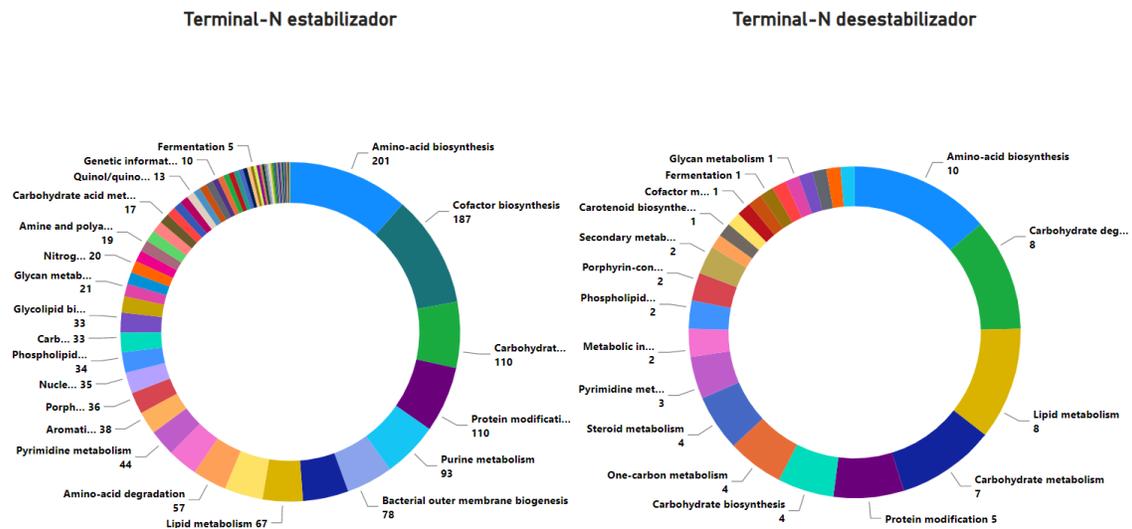


Figura 59. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica sin distinción del proteoma. A la izquierda se ilustra el número de enzimas, por cada ruta metabólica, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada ruta metabólica, con un terminal-N desestabilizador.

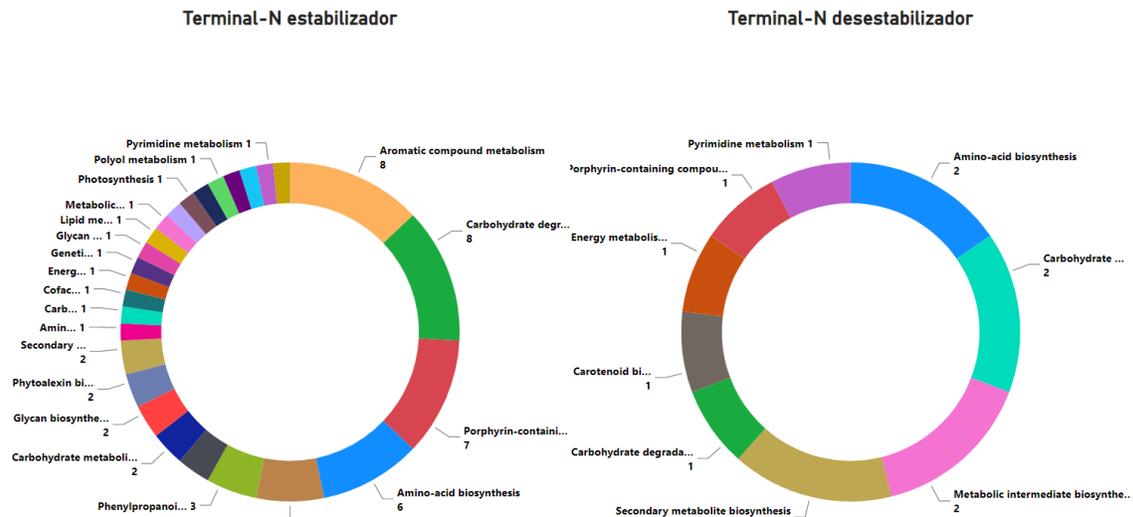


Figura 60. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en *Nicotiana tabacum*. A la izquierda se ilustra el número de enzimas, por cada ruta metabólica para el proteoma *Nicotiana tabacum*, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada ruta metabólica para el mismo proteoma, con un terminal-N desestabilizador.

Ahora bien, analizando los datos individualmente por cada uno de los cuatro proteomas, se observa que el proteoma *Escherichia coli* tiene la ruta metabólica (biosíntesis de cofactores) donde participa el mayor número de enzimas con tiempos de vida estable, 124 enzimas, seguido del proteoma *Saccharomyces cerevisiae* (modificación de proteínas) con 88 enzimas; el proteoma *Nicotiana tabacum* (metabolismo de compuestos aromáticos) con 8 enzimas y, por último, el proteoma *Oryctolagus cuniculus* con 7 enzimas en la ruta metabólica modificación de proteínas. Mientras que la ruta metabólica donde participa un mayor número de enzimas con tiempo de vida inestable es la biosíntesis de aminoácidos (con 8 enzimas) del proteoma *Saccharomyces cerevisiae*, seguido de la degradación de carbohidratos (con 7 enzimas), biosíntesis de aminoácidos (con 2 enzimas) y metabolismo de fosfolípidos (con 2 enzimas) de los proteomas *Oryctolagus cuniculus*, *Nicotiana tabacum* y *Escherichia coli*, respectivamente (Figura 60, en anexos la Figura 83 a la Figura 85).

8 CONCLUSIONES

Al analizar la estabilidad de los tiempos de vida de las enzimas de los proteomas de estudio aplicando la regla del terminal-N, se concluye que el tiempo de vida de la mayoría de las enzimas en los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus* tiende a ser estable.

Al comparar las tendencias de la estabilidad de los tiempos de vida de las enzimas pertenecientes a cada uno de los proteomas, se concluye que el proteoma de la *Escherichia coli* tiene el mayor número de enzimas con tiempo de vida estable, seguido de los proteomas de *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus*.

La mayoría de las enzimas en los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus* pertenecen a las clases de oxidorreductasas (EC:1), transferasas (EC:2) e hidrolasas EC:3) y la mayoría de estas presentan un residuo terminal-N estabilizante.

En las proteínas con actividad catalítica oxidorreductasas (EC:1), transferasas (EC:2) e hidrolasas EC:3) se presenta la mayor diversidad de superfamilias proteicas y están asociadas a la mayoría de las rutas metabólicas.

La mayoría de las enzimas de los proteomas de *Escherichia coli* y *Saccharomyces cerevisiae* tienden a participar en rutas metabólicas relacionadas con la biosíntesis de aminoácidos, mientras que la mayoría de las enzimas de los proteomas *Nicotiana tabacum* y *Oryctolagus cuniculus* tienden a participar en rutas metabólicas relacionadas con la degradación de carbohidratos.

Todas las enzimas en los proteomas de *Escherichia coli*, *Nicotiana tabacum* y *Oryctolagus cuniculus* que se clasifican como translocasas (EC:7) tienen un terminal-N estabilizante.

Los tres terminales-N más comunes en los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae*, *Nicotiana tabacum* y *Oryctolagus cuniculus* son Metionina (M), seguido de la Serina (S) y Alanina (A). La M es un terminal-N estabilizante en los proteomas en estudio; la S y la A son residuos terminal-N desestabilizantes únicamente en el proteoma de *Oryctolagus cuniculus*.

Las proteínas *revisadas* con actividad catalítica del proteoma de *Nicotiana tabacum* comparte al menos el 80% de las superfamilias proteicas con las enzimas *revisadas* en los proteomas de *Escherichia coli* y *Saccharomyces cerevisiae*.

La mayoría de las enzimas de los proteomas de *Escherichia coli*, *Saccharomyces cerevisiae* y *Nicotiana tabacum* pertenecen a la superfamilia proteica SSF52540 (Hidrolasas de trifosfato de nucleósido que contienen bucles P), mientras que la mayoría de las enzimas de del proteoma *Oryctolagus cuniculus* tienden a pertenecer a la superfamilia proteica SSF56112 (similar a la proteína quinasa (*PK-like*)).

9 REFERENCIAS BIBLIOGRÁFICAS

- Alvarez-Castelao, B., & Schuman, E. M. (2015). The regulation of synaptic protein turnover. *Journal of Biological Chemistry*, 290(48), 28623–28630. <https://doi.org/10.1074/jbc.R115.657130>
- Bachmair, A., Finley, D., & Varshavsky, A. (1986). In Vivo Half-Life of a Protein Is a Function of Its Amino-Terminal Residue, 523(1985). <https://doi.org/10.1126/science.3018930>
- Bairoch, A. (2000). The ENZYME database in 2000. *Nucleic Acids Research*, 28(1), 304–305. <https://doi.org/10.1093/nar/28.1.304>
- Balaji, S. (2021). The transferred translocases: An old wine in a new bottle. *Biotechnology and Applied Biochemistry*. <https://doi.org/10.1002/BAB.2230>
- Basisty, N., Holtz, A., & Schilling, B. (2019). Accumulation of “Old Proteins” and the Critical Need for MS-based Protein Turnover Measurements in Aging and Longevity. *Proteomics*, 1800403, 1800403. <https://doi.org/10.1002/pmic.201800403>
- Balaji, S. (2021). The transferred translocases: An old wine in a new bottle. *Biotechnology and Applied Biochemistry*. <https://doi.org/10.1002/BAB.2230>
- Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., da Silva, A., Denny, P., Dogan, T., Ebenezer, T. G., Fan, J., Castro, L. G., ... Teodoro, D. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, 49(D1), D480–D489. <https://doi.org/10.1093/NAR/GKAA1100>
- Bozaykut, P., Ozer, N. K., & Karademir, B. (2014). Regulation of protein turnover by heat shock proteins. *Free Radical Biology and Medicine*, 77, 195–209. <https://doi.org/10.1016/j.freeradbiomed.2014.08.012>
- Cheignon, C., Tomas, M., Bonnefont-Rousselot, D., Faller, P., Hureau, C., & Collin, F. (2018). Oxidative stress and the amyloid beta peptide in Alzheimer’s disease. *Redox Biology*, 14(October 2017), 450–464. <https://doi.org/10.1016/j.redox.2017.10.014>

- Chen, S.-J., Wu, X., Wadas, B., Oh, J.-H., & Varshavsky, A. (2017). An N-end rule pathway that recognizes proline and destroys gluconeogenic enzymes. *Science*, *355*(6323), eaal3655. <https://doi.org/10.1126/science.aal3655>
- Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology*, *6*(1), 79–87. <https://doi.org/10.1038/nrm1552>
- Collins, G. A., & Goldberg, A. L. (2017). The Logic of the 26S Proteasome. *Cell*, *169*(5), 792–806. <https://doi.org/10.1016/j.cell.2017.04.023>
- De Groot, R. J., Rumenapf, T., Kuhn, R. J., Strauss, E. G., & Strauss, J. H. (1991). Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(20), 8967–8971. <https://doi.org/10.1073/pnas.88.20.8967>
- Dissmeyer, N. (2019). Conditional Protein Function via N-Degron Pathway–Mediated Proteostasis in Stress Physiology. *Annual Review of Plant Biology*, *70*(1), 83–117. <https://doi.org/10.1146/annurev-arplant-050718-095937>
- Dissmeyer, N., Rivas, S., & Graciet, E. (2018). Life and death of proteins after protease cleavage: protein degradation by the N-end rule pathway. *New Phytologist*, *218*(3), 929–935. <https://doi.org/10.1111/nph.14619>
- Dohmen, J. (2015). Starting with a degron: N-terminal formyl-methionine of nascent bacterial proteins contributes to their proteolytic control. *Microbial Cell*, *2*(10), 356–359. <https://doi.org/10.15698/mic2015.10.235>
- Dong, C., Zhang, H., Li, L., Tempel, W., Loppnau, P., & Min, J. (2018). Molecular basis of GID4-mediated recognition of degrons for the Pro/N-end rule pathway. *Nature Chemical Biology*, *14*(5), 466–473. <https://doi.org/10.1038/s41589-018-0036-1>
- Donida, B., Jacques, C. E. D., Mescka, C. P., Rodrigues, D. G. B., Marchetti, D. P., Ribas, G., ... Vargas, C. R. (2017). Oxidative damage and redox in Lysosomal Storage Disorders: Biochemical markers. *Clinica Chimica Acta*, *466*, 46–53. <https://doi.org/10.1016/j.cca.2017.01.007>
- Dougan, D. A., Micevski, D., & Truscott, K. N. (2012). The N-end rule pathway: From recognition by N-recognins, to destruction by AAA+proteases. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1823*(1), 83–91. <https://doi.org/10.1016/j.bbamcr.2011.07.002>
- Dougan, D. A., Truscott, K. N., & Zeth, K. (2010). The bacterial N-end rule pathway: Expect

the unexpected. *Molecular Microbiology*. Blackwell Publishing Ltd.
<https://doi.org/10.1111/j.1365-2958.2010.07120.x>

Dougan, David A., & Varshavsky, A. (2018). Understanding the Pro/N-end rule pathway. *Nature Chemical Biology*, *14*(5), 415–416. <https://doi.org/10.1038/s41589-018-0045-0>

Download Microsoft Power Query para Excel from Official Microsoft Download Center. (n.d.). Retrieved June 26, 2022, from <https://www.microsoft.com/es-co/download/details.aspx?id=39379>

Eldeeb, M., & Fahlman, R. (2016). The-N-End Rule: The Beginning Determines the End. *Protein & Peptide Letters*, *23*(4), 343–348.
<https://doi.org/10.2174/0929866523666160108115809>

Gibbs, D. J., Lee, S. C., Md Isa, N., Gramuglia, S., Fukao, T., Bassel, G. W., ... Holdsworth, M. J. (2011). Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature*, *479*(7373), 415–418. <https://doi.org/10.1038/nature10534>

Gonda, D. K., Bachmair, A., Wunning, I., Tobias, J. W., Lane, W. S., & Varshavsky, A. (1989). Universality and structure of the N-end rule. *Journal of Biological Chemistry*, *264*(28), 16700–16712.

Gough J., Karplus K., Hughey R., Chordia C. Genome sequences using a library of hidden Markov models that represent all proteins of known structure. (2001). *Journal of Molecular Biology*, 903-919, 313(4). <https://doi.org/10.1006/jmbi.2001.5080>

Graciet, E., Mesiti, F., & Wellmer, F. (2010). Structure and evolutionary conservation of the plant N-end rule pathway. *The Plant Journal*, *61*(5), 741–751.
<https://doi.org/10.1111/j.1365-313X.2009.04099.x>

Greer, J. B., Early, B. P., Durbin, K. R., Patrie, S. M., Thomas, P. M., Kelleher, N. L., LeDuc, R. D., & Fellers, R. T. (2022). ProSight Annotator: Complete control and customization of protein entries in UniProt XML files. *PROTEOMICS*, *22*(11–12), 2100209.
<https://doi.org/10.1002/PMIC.202100209>

Hanson, A. D., Henry, C. S., Fiehn, O., & de Crécy-Lagard, V. (2016). Metabolite Damage and Metabolite Damage Control in Plants. *Annual Review of Plant Biology*, *67*(1), 131–152. <https://doi.org/10.1146/annurev-arplant-043015-111648>

Hetzer, M. W., & Toyama, B. H. (2013). Protein homeostasis: live long, won't prosper. *Nature Reviews Molecular Cell Biology*, *14*(1), 55–61.
<https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted>

- Hu, R.-G., Sheng, J., Qi, X., Xu, Z., Takahashi, T. T., & Varshavsky, A. (2005). The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature*, *437*(7061), 981–986. <https://doi.org/10.1038/nature04027>
- Izard, J. W., & Kendall, D. A. (1994). Signal peptides: exquisitely designed transport promoters. *Molecular Microbiology*, *13*(5), 765–773. <https://doi.org/10.1111/j.1365-2958.1994.tb00469.x>
- Kim, H.-K., Kim, R.-R., Oh, J.-H., Cho, H., Varshavsky, A., & Hwang, C.-S. (2014). The N-Terminal Methionine of Cellular Proteins as a Degradation Signal. *Cell*, *156*(1–2), 158–169. <https://doi.org/10.1016/j.cell.2013.11.031>
- Kim, J. M., & Hwang, C. S. (2014). Crosstalk between the Arg/N-end and Ac/N-end rule. *Cell Cycle*, *13*(9), 1366–1367. <https://doi.org/10.4161/cc.28751>
- Kim, J. M., Seok, O. H., Ju, S., Heo, J. E., Yeom, J., Kim, D. S., ... Hwang, C. S. (2018). Formyl-methionine as an N-degron of a eukaryotic N-end rule pathway. *Science*, *362*(6418), 6418 (1-28). <https://doi.org/10.1126/science.aat0174>
- Klionsky, D. J. (2005). The molecular machinery of autophagy: unanswered questions. *Journal of Cell Science*, *118*(1), 7–18. <https://doi.org/10.1242/jcs.01620>
- Kwon, Y. T., & Ciechanover, A. (2017). The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends in Biochemical Sciences*, *42*(11), 873–886. <https://doi.org/10.1016/j.tibs.2017.09.002>
- Lange, P. F., Huesgen, P. F., Nguyen, K., & Overall, C. M. (2014). Annotating N termini for the human proteome project: N termini and N α -acetylation status differentiate stable cleaved protein species from degradation remnants in the human erythrocyte proteome. *Journal of Proteome Research*, *13*(4), 2028–2044. <https://doi.org/10.1021/pr401191w>
- Li, L., Yan, G., & Zhang, X. (2019). A rapid and efficient method for N-termini analysis in short-lived proteins. *Talanta*, *204*, 367–371. <https://doi.org/10.1016/j.talanta.2019.06.025>
- Martinez, A., Traverso, J. A., Valot, B., Ferro, M., Espagne, C., Ephritikhine, G., ... Meinel, T. (2008). Extent of N-terminal modifications in cytosolic proteins from eukaryotes. *Proteomics*, *8*(14), 2809–2831. <https://doi.org/10.1002/pmic.200701191>
- Martoglio, B., & Dobberstein, B. (1998). Signal sequences: More than just greasy peptides. *Trends in Cell Biology*, *8*(10), 410–415. [https://doi.org/10.1016/S0962-8924\(98\)01360-9](https://doi.org/10.1016/S0962-8924(98)01360-9)

- Mccaffrey, K., & Braakman, I. (2016). Protein quality control at the endoplasmic reticulum Protein folding at the endoplasmic reticulum Targeting and translocation. *Essays in Biochemistry*, 227–235. <https://doi.org/10.1042/EBC20160003>
- Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B., & Dosztányi, Z. (2017). Degrons in cancer. *Science Signaling*, 10(470), eaak9982. <https://doi.org/10.1126/scisignal.aak9982>
- Mulder, L. C. F., & Muesing, M. A. (2000). Degradation of HIV-1 integrase by the N-end rule pathway. *Journal of Biological Chemistry*, 275(38), 29749–29753. <https://doi.org/10.1074/jbc.M004670200>
- Nguyen, K. T., Mun, S. H., Lee, C. S., & Hwang, C. S. (2018). Control of protein degradation by N-terminal acetylation and the N-end rule pathway. *Experimental and Molecular Medicine*, 50(7), 1–8. <https://doi.org/10.1038/s12276-018-0097-y>
- Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., & Demchak, B. (2019). Cytoscape Automation: Empowering workflow-based network analysis. *Genome Biology*, 20(1). <https://doi.org/10.1186/S13059-019-1758-4>
- Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., & Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. *European Journal of Cell Biology*, 97(6), 422–441. <https://doi.org/10.1016/j.ejcb.2018.06.003>
- Piatkov, K., Vu, T., Hwang, C.-S., & Varshavsky, A. (2015). Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. *Microbial Cell*, 2(10), 376–393. <https://doi.org/10.15698/mic2015.10.231>
- Porto de Souza Vandenberghe, L., Karp, S. G., Binder Pagnoncelli, M. G., von Linsingen Tavares, M., Libardi Junior, N., Valladares Diestra, K., Viesser, J. A., & Soccol, C. R. (2020). Classification of enzymes and catalytic properties. *Biomass, Biofuels, Biochemicals*, 11–30. <https://doi.org/10.1016/B978-0-12-819820-9.00002-8>
- Power BI Desktop: informes interactivos | Microsoft Power BI.* (n.d.). Retrieved June 26, 2022, from <https://powerbi.microsoft.com/es-es/desktop/>
- Saez, I., & Vilchez, D. (2014). The Mechanistic Links Between Proteasome Activity, Aging and Agerelated Diseases. *Current Genomics*, 15(1), 38–51. <https://doi.org/10.2174/138920291501140306113344>
- Schmidt, R., Zahn, R., Bukau, B., & Mogk, A. (2009). ClpS is the recognition component for Escherichia coli substrates of the N-end rule degradation pathway. *Molecular Microbiology*, 72(2), 506–517. <https://doi.org/10.1111/j.1365-2958.2009.06666.x>

- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Model of Biomolecular Interaction Networks. *Genome Research*, 13(22), 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Sharpless, N. E., & Sherr, C. J. (2015). Forging a signature of in vivo senescence. *Nature Reviews Cancer*, 15(7), 397–408. <https://doi.org/10.1038/nrc3960>
- Striebel, F., Imkamp, F., Özcelik, D., & Weber-Ban, E. (2014). Pupylation as a signal for proteasomal degradation in bacteria. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1843(1), 103–113. <https://doi.org/10.1016/j.bbamcr.2013.03.022>
- Tamás, M. J., Fauvet, B., Christen, P., & Goloubinoff, P. (2018). Misfolding and aggregation of nascent proteins: a novel mode of toxic cadmium action in vivo. *Current Genetics*, 64(1), 177–181. <https://doi.org/10.1007/s00294-017-0748-x>
- Tao, Z., Dong, B., Teng, Z., & Zhao, Y. (2020). The Classification of Enzymes by Deep Learning. *IEEE Access*, 8, 89802–89811. <https://doi.org/10.1109/ACCESS.2020.2992468>
- Tasaki, T., Sriram, S. M., Park, K. S., & Kwon, Y. T. (2012). The N-End Rule Pathway. *Annual Review of Biochemistry*, 81(1), 261–289. <https://doi.org/10.1146/annurev-biochem-051710-093308>
- Tasaki, T., Zakrzewska, A., Dudgeon, D. D., Jiang, Y., Lazo, J. S., & Kwon, Y. T. (2009). The substrate recognition domains of the N-end rule pathway. *Journal of Biological Chemistry*, 284(3), 1884–1895. <https://doi.org/10.1074/jbc.M803641200>
- Tobias, J., Shrader, T., Rocap, G., & Varshavsky, A. (1991). The N-end rule in bacteria. *Science*, 254(5036), 1374–1377. <https://doi.org/10.1126/science.1962196>
- Ubíeta, R., & Santiago, N. (1993). Degradacion de proteínas en *Escherichia coli*: papel de la estructura proteica. *Biotecnologia Aplicada*, 1(3).
- Varshavsky, A. (1992). The N-end rule. *Cell*, 69(5), 725–735. [https://doi.org/10.1016/0092-8674\(92\)90285-K](https://doi.org/10.1016/0092-8674(92)90285-K)
- Varshavsky, A. (1997). The N-end rule pathway of protein degradation. *Genes to Cells*, 2(1), 13–28. <https://doi.org/10.1046/j.1365-2443.1997.1020301.x>
- Varshavsky, A. (2008). Discovery of Cellular Regulation by Protein Degradation, 283(50), 34469–34489. <https://doi.org/10.1074/jbc.X800009200>

- Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. *Protein Science*, 20(8), 1298–1345. <https://doi.org/10.1002/pro.666>
- Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. *Proceedings of the National Academy of Sciences*, 116(2), 358–366. <https://doi.org/10.1073/pnas.1816596116>
- Vittal, V., Stewart, M. D., Brzovic, P. S., & Klevit, R. E. (2015). Regulating the regulators: Recent revelations in the control of E3 ubiquitin ligases. *Journal of Biological Chemistry*, 290(35), 21244–21251. <https://doi.org/10.1074/jbc.R115.675165>
- Wouters, E. F. M., & Baarends, E. M. (2006). ENERGY METABOLISM. *Encyclopedia of Respiratory Medicine, Four-Volume Set*, 86–90. <https://doi.org/10.1016/B0-12-370879-6/00126-5>
- Zeiler, E., List, A., Alte, F., Gersch, M., Wachtel, R., Poreba, M., ... Sieber, S. A. (2013). Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28), 11302–11307. <https://doi.org/10.1073/pnas.1219125110>

GRÁFIAS

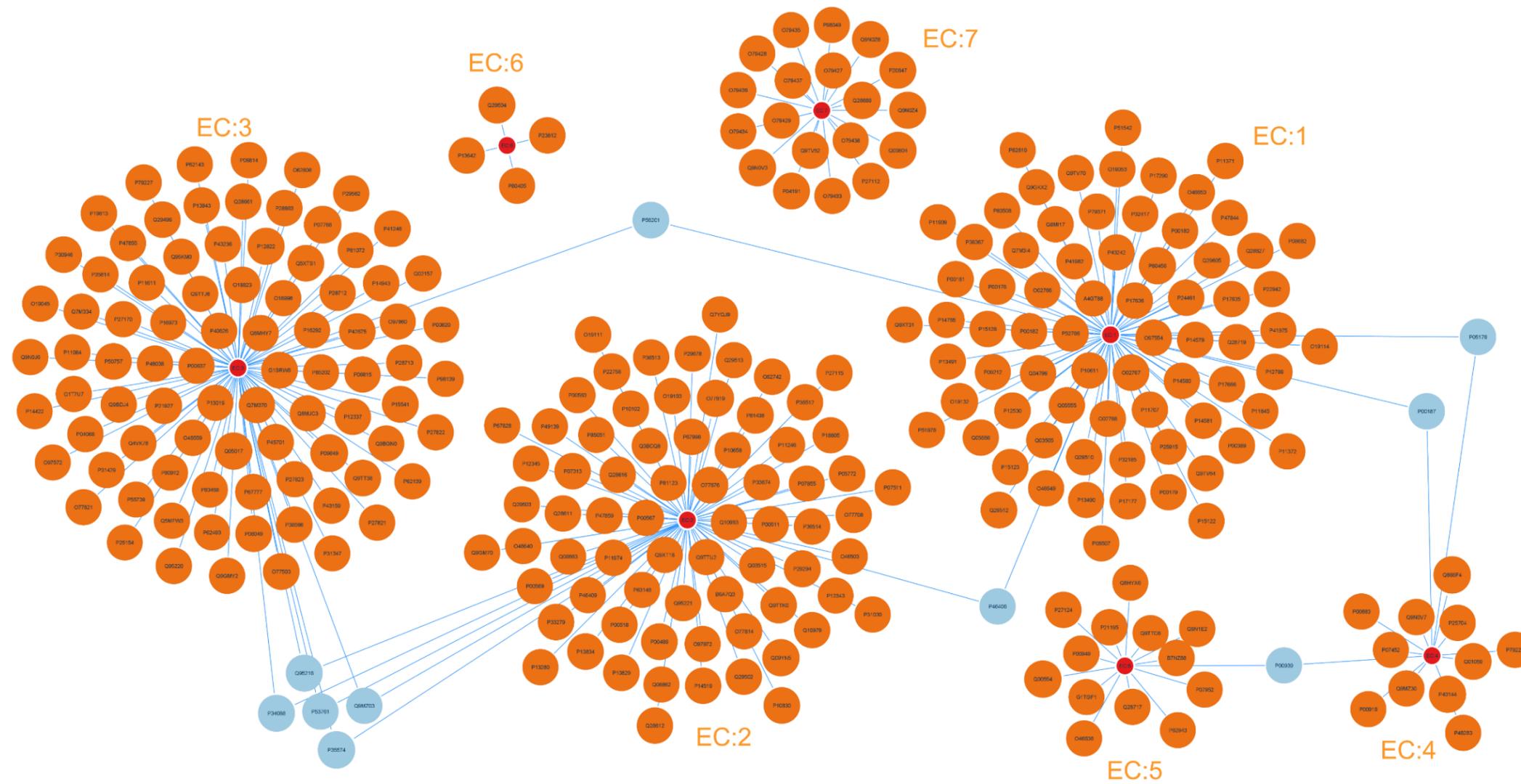


Figura 61. Proteínas con más de una actividad enzimática en *Oryctolagus cuniculus*. Se presentan 7 clúster, uno por cada clase de enzima (EC). Los nodos color azul son enzimas que tienen 2 aristas y los nodos naranjas solo tienen una arista.

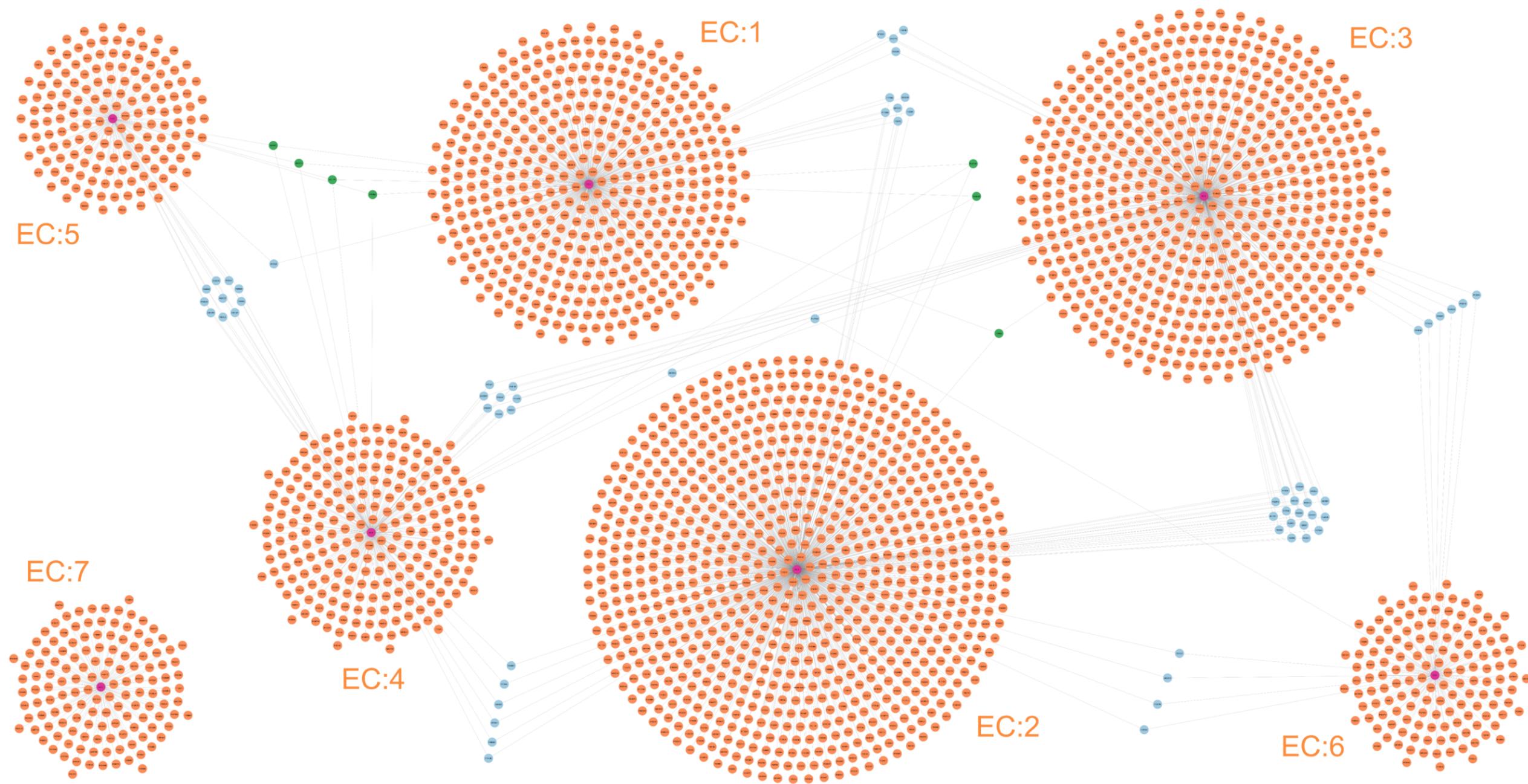


Figura 62. Proteínas con más de una actividad enzimática en *Escherichia coli*. Se presenta 7 agrupamientos, uno por cada clase de enzima (EC). Los nodos color verde representan las enzimas con 3 aristas, los nodos color azul representan a las enzimas con dos aristas y los nodos color naranjas son enzimas que tienen una arista.



Figura 63. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador en *Escherichia coli*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En esta gráfica, se representa los enzimas con un Terminal-N desestabilizador (nodos color rosa) y las enzimas con un Terminal-N estabilizador (nodos color gris). El tamaño del nodo es directamente proporcional al número de residuos aminoácidos de la secuencia de la enzima madura.

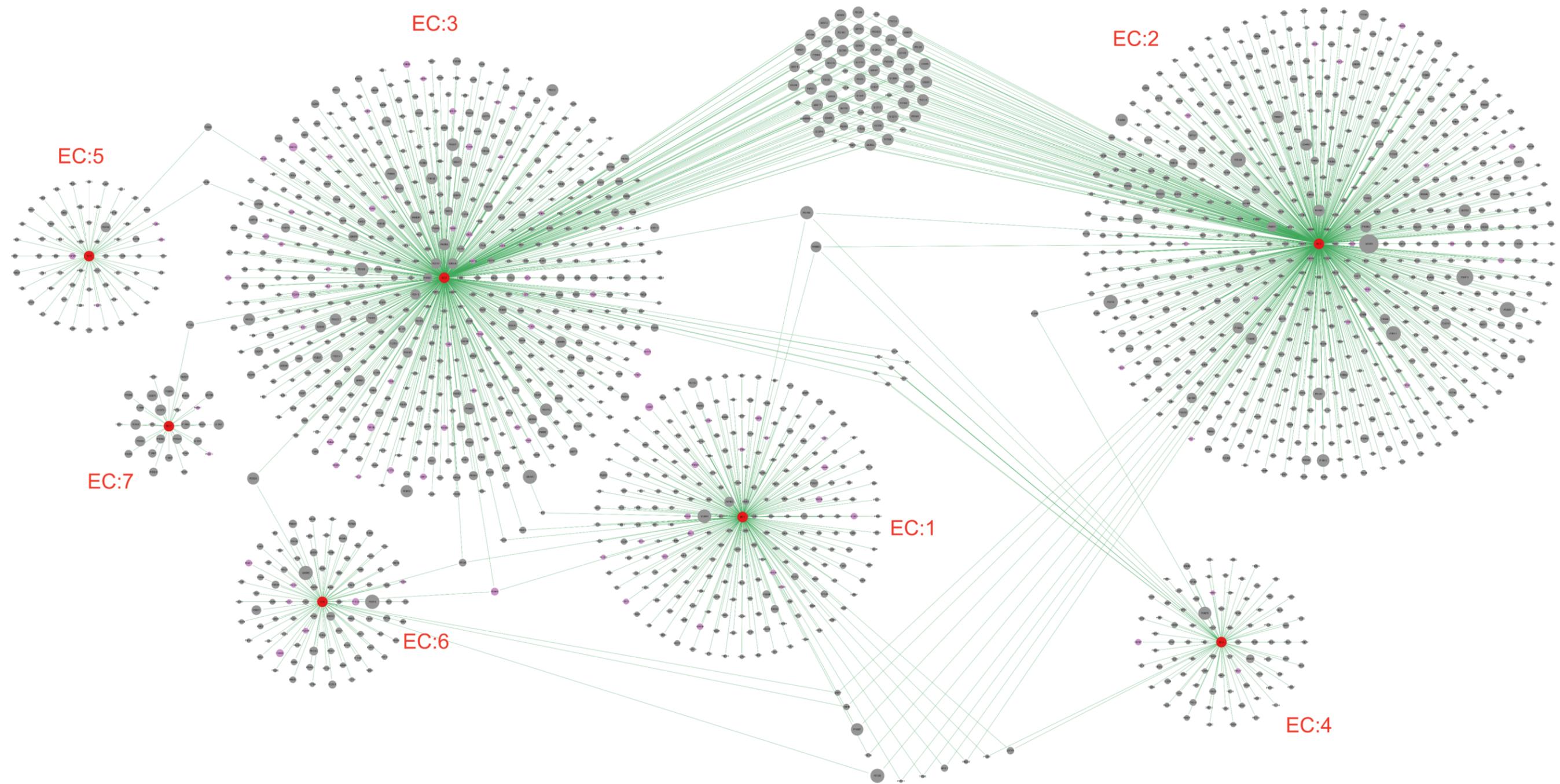


Figura 64. Enzimas con un terminal-N estabilizador y enzimas con un terminal-N desestabilizador en *Saccharomyces cerevisiae*. Oxidorreductasas (EC:1), Transferasas (EC:2); Hidrolasas (EC:3), Liasas (EC:4), Isomerasas (EC:5), Ligasas (EC:6) y Translocasas (EC:7). En esta gráfica, se representa los enzimas con un Terminal-N desestabilizador (nodos color rosa) y las enzimas con un Terminal-N estabilizador (nodos color gris). El tamaño del nodo es directamente proporcional al número de residuos aminoácidos de la secuencia de la enzima madura.

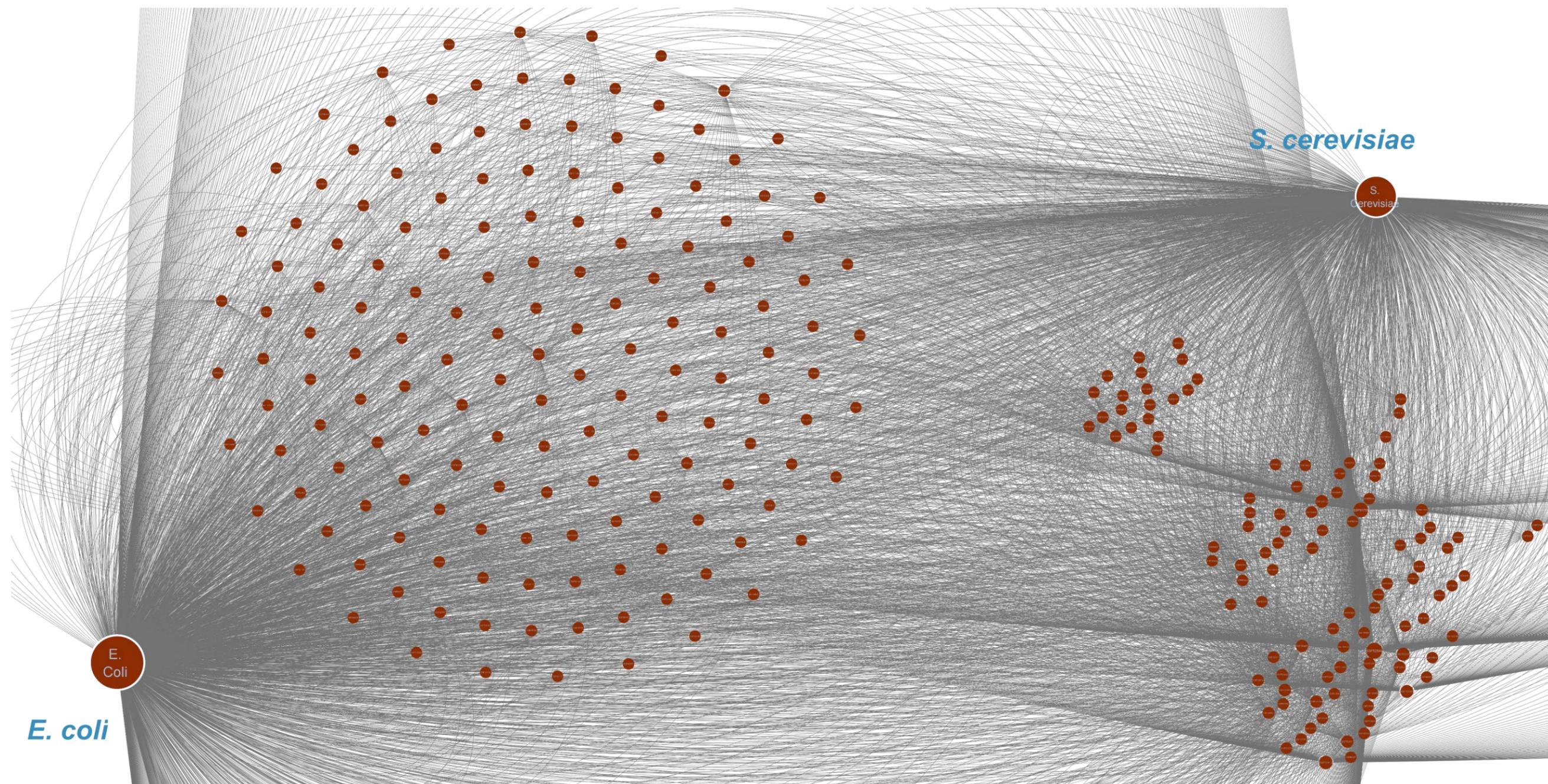


Figura 65. Superfamilias en común en los proteomas *Escherichia coli* y *Saccharomyces cerevisiae*. Estas son las superfamilias que solo tienen asignada enzimas de los dos proteomas.

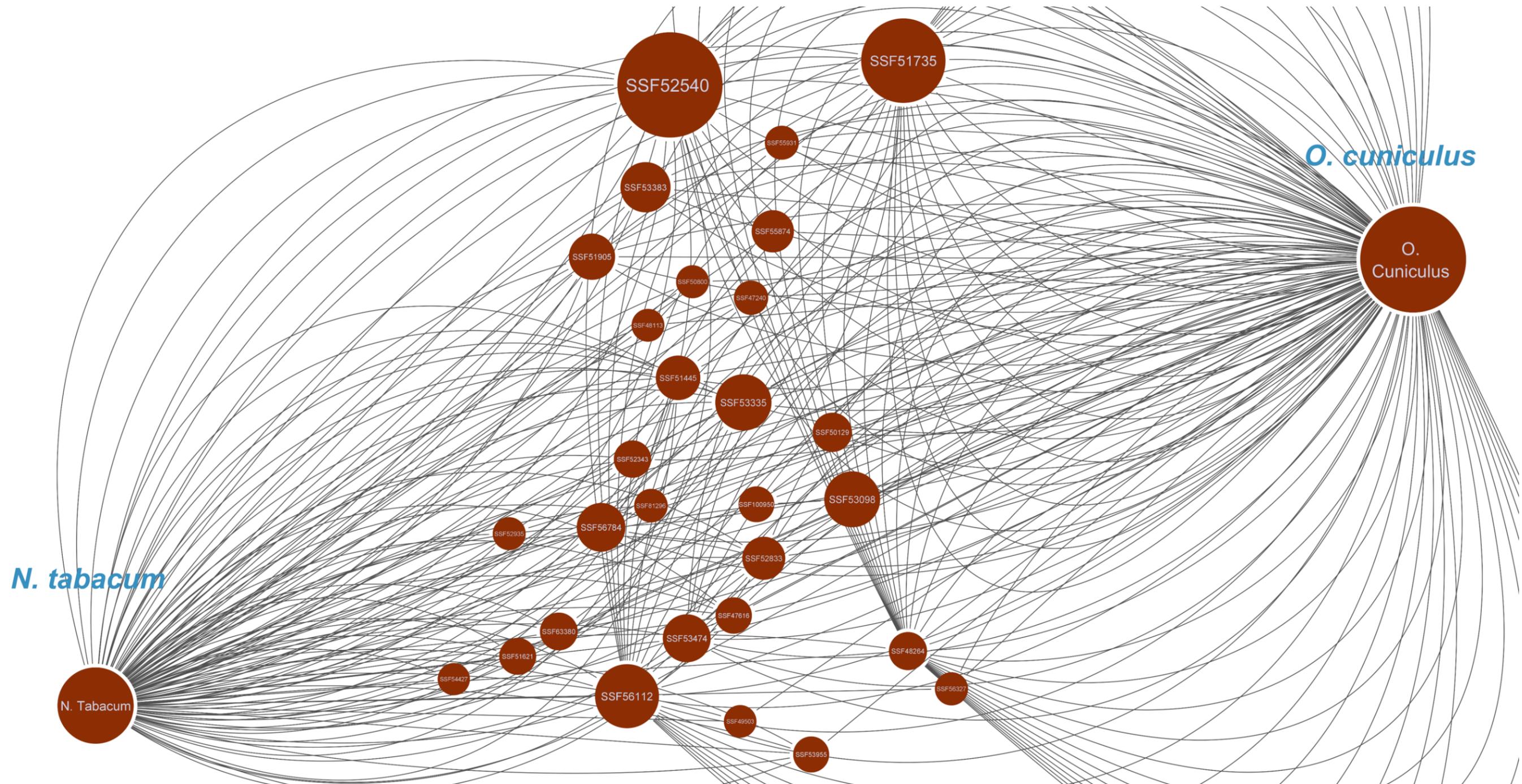


Figura 66. Superfamilias en común en los proteomas *Nicotiana tabacum* y *Oryctolagus cuniculus*. Estas son las superfamilias que solo tienen asignada enzimas de los dos proteomas. La cantidad de aristas que contiene cada nodo es proporcional a su tamaño.

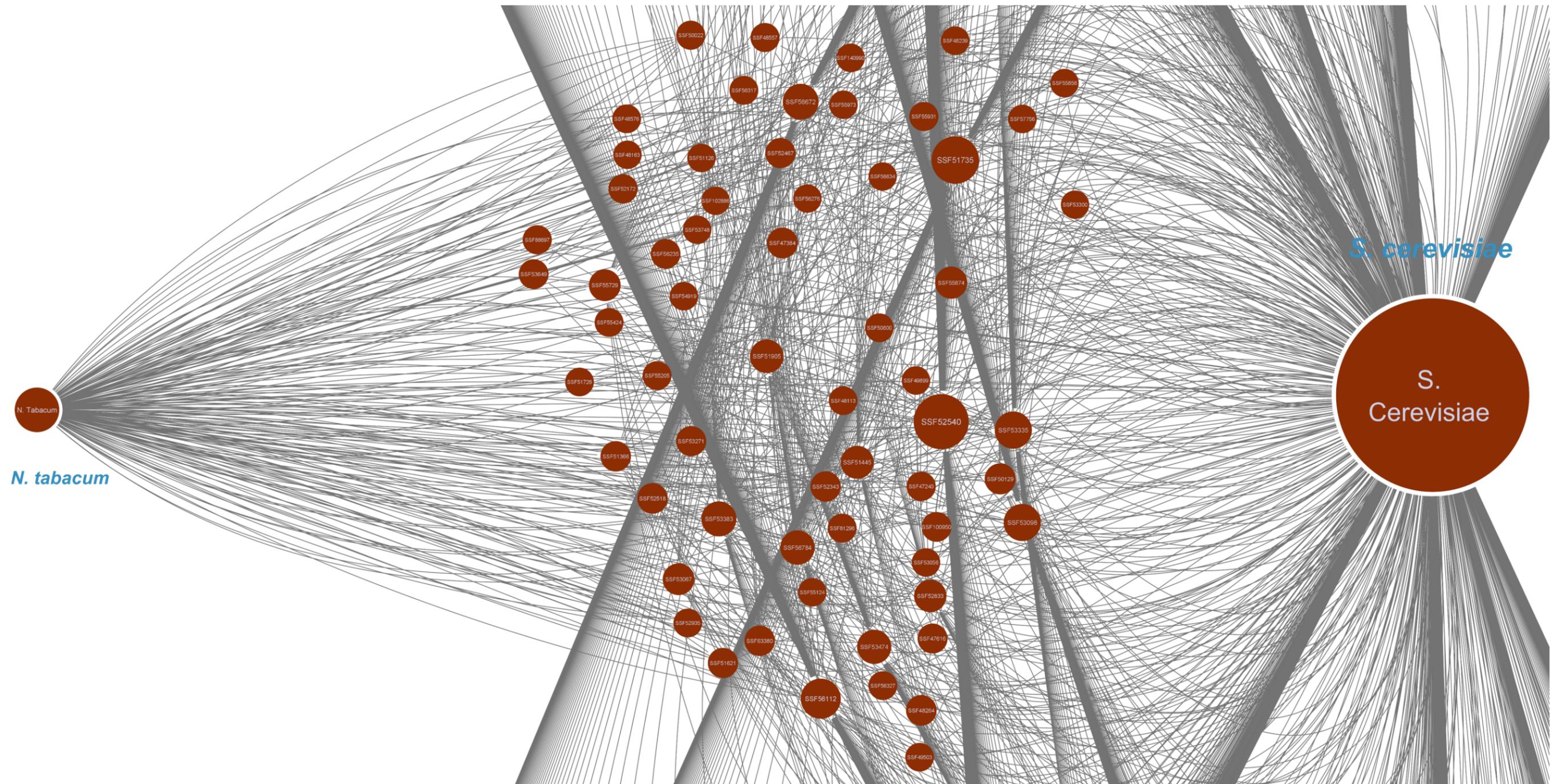


Figura 67. Superfamilias en común en los proteomas *Nicotiana tabacum* y *Saccharomyces cerevisiae*. Estas son las superfamilias que solo tienen asignada enzimas de los dos proteomas. La cantidad de aristas que contiene cada nodo es proporcional a su tamaño.

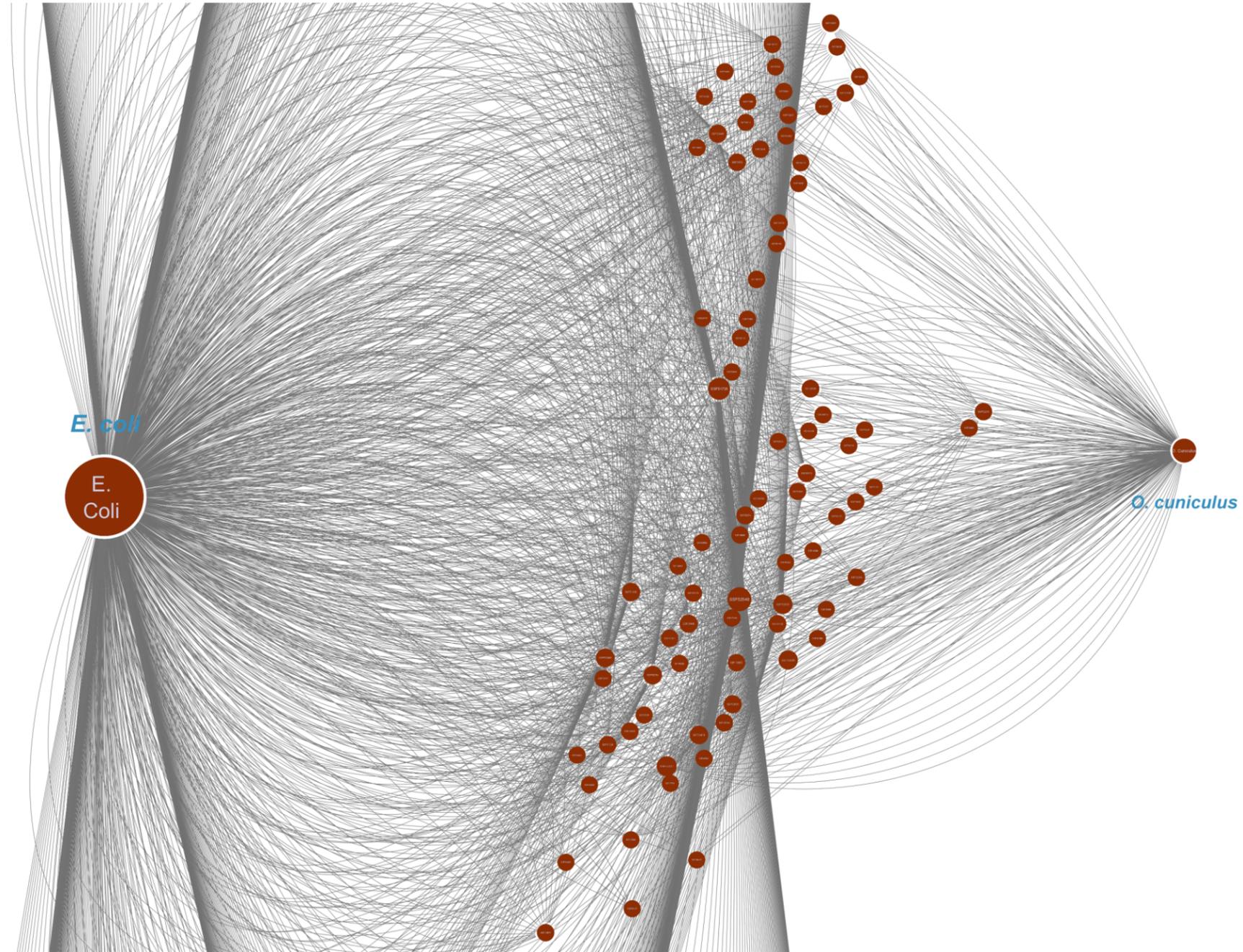


Figura 68. Superfamilias que interactúan solamente con los proteomas *Escherichia coli* y *Oryctolagus cuniculus*. Estas son las superfamilias que solo tienen asignada enzimas los dos proteomas. La cantidad de aristas que contiene cada nodo es proporcional a su tamaño.

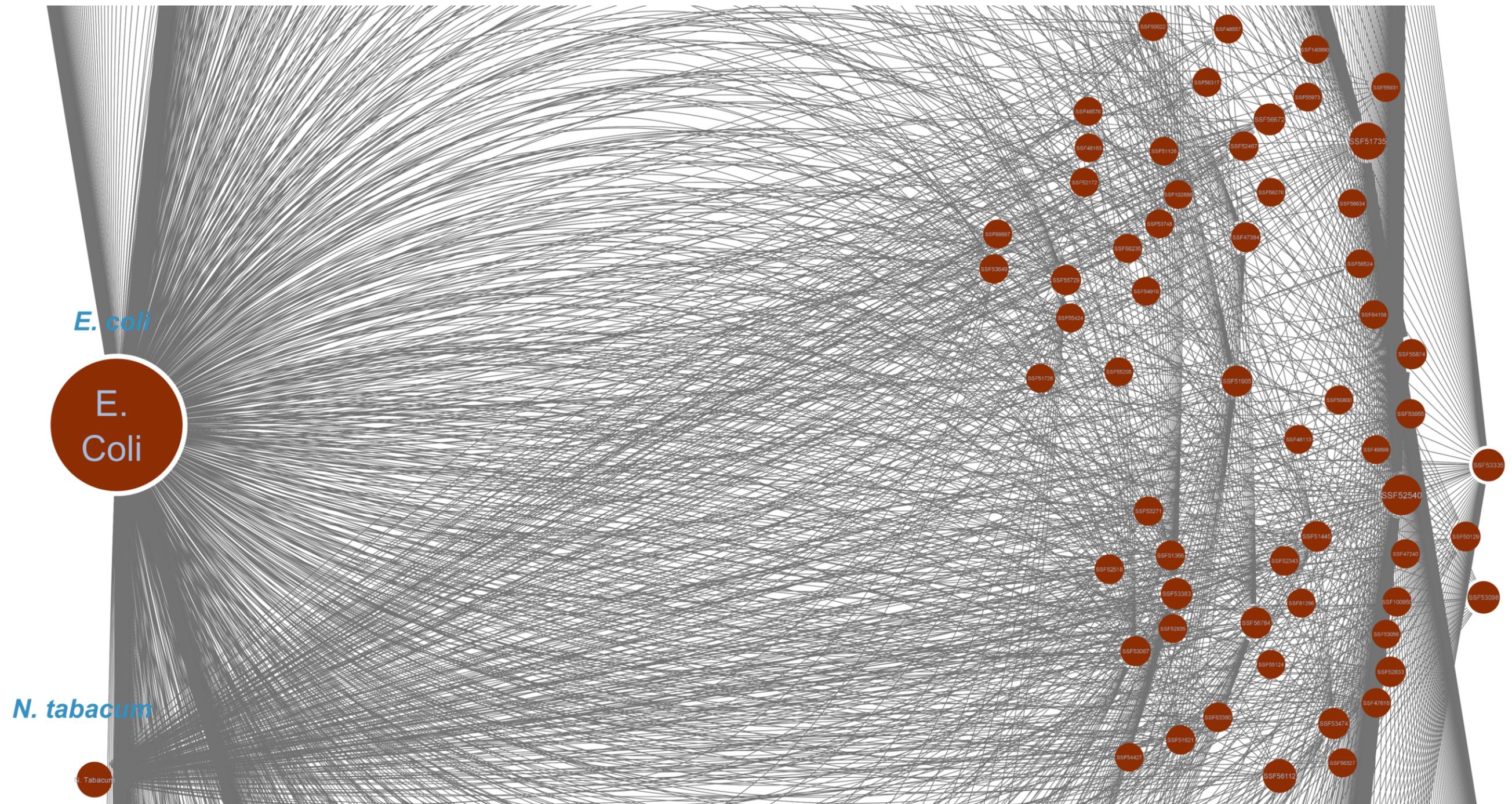


Figura 69. Superfamilias que interactúan solamente con los proteomas *Escherichia coli* y *Nicotiana tabacum*. Estas son las superfamilias que solo tienen asignada enzimas los dos proteomas. La cantidad de aristas que contiene cada nodo es proporcional a su tamaño.

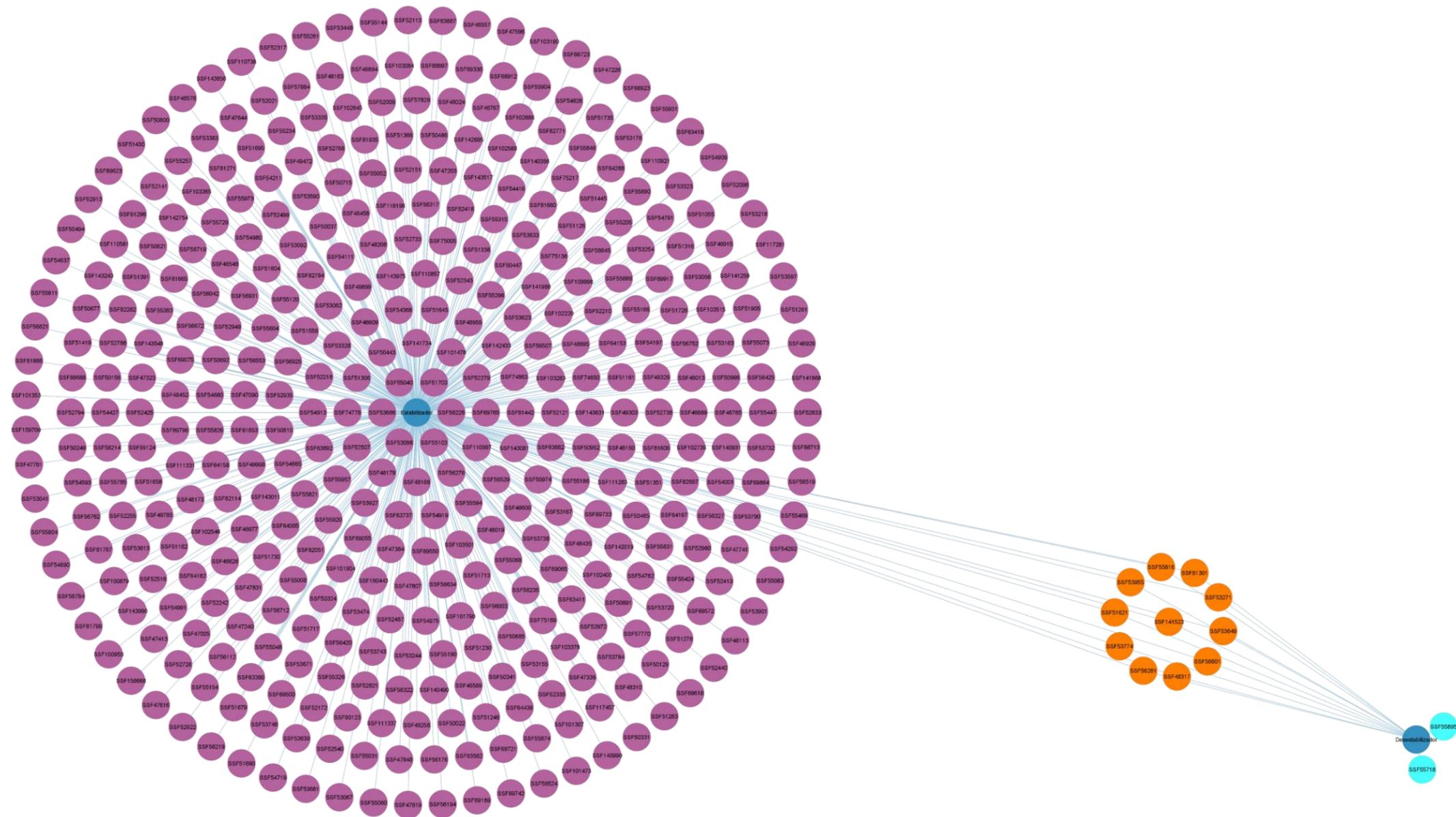


Figura 71. Superfamilias que tienen asignada enzimas con terminal-N estabilizador y desestabilizador en *Escherichia coli*. En la gráfica muestra el agrupamiento de las superfamilias que tienen enzimas con terminal-N estabilizador y terminal-N desestabilizador. Los nodos color naranja representa las superfamilias que tienen enzimas con terminal-N estabilizador y desestabilizador, los nodos color morado representa las superfamilias con solo enzimas que tienen un terminal-N estabilizador y Los nodos color azul representa las superfamilias con solo enzimas que tienen un terminal-N desestabilizador.

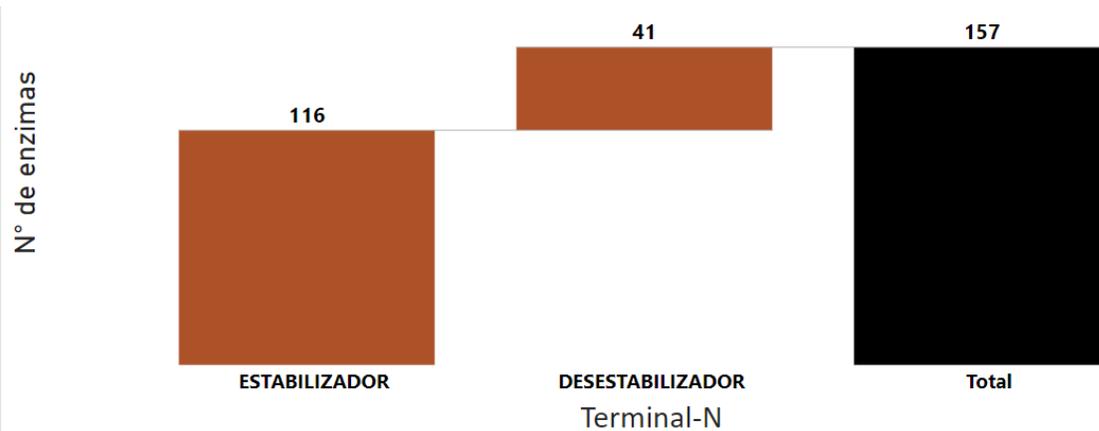


Figura 72. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma *Nicotiana tabacum*. De las 158 enzimas que tiene asignada alguna superfamilia, solo 157 están caracterizadas como enzimas con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 157 enzimas, 116 se clasifican como enzimas con tiempos de vida estable y 41 como enzimas con tiempos de vida inestable.

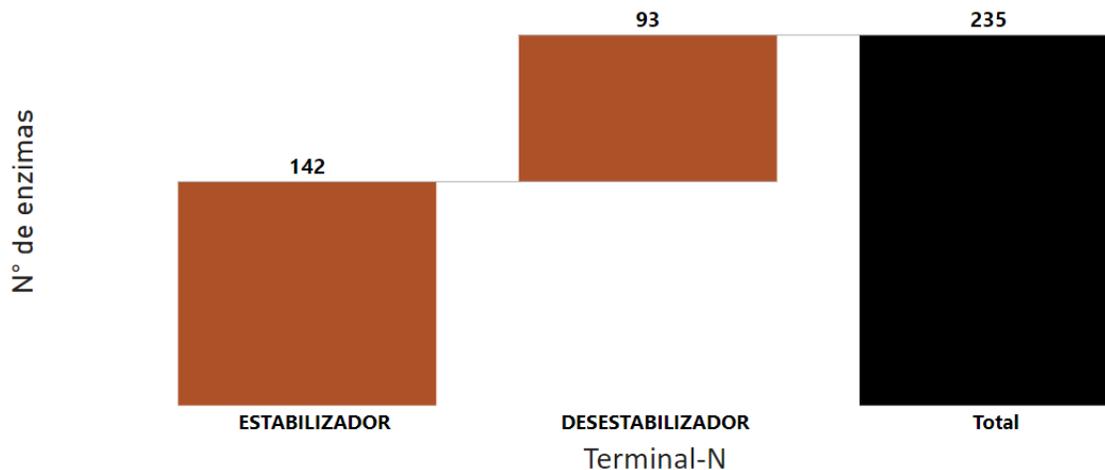


Figura 73. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma *Oryctolagus cuniculus*. De las 235 enzimas que tiene asignada alguna superfamilia, las 235 están caracterizadas como enzimas con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 235 enzimas, 142 se clasifican como enzimas con tiempos de vida estable y 93 como enzimas con tiempos de vida inestable.

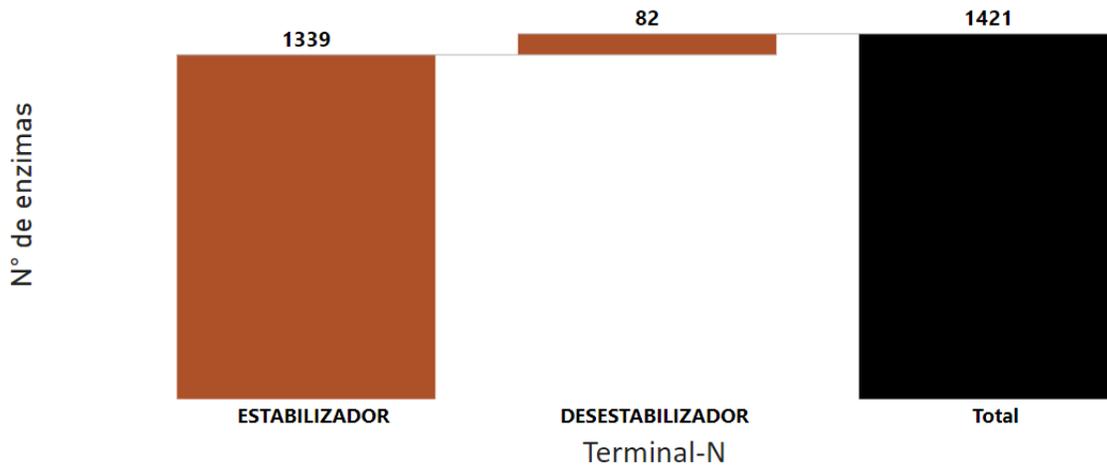


Figura 74. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N desestabilizador en el proteoma *Saccharomyces cerevisiae*. De las 1421 enzimas que tiene asignada alguna superfamilia, las 1421 están caracterizadas como enzimas con tiempo de vida estable o inestable, dependiendo de residuo aminoácido terminal-N que posea. De las 1421 enzimas, 1339 se clasifican como enzimas con tiempos de vida estable y 82 como enzimas con tiempos de vida inestable.

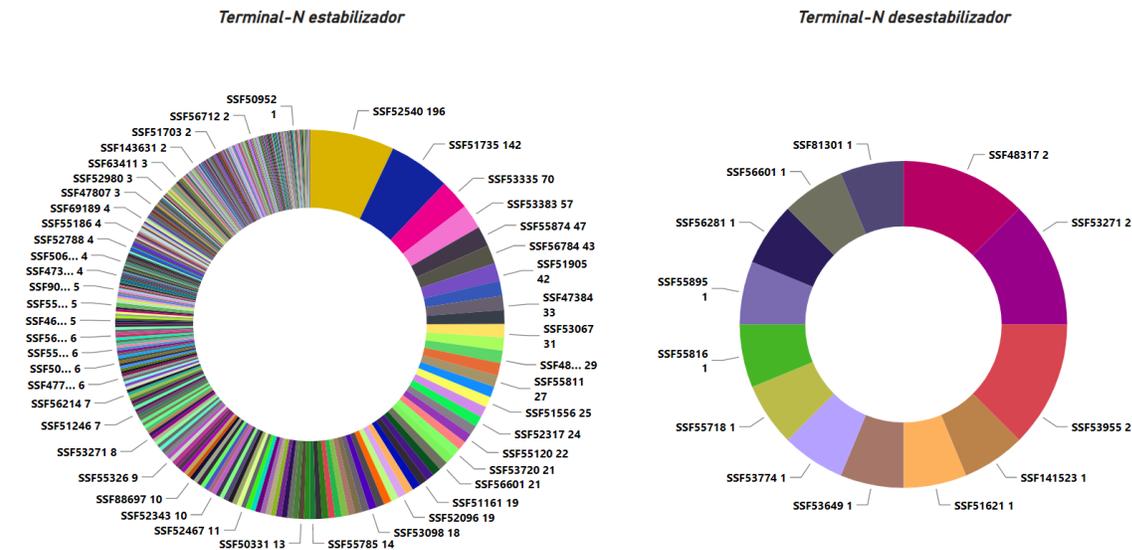


Figura 75. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para *Escherichia coli*. A la izquierda se ilustra el número de enzimas, por cada superfamilia, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada superfamilia, con un terminal-N desestabilizador.

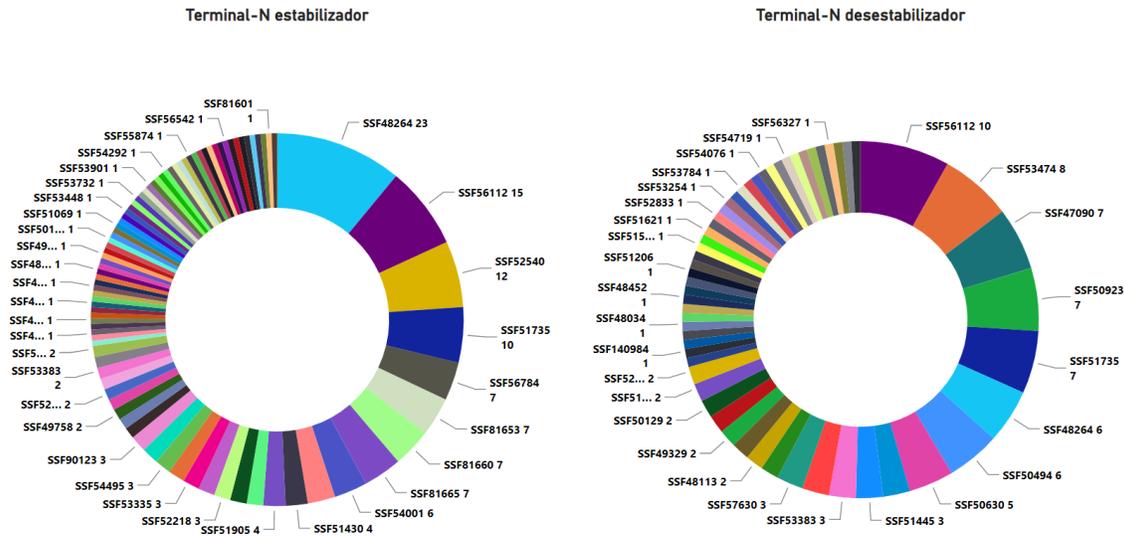


Figura 76. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para *Oryctolagus cuniculus*. A la izquierda se ilustra el número de enzimas, por cada superfamilia, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada superfamilia, con un terminal-N desestabilizador.

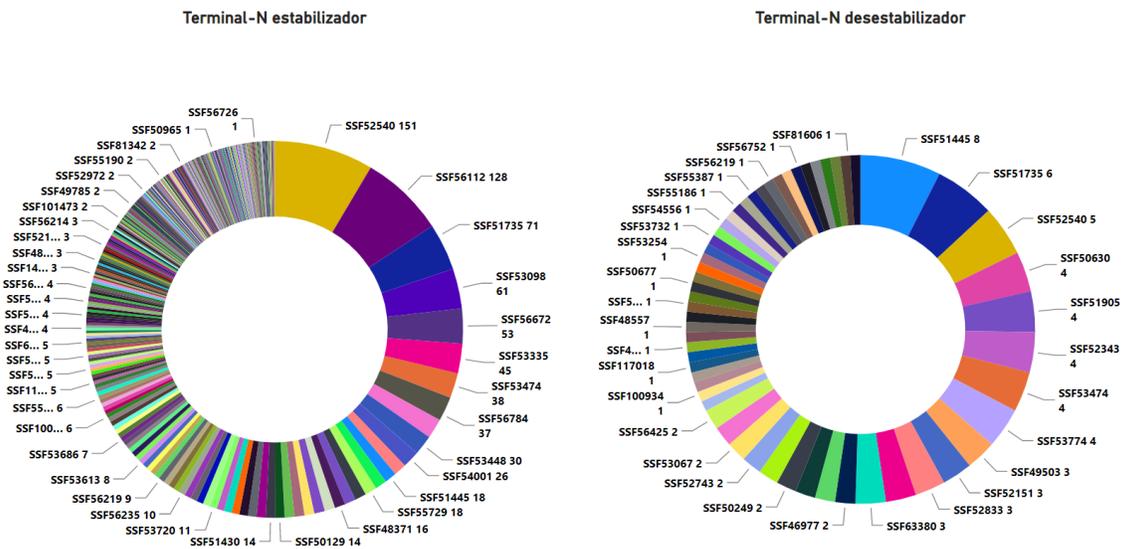


Figura 77. Conteo de enzimas (con terminal-N estabilizador y desestabilizador) por cada superfamilia para *Saccharomyces cerevisiae*. A la izquierda se ilustra el número de enzimas, por cada superfamilia, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada superfamilia, con un terminal-N desestabilizador.

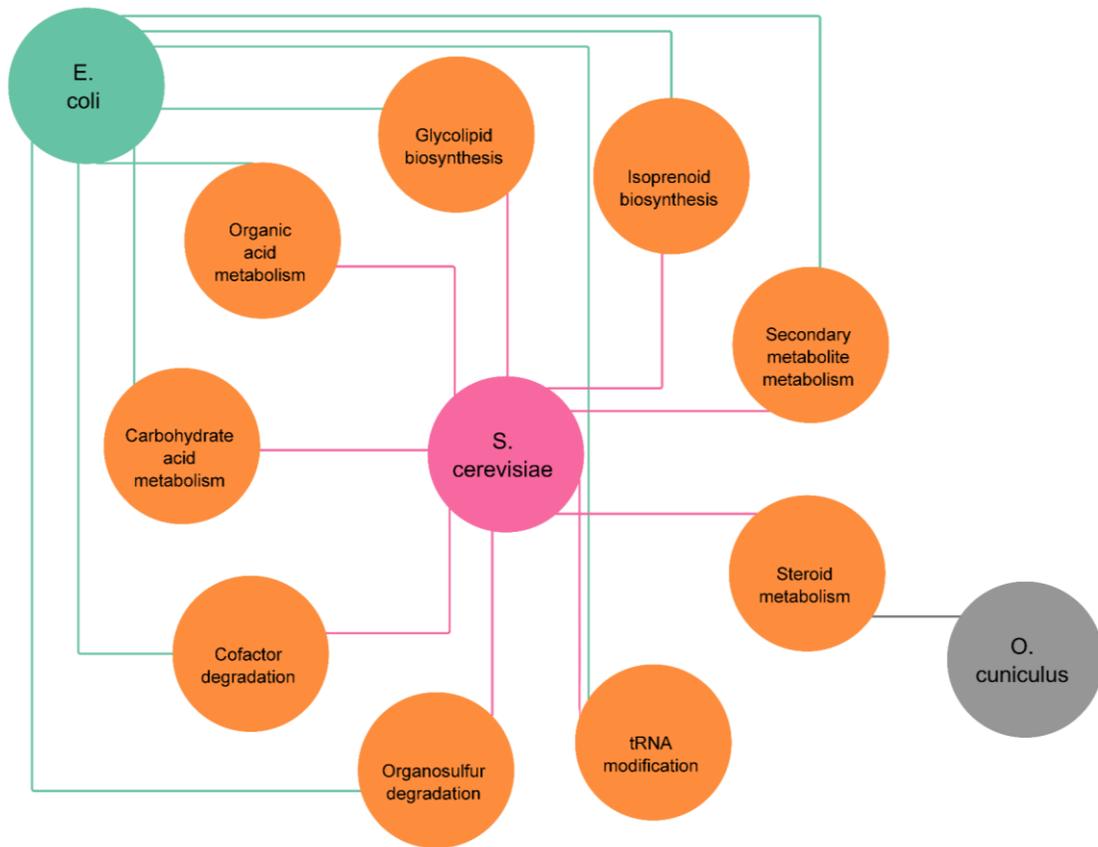


Figura 78. Rutas metabólicas que coinciden solo con los proteomas *Escherichia coli*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus*. En la gráfica se observa las rutas metabólicas que están asociadas a solo dos de los cuatro proteomas.

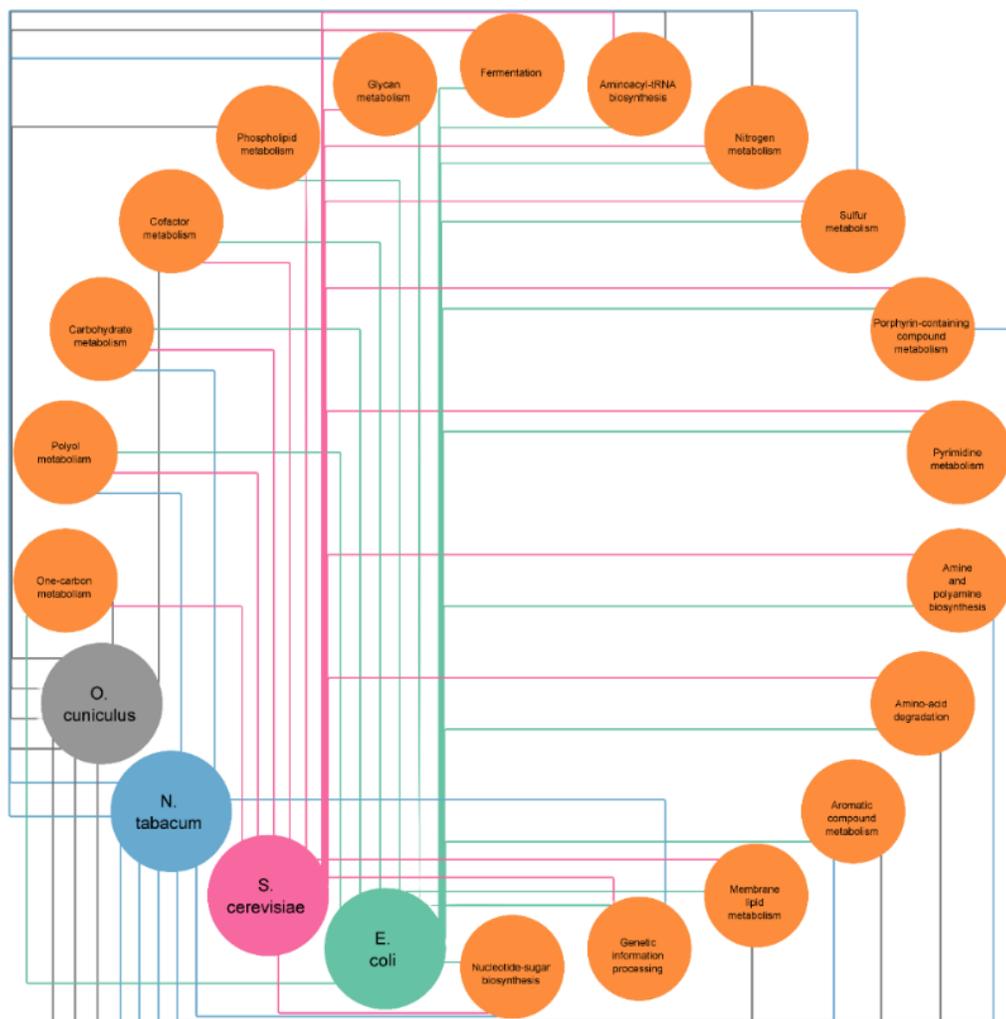


Figura 79. Rutas metabólicas que coinciden solo con los proteomas *Escherichia coli*, *Saccharomyces cerevisiae* y *Oryctolagus cuniculus*. En la gráfica se observa las rutas metabólicas que están asociadas a solo tres de los cuatro proteomas.

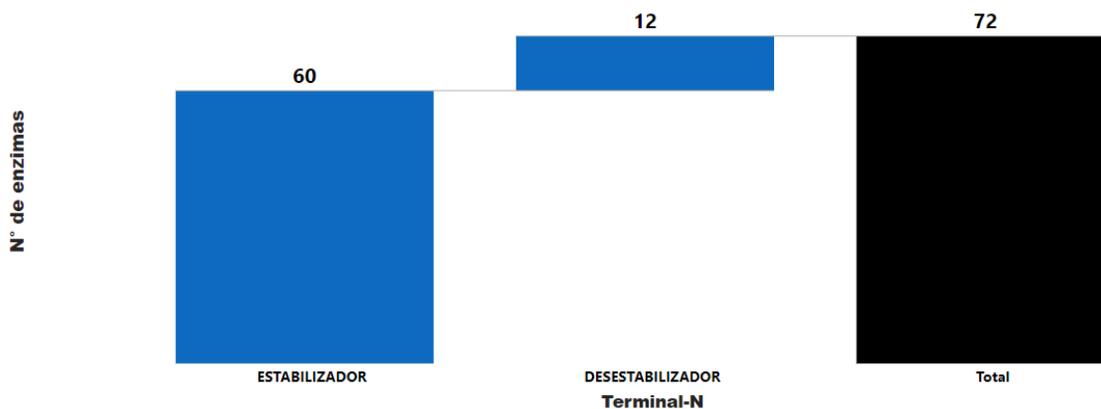


Figura 80. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N destabilizador para *Nicotiana tabacum* que participan en alguna ruta metabólica. De las 72 enzimas que están caracterizadas como enzimas con tiempo de vida estable o inestable, 60 se clasifican como enzimas con tiempos de vida estable y 12 como enzimas con tiempos de vida inestable.

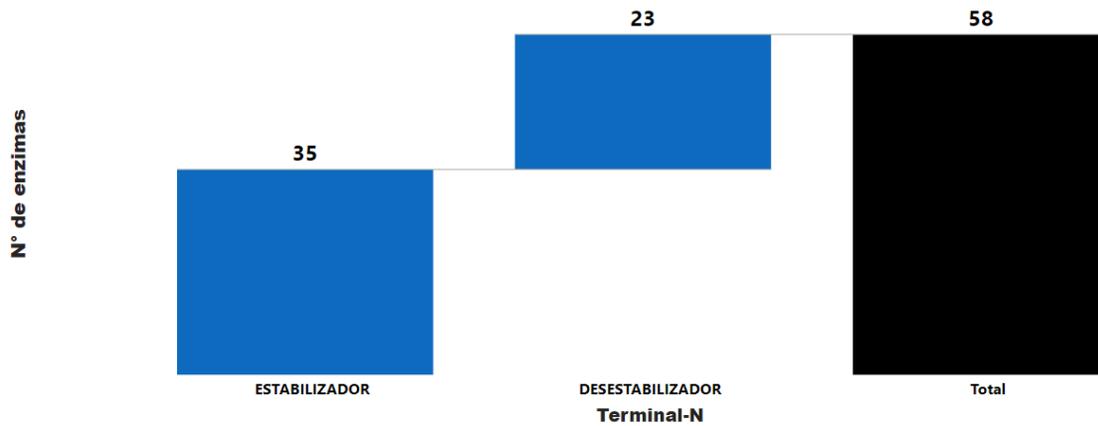


Figura 81. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N destabilizador para *Oryctolagus cuniculus* que participan en alguna ruta metabólica. De las 58 enzimas que están caracterizadas como enzimas con tiempo de vida estable o inestable, 35 se clasifican como enzimas con tiempos de vida estable y 23 como enzimas con tiempos de vida inestable.

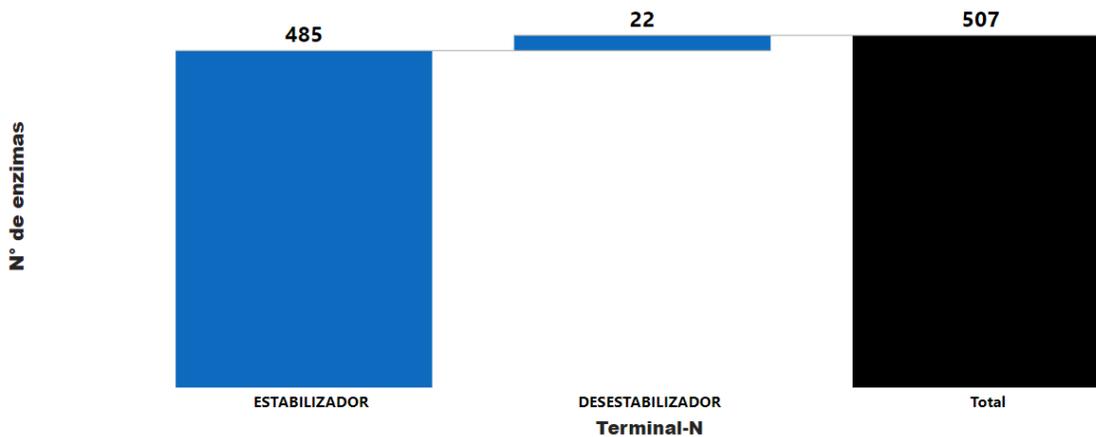


Figura 82. Conteo de enzimas que tienen un terminal-N estabilizador y un terminal-N destabilizador para *Saccharomyces cerevisiae* que participan en alguna ruta metabólica. En el *Saccharomyces cerevisiae*, de las 507 enzimas que están caracterizadas como enzimas con tiempo de vida estable o inestable, 485 se clasifican como enzimas con tiempos de vida estable y 22 como enzimas con tiempos de vida inestable.

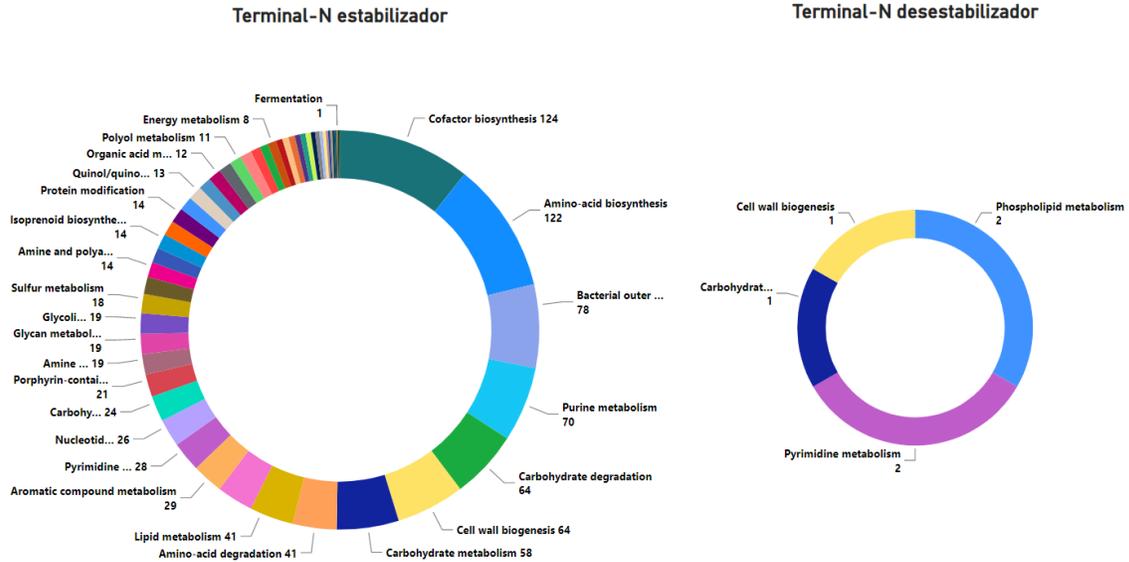


Figura 83. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en *Escherichia coli*. A la izquierda se ilustra el número de enzimas, por cada ruta metabólica para el proteoma *Escherichia coli*, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada ruta metabólica para el mismo proteoma, con un terminal-N desestabilizador.

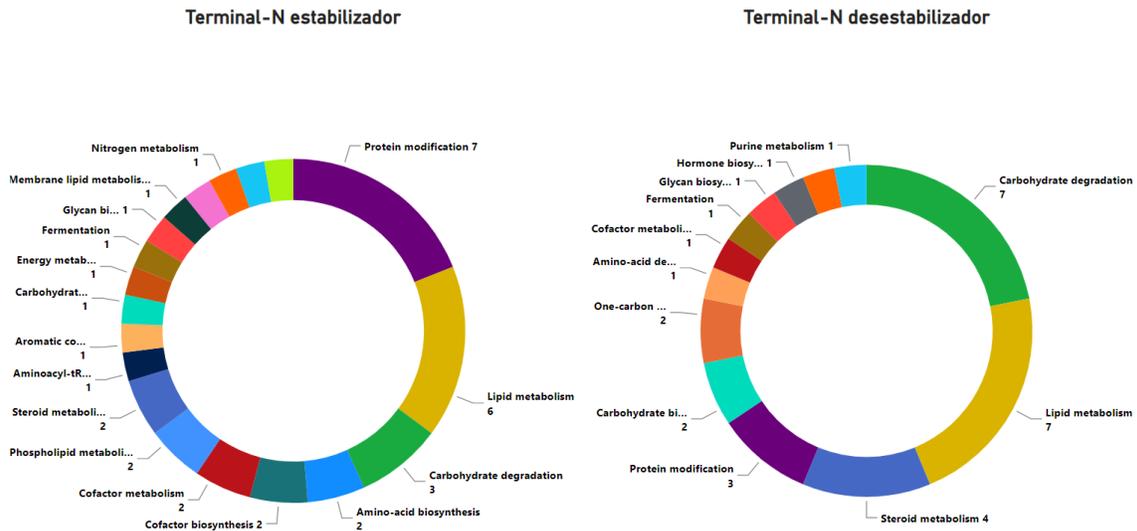


Figura 84. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en *Oryctolagus cuniculus*. A la izquierda se ilustra el número de enzimas, por cada ruta metabólica para el proteoma *Oryctolagus cuniculus*, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada ruta metabólica para el mismo proteoma, con un terminal-N desestabilizador.

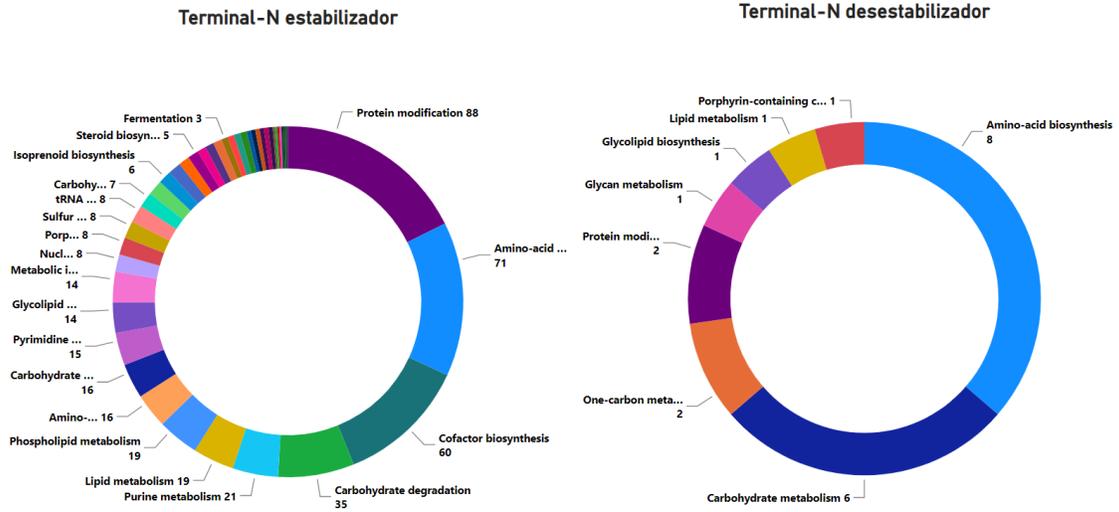


Figura 85. Conteo de proteínas (con terminal-N estabilizador y desestabilizador) que participan en una ruta metabólica en *Saccharomyces cerevisiae*. A la izquierda se ilustra el número de enzimas, por cada ruta metabólica para el proteoma *Saccharomyces cerevisiae*, con un terminal-N estabilizador y a la derecha se ilustra el número de enzimas, por cada ruta metabólica para el mismo proteoma, con un terminal-N desestabilizador.

TABLAS

Tabla 8. Proteomas con las proteínas que tienen más de una actividad enzimática.

PROTEOMAS							
<i>Escherichia coli</i>		<i>Saccharomyces cerevisiae</i>		<i>Oryctolagus cuniculus</i>		<i>Nicotiana tabacum</i>	
Nodo	Aristas	Nodo	Aristas	Nodo	Aristas	Nodo	Aristas
P77399	3	P07149	4	P00187	2	Q42942	2
P0AEA8	3	P08566	3	P53761	2	P10978	2
P21177	3	P09440	3	P05176	2	Q42965	2
Q8XCP2	3	P07245	3	P56201	2	x	x
Q8X8I2	3	O13527	2	Q9MZ03	2	x	x
P0AEA9	3	P0C2J7	2	Q95216	2	x	x
P33644	3	P17255	2	P35574	2	x	x
P00561	2	Q99315	2	P34088	2	x	x
P31119	2	P0CX64	2	P00939	2	x	x
P0ADI0	2	Q03144	2	P00187	2	x	x
P15639	2	Q12362	2	x	x	x	x
P23837	2	O13535	2	x	x	x	x
P00904	2	P0C2I2	2	x	x	x	x
Q8XCT6	2	P38088	2	x	x	x	x
Q8X9C6	2	P38086	2	x	x	x	x
P06961	2	P19097	2	x	x	x	x
P18956	2	Q12491	2	x	x	x	x
P77182	2	P25384	2	x	x	x	x
P0A9J9	2	P38295	2	x	x	x	x
Q8X6J8	2	P40471	2	x	x	x	x
P31658	2	P53848	2	x	x	x	x
P0ADI5	2	P47098	2	x	x	x	x
P31806	2	Q12472	2	x	x	x	x
P0AFB5	2	P36156	2	x	x	x	x
Q8X8Y6	2	Q04214	2	x	x	x	x
P58237	2	P53318	2	x	x	x	x
Q8X739	2	P48239	2	x	x	x	x
Q9S5G5	2	Q12113	2	x	x	x	x
P09835	2	Q12193	2	x	x	x	x
P64148	2	Q12196	2	x	x	x	x
P75726	2	P0C2J1	2	x	x	x	x
Q8XDZ3	2	P38009	2	x	x	x	x
P0ABK5	2	P25373	2	x	x	x	x
P0ADI4	2	P0C2I5	2	x	x	x	x

Q8X611	2	Q07793	2	x	x	x	x
P05523	2	P31378	2	x	x	x	x
P11071	2	Q12269	2	x	x	x	x
P60595	2	Q07791	2	x	x	x	x
P00562	2	P07259	2	x	x	x	x
P0ABQ0	2	P47024	2	x	x	x	x
Q7AD14	2	Q01217	2	x	x	x	x
P77444	2	Q06137	2	x	x	x	x
Q8X7B7	2	Q12501	2	x	x	x	x
P0AC33	2	P15436	2	x	x	x	x
Q8X4P8	2	Q02891	2	x	x	x	x
P25539	2	Q06817	2	x	x	x	x
P0A6Q4	2	Q12316	2	x	x	x	x
P77398	2	P53823	2	x	x	x	x
P10378	2	P53397	2	x	x	x	x
P0A6Q3	2	Q12141	2	x	x	x	x
P0AG24	2	P54113	2	x	x	x	x
P77455	2	Q06625	2	x	x	x	x
P25960	2	P17695	2	x	x	x	x
Q8XBL4	2	Q04670	2	x	x	x	x
P02919	2	P32361	2	x	x	x	x
P27249	2	P15807	2	x	x	x	x
P06987	2	Q02207	2	x	x	x	x
P0AES0	2	P33734	2	x	x	x	x
P02918	2	Q08214	2	x	x	x	x
P0A9J8	2	P0CX63	2	x	x	x	x
Q8X607	2	P0C2I6	2	x	x	x	x
P07023	2	Q05902	2	x	x	x	x
P03014	2	Q04711	2	x	x	x	x
Q8XBV3	2	P40353	2	x	x	x	x
P0AFB6	2	P32528	2	x	x	x	x
P24186	2	Q12088	2	x	x	x	x
P0AC16	2	Q92393	2	x	x	x	x
P00909	2	P0C2I9	2	x	x	x	x
P50465	2	P43544	2	x	x	x	x
Q8XCQ7	2	Q03612	2	x	x	x	x
Q46925	2	Q7LHG5	2	x	x	x	x
P77170	2	Q03494	2	x	x	x	x
x	x	Q12490	2	x	x	x	x
x	x	P47100	2	x	x	x	x
x	x	Q12112	2	x	x	x	x
x	x	P00815	2	x	x	x	x

x	x	Q12273	2	x	x	x	x
x	x	Q99337	2	x	x	x	x
x	x	P53323	2	x	x	x	x
x	x	Q99231	2	x	x	x	x
x	x	P0C2J3	2	x	x	x	x
x	x	Q03434	2	x	x	x	x
x	x	P0C2I3	2	x	x	x	x
x	x	Q12414	2	x	x	x	x
x	x	Q03855	2	x	x	x	x
x	x	P21951	2	x	x	x	x
x	x	P0C2J0	2	x	x	x	x
x	x	Q03619	2	x	x	x	x
x	x	P53208	2	x	x	x	x

Tabla 9. Terminal-N presentes en las superfamilias proteicas.

Superfamilia con enzimas que tienen terminal-N estabilizador y desestabilizador	Superfamilia con enzimas que tienen solo terminal-N estabilizador	Superfamilia con enzimas que tienen solo terminal-N desestabilizador
SSF100920	SSF100879	SSF101912
SSF100934	SSF100939	SSF117018
SSF100950	SSF101224	SSF47862
SSF101353	SSF101307	SSF48619
SSF110921	SSF101473	SSF50923
SSF140984	SSF101478	SSF54076
SSF141523	SSF101690	SSF54556
SSF46609	SSF101790	SSF55387
SSF46977	SSF101904	SSF56266
SSF47090	SSF102220	SSF57016
SSF47240	SSF102400	SSF57440
SSF47323	SSF102462	SSF57630
SSF48034	SSF102546	x
SSF48113	SSF102588	x
SSF48179	SSF102645	x
SSF48208	SSF102735	x
SSF48225	SSF102860	x
SSF48239	SSF102886	x
SSF48264	SSF103084	x
SSF48317	SSF103190	x

SSF48452	SSF103243	x
SSF48557	SSF103263	x
SSF48576	SSF103365	x
SSF48613	SSF103378	x
SSF49329	SSF103501	x
SSF49503	SSF103515	x
SSF49723	SSF109998	x
SSF49899	SSF110019	x
SSF50022	SSF110581	x
SSF50129	SSF110738	x
SSF50249	SSF110849	x
SSF50324	SSF110857	x
SSF50447	SSF110942	x
SSF50494	SSF110997	x
SSF50630	SSF111278	x
SSF50677	SSF111283	x
SSF50800	SSF111321	x
SSF50891	SSF111331	x
SSF51069	SSF111337	x
SSF51126	SSF117281	x
SSF51206	SSF117457	x
SSF51351	SSF118196	x
SSF51366	SSF140356	x
SSF51430	SSF140490	x
SSF51445	SSF140586	x
SSF51556	SSF140931	x
SSF51604	SSF140959	x
SSF51621	SSF140990	x
SSF51735	SSF141259	x
SSF51905	SSF141734	x
SSF52129	SSF141868	x
SSF52151	SSF141986	x
SSF52210	SSF142019	x
SSF52343	SSF142433	x
SSF52467	SSF142695	x
SSF52518	SSF142754	x
SSF52540	SSF143011	x
SSF52743	SSF143081	x
SSF52833	SSF143243	x
SSF52935	SSF143517	x
SSF53056	SSF143548	x
SSF53067	SSF143631	x

SSF53254	SSF143791	x
SSF53271	SSF143856	x
SSF53335	SSF143975	x
SSF53383	SSF143990	x
SSF53448	SSF144005	x
SSF53474	SSF144206	x
SSF53649	SSF158668	x
SSF53732	SSF159065	x
SSF53774	SSF159709	x
SSF53784	SSF160350	x
SSF53955	SSF160443	x
SSF54001	SSF46458	x
SSF54197	SSF46548	x
SSF54211	SSF46585	x
SSF54427	SSF46589	x
SSF54719	SSF46626	x
SSF54826	SSF46689	x
SSF54975	SSF46767	x
SSF55186	SSF46785	x
SSF55205	SSF46894	x
SSF55315	SSF46915	x
SSF55424	SSF46929	x
SSF55718	SSF46934	x
SSF55816	SSF46955	x
SSF55856	SSF47005	x
SSF55895	SSF47060	x
SSF55931	SSF47203	x
SSF56112	SSF47212	x
SSF56219	SSF47226	x
SSF56281	SSF47336	x
SSF56327	SSF47370	x
SSF56425	SSF47384	x
SSF56601	SSF47413	x
SSF56672	SSF47473	x
SSF56752	SSF47598	x
SSF63380	SSF47616	x
SSF63737	SSF47644	x
SSF69572	SSF47648	x
SSF74650	SSF47672	x
SSF75217	SSF47741	x
SSF81296	SSF47769	x
SSF81301	SSF47781	x

SSF81464	SSF47802	x
SSF81606	SSF47807	x
SSF88697	SSF47819	x
x	SSF47831	x
x	SSF48013	x
x	SSF48019	x
x	SSF48024	x
x	SSF48108	x
x	SSF48150	x
x	SSF48163	x
x	SSF48168	x
x	SSF48173	x
x	SSF48256	x
x	SSF48310	x
x	SSF48371	x
x	SSF48403	x
x	SSF48435	x
x	SSF48484	x
x	SSF48600	x
x	SSF48695	x
x	SSF48726	x
x	SSF49265	x
x	SSF49303	x
x	SSF49309	x
x	SSF49472	x
x	SSF49599	x
x	SSF49758	x
x	SSF49764	x
x	SSF49785	x
x	SSF49879	x
x	SSF49998	x
x	SSF50037	x
x	SSF50156	x
x	SSF50331	x
x	SSF50341	x
x	SSF50443	x
x	SSF50465	x
x	SSF50486	x
x	SSF50615	x
x	SSF50621	x
x	SSF50685	x
x	SSF50692	x

x	SSF50715	x
x	SSF50814	x
x	SSF50952	x
x	SSF50965	x
x	SSF50974	x
x	SSF50978	x
x	SSF50998	x
x	SSF51045	x
x	SSF51055	x
x	SSF51101	x
x	SSF51161	x
x	SSF51182	x
x	SSF51230	x
x	SSF51246	x
x	SSF51261	x
x	SSF51278	x
x	SSF51283	x
x	SSF51294	x
x	SSF51306	x
x	SSF51316	x
x	SSF51338	x
x	SSF51391	x
x	SSF51419	x
x	SSF51645	x
x	SSF51658	x
x	SSF51679	x
x	SSF51690	x
x	SSF51695	x
x	SSF51703	x
x	SSF51713	x
x	SSF51717	x
x	SSF51726	x
x	SSF51730	x
x	SSF52009	x
x	SSF52021	x
x	SSF52029	x
x	SSF52096	x
x	SSF52113	x
x	SSF52121	x
x	SSF52141	x
x	SSF52156	x
x	SSF52172	x

x	SSF52218	x
x	SSF52242	x
x	SSF52255	x
x	SSF52279	x
x	SSF52317	x
x	SSF52335	x
x	SSF52413	x
x	SSF52418	x
x	SSF52425	x
x	SSF52440	x
x	SSF52499	x
x	SSF52507	x
x	SSF52728	x
x	SSF52733	x
x	SSF52738	x
x	SSF52768	x
x	SSF52788	x
x	SSF52794	x
x	SSF52799	x
x	SSF52821	x
x	SSF52913	x
x	SSF52922	x
x	SSF52949	x
x	SSF52972	x
x	SSF52980	x
x	SSF53032	x
x	SSF53041	x
x	SSF53062	x
x	SSF53092	x
x	SSF53098	x
x	SSF53155	x
x	SSF53163	x
x	SSF53167	x
x	SSF53178	x
x	SSF53218	x
x	SSF53244	x
x	SSF53300	x
x	SSF53323	x
x	SSF53328	x
x	SSF53590	x
x	SSF53597	x
x	SSF53613	x

x	SSF53623	x
x	SSF53633	x
x	SSF53639	x
x	SSF53671	x
x	SSF53681	x
x	SSF53686	x
x	SSF53720	x
x	SSF53738	x
x	SSF53743	x
x	SSF53748	x
x	SSF53790	x
x	SSF53901	x
x	SSF53927	x
x	SSF54111	x
x	SSF54160	x
x	SSF54236	x
x	SSF54292	x
x	SSF54368	x
x	SSF54416	x
x	SSF54495	x
x	SSF54585	x
x	SSF54593	x
x	SSF54626	x
x	SSF54637	x
x	SSF54665	x
x	SSF54680	x
x	SSF54690	x
x	SSF54782	x
x	SSF54791	x
x	SSF54909	x
x	SSF54913	x
x	SSF54919	x
x	SSF54980	x
x	SSF54991	x
x	SSF55003	x
x	SSF55008	x
x	SSF55031	x
x	SSF55035	x
x	SSF55040	x
x	SSF55048	x
x	SSF55052	x
x	SSF55060	x

x	SSF55068	x
x	SSF55073	x
x	SSF55083	x
x	SSF55103	x
x	SSF55120	x
x	SSF55124	x
x	SSF55144	x
x	SSF55154	x
x	SSF55166	x
x	SSF55190	x
x	SSF55234	x
x	SSF55248	x
x	SSF55257	x
x	SSF55261	x
x	SSF55298	x
x	SSF55326	x
x	SSF55383	x
x	SSF55447	x
x	SSF55469	x
x	SSF55594	x
x	SSF55604	x
x	SSF55608	x
x	SSF55658	x
x	SSF55666	x
x	SSF55729	x
x	SSF55785	x
x	SSF55804	x
x	SSF55811	x
x	SSF55821	x
x	SSF55826	x
x	SSF55831	x
x	SSF55846	x
x	SSF55874	x
x	SSF55890	x
x	SSF55904	x
x	SSF55920	x
x	SSF55957	x
x	SSF55973	x
x	SSF56003	x
x	SSF56042	x
x	SSF56176	x
x	SSF56194	x

x	SSF56204	x
x	SSF56214	x
x	SSF56228	x
x	SSF56235	x
x	SSF56276	x
x	SSF56317	x
x	SSF56322	x
x	SSF56349	x
x	SSF56420	x
x	SSF56507	x
x	SSF56512	x
x	SSF56519	x
x	SSF56524	x
x	SSF56529	x
x	SSF56534	x
x	SSF56542	x
x	SSF56553	x
x	SSF56634	x
x	SSF56645	x
x	SSF56712	x
x	SSF56719	x
x	SSF56726	x
x	SSF56741	x
x	SSF56762	x
x	SSF56784	x
x	SSF56821	x
x	SSF56925	x
x	SSF56931	x
x	SSF57667	x
x	SSF57756	x
x	SSF57770	x
x	SSF57829	x
x	SSF57884	x
x	SSF57903	x
x	SSF63411	x
x	SSF63418	x
x	SSF63562	x
x	SSF63862	x
x	SSF63867	x
x	SSF63882	x
x	SSF63892	x
x	SSF63999	x

x	SSF64005	x
x	SSF64153	x
x	SSF64158	x
x	SSF64167	x
x	SSF64182	x
x	SSF64268	x
x	SSF64288	x
x	SSF64356	x
x	SSF64438	x
x	SSF64586	x
x	SSF68906	x
x	SSF68912	x
x	SSF68923	x
x	SSF69000	x
x	SSF69012	x
x	SSF69055	x
x	SSF69065	x
x	SSF69075	x
x	SSF69189	x
x	SSF69336	x
x	SSF69500	x
x	SSF69618	x
x	SSF69721	x
x	SSF69742	x
x	SSF69765	x
x	SSF69864	x
x	SSF69917	x
x	SSF74778	x
x	SSF74863	x
x	SSF75005	x
x	SSF75138	x
x	SSF75169	x
x	SSF75304	x
x	SSF81271	x
x	SSF81342	x
x	SSF81442	x
x	SSF81452	x
x	SSF81496	x
x	SSF81601	x
x	SSF81648	x
x	SSF81653	x
x	SSF81660	x

x	SSF81665	x
x	SSF81767	x
x	SSF81799	x
x	SSF81822	x
x	SSF81886	x
x	SSF81935	x
x	SSF82051	x
x	SSF82109	x
x	SSF82114	x
x	SSF82215	x
x	SSF82282	x
x	SSF82697	x
x	SSF82771	x
x	SSF82784	x
x	SSF88688	x
x	SSF88713	x
x	SSF88723	x
x	SSF89095	x
x	SSF89550	x
x	SSF89562	x
x	SSF89623	x
x	SSF89733	x
x	SSF89796	x
x	SSF90123	x

Tabla 10. Terminales-N presentes en las rutas metabólicas.

Rutas metabólicas en las que participan solo enzimas con terminal-N estabilizador	Rutas metabólicas en las que participan enzimas con terminal-N estabilizador y desestabilizador	Rutas metabólicas en las que participan solo enzimas con terminal-N desestabilizador
<i>Phenylpropanoid metabolism</i>	<i>Glycan biosynthesis</i>	<i>Hormone biosynthesis</i>
<i>Cell wall degradation</i>	<i>Protein modification</i>	<i>Carotenoid biosynthesis</i>
<i>Nucleoside biosynthesis</i>	<i>Lipid metabolism</i>	x
<i>Aromatic compound metabolism</i>	<i>Metabolic intermediate biosynthesis</i>	x
<i>Organosulfur degradation</i>	<i>Porphyrin-containing compound metabolism</i>	x

<i>Amine and polyamine biosynthesis</i>	<i>Carbohydrate metabolism</i>	X
<i>Aminoacyl-tRNA biosynthesis</i>	<i>Carbohydrate degradation</i>	X
<i>Steroid biosynthesis</i>	<i>One-carbon metabolism</i>	X
<i>Alkaloid biosynthesis</i>	<i>Secondary metabolite biosynthesis</i>	X
<i>Phytoalexin biosynthesis</i>	<i>Purine metabolism</i>	X
<i>Sphingolipid metabolism</i>	<i>Fermentation</i>	X
<i>Glycolipid metabolism</i>	<i>Cell wall biogenesis</i>	X
<i>Organic acid metabolism</i>	<i>Steroid metabolism</i>	X
<i>Bacterial outer membrane biogenesis</i>	<i>Phospholipid metabolism</i>	X
<i>Secondary metabolite metabolism</i>	<i>Pyrimidine metabolism</i>	X
<i>Glycan degradation</i>	<i>Amino-acid degradation</i>	X
<i>Nitrogen metabolism</i>	<i>Energy metabolism</i>	X
<i>Genetic information processing</i>	<i>Cofactor metabolism</i>	X
<i>Plant hormone biosynthesis</i>	<i>Glycolipid biosynthesis</i>	X
<i>Photosynthesis</i>	<i>Amino-acid biosynthesis</i>	X
<i>Alkene biosynthesis</i>	<i>Carbohydrate biosynthesis</i>	X
<i>Terpene metabolism</i>	<i>Glycan metabolism</i>	X
<i>Cofactor degradation</i>	X	X
<i>Membrane lipid metabolism</i>	X	X
<i>Sulfur metabolism</i>	X	X
<i>Quinol/quinone metabolism</i>	X	X
<i>Exopolysaccharide biosynthesis</i>	X	X
<i>Metabolic intermediate degradation</i>	X	X
<i>Steroid hormone biosynthesis</i>	X	X
<i>Slime biogenesis</i>	X	X
<i>Flavonoid metabolism</i>	X	X
<i>Siderophore biosynthesis</i>	X	X
<i>Metabolic intermediate metabolism</i>	X	X
<i>Carbohydrate acid metabolism</i>	X	X
<i>Amino-sugar metabolism</i>	X	X

<i>Polyol metabolism</i>	X	X
<i>Glycerolipid metabolism</i>	X	X
<i>Cofactor biosynthesis</i>	X	X
<i>Isoprenoid biosynthesis</i>	X	X
<i>tRNA modification</i>	X	X
<i>Amine and polyamine degradation</i>	X	X
<i>Protein biosynthesis</i>	X	X
<i>Lipopolysaccharide metabolism</i>	X	X
<i>Amine and polyamine metabolism</i>	X	X
<i>Nucleotide-sugar biosynthesis</i>	X	X
<i>Alcohol metabolism</i>	X	X

Tabla 11. Información referente al virus VIH-1

Organismo	VIH-1
N° de proteínas	9
N° de proteínas con estatus de revisadas	9
Actividad catalítica	1
EC	EC:2: EC:3
Terminal-N	G
Estabilidad	Estable
Superfamilia	SSF46919; SSF47836; SSF47943; SSF50122; SSF50630; SSF53098; SSF56672; SSF57756
Rutas metabólicas	x

Tabla 12. Tiempos de tiempo de vida de la proteína β -galactosidasa, con los diferentes Terminal-N, en distintos organismos.

Aminoácido	Conejo	Método	Levadura	Método	Bacterias	Método
	<i>(Oryctolagus cuniculus)</i>		<i>(Saccharomyces cerevisiae)</i>		<i>(Escherichia coli)</i>	
	(min)		(min)		(min)	
Ala (A)	264	Unión lineal de aminoácidos	> 1200	Unión lineal de aminoácidos	> 600	Unión lineal de aminoácidos
Cys (C)	72		> 1200		> 600	

Asp (D)	66	a la proteína β -galactosidasa usando muta	3	a la proteína β -galactosidasa usando muta	> 600	a la proteína β -galactosidasa usando muta
Glu (E)	60	génesis en el gen que	30	génesis en el gen que	> 600	génesis en el gen que
Phe (F)	66	codifica la fusión de Ub en levadura.	3	codifica la fusión de Ub en levadura.	2	codifica la fusión de Ub en levadura.
Gly (G)	1800		>1200		> 600	
His (H)	210		10		> 600	
Ile (I)	1200		30		> 600	
Lys (K)	78		3		2	
Leu (L)	90		3		2	
Met (M)	1800		> 1200		> 600	
Asn (N)	84		3		> 600	
Pro (P)	> 1200		> 1200		x	
Gln (Q)	48		10		> 600	
Arg (R)	60		2		2	
Ser (S)	114		> 1200		> 600	
Thr (T)	432		> 1200		> 600	
Val (V)	6000		> 1200		> 600	
Trp (W)	168		3		2	
Tyr (Y)	168		10		2	
Referencia	(Gonda et al., 1989)		(Bachmair et al., 1986; Varshavsky, 1992)		(Tobias et al., 1991; Varshavsky, 1992)	