الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعـليـم العالي والـبحث الـعلمي

Ministère de l'enseignement supérieur et de la recherche scientifiq

جامعة الشهيد حمة لخضر الوادي

Université Echahid Hamma Lakdhar EL-OUED

كلية العلوم الطبيعة والحياة

Faculté des Sciences de la Nature et de la Vie

قسم البيولوجيا الخلوية والجزيئية

Département de Biologie Cellulaire et Moléculaire

Mèmoire de fin d'ètude

En vue de l'obtention du diplôme de Master Académique en Sciences biologiques

Spécialité : Biochimie appliquée

THÈME

منهجية حاسوبية صارمة لتحليل مركبات المجموعة المرضية 3 المرتبطة باضطرابات النواقل العصبية والخلل الميتوكوندري ضمن إطار الشبكة الصيدلانية

Présenté Par :CHIHI HOUDA

Devant le jury composé de :

Président : Mr.Zaater Abdelmalek	**M.C.A**	**Université d'El Oued.**
Examinateur :Mr.Laiche Khaled	**M.A.A**	**Université d'El Oued**
Promoteur : Mr. ALLALI AHMED	**(Pr)**	**Université d'El Oued**

Année universitaire : 2025/2026

إهـــــداء

بسم الله الرحمان الرحيم

اللهم صل وسلم على سيدنا محمد خاتم الأنبياء والمرسلين

اهدي عملي هذا الى اغلى ما في الوجود

الى ابي الغالي ..رحمة الله عليه

والى معنى الحب والحنان ومن كان دعائها سر نجاحي

امي الحبيبة اطال الله في عمرها الى اخوتي واخواتي

سندي وعزوتي الى صديقاتي اللواتي شاركني

لحظات التعب والفرح خاصة الدكتورة يوسفي خيرة التي لم تبخل عليا باي جهد

الى كل أساتذتي من الابتدائي الى الجامعي

الى كل من علمني حرفا الى كل باحث عن العلم والمعرفة

اهديكم ثمره جهدي هذا

الطالبة: هدى شيحي

السنة الجامعية 2026/2025

شكـــــر وعرفـــــان

الحمد والشكر لله الحي القيوم أولا وأخيرا وامتثالا لقوله صل الله عليه وسلم

من لا يشكر الناس لا يشكر الله

أتوجه بجزيل الشكر وجميل العرفان للدكتور علالي احمد الذي تكرم بقبول الاشراف على هذه المذكرة

وعلى جميع التوجيهات والملاحظات والنصائح.

كما لا يفوتنا ان نتقدم بوافر التقدير والاحترام لاعضاء اللجنة المحترمين على عناء قراءتها

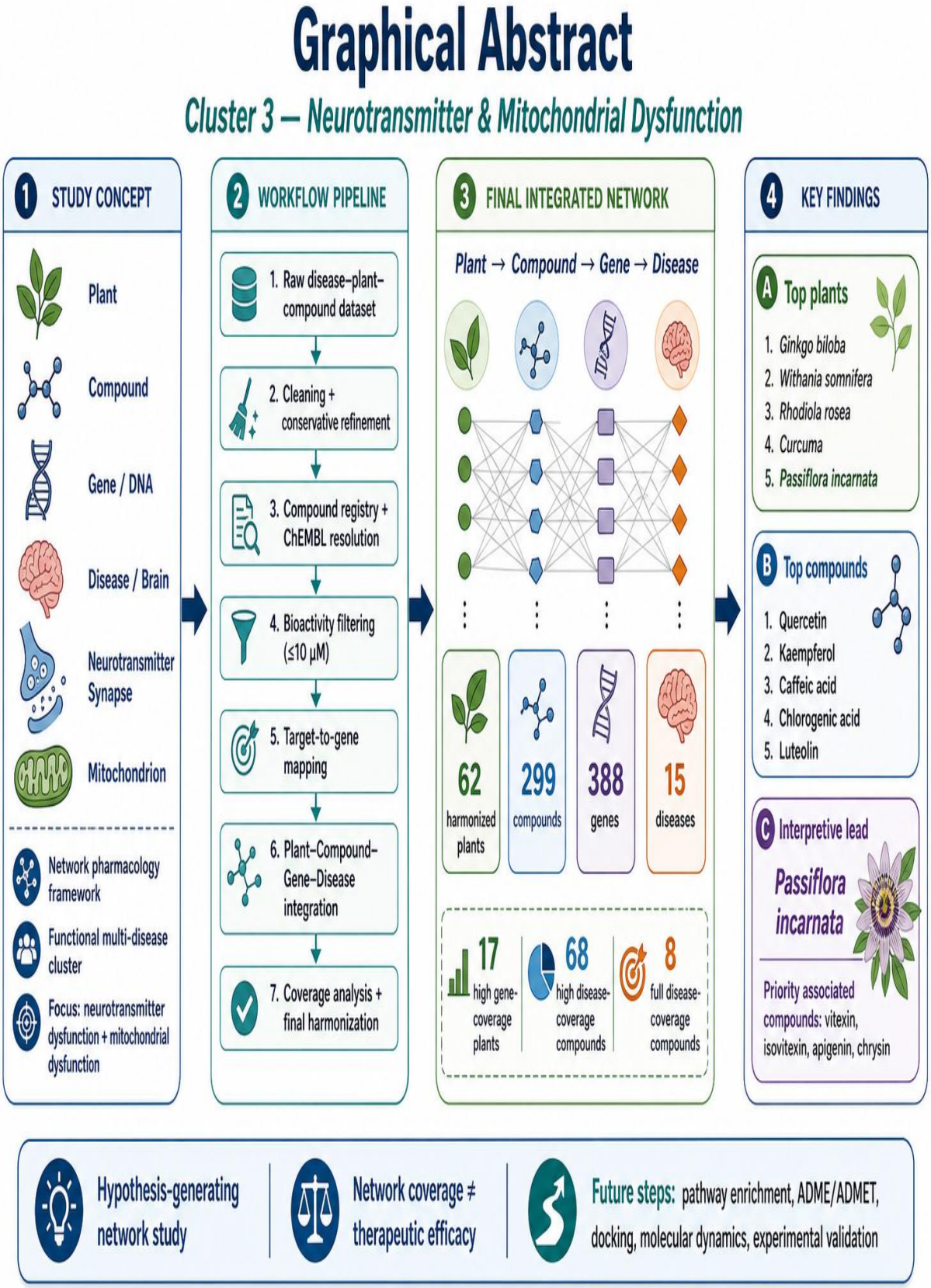
وقبولها وتصويبها

وكذالك نتقدم بخالص الشكر الى كل من درسنا من أساتذة كلية علوم الطبيعة والحياة بجامعة الشهيد حمة الأخضر

وفي الأخير نشكر كل من قدم لنا يد العون والمساعدة من قريب او من بعيد

ونسال الله عز وجل ان يجعل ذالك في ميزان حسناتهم انه قريب مجيب

Figure GA.1. Graphical abstract of the Cluster 3 network pharmacology workflow and corrected final findings



ملخص:

تهدف هذه الدراسة إلى تحليل Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction باعتباره مجموعة أمراض وظيفية متعددة، وليس مرضًا منفردًا أو تصنيفًا سريريًا رسميًا. اعتمدت الدراسة على منهجية network pharmacology لبناء شبكة متعددة الطبقات تربط بين النباتات، المركبات الكيميائية، الجينات، والأمراض وفق نموذج Plant–Compound–Gene–Disease. شملت المنهجية فحص البيانات الخام، تنظيف الإدخالات، بناء سجل للمركبات، مطابقة المركبات مع قاعدة ChEMBL، استخراج بيانات النشاط الحيوي، فلترة العلاقات، ربط الأهداف بالجينات، ثم دمج الشبكة وتحليل مؤشرات التغطية.

بعد المراجعة التصحيحية النهائية في Stage 10E — Final Passiflora Harmonization، تم حذف الإدخالات غير النباتية من جداول النباتات، مثل المركبات والمستخلصات والصيغ العشبية المركبة، كما تم توحيد الصيغ المرتبطة بـ Passiflora incarnata. أظهرت النتائج النهائية شبكة تضم 62 كيانًا نباتيًا موحّدًا، 299 مركبًا، 388 جينًا، و15 مرضًا. كما حدد التحليل 17 نباتًا بتغطية جينية مرتفعة، و68 مركبًا بتغطية مرضية عالية، إضافة إلى 8 مركبات ذات تغطية مرضية كاملة داخل Cluster 3.

أظهرت النتائج أن الشبكة لا تعتمد على نبات واحد أو مركب واحد أو جين واحد، بل تمتلك بنية موزعة متعددة الطبقات. وقد برزت نباتات معروفة مثل Ginkgo biloba، في حين مثّل ظهور Passiflora incarnata نتيجة ذات قيمة تفسيرية خاصة بسبب ارتباطه المحتمل بمحور النواقل العصبية، خصوصًا النظام GABAergic. كما أظهرت طبقة المركبات حضور مركبات واسعة التمثيل مثل quercetin، مع أهمية مستقبلية لمركبات Passiflora مثل vitexin, isovitexin, apigenin, و.chrysin

تؤكد الدراسة أن مؤشرات network coverage تمثل أولوية بحثية وتوليدًا للفرضيات، ولا تعني فعالية علاجية مباشرة. ولذلك، توصي الدراسة بمراحل لاحقة تشمل تحليل الحساسية للأوزان، تحليل تداخل الجينات بين الأمراض، pathway enrichment، ADME/ADMET، molecular docking، molecular dynamics، والتحقق التجريبي.

الكلمات المفتاحية: **Cluster 3 Network pharmacology؛ ChEMBL؛ المركبات النباتية؛ خلل الميتوكوندريا؛ النواقل العصبية؛ Passiflora incarnata؛ التغطية الشبكية.**

Résumé

Cette étude vise à analyser le Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction en tant que groupe fonctionnel multi-maladies, et non comme une maladie unique ou une classification clinique officielle. L'approche adoptée repose sur le network pharmacology afin de construire un réseau multicouche reliant les plantes, les composés chimiques, les gènes et les maladies selon le modèle Plant–Compound–Gene–Disease. La méthodologie a inclus le contrôle des données brutes, le nettoyage des entrées, la construction d'un registre de composés, la résolution des composés via ChEMBL, l'extraction des données de bioactivité, le filtrage des relations, le mapping des cibles vers les gènes, puis l'intégration du réseau et l'analyse de la couverture.

Après la correction finale réalisée lors du Stage 10E — Final Passiflora Harmonization, les entrées non végétales ont été supprimées des tableaux de plantes, notamment les composés, les extraits et les formules herbacées complexes. Les variantes associées à Passiflora incarnata ont également été harmonisées. Le réseau final corrigé comprend 62 entités végétales harmonisées, 299 composés, 388 gènes et 15 maladies. L'analyse a identifié 17 plantes présentant une couverture génique élevée, 68 composés avec une forte couverture des maladies, ainsi que 8 composés couvrant l'ensemble des maladies du Cluster 3.

Les résultats montrent que le réseau ne repose pas sur une seule plante, un seul composé ou un seul gène, mais sur une structure distribuée et multicouche. Des plantes bien connues comme Ginkgo biloba apparaissent parmi les entités les mieux classées, tandis que Passiflora incarnata constitue une contribution interprétative importante en raison de son lien potentiel avec la modulation des neurotransmetteurs, notamment l'axe GABAergique. La couche des composés met également en évidence des composés largement représentés tels que quercetin, tout en suggérant un intérêt futur pour des composés associés à Passiflora, tels que vitexin, isovitexin, apigenin et chrysin.

Cette étude souligne que la network coverage doit être interprétée comme un indicateur de priorité de recherche et de génération d'hypothèses, et non comme une preuve directe d'efficacité thérapeutique. Des analyses futures sont nécessaires, notamment une analyse de sensibilité des pondérations, une analyse du chevauchement gène–maladie, un enrichissement des voies biologiques, l'ADME/ADMET, le docking moléculaire, la dynamique moléculaire et la validation expérimentale.

Mots-clés : Cluster 3 ; Network pharmacology ; ChEMBL ; composés végétaux ; dysfonction mitochondriale ; neurotransmetteurs ; Passiflora incarnata ; couverture réseau.

Abstract

This study aimed to analyze Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction as a multi-disease functional group rather than a single disease or an official clinical classification. A network pharmacology approach was applied to build a multilayer network linking plants, chemical compounds, genes, and diseases according to a Plant–Compound–Gene–Disease framework. The workflow included raw data inspection, conservative data cleaning, compound registry construction, ChEMBL compound resolution, bioactivity extraction, evidence filtering, target-to-gene mapping, network integration, and coverage-based analysis. Following the final corrective step, Stage 10E — Final Passiflora Harmonization, non-plant entries were removed from plant tables, including compounds, extracts, and multi-herbal formulas. Different textual variants related to Passiflora incarnata were also harmonized. The corrected final network included 62 harmonized plant entities, 299 compounds, 388 genes, and 15 diseases. The analysis identified 17 plants with high gene coverage, 68 compounds with high disease coverage, and 8 compounds with full disease coverage across all diseases included in Cluster 3.

The results indicate that the network is not driven by a single plant, compound, or gene, but rather by a distributed multilayer structure. Well-known entities such as Ginkgo biloba and quercetin were captured by the network, supporting methodological consistency. However, the emergence of Passiflora incarnata represents a particularly relevant interpretive finding because of its potential association with neurotransmitter modulation, especially the GABAergic axis. In addition, Passiflora-related compounds such as vitexin, isovitexin, apigenin, and chrysin may represent more specific candidates for future investigation than broadly studied compounds alone.

The study emphasizes that network coverage should be interpreted as a research-prioritization and hypothesis-generating indicator, not as direct therapeutic evidence. Future work should include weight sensitivity analysis, gene–disease overlap analysis, pathway enrichment, ADME/ADMET evaluation, molecular docking, molecular dynamics simulations, and experimental validation.

Keywords: Cluster 3; Network pharmacology; ChEMBL; plant compounds; mitochondrial dysfunction; neurotransmitters; Passiflora incarnata; network coverage.

مقـــدمـــــة

1.1.مقدمة عامة

تُعد الاضطرابات العصبية والوظيفية من أكثر الحالات البيولوجية تعقيدًا، لأنها لا ترتبط غالبًا بآلية مرضية واحدة أو هدف جزيئي منفرد، بل تنتج عن تداخل عدة مسارات خلوية وجزيئية. إذ لا تقتصر هذه الاضطرابات على مسار مرضي واحد، بل تنطوي على شبكة مترابطة من التفاعلات الجزيئية التي تعجز المقاربات التقليدية أحادية الهدف عن استيعاب تعقيداتها. ومن بين هذه المسارات، يبرز خلل النواقل العصبية واضطراب وظائف الميتوكوندريا كعاملين مهمين في عدد من الأمراض ذات الطابع العصبي أو النفسي أو التنكسي [1]، [2].

تقوم النواقل العصبية بدور محوري في تنظيم الاتصال بين الخلايا العصبية، بينما تمثل الميتوكوندريا المصدر الرئيسي للطاقة الخلوية ومنظمًا مهمًا للإجهاد التأكسدي وموت الخلايا المبرمج. ولذلك، فإن أي اضطراب في هذين المحورين قد ينعكس على وظائف الدماغ، الاتزان العصبي، الاستجابة الالتهابية، واستقرار الخلايا العصبية [2]، [3].

في المقابل، تحتوي النباتات الطبية على عدد واسع من المركبات الكيميائية الحيوية التي قد تتفاعل مع عدة أهداف جزيئية في الوقت نفسه. غير أن دراسة هذه المركبات بمنطق "مركب واحد–هدف واحد" قد لا تكون كافية لفهم تعقيدها البيولوجي. لذلك، ظهرت منهجية network pharmacology كأداة مناسبة لدراسة العلاقات متعددة الطبقات بين النباتات، المركبات، الجينات، والأمراض [4]. وتكمن قوة هذا النهج في تكامله الرأسي والأفقي: بدءاً بتحليل النسخ الجيني لتحديد الوحدات المرضية، مروراً بتحليل الشبكات لاستخلاص الأهداف الحرجة، وانتهاءً بالفطْر الافتراضي والتحقق بالرسو الجزيئي، وكلها خطوات يمكن تطبيقها بصرامة لاستكشاف مركبات المجموعة المرضية 3 التي تعمل على تقاطع الخلل الميتوكوندري واضطراب النواقل العصبية، مما يمهد الطريق لفهم أعمق للأمراض وتطوير استراتيجيات علاجية متعددة الأهداف وأكثر فعالية وأقل سمية.

في هذا السياق، تركز هذه الدراسة على Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction باعتباره مجموعة أمراض وظيفية متعددة، وليس تصنيفًا سريريًا رسميًا أو دراسة موجهة نحو مرض واحد. يهدف هذا الإطار إلى تحليل التداخلات المحتملة بين النباتات، المركبات، الجينات، والأمراض من منظور شبكي، مع الالتزام بعدم الخلط بين التغطية الشبكية والفعالية العلاجية.

1.2. خلفية الدراسة

تعتمد الدراسة الحالية على فكرة أن بعض الأمراض، رغم اختلافها السريري، قد تتقاطع في محاور بيولوجية مشتركة. وفي حالة Cluster 3، تتمثل هذه المحاور أساسًا في خلل النواقل العصبية، خلل الميتوكوندريا، الإجهاد التأكسدي، الالتهاب العصبي، اضطراب proteostasis، والمسارات المرتبطة بموت الخلايا المبرمج.

لا يعني هذا أن جميع الأمراض داخل Cluster 3 تمتلك نفس السبب المرضي أو نفس المسار البيولوجي، بل يعني أن هناك إمكانية لدراسة هذه الأمراض ضمن إطار وظيفي مشترك يسمح بالكشف عن نقاط التداخل الشبكي.

من جهة أخرى، تمثل النباتات الطبية مصدرًا مهمًا للمركبات الطبيعية متعددة الأهداف. فقد تحتوي النبتة الواحدة على عدة مركبات، كما يمكن للمركب الواحد أن يوجد في عدة نباتات وأن يرتبط بعدة أهداف جينية. لذلك، فإن تحليل هذه العلاقات يتطلب إطارًا يتجاوز العلاقة الخطية التقليدية بين نبات ومرض.

تعتمد الدراسة على بناء شبكة من نوع Plant–Compound–Gene–Disease بهدف تحليل مدى تغطية النباتات والمركبات والجينات لأمراض Cluster 3، مع ترتيب الكيانات حسب مؤشرات شبكية قابلة للتفسير. وتُعامل هذه المؤشرات باعتبارها أدوات ترتيب وأولوية بحثية، وليست مقاييس فعالية علاجية.

1.3. إشكالية الدراسة

رغم تزايد الاهتمام بالمركبات النباتية في دراسة الاضطرابات العصبية والوظيفية، لا تزال العلاقات بين النباتات، مركباتها الكيميائية، الأهداف الجينية، ومجموعات الأمراض المرتبطة بخلل النواقل العصبية والميتوكوندريا غير منظمة بشكل كافٍ ضمن إطار شبكي واضح. فكثير من الدراسات تميل إلى التركيز على مرض واحد، أو نبات واحد، أو مركب واحد، مما يحد من القدرة على فهم التداخلات متعددة الأمراض ومتعددة الأهداف.

في المقابل، فإن Cluster 3 يمثل مجموعة أمراض وظيفية يُحتمل أن تتقاطع في محاور بيولوجية متعددة، مثل اضطراب النواقل العصبية، خلل الميتوكوندريا، الإجهاد التأكسدي، الالتهاب العصبي، واضطراب proteostasis. غير أن هذا التداخل لا يمكن فهمه بشكل دقيق من خلال المقاربات الخطية التقليدية، لأن العلاقة بين النبات والمركب والجين والمرض غالبًا ما تكون متعددة الطبقات.

من جهة أخرى، فإن اعتماد قواعد بيانات مثل ChEMBL يوفر فرصة لبناء روابط بين المركبات والأهداف الحيوية، لكنه في الوقت نفسه يتطلب تنظيفًا صارمًا للبيانات، وتمييزًا واضحًا بين المركبات الدقيقة، الفئات الكيميائية العامة، والمطابقات غير المؤكدة. كما أن ارتفاع تغطية مركب أو جين داخل الشبكة لا يعني بالضرورة فعالية علاجية، بل قد يعكس فقط كثافة التمثيل في قواعد البيانات أو الأدبيات.

من هنا تنطلق إشكالية هذه الدراسة في السؤال المركزي الآتي:

كيف يمكن بناء وتحليل شبكة متعددة الطبقات تربط النباتات والمركبات والجينات والأمراض داخل **Cluster** 3، بما يسمح بتحديد أنماط التغطية الشبكية والأولويات البحثية، مع تجنب الخلط بين **network coverage** والفعالية العلاجية؟

وتسعى هذه الدراسة إلى الإجابة عن هذه الإشكالية من خلال تطبيق pipeline منظم يعتمد على تنظيف البيانات، مطابقة المركبات، استخراج النشاط الحيوي، ربط الأهداف بالجنيات، ثم تحليل التغطية والترتيب الشبكي.

1.4. أسئلة الدراسة

انطلاقًا من الإشكالية السابقة، تسعى هذه الدراسة إلى الإجابة عن الأسئلة الآتية:

- ما حجم شبكة Plant–Compound–Gene–Disease الناتجة عن تحليل Cluster 3؟
- ما النباتات ذات أعلى تغطية شبكية داخل Cluster 3؟
- ما المركبات ذات التغطية الواسعة عبر أمراض Cluster 3؟
- ما الجينات ذات أعلى degree-like coverage داخل الشبكة؟
- هل تعتمد شبكة Cluster 3 على نبات أو مركب أو جين واحد، أم أنها تمتلك بنية موزعة متعددة الكيانات؟
- كيف يمكن التمييز بين network coverage وtherapeutic efficacy عند تفسير النتائج؟
- ما حدود الاعتماد على قواعد بيانات مثل ChEMBL في بناء شبكة متعددة الطبقات؟

1.5. فرضية الدراسة

تفترض هذه الدراسة أن أمراض Cluster 3، رغم اختلافها السريري والوظيفي، قد تُظهر تداخلًا شبكيًا على مستوى النباتات، المركبات، والجينات. كما تفترض أن منهجية network pharmacology قادرة على كشف طبقات تغطية مشتركة داخل هذه المجموعة المرضية، من خلال بناء شبكة متعددة الطبقات تربط Plant–Compound–Gene–Disease.

وتفترض الدراسة أيضًا أن أعلى الكيانات ترتيبًا داخل الشبكة لا تمثل بالضرورة عناصر علاجية مثبتة، بل تمثل مرشحات بحثية ذات تغطية شبكية مرتفعة يمكن اختبارها لاحقًا من خلال تحليلات إضافية مثل pathway enrichment، ADME/ADMET،docking ، والتحقق التجريبي.

بناءً على ذلك، فإن الفرضية العامة للدراسة هي:

يمكن لتحليل **network pharmacology** أن يكشف بنية شبكية موزعة داخل **Cluster** 3، تسمح بتحديد نباتات ومركبات وجينات ذات أولوية بحثية، دون اعتبار هذه الأولوية دليلًا مباشرًا على الفعالية العلاجية.

1.6. أهداف الدراسة

1.6.1. الهدف العام

تهدف هذه الدراسة إلى بناء وتحليل شبكة متعددة الطبقات داخل Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction تربط بين النباتات، المركبات الكيميائية، الجينات، والأمراض، وذلك باستخدام منهجية network pharmacology بهدف تحديد أنماط التغطية الشبكية والأولويات البحثية داخل هذه المجموعة المرضية الوظيفية.

1.6.2. الأهداف الخاصة

- جمع وتنظيم بيانات النباتات، المركبات، والأمراض المرتبطة بـ Cluster 3.
- تنظيف بيانات المركبات وتمييز المركبات الواضحة عن المستخلصات، الخلطات، والفئات الكيميائية العامة.
- بناء سجل مركبات موحد اعتمادًا على مفاتيح مطابقة معيارية.
- ترتيب المركبات حسب درجة تمثيلها في النباتات والأمراض.
- مطابقة المركبات ذات الأولوية مع قاعدة بيانات ChEMBL.
- استخراج بيانات النشاط الحيوي وربط المركبات بالأهداف البيولوجية البشرية.
- فلترة علاقات المركب–الهدف اعتمادًا على معايير نشاط حيوي صارمة.
- تحويل أهداف ChEMBL إلى رموز جينية قابلة للتفسير.
- بناء شبكة متكاملة من نوع Plant–Compound–Gene–Disease.
- حساب مؤشرات التغطية والترتيب على مستوى النباتات، المركبات، والجينات.
- تحليل فئات التغطية لتجنب التركيز على أعلى نبات أو مركب فقط.
- إنتاج جداول وأشكال تساعد على عرض النتائج بطريقة كمية وبصرية.
- التمييز منهجيًا بين network coverage و therapeutic efficacy.
- تحديد حدود النتائج واقتراح خطوات مستقبلية للتحقق منها.

Table 1.1. Summary of the general and specific objectives of the study.

المحتوى	نوع الهدف
بناء وتحليل شبكة متعددة الطبقات داخل Cluster 3 تربط النباتات، المركبات، الجينات، والأمراض.	الهدف العام
جمع وتنظيف وتوحيد بيانات النباتات والمركبات والأمراض.	تنظيم البيانات
مطابقة المركبات مع ChEMBL واستخراج النشاط الحيوي وربطه بالأهداف البشرية.	بناء الأدلة الحيوية
بناء شبكة Plant–Compound–Gene–Disease وحساب مؤشرات التغطية والترتيب.	التحليل الشبكي
تمييز network coverage عن therapeutic efficacy وتحديد حدود النتائج وخطوات التحقق المستقبلية.	التحقق المنهجي

1.7. أهمية الدراسة

تكتسب هذه الدراسة أهميتها من كونها تجمع بين التحليل النباتي–الكيميائي والتحليل الجيني–الشبكي ضمن إطار واحد، بهدف فهم العلاقات المحتملة بين النباتات، المركبات، الجينات، ومجموعة الأمراض المرتبطة بخلل النواقل العصبية والميتوكوندريا. ولا تقتصر أهمية الدراسة على تحديد نباتات أو مركبات معينة، بل تمتد إلى بناء نموذج تحليلي يسمح بفهم التداخلات متعددة الطبقات داخل Cluster 3.

1.7.1. الأهمية العلمية

تساهم الدراسة في تنظيم بيانات متعددة المصادر ضمن شبكة واحدة من نوع Plant–Compound–Gene–Disease. ويسمح هذا التنظيم بفهم أوسع للتداخلات المحتملة بين الأمراض داخل Cluster 3، بدل التعامل مع كل مرض بمعزل عن الآخر. كما تتيح الدراسة تحديد الكيانات ذات التمثيل الشبكي المرتفع، سواء على مستوى النباتات أو المركبات أو الجينات.

1.7.2. الأهمية المنهجية

تتميز الدراسة باعتماد pipeline مرحلي قابل للتتبع، يشمل فحص البيانات، تنظيفها، مطابقة المركبات، استخراج النشاط الحيوي، mapping الجيني، ثم تحليل التغطية والترتيب. كما تم اعتماد ملفات QC وlogs في كل مرحلة، مما يعزز قابلية إعادة التنفيذ والمراجعة.

1.7.3. الأهمية التطبيقية

توفر الدراسة قائمة أولية من النباتات والمركبات والجينات ذات الأولوية البحثية، والتي يمكن استخدامها لاحقًا في مراحل أكثر تخصصًا مثل pathway enrichment، ADME/ADMET، molecular docking، molecular dynamics، وexperimental validation. وبذلك، تمثل النتائج نقطة انطلاق لاختيار مرشحات بحثية قابلة للاختبار، لا نتائج علاجية نهائية.

1.7.4. الأهمية النقدية

تكمن أهمية الدراسة أيضًا في حرصها على التمييز بين network coverage وtherapeutic efficacy. فارتفاع ترتيب نبات أو مركب أو جين داخل الشبكة لا يُفسر كدليل علاجي، بل كمؤشر تغطية شبكية يحتاج إلى تحقق لاحق. ويساعد هذا التمييز على تجنب المبالغة في تفسير نتائج network pharmacology.

1.8. حدود الدراسة

رغم أن هذه الدراسة تعتمد على منهجية منظمة وقابلة للتتبع، إلا أنها تبقى دراسة حاسوبية استكشافية، وليست دراسة تجريبية أو سريرية. لذلك، تم تحديد حدودها منذ البداية لتجنب المبالغة في تفسير النتائج.

- Cluster • 3 إطار تحليلي وظيفي وليس تصنيفًا سريريًا رسميًا.
- مخرجات الشبكة تُستخدم لتحديد أولويات بحثية، ولا تمثل دليلًا مباشرًا على الفعالية العلاجية.
- لم يتم تقييم ADME/ADMET مثل الامتصاص، التوافر الحيوي، السمية، أو عبور الحاجز الدموي الدماغي.
- لم يتم تنفيذ molecular docking أو molecular dynamics لاختبار الارتباط الجزيئي بين المركبات والأهداف البروتينية.
- لم تُجرَ تجارب مخبرية أو خلوية أو حيوانية للتحقق من الفرضيات الناتجة.
- قد تتأثر النتائج بتحيز قواعد البيانات، كثافة النشر، وتوفر البيانات حول مركبات أو أهداف معينة.
- تم الاحتفاظ ببعض non-exact ChEMBL mappings مع وسمها، مما يمثل مصدرًا محتملًا لعدم اليقين.

1.9. بنية المذكرة

تم تنظيم هذه المذكرة في ستة فصول رئيسية، إضافة إلى المراجع والملحقات، كما يلي:

الفصل	المحتوى العام
الفصل الأول	المقدمة العامة، الإشكالية، وأهداف الدراسة.
الفصل الثاني	الإطار النظري حول network pharmacology، المركبات النباتية، خلل النواقل العصبية، خلل الميتوكوندريا، والإجهاد التأكسدي والالتهاب العصبي.
الفصل الثالث	المواد والطرق، ويعرض منهجية تنظيف البيانات، مطابقة ChEMBL، استخراج النشاط الحيوي، mapping الجيني، دمج الشبكة، وتحليل التغطية.
الفصل الرابع	النتائج، ويقدم المخرجات الرقمية والبصرية، بما في ذلك حجم الشبكة، ترتيب النباتات، المركبات، الجينات، وفئات التغطية.
الفصل الخامس	المناقشة، ويحلل النتائج في ضوء network pharmacology وpolypharmacology وحدود قواعد البيانات.
الفصل السادس	الحدود، الاستنتاجات، والتوصيات، ويعرض الخلاصة العامة والخطوات المستقبلية.

حوصلة للفصل الأول

بناءً على ما سبق، تنطلق هذه الدراسة من الحاجة إلى مقاربة شبكية قادرة على تنظيم العلاقات بين النباتات، المركبات، الجينات، والأمراض داخل Cluster 3. ويُعد هذا العمل محاولة لبناء إطار تحليلي متعدد الطبقات يسمح بفهم أنماط

التغطية الشبكية وتحديد الأولويات البحثية، مع الالتزام الواضح بعدم الخلط بين نتائج network coverage والفعالية العلاجية. وتمثل الفصول اللاحقة امتدادًا لهذا الإطار، بدءًا من الخلفية النظرية، مرورًا بالمنهجية، وصولًا إلى النتائج والمناقشة والاستنتاجات.

الفصـــل النـــظري

2.1. تمهيد الفصل

يعرض هذا الفصل الإطار النظري الذي تستند إليه دراسة Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction. وبما أن الدراسة لا تركز على مرض واحد منفرد، بل على مجموعة أمراض وظيفية متعددة، فإن هذا الفصل يهدف إلى توضيح المفاهيم البيولوجية والمنهجية التي تسمح بفهم هذا التجميع من منظور شبكي [1]–[3].

يركز الفصل على المحاور البيولوجية الأساسية المرتبطة بـ Cluster 3، خاصة خلل الميتوكوندريا، اضطراب النواقل العصبية، الإجهاد التأكسدي، الالتهاب العصبي، واضطراب توازن البروتينات. كما يتناول المفاهيم المرتبطة بـ network pharmacology، والمركبات النباتية متعددة الأهداف، وتحليل الشبكات الحيوية.

يمثل هذا الإطار النظري قاعدة لفهم الفصول اللاحقة، خصوصًا فصل المواد والطرق، حيث يتم بناء شبكة Plant–Compound–Gene–Disease، وكذلك فصل النتائج والمناقشة، حيث يتم تفسير مؤشرات التغطية الشبكية دون الخلط بينها وبين الفعالية العلاجية.

2.2. مفهوم Network Pharmacology

تُعد network pharmacology مقاربة حديثة تجمع بين علم الأدوية، البيولوجيا الحاسوبية، وتحليل الشبكات الحيوية. تهدف هذه المقاربة إلى فهم تأثير المركبات الكيميائية من خلال شبكة من العلاقات بين المركبات، الأهداف الجزيئية، الجينات، المسارات البيولوجية، والأمراض، بدل الاقتصار على نموذج خطي بسيط يقوم على علاقة مركب واحد بهدف واحد [1]–[4].

يعتمد هذا المفهوم على أن الأنظمة البيولوجية، خصوصًا في الأمراض المعقدة، لا تعمل عادة عبر مسار واحد أو هدف منفرد، بل عبر تداخل عدد كبير من البروتينات، الجينات، والإشارات الخلوية. لذلك، فإن تحليل التأثيرات الدوائية أو النباتية يتطلب النظر إلى العلاقات المتعددة بين الكيانات المختلفة داخل شبكة واحدة.

في سياق النباتات الطبية، تكتسب network pharmacology أهمية خاصة، لأن النبات الواحد قد يحتوي على عشرات أو مئات المركبات، وقد يرتبط كل مركب بعدة أهداف حيوية. لذلك، تسمح هذه المقاربة بدراسة النبات كنظام كيميائي متعدد المكونات، بدل اختزاله في مركب واحد فقط.

ومع ذلك، يجب التأكيد أن نتائج network pharmacology لا تمثل دليلًا علاجيًا مباشرًا، بل تُعد نتائج مولدة للفرضيات. فارتفاع تغطية مركب أو نبات أو جين داخل الشبكة يعني أنه واسع التمثيل في البيانات، ولا يعني بالضرورة أنه فعال علاجيًا.

2.3. من نموذج Single-Target إلى Multi-Target

اعتمد علم الأدوية التقليدي تاريخيًا على نموذج Single-Target، الذي يفترض أن المركب الدوائي يؤثر بشكل أساسي على هدف جزيئي واحد محدد، والذي بدوره يرتبط بمرض معين. ورغم نجاح هذا النموذج في بعض الحالات، إلا أنه يواجه محدودية واضحة عند التعامل مع الأمراض المعقدة، خصوصًا الاضطرابات العصبية والوظيفية [3]–[6].

في المقابل، يعتمد نموذج Multi-Target Pharmacology على فكرة أن المركب أو مجموعة المركبات يمكن أن تؤثر في عدة أهداف جزيئية ومسارات بيولوجية في آن واحد. ويصبح هذا النموذج أكثر ملاءمة عند دراسة الأمراض متعددة العوامل والأنظمة البيولوجية المعقدة والمركبات النباتية متعددة المكونات.

تُعد المركبات النباتية مثالًا واضحًا لهذا النموذج، حيث يحتوي النبات الواحد على عدة مركبات، وقد يرتبط كل مركب بعدة أهداف جينية. كما يمكن أن تتفاعل هذه المركبات معًا بطريقة تآزرية، مما يؤدي إلى تأثيرات بيولوجية لا يمكن تفسيرها من خلال مركب منفرد.

في هذا السياق، توفر منهجية network pharmacology إطارًا مناسبًا لدراسة هذه العلاقات متعددة المستويات من خلال شبكة Plant–Compound–Gene–Disease، مع التأكيد أن تعدد الأهداف لا يعني بالضرورة فعالية علاجية مثبتة.

2.4. المركبات النباتية في علم الأدوية

تُعد النباتات مصدرًا غنيًا للمركبات الكيميائية الحيوية التي تمتلك خصائص بيولوجية متنوعة، وقد استُخدمت منذ القدم في العلاجات التقليدية. في السياق الحديث، أصبحت هذه المركبات محل اهتمام كبير في علم الأدوية، خاصة عند دراستها ضمن إطار network pharmacology [5]–[8].

تحتوي النباتات على عدة فئات من المركبات الثانوية مثل flavonoids، alkaloids، terpenoids ، phenolic acids، glycosides، saponins ، وtannins. تختلف هذه المركبات في بنيتها الكيميائية، خواصها الفيزيائية، وآليات تأثيرها البيولوجي.

تتميز العديد من المركبات النباتية بقدرتها على التفاعل مع أكثر من هدف جزيئي داخل الخلية، وهي خاصية تُعرف باسم polypharmacology. وتمثل هذه الخاصية أحد الأسس التي تجعل المركبات الطبيعية مناسبة لدراسة الأمراض متعددة العوامل.

في العديد من الحالات، لا تعمل المركبات النباتية بشكل مستقل، بل قد تتفاعل معًا داخل نفس المستخلص النباتي. يمكن أن يؤدي هذا التفاعل إلى تعزيز التأثير البيولوجي أو تعديل استجابة الهدف، غير أن تحليل هذه التفاعلات يمثل تحديًا لأن البيانات المتاحة غالبًا تكون على مستوى المركب الفردي.

رغم الإمكانات الكبيرة للمركبات النباتية، إلا أن استعمالها يواجه حدودًا مثل عدم توازن البيانات، نقص معلومات التوافر الحيوي، صعوبة تقييم السمية، واحتمال أن تمثل بعض الإدخالات فئات كيميائية عامة لا مركبات محددة.

2.5المحاور البيولوجية الأساسية المرتبطة ب **cluster3**

2.5.1 خلل الميتوكوندريا (**Mitochondrial Dysfunction**)

تُعد الميتوكوندريا من أهم العضيات الخلوية، حيث تلعب دورًا محوريًا في إنتاج الطاقة، تنظيم الإجهاد التأكسدي، والتحكم في مسارات موت الخلايا. لذلك، فإن أي اضطراب في وظائف الميتوكوندريا يمكن أن يؤدي إلى اختلالات واسعة في توازن الخلية، خاصة في الخلايا العصبية التي تعتمد بشكل كبير على الطاقة [6]–[9].

تقوم الميتوكوندريا بتحويل الطاقة الكيميائية المخزنة في المغذيات إلى ATP عبر الفسفرة التأكسدية. وتعتمد هذه العملية على سلسلة نقل الإلكترون داخل الغشاء الداخلي للميتوكوندريا. أي خلل في هذه السلسلة قد يؤدي إلى انخفاض ATP، زيادة تسرب الإلكترونات، وارتفاع إنتاج ROS.

تُعد الميتوكوندريا مصدرًا رئيسيًا لإنتاج Reactive Oxygen Species. في الظروف الطبيعية، يتم التحكم في هذه الجزيئات عبر أنظمة مضادة للأكسدة. لكن عند حدوث خلل ميتوكوندري، قد يرتفع إنتاج ROS بشكل غير متحكم فيه، مما يؤدي إلى تلف البروتينات والدهون والحمض النووي.

تلعب الميتوكوندريا دورًا أساسيًا أيضًا في تنظيم apoptosis. فعند تعرض الخلية لإجهاد شديد، يمكن للميتوكوندريا إطلاق عوامل مثل cytochrome c، مما يؤدي إلى تفعيل مسارات تنتهي بموت الخلية.

في سياق Cluster 3 ، يُعد خلل الميتوكوندريا محورًا مركزيًا يمكن أن يفسر جزءًا من التداخل بين عدة أمراض ضمن هذا الإطار الوظيفي.

2.5.2خلل النواقل العصبية (Neurotransmitter Dysfunction)

تُعد النواقل العصبية عناصر أساسية في نقل الإشارات بين الخلايا العصبية، حيث تضمن تنظيم التواصل داخل الجهاز العصبي المركزي والمحيطي. لذلك، فإن أي اضطراب في توازن هذه النواقل أو في مستقبلاتها قد يؤدي إلى خلل في الوظائف العصبية والمعرفية والسلوكية [7]–[10].

النواقل العصبية هي جزيئات كيميائية تُفرز من النهايات العصبية إلى الشق المشبكي، حيث ترتبط بمستقبلات محددة على الخلية الهدف. تعتمد هذه العملية على توازن دقيق بين التخليق، التخزين، الإفراز، إعادة الامتصاص، والتحلل.

تشمل أهم النواقل العصبية المرتبطة بوظائف الجهاز العصبي dopamine، serotonin، GABA، glutamate ، وacetylcholine. ويرتبط كل منها بوظائف مختلفة مثل الحركة، المزاج، النوم، التثبيط العصبي، التعلم، والذاكرة.

يمكن أن يحدث خلل النواقل العصبية عند اختلال التوازن بين النواقل المثبطة والمنبهة، أو تغير مستويات ناقل معين، أو تغير حساسية المستقبلات. في إطار Cluster 3، يمثل اضطراب النواقل العصبية محورًا أساسيًا يتقاطع مع خلل الميتوكوندريا والإجهاد التأكسدي.

2.5.3الإجهاد التأكسدي (Oxidative Stress)

يُعد الإجهاد التأكسدي من المحاور البيولوجية المهمة في تفسير عدد من الأمراض العصبية والوظيفية، خاصة تلك المرتبطة بخلل الميتوكوندريا واضطراب النواقل العصبية. يحدث الإجهاد التأكسدي عندما يختل التوازن بين إنتاج الجذور الحرة والقدرة المضادة للأكسدة داخل الخلية [8]–[11].

تشير Reactive Oxygen Species إلى مجموعة من الجزيئات التفاعلية المشتقة من الأكسجين، مثل superoxide anion، hydrogen peroxide ، وhydroxyl radical. تُنتج هذه الجزيئات بكميات محدودة في الظروف الطبيعية، لكنها تسبب تلفًا واسعًا عند زيادتها المفرطة.

تشمل مصادر الإجهاد التأكسدي خلل الميتوكوندريا، اضطراب سلسلة نقل الإلكترون، الالتهاب المزمن، التعرض للسموم، نقص مضادات الأكسدة، وخلل إنزيمات الدفاع المضاد للأكسدة.

تمتلك الخلية أنظمة دفاعية للحد من تراكم ROS، مثل superoxide dismutase، catalase ، glutathione peroxidase، وglutathione. في الخلايا العصبية، تكون الحساسية للإجهاد التأكسدي مرتفعة بسبب الحاجة العالية للطاقة وغنى الأغشية بالدهون غير المشبعة.

في سياق Cluster 3، يمثل الإجهاد التأكسدي محورًا رابطًا بين خلل الميتوكوندريا، الالتهاب العصبي، واضطراب النواقل العصبية.

2.5.4الالتهاب العصبي (Neuroinflammation)

يُعد الالتهاب العصبي أحد المحاور البيولوجية المهمة في تفسير العديد من الاضطرابات العصبية والوظيفية، حيث يمثل استجابة مناعية داخل الجهاز العصبي المركزي نتيجة عوامل مثل الإجهاد التأكسدي، خلل الميتوكوندريا، أو الضرر الخلوي [9]–[12].

يشير الالتهاب العصبي إلى تنشيط الجهاز المناعي داخل الدماغ والنخاع الشوكي، ويشمل ذلك تفاعل الخلايا الداعمة وإفراز وسائط التهابية تؤثر على البيئة العصبية.

تلعب microglia دور خط الدفاع الأول داخل الجهاز العصبي، بينما تساهم astrocytes في دعم الخلايا العصبية وتنظيم البيئة المحيطة. وعند تنشيط هذه الخلايا بشكل مفرط أو مزمن، قد تتحول من دور وقائي إلى دور ضار بالخلايا العصبية.

يرتبط الالتهاب العصبي بإفراز وسائط مثل TNF-alpha، IL-1beta، IL-6، chemokines ، وROS. وقد تؤثر هذه الجزيئات على وظيفة الخلايا العصبية، نقل الإشارات، واستقرار الشبكات العصبية.

في إطار Cluster 3، يمثل الالتهاب العصبي نقطة تقاطع بين خلل الميتوكوندريا، الإجهاد التأكسدي، واضطراب النواقل العصبية.

2.5.5 اضطراب **Proteostasis**

يشير proteostasis إلى المنظومة الخلوية المسؤولة عن الحفاظ على سلامة البروتينات من حيث الطي الصحيح، الاستقرار، والتخلص من البروتينات التالفة. يعتمد هذا التوازن على آليات تنسق بين تصنيع البروتينات وطيها وإعادة طيها أو تفكيكها عند الضرورة [10]–[13].

عند حدوث خلل في هذه المنظومة، يتراكم البروتين غير المطوي أو سيئ الطي، مما قد يؤدي إلى اضطرابات خلوية، خاصة في الخلايا العصبية الحساسة. وتشمل آليات الحفاظ على proteostasis molecular chaperones، ubiquitin-proteasome system، وautophagy.

يمثل نظام ubiquitin-proteasome أحد أهم آليات التخلص من البروتينات التالفة، إذ يتم وسم البروتين غير المرغوب فيه بجزيء ubiquitin ثم توجيهه إلى proteasome للتفكيك. أما autophagy فتسمح بالتخلص من البروتينات التالفة أو العضيات غير الوظيفية، بما في ذلك الميتوكوندريا.

في سياق Cluster 3، يرتبط اضطراب proteostasis بعدة محاور أخرى مثل الإجهاد التأكسدي، خلل الميتوكوندريا، والالتهاب العصبي، مما يعزز فكرة وجود شبكة بيولوجية مترابطة بدل مسار مرضي منفرد.

2.6 التداخل بين المحاور البيولوجية

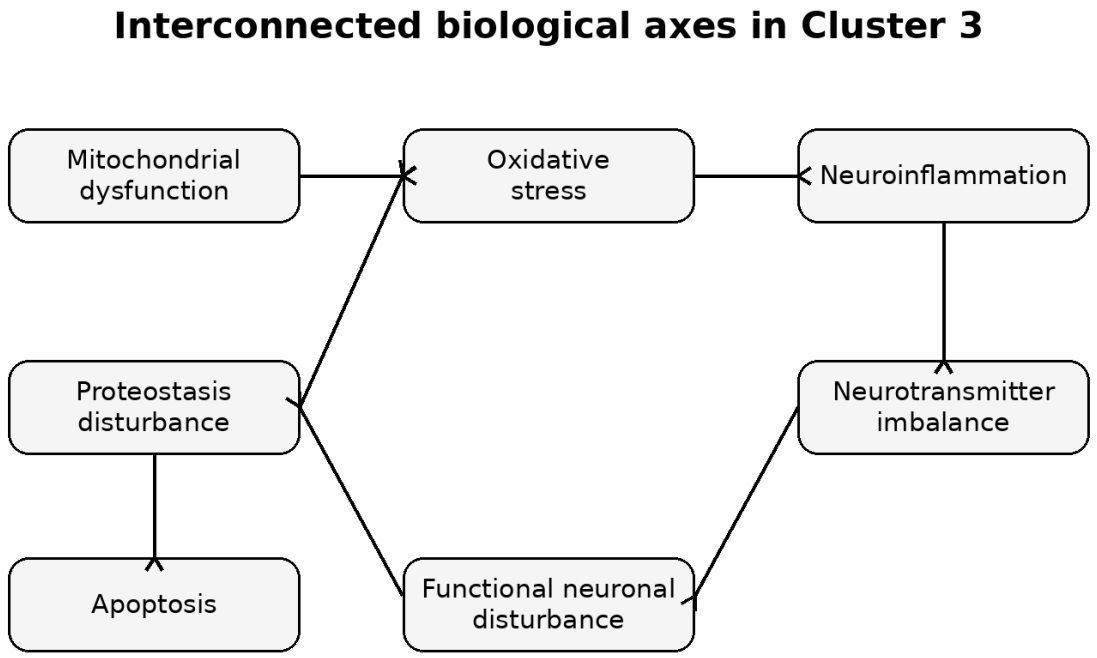
لا تعمل المحاور البيولوجية المرتبطة بـ Cluster 3 بشكل منفصل، بل توجد بينها علاقات متداخلة ومعقدة. فخلل الميتوكوندريا قد يؤدي إلى زيادة الإجهاد التأكسدي، والإجهاد التأكسدي قد يحفز الالتهاب العصبي، كما أن الالتهاب قد يؤثر في النواقل العصبية ووظيفة الخلايا العصبية [11]–[14].

تمثل الميتوكوندريا مصدرًا رئيسيًا لإنتاج ROS داخل الخلية. وعند حدوث خلل في الميتوكوندريا أو في سلسلة نقل الإلكترون، قد يزداد إنتاج ROS، مما يؤدي إلى إجهاد تأكسدي وتلف خلوي. وفي المقابل، يمكن للإجهاد التأكسدي أن يزيد من تلف الميتوكوندريا، مما يخلق حلقة مفرغة.

يمكن للإجهاد التأكسدي أن ينشط الخلايا الدبقية مثل microglia وastrocytes، مما يؤدي إلى إفراز وسائط التهابية. وفي المقابل، قد تؤدي الالتهابات المزمنة إلى زيادة إنتاج ROS، مما يعمق الضرر الخلوي.

يمكن للالتهاب العصبي أن يؤثر في توازن النواقل العصبية من خلال تغيير نشاط الإنزيمات أو المستقبلات أو آليات إعادة الامتصاص. كما أن تراكم البروتينات غير المطوية قد يؤدي إلى تفعيل apoptosis. لذلك، فإن Cluster 3 يجب أن يُفهم كشبكة محاور مترابطة.

Figure 2.1. Interconnection between biological mechanisms involved in Cluster 3.



Network interpretation: shared axes support hypothesis generation, not therapeutic proof.

الشكل 2.1. التداخل المفاهيمي بين خلل الميتوكوندريا، الإجهاد التأكسدي، الالتهاب العصبي، اضطراب النواقل العصبية، اضطراب **proteostasis**، و**apoptosis**.

Table 2.1. Biological axes related to Cluster 3.

علاقته بـ **Cluster 3**	الدور العام	المحور البيولوجي
محور مركزي في ضعف الخلايا العصبية	اضطراب إنتاج الطاقة وATP	Mitochondrial dysfunction
يؤثر في المزاج والحركة والذاكرة والسلوك	اضطراب الإشارات العصبية	Neurotransmitter dysfunction
رابط بين خلل الميتوكوندريا والالتهاب	زيادة ROS وتلف الخلايا	Oxidative stress
يساهم في اضطراب البيئة العصبية	تنشيط microglia وcytokines	Neuroinflammation
يرتبط بالتنكس العصبي والإجهاد الخلوي	تراكم البروتينات غير المطوية	Proteostasis disturbance
نتيجة محتملة لخلل الطاقة والإجهاد الخلوي	موت خلوي مبرمج	Apoptosis

2.7 مفهوم Multi-Disease Networks

يشير مفهوم Multi-Disease Networks إلى دراسة مجموعة من الأمراض ضمن شبكة واحدة بدل تحليل كل مرض بشكل منفصل. يعتمد هذا المفهوم على فكرة أن بعض الأمراض، رغم اختلافها السريري، قد تشترك في جينات، مسارات بيولوجية، أو آليات خلوية متداخلة [12]–[15].

الأمراض متعددة العوامل لا تنتج عن سبب واحد فقط، بل عن تفاعل عوامل جينية وبيئية واستقلابية وعصبية ومناعية. لذلك، يصعب تفسيرها بنموذج خطي بسيط، ويصبح تحليل الشبكات مناسبًا لفهم العلاقات المتداخلة بين مكوناتها.

قد تشترك عدة أمراض في مسارات مثل mitochondrial dysfunction،oxidative stress ، neuroinflammation، synaptic dysfunction، protein aggregation ، وapoptosis. وجود هذه المسارات المشتركة لا يعني أن الأمراض متطابقة، بل يعني وجود functional overlap بينها.

تُعد هذه الفكرة أساسية في الدراسة الحالية، لأن Cluster 3 لا يُعامل كمرض واحد، بل كمجموعة أمراض وظيفية قد تشترك في محاور مرتبطة بخلل النواقل العصبية والميتوكوندريا.

2.8 تحليل الشبكات الحيوية (Biological Network Analysis)

يُعد تحليل الشبكات الحيوية من الأدوات المهمة في فهم الأنظمة البيولوجية المعقدة، لأنه يسمح بتمثيل العلاقات بين الكيانات الحيوية في شكل شبكة مكونة من عقد وروابط [13]–[16].

في سياق هذه الدراسة، تمثل العقد النباتات، المركبات، الجينات، والأمراض، بينما تمثل الروابط علاقات مثل نبات–مركب، مركب–جين، وجين–مرض. وكلما زاد عدد الروابط المرتبطة بعقدة معينة، زاد تمثيلها داخل الشبكة.

يشير degree إلى عدد الروابط المتصلة بعقدة معينة. أما centrality فتشير إلى مجموعة من المقاييس التي تحدد أهمية العقد داخل الشبكة، مثل degree centrality، betweenness centrality، closeness centrality ، و.eigenvector centrality

في هذه الدراسة، تم اعتماد مؤشر degree-like hub score فقط، وهو مؤشر تغطية شبكية مبني على عدد المركبات والنباتات والأمراض المرتبطة بالجين. لذلك، لا يجب تفسيره كمقياس centrality متقدم.

2.9 مفهوم Network Coverage

يُعد network coverage من المفاهيم الأساسية في هذه الدراسة، حيث يُستخدم لقياس مدى انتشار كيان معين داخل الشبكة، سواء كان نباتًا، مركبًا، أو جينًا، عبر مجموعة الأمراض ضمن Cluster 3 [14]–[17].

يمكن تعريف coverage على أنه نسبة تمثيل كيان معين داخل مجموعة من العناصر المرجعية. فـ disease coverage تشير إلى عدد الأمراض التي يغطيها كيان معين مقارنة بإجمالي عدد الأمراض، بينما تشير gene coverage إلى عدد الجينات المرتبطة بكيان معين مقارنة بإجمالي الجينات في الشبكة.

تم حساب التغطية على مستويات متعددة تشمل plant disease coverage، plant gene coverage ، compound disease coverage، compound gene coverage، gene disease coverage ، وغيرها. يسمح هذا التعدد بفهم دور كل كيان داخل الشبكة بدل الاعتماد على مؤشر واحد فقط.

رغم أهمية coverage كمؤشر تحليلي، إلا أنه قد يعكس كثافة البيانات في قواعد البيانات وليس الأهمية البيولوجية الحقيقية. كما لا يعكس قوة التأثير أو الخصائص الدوائية أو الفعالية العلاجية.

2.10 التمييز بين Network Coverage و Therapeutic Effect

يُعد التمييز بين network coverage وtherapeutic effect من أهم النقاط المنهجية في هذه الدراسة. فارتفاع تغطية نبات أو مركب أو جين داخل الشبكة لا يعني بالضرورة وجود تأثير علاجي مباشر، بل يعني فقط أن هذا الكيان يمتلك تمثيلًا واسعًا داخل البيانات المستعملة [15]–[18].

يشير therapeutic effect إلى تأثير علاجي مثبت أو محتمل يتم دعمه عادةً بأدلة أقوى، مثل تجارب خلوية أو حيوانية أو سريرية، تحليل ADME/ADMET، docking، molecular dynamics ، أو تحقق من الجرعة والسلامة.

لذلك، تم اعتماد التفسير الآتي في هذه الدراسة: High coverage = research priority، بينما High coverage لا تعني therapeutic efficacy. هذا التمييز سيُطبق في فصل النتائج والمناقشة والحدود.

Table 2.2. Difference between network coverage and therapeutic efficacy.

المفهوم	المعنى	ماذا لا يعني؟
Network coverage	مدى تمثيل نبات أو مركب أو جين داخل الشبكة	لا يعني فعالية علاجية
Disease coverage	عدد الأمراض التي يغطيها الكيان داخل Cluster 3	لا يعني علاج هذه الأمراض
Gene coverage	عدد الجينات المرتبطة بكيان معين	لا يعني أن هذه الجينات targets مثبتة
Therapeutic efficacy	تأثير علاجي مثبت بتجارب أو أدلة سريرية	لا يُستنتج من الشبكة وحدها
Hypothesis-generating result	نتيجة تقترح فرضية قابلة للاختبار	ليست دليلًا نهائيًا

2.11 دور قواعد البيانات الحيوية

تلعب قواعد البيانات الحيوية دورًا أساسيًا في دراسات network pharmacology، لأنها توفر المعلومات اللازمة لربط المركبات الكيميائية بالأهداف الحيوية، الجينات، الأمراض، والمسارات البيولوجية [16]–[20].

تُعد ChEMBL قاعدة بيانات حيوية متخصصة في المعلومات الكيميائية–الحيوية، وتحتوي على بيانات حول المركبات، الأهداف البيولوجية، النشاط الحيوي، والاختبارات التجريبية. في هذه الدراسة، استُخدمت ChEMBL لمطابقة المركبات واستخراج بيانات النشاط الحيوي وبناء طبقة Compound–Target evidence.

تُعد UniProt من أهم قواعد بيانات البروتينات، حيث توفر معلومات حول البروتينات ومعرفاتها ووظائفها. كما تمثل Open Targets منصة مهمة لربط الجينات والأهداف الدوائية بالأمراض، ويمكن استعمالها لدعم تفسير العلاقات الجين–مرض في المراحل المستقبلية.

رغم أهمية قواعد البيانات، إلا أنها قد تحمل database bias، publication bias، missing data، mapping ambiguity، وassay variability. لذلك، فإن النتائج المبنية على قواعد البيانات يجب أن تُفسر كأدلة حاسوبية مولدة للفرضيات.

2.12 حدود المقاربات الحاسوبية

رغم أن المقاربات الحاسوبية مثل network pharmacology توفر أدوات قوية لتنظيم البيانات وتحليل العلاقات المعقدة بين النباتات، المركبات، الجينات، والأمراض، إلا أنها تبقى محدودة بطبيعة البيانات المستعملة وبنوع التحليل المنفذ [18]–[21].

تعتمد دقة التحليل الحاسوبي على جودة البيانات المدخلة. فإذا احتوت البيانات على أسماء مركبات غير دقيقة أو مستخلصات أو فئات كيميائية عامة أو تكرارات غير مضبوطة، فقد يؤدي ذلك إلى نتائج مضللة.

لا تستطيع المقاربات الحاسوبية وحدها تمثيل السياق الحيوي الكامل مثل تركيز المركب داخل النبات، التوافر الحيوي، الاستقلاب، السمية، عبور الحاجز الدموي الدماغي، نوع النسيج المستهدف، والجرعة الفعالة.

تحتاج النتائج الحاسوبية إلى تحقق لاحق عبر ADME/ADMET، molecular docking، molecular dynamics، pathway enrichment، cell-based assays، أو animal models. وبما أن هذه الدراسة لا تشمل كل هذه المراحل، فإن نتائجها تبقى ضمن مستوى hypothesis generation.

2.13 حوصلة الفصل الثاني

قدّم هذا الفصل الإطار النظري الذي تستند إليه الدراسة، من خلال عرض المفاهيم الأساسية المرتبطة بـ network pharmacology، والتحول من نموذج الهدف الواحد إلى نموذج متعدد الأهداف، إضافة إلى دور المركبات النباتية وخصائصها متعددة المكونات.

كما تم استعراض المحاور البيولوجية المرتبطة بـ Cluster 3، بما في ذلك خلل الميتوكوندريا، اضطراب النواقل العصبية، الإجهاد التأكسدي، الالتهاب العصبي، واضطراب توازن البروتينات، مع توضيح التداخل بين هذه المحاور ضمن شبكة بيولوجية مترابطة.

كذلك، تم تقديم مفاهيم تحليل الشبكات الحيوية، خاصة مفهوم network coverage، مع التأكيد على ضرورة التمييز بين التغطية الشبكية والفعالية العلاجية. وتمت مناقشة دور قواعد البيانات الحيوية وحدود المقاربات الحاسوبية، مما يوفر أساسًا نقديًا لتفسير النتائج.

بناءً على ذلك، يشكل هذا الفصل قاعدة نظرية لفهم المنهجية المعتمدة في الفصل الثالث، كما يهيئ لتفسير النتائج في الفصل الرابع ضمن إطار علمي متوازن، يراعي تعقيد الأنظمة البيولوجية ويجنب المبالغة في الاستنتاجات.

الفـــصل التطبيـــقي

المـــــواد والطرق

3.1. تمهيد الفصل

يعرض هذا الفصل المنهجية المعتمدة في بناء وتحليل شبكة Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction وفق مقاربة network pharmacology. وقد تم التعامل مع هذا الـ Cluster باعتباره مجموعة أمراض وظيفية متعددة، وليس مرضًا منفردًا أو تصنيفًا سريريًا رسميًا. لذلك، صُممت المنهجية لفحص العلاقات بين النباتات، المركبات الكيميائية، الجينات، والأمراض ضمن إطار شبكي متعدد الطبقات يسمح بتتبع التداخلات المحتملة بين محاور خلل النواقل العصبية، اضطراب الميتوكوندريا، الإجهاد التأكسدي، الالتهاب العصبي، واضطراب توازن البروتينات.

اعتمدت الدراسة على نموذج تحليلي من نوع:

```
Plant → Compound → Gene → Disease
```

حيث تمثل النباتات طبقة المصدر الطبيعي، وتمثل المركبات الطبقة الكيميائية الوسيطة، بينما تمثل الجينات والأمراض الطبقات البيولوجية والتفسيرية للشبكة. ويهدف هذا التصميم إلى تجاوز المقاربة الخطية البسيطة التي تربط نباتًا واحدًا أو مركبًا واحدًا بمرض واحد، والانتقال إلى تحليل متعدد المستويات يتناسب مع الطبيعة المعقدة للأمراض المرتبطة بـ Cluster 3.

تم تنفيذ العمل عبر خط سير منهجي مرحلي شمل فحص البيانات الخام، تنظيف الإدخالات، بناء سجل المركبات، مطابقة المركبات مع قاعدة ChEMBL، استخراج بيانات النشاط الحيوي، فلترة العلاقات مركب–هدف، تحويل الأهداف إلى رموز جينية، دمج الشبكة، ثم حساب مؤشرات التغطية والترتيب. كما تم اعتماد ملفات ضبط جودة QC وسجلات تنفيذ في مختلف المراحل لضمان قابلية التتبع وإعادة الفحص.

ونظرًا لأن التحليل الشبكي حساس جدًا لأخطاء التصنيف والتطبيع النصي، فقد أُضيفت مراحل تصحيح نهائية لاحقة لمراجعة الجداول الناتجة قبل اعتمادها في فصل النتائج. شملت هذه المراحل حذف الإدخالات غير النباتية من طبقة النباتات، ودمج الأسماء المتكررة أو المترادفة، خصوصًا الصيغ المرتبطة بـ Passiflora incarnata. وقد نتج عن التصحيح النهائي اعتماد شبكة تضم 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا، وهي الأرقام التي استُخدمت لاحقًا في عرض النتائج والمناقشة والاستنتاجات.

يركز هذا الفصل على وصف المنهجية والأدوات والمعايير المستعملة فقط، دون مناقشة النتائج أو تفسير دلالاتها البيولوجية. وتُعرض جميع مراحل العمل باعتبارها خطوات قابلة للتتبع، بحيث يمكن ربط كل جدول أو شكل في فصل النتائج بمرحلة محددة من خط سير التحليل. وبذلك، يشكل هذا الفصل الأساس المنهجي الذي يضمن أن النتائج النهائية ليست مجرد مخرجات رقمية، بل ناتجة عن مسار مضبوط، قابل للمراجعة، ومصحح منهجيًا قبل اعتماده.

3.2. تصميم الدراسة

صُممت هذه الدراسة كتحليل حاسوبي استكشافي قائم على منهجية network pharmacology، بهدف بناء شبكة متعددة الطبقات تربط بين النباتات، المركبات الكيميائية، الجينات، والأمراض داخل Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction. ويُعد هذا التصميم مناسبًا لطبيعة الدراسة، لأن Cluster 3 لا يمثل مرضًا واحدًا، بل مجموعة أمراض وظيفية يُفترض أنها تتقاطع في محاور بيولوجية مشتركة، مثل خلل النواقل العصبية، اضطراب الميتوكوندريا، الإجهاد التأكسدي، والالتهاب العصبي.

اعتمد التصميم العام على الانتقال التدريجي من البيانات الخام إلى شبكة قابلة للتحليل. بدأت الدراسة بملف أولي يضم علاقات بين الأمراض والنباتات والمركبات، ثم خضعت هذه البيانات لمراحل تنظيف وتوحيد قبل ربط المركبات بقواعد

بيانات النشاط الحيوي. بعد ذلك، تم استخراج علاقات المركب–الهدف من ChEMBL، وتحويل الأهداف إلى رموز جينية، ثم دمج هذه الطبقات ضمن شبكة موحدة من نوع:

```
Plant → Compound → Gene → Disease
```

يقوم التصميم على مبدأ الفصل بين مراحل إعداد البيانات ومراحل تحليل الشبكة. لذلك، لم يتم إدخال بيانات ADME/ADMET أو molecular docking أو molecular dynamics في هذه المرحلة، لأن الهدف كان أولًا بناء شبكة نظيفة وقابلة للتتبع، ثم حساب مؤشرات التغطية والترتيب داخلها. هذا الفصل المنهجي يمنع الخلط بين مرحلة network coverage analysis ومرحلة التحقق الدوائي أو البنيوي التي يمكن تنفيذها لاحقًا.

كما اعتمدت الدراسة على مبدأ التحقق المرحلي عبر ملفات QC وسجلات التنفيذ، بحيث تنتج كل مرحلة ملفات مستقلة يمكن مراجعتها قبل الانتقال إلى المرحلة التالية. وقد كان هذا التصميم مهمًا بشكل خاص بعد ظهور بعض أخطاء التصنيف في الجداول الأولية، مما استدعى إضافة مراحل تصحيح نهائية Stages 10C–10E قبل اعتماد النتائج النهائية. وبذلك أصبح التصميم لا يقتصر على إنتاج شبكة فقط، بل يشمل أيضًا آلية مراجعة وتصحيح تضمن اتساق النتائج مع تعريفات المنهجية.

تُعد الدراسة في جوهرها hypothesis-generating study؛ أي أنها تهدف إلى توليد فرضيات بحثية حول النباتات والمركبات والجينات ذات التمثيل الشبكي المرتفع داخل Cluster 3. ولا تهدف هذه الدراسة إلى إثبات فعالية علاجية مباشرة لأي نبات أو مركب، بل إلى تحديد أولويات بحثية يمكن اختبارها لاحقًا عبر pathway enrichment، ADME/ADMET، docking، molecular dynamics، والتحقق التجريبي.

Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction

Table 3.1. ملخص خط سير التحليل المنهجي

المخرجات الرئيسية	الوظيفة المنهجية	المرحلة
ملفات intake وQC	فحص الإدخال والتطبيع الأولي	Stage 0
قوائم مقبولة ومراجعة يدوية	تنظيف محافظ ومتقدم	Stage 1–1.5
compound registry وpriority tiers	بناء سجل المركبات وترتيب الأولويات	Stage 2
Compound–Target evidence	مطابقة ChEMBL واستخراج/فلترة النشاط الحيوي	Stage 3–5
Plant–Compound–Gene–Disease network	الربط الجيني ودمج الشبكة وتحليل التغطية	Stage 6–8
الجداول والأشكال النهائية المصححة	تحليل فئات التغطية والتصحيح النهائي	Stage 9B–10E

3.3. تعريف **Cluster3** ونطاق الدراسة

يمثل Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction إطارًا تحليليًا وظيفيًا يضم مجموعة أمراض يُفترض أنها تتقاطع في محاور بيولوجية مرتبطة بخلل النواقل العصبية واضطراب وظائف الميتوكوندريا. ولا يُعامل هذا الـ Cluster في هذه الدراسة كتصنيف سريري رسمي أو كمرض واحد، بل كإطار شبكي يسمح بفحص التداخلات المحتملة بين عدة أمراض من منظور النباتات، المركبات، الجينات، والأهداف الحيوية.

يرتكز نطاق Cluster 3 على مجموعة من المحاور البيولوجية الأساسية، أهمها اضطراب النواقل العصبية، خلل إنتاج الطاقة داخل الخلية، الإجهاد التأكسدي، الالتهاب العصبي، اضطراب توازن البروتينات، ومسارات موت الخلايا المبرمج.

وقد تم اختيار هذه المحاور لأنها تمثل نقاط تقاطع محتملة بين عدد من الاضطرابات العصبية والوظيفية، ولأنها تتناسب مع منطق multi-disease network analysis الذي يقوم عليه هذا العمل.

في هذه الدراسة، تم التعامل مع الأمراض المدرجة داخل Cluster 3 كوحدة تحليلية متعددة الأمراض، وليس كأمراض مستقلة تمامًا. لذلك، فإن مؤشرات مثل disease coverage وmulti-disease coverage لا تعني أن نباتًا أو مركبًا معينًا يؤثر علاجيًا في جميع هذه الأمراض، بل تعني فقط مدى تمثيله الشبكي داخل مجموعة الأمراض المدروسة.

يسمح هذا التعريف بتجنب خطأ منهجي مهم، وهو تعميم نتائج الشبكة على مرض واحد محدد أو تحويل التغطية الشبكية إلى ادعاء علاجي مباشر. فالدراسة لا تهدف إلى القول إن نباتًا معينًا يعالج جميع أمراض Cluster 3، بل تهدف إلى تحديد النباتات والمركبات والجينات التي تظهر بتغطية عالية داخل هذا الإطار الوظيفي، ثم اقتراحها كأولويات بحثية قابلة للتحقق في مراحل لاحقة.

وبذلك، يشكل تعريف Cluster 3 ونطاقه أساسًا مهمًا لبقية المنهجية. فكل مراحل التنظيف، المطابقة، استخراج النشاط الحيوي، الربط الجيني، وحساب التغطية بُنيت على هذا التصور: مجموعة أمراض وظيفية متعددة يتم تحليلها من خلال شبكة Plant–Compound–Gene–Disease، مع الحفاظ على الفصل بين التمثيل الشبكي والدلالة العلاجية.

3.4. مصادر البيانات

اعتمدت هذه الدراسة على ملف بيانات أولي يمثل طبقة المصدر في شبكة Cluster 3، حيث يربط بين الأمراض، النباتات، والمركبات الكيميائية المرتبطة بها. وقد استُخدم هذا الملف كنقطة انطلاق لبناء شبكة متعددة الطبقات من نوع Plant–Compound–Gene–Disease. احتوى ملف الإدخال الخام على أعمدة أساسية تمثل اسم الـ cluster، المرض، النبات، والمركب، ثم خضعت هذه الأعمدة لاحقًا لعمليات توحيد وتطبيع نصي حتى تصبح قابلة للمعالجة البرمجية.

تمثل بيانات المصدر الأولية العلاقة الأساسية:

```
Disease → Plant → Compound
```

ثم تم توسيع هذه العلاقة لاحقًا عبر ربط المركبات بقواعد بيانات النشاط الحيوي، للحصول على علاقات إضافية بين المركبات والأهداف الحيوية والجينات. وبذلك، لم تبقَ الدراسة محصورة في بيانات نباتية وصفية، بل انتقلت إلى بناء شبكة بيولوجية–كيميائية تعتمد على أدلة نشاط حيوي مستخرجة من قواعد بيانات متخصصة.

استُخدمت قاعدة ChEMBL كمصدر رئيسي لمطابقة أسماء المركبات واستخراج بيانات النشاط الحيوي المرتبطة بالأهداف البيولوجية. وقد شملت البيانات المستخرجة معرفات المركبات، معرفات الأهداف، نوع النشاط الحيوي، القيم القياسية للنشاط، وحدات القياس، ومعلومات التجارب أو الوثائق المرتبطة بها. سمح ذلك ببناء طبقة Compound–Target evidence، والتي استُعملت لاحقًا لتحويل الأهداف إلى رموز جينية قابلة للتفسير.

كما استُخدمت معلومات الأهداف والبروتينات المرتبطة بـ ChEMBL من أجل دعم مرحلة الربط الجيني، حيث تم تحويل معرفات الأهداف إلى رموز جينية بشرية كلما كان ذلك ممكنًا. وقد تم اعتماد الأهداف البشرية فقط في مراحل الفلترة، حتى تكون النتائج مرتبطة بسياق بيولوجي مناسب للدراسة.

تم الالتزام بقاعدة منهجية أساسية مفادها أن الدراسة تنقل المنهجية المطورة من الأعمال السابقة، ولا تنقل النتائج أو النباتات أو المركبات أو الجينات من clusters أخرى. لذلك، فإن كل كيان نباتي أو مركب أو جين ظهر في النتائج النهائية كان ناتجًا عن بيانات Cluster 3 نفسها ومراحل التحليل الخاصة بها.

بناءً على ذلك، يمكن تلخيص مصادر البيانات في ثلاثة مستويات رئيسية: ملف الإدخال الأولي الخاص بعلاقات المرض–النبات–المركب، قاعدة ChEMBL الخاصة بالمركبات والنشاط الحيوي، ومعلومات الأهداف المستعملة في الربط الجيني. وقد تم دمج هذه المصادر وفق خط سير مرحلي مضبوط لضمان قابلية التتبع وتقليل أخطاء التصنيف أو المطابقة.

3.5. بنية المنهجية العامة

اعتمدت الدراسة على بنية منهجية مرحلية تهدف إلى تحويل بيانات أولية غير متجانسة حول الأمراض، النباتات، والمركبات إلى شبكة متعددة الطبقات قابلة للتحليل. وقد صُممت هذه البنية بحيث تمر البيانات عبر مراحل متتابعة من الفحص، التنظيف، المطابقة، استخراج النشاط الحيوي، الربط الجيني، دمج الشبكة، ثم حساب مؤشرات التغطية والترتيب. يمكن تلخيص البنية العامة للمنهجية كما يلي:

Raw disease–plant–compound data

↓

Data intake and preflight

↓

Conservative cleaning and refinement

↓

Compound registry and prioritization

↓

ChEMBL compound resolution

↓

Bioactivity extraction and filtering

↓

Target-to-gene mapping

↓

Plant–Compound–Gene–Disease integration

↓

Coverage and ranking analysis

↓

Corrective QC and final harmonization

↓

Final tables and figures

تم تقسيم العمل إلى مراحل مستقلة، بحيث تنتج كل مرحلة ملفات مخرجات خاصة بها، إضافة إلى ملفات ضبط جودة QC وسجلات تنفيذ. يسمح هذا التنظيم بتتبع مصدر كل نتيجة، ومعرفة المرحلة التي أنتجت كل جدول أو شكل في فصل النتائج. كما يسمح بإعادة تشغيل مرحلة محددة دون الحاجة إلى إعادة تنفيذ كامل خط سير التحليل.

تبدأ المنهجية بفحص ملف الإدخال الخام والتأكد من وجود الأعمدة الأساسية، ثم تنظيف الإدخالات غير المناسبة مثل المستخلصات، الخلطات، الفئات الكيميائية العامة، أو الأوصاف غير الدقيقة. بعد ذلك، يتم بناء سجل مركبات موحد وترتيب المركبات حسب أولويتها الشبكية قبل مطابقتها مع قاعدة ChEMBL. وتُستعمل نتائج ChEMBL لاحقًا لاستخراج علاقات النشاط الحيوي بين المركبات والأهداف، ثم تُفلتر هذه العلاقات وفق معايير النشاط الحيوي والأهداف البشرية.

بعد مرحلة الفلترة، يتم تحويل الأهداف الحيوية إلى رموز جينية، ثم دمج طبقات النبات، المركب، الجين، والمرض في شبكة واحدة. بعد ذلك، تُحسب مؤشرات التغطية على مستوى النباتات والمركبات والجينات، مثل disease coverage وgene coverage وintegrated coverage score. وقد أضيفت في نهاية المنهجية مراحل تصحيح نهائية Stages 10C–10E لمراجعة الجداول الناتجة، حذف الإدخالات غير النباتية، وتوحيد الأسماء النباتية قبل اعتماد النتائج النهائية.

تضمن هذه البنية أن النتائج النهائية لا تمثل مجرد مخرجات مباشرة من قواعد البيانات، بل ناتجة عن خط سير منهجي قابل للتتبع والتصحيح. وبذلك، تصبح الشبكة النهائية أساسًا مناسبًا لتوليد الفرضيات البحثية، مع الحفاظ على الفصل بين network coverage كأداة ترتيب، وبين الفعالية العلاجية التي تتطلب تحققًا لاحقًا.

3.6. فحص البيانات الأولي والإدخال (Stage 0)

تم تنفيذ Stage 0 — Intake and Preflight كمرحلة أولى في خط سير التحليل، بهدف فحص ملف الإدخال الخام قبل إدخاله في مراحل التنظيف والمطابقة. كان الهدف من هذه المرحلة التأكد من أن البيانات الأولية صالحة للمعالجة البرمجية، وأنها تحتوي على الأعمدة الضرورية لبناء طبقة المصدر الخاصة بـ Cluster 3.

شملت هذه المرحلة قراءة ملف الإدخال، التحقق من بنيته، وفحص الأعمدة الأساسية التي تمثل cluster، disease، plant، وcompound. بعد ذلك، تم توحيد أسماء الأعمدة وتحويلها إلى صيغ معيارية قابلة للاستعمال في السكريبتات اللاحقة. كما تم إنشاء مفاتيح مطابقة نصية matchkeys لكل من الأمراض، النباتات، والمركبات، بهدف تقليل أثر اختلافات الكتابة، المسافات، الحروف الكبيرة والصغيرة، وبعض الضوضاء النصية البسيطة.

سمحت مفاتيح المطابقة بالكشف عن نوعين من التكرارات: التكرارات المباشرة الناتجة عن تكرار نفس الصف، والتكرارات الناتجة عن اختلافات كتابية بسيطة بعد التطبيع النصي. كما تم فصل الصفوف إلى صفوف مقبولة قابلة للمتابعة في التحليل، وصفوف تحتاج إلى مراجعة يدوية بسبب وجود ضوضاء تنسيقية، أخطاء كتابية محتملة، إدخالات غير واضحة، أو عناصر لا تمثل نباتًا أو مركبًا محددًا.

لم يكن الهدف من Stage 0 حذف أكبر عدد ممكن من الصفوف، بل فحص جودة البيانات وتحديد مستوى الضوضاء قبل الانتقال إلى التنظيف المحافظ. لذلك، تم الاحتفاظ بسجل واضح للصفوف المقبولة والصفوف الموسومة للمراجعة، حتى تبقى عملية التنظيف قابلة للتتبع ولا تعتمد على حذف غير موثق.

3.7.التنظيف المحافظ للبيانات (Stage 1)

بعد فحص ملف الإدخال في Stage 0، تم تنفيذ Stage 1 — Conservative Cleaning بهدف تنظيف البيانات بطريقة محافظة قبل الانتقال إلى بناء سجل المركبات ومطابقتها مع قواعد البيانات. ركزت هذه المرحلة على تقليل الضوضاء في طبقة المركبات والنباتات دون حذف غير مبرر للبيانات، مع الحفاظ على إمكانية تتبع كل قرار تنظيف.

اعتمد التنظيف المحافظ على مبدأ أن التحليل اللاحق، خاصة مطابقة المركبات مع ChEMBL، يجب أن يعتمد على مركبات كيميائية واضحة ومحددة قدر الإمكان. لذلك، تم وسم أو استبعاد الإدخالات التي لا تمثل مركبات مفردة قابلة للمطابقة، مثل المستخلصات، الخلطات، التحضيرات، الأجزاء النباتية، الفئات الكيميائية العامة، أو العبارات الوصفية غير الدقيقة.

شملت الإدخالات التي خضعت للوسم أو المراجعة فئات مثل extracts، mixtures، plant parts، generic formatting، و preparations، enumerated combinations، vague descriptors، chemical classes noise. تم تصنيف الصفوف بعد التنظيف إلى صفوف مقبولة يمكن استعمالها في المراحل اللاحقة، وصفوف تحتاج إلى مراجعة يدوية.

كما ساعدت هذه المرحلة على حماية الفصل بين طبقات الشبكة. فالنبات يجب أن يبقى ضمن طبقة النباتات، والمركب ضمن طبقة المركبات، والمستخلص أو الصيغة العشبية المركبة لا ينبغي أن يُعامل كنبات مفرد أو مركب كيميائي محدد. وقد برزت أهمية هذا المبدأ لاحقًا عند تنفيذ التصحيحات النهائية في Stages 10C–10E.

3.8 التنظيف المتقدم والتحسين المحافظ (Stage 1.5)

بعد تنفيذ التنظيف المحافظ في Stage 1، تم اعتماد مرحلة إضافية تحت اسم Stage 1.5 — Deep Conservative Refinement بهدف تحسين جودة قائمة المركبات قبل بناء السجل النهائي للمركبات. جاءت هذه المرحلة لأن التنظيف الأولي، رغم أهميته، قد لا يكون كافيًا لاكتشاف جميع الإدخالات العامة أو غير الدقيقة التي يمكن أن تمر ظاهريًا كمركبات. ركزت هذه المرحلة على التمييز بين المركب الكيميائي الفردي والفئة الكيميائية العامة. فمصطلحات مثل flavonoids أو alkaloids أو polyphenols أو terpenoids لا تمثل مركبات منفردة، بل عائلات كيميائية واسعة تضم عددًا كبيرًا من الجزيئات. إدخال هذه المصطلحات في مراحل ChEMBL كما لو كانت مركبات محددة قد يؤدي إلى نتائج غير دقيقة.

لذلك، تم وسم أو استبعاد الإدخالات التي تمثل فئات كيميائية عامة أو أوصافًا غير محددة مثل flavonoids، alkaloids، polyphenols، terpenoids، saponins، tannins، glycosides، essential oils، bioactive compounds، و phytochemicals.

لم يكن الهدف من هذه المرحلة تقليل عدد المركبات فقط، بل تحسين نوعية المركبات التي ستدخل في مراحل المطابقة والتحليل. لذلك اعتمدت هذه المرحلة مبدأ .Quality before quantity

3.9.بناء سجل المركبات وترتيب الأولويات (Stage 2)

بعد الانتهاء من مراحل الفحص والتنظيف والتحسين المحافظ، تم تنفيذ Stage 2 — Compound Registry and Prioritization بهدف بناء سجل موحد للمركبات المقبولة داخل Cluster 3. تمثل هذه المرحلة نقطة انتقال مهمة من بيانات خام تحتوي على تكرارات واختلافات كتابية إلى قائمة منظمة من المركبات القابلة للمتابعة في مراحل المطابقة مع ChEMBL.

اعتمد بناء سجل المركبات على مفاتيح المطابقة النصية compound matchkeys، حيث تم تجميع الصيغ الكتابية المتقاربة للمركب نفسه تحت مفتاح موحد. سمح ذلك بتقليل التكرارات الناتجة عن اختلافات في الكتابة، المسافات، علامات الترقيم، أو صيغ التسمية.

تم حساب مؤشرات أولية لكل مركب، أهمها source_rows، source_plant_count، source_disease_count، coverage_priority_score، وpriority_tier. استُخدم coverage_priority_score لترتيب المركبات حسب درجة تمثيلها داخل بيانات .Cluster 3

بناءً على هذه المؤشرات، تم تقسيم المركبات إلى فئات أولوية مثل P1_critical، P2_high، P3_medium، وP4_low. مثّلت فئات 1P و2P المركبات الأعلى أولوية للمتابعة في مراحل المطابقة مع ChEMBL، لأنها تمتلك تمثيلًا أكبر داخل النباتات والأمراض.

3.10.مطابقة المركبات معChEMBL (Stage3)

بعد بناء سجل المركبات وترتيبها حسب الأولوية في Stage 2، تم تنفيذ Stage 3 — ChEMBL Compound Resolution بهدف ربط أسماء المركبات المقبولة بمعرفات قياسية داخل قاعدة ChEMBL. وتُعد هذه المرحلة أساسية لأنها تنقل المركبات من مستوى الأسماء النصية المستخرجة من ملف الإدخال إلى مستوى معرفات كيميائية قابلة للاستعمال في استخراج بيانات النشاط الحيوي.

تم تصنيف نتائج المطابقة إلى فئات رئيسية منها resolved_exact، resolved_best_hit_non_exact، unresolved_no_hit، وambiguous_multiple_hits. تمثل فئة resolved_exact الحالات التي وُجد فيها تطابق

واضح بين اسم المركب ومعرف ChEMBL، بينما تمثل resolved_best_hit_non_exact الحالات القريبة ولكن غير الدقيقة تمامًا.

اعتمدت الدراسة استراتيجية محافظة في التعامل مع المطابقات غير الدقيقة. فقد تم الاحتفاظ ببعضها مع وسم حالة المطابقة بوضوح، حتى لا تختلط بالمطابقات الدقيقة في التحليل اللاحق. كما تم استعمال exact edge ratio لاحقًا ضمن حساب الدرجات التكاملية، بهدف إعطاء وزن إضافي للروابط ذات المطابقة الأكثر موثوقية.

لم يكن الهدف من هذه المرحلة تعظيم عدد المركبات المطابقة بأي ثمن، بل تحقيق توازن بين التغطية والدقة. وبذلك، مثّلت Stage 3 خطوة محورية في بناء طبقة المركبات القياسية داخل الشبكة.

3.11.استخراج بيانات النشاط الحيوي من**ChEMBL (Stage4)**

بعد إتمام مطابقة المركبات مع قاعدة ChEMBL في Stage 3، تم تنفيذ Stage 4 — Bioactivity Extraction بهدف استخراج بيانات النشاط الحيوي المرتبطة بالمركبات التي تم ربطها بمعرفات ChEMBL. تمثل هذه المرحلة نقطة الانتقال من مجرد مطابقة أسماء المركبات إلى بناء طبقة أدلة حيوية تربط المركبات بالأهداف البيولوجية.

اعتمدت هذه المرحلة على استرجاع سجلات النشاط الحيوي المتاحة لكل مركب، مع التركيز على السجلات التي تحتوي على قياسات كمية قابلة للتفسير. شملت البيانات المستخرجة معرف المركب، معرف الهدف، اسم الهدف، الكائن الحي، نوع النشاط الحيوي، قيمة النشاط، وحدة القياس، معرف الاختبار، ومعرف الوثيقة المرتبطة بالسجل.

من بين أنواع النشاط الحيوي التي تم استخراجها IC50، Ki، Kd، EC50، 50AC، وPotency. غير أن وجود سجل نشاط حيوي في ChEMBL لا يعني قبوله مباشرة في الشبكة النهائية؛ لذلك تم التعامل مع مخرجات Stage 4 كبيانات خام أو شبه خام تحتاج إلى فلترة لاحقة في Stage 5.

تُعد هذه المرحلة ضرورية لأنها وفرت الأساس الذي بُنيت عليه طبقة Compound–Target evidence. ومن دون هذه الطبقة، لا يمكن الانتقال إلى الربط الجيني أو بناء شبكة Plant–Compound–Gene–Disease.

3.12.فلترة الأدلة الحيوية وبناء روابط المركب–الهدف **(Stage 5)**

بعد استخراج بيانات النشاط الحيوي الخام من ChEMBL في Stage 4، تم تنفيذ Stage 5 — Strict Core Bioactivity Filtering بهدف تحويل السجلات الخام إلى مجموعة أكثر صرامة من علاقات المركب–الهدف. وقد جاءت هذه المرحلة لتقليل الضوضاء الناتجة عن اختلاف أنواع الاختبارات، اختلاف الكائنات الحية، تفاوت وحدات القياس، ووجود سجلات نشاط ضعيفة أو غير قابلة للتفسير المباشر.

تم التركيز على الأهداف البشرية فقط، واختيار أنواع النشاط الحيوي الأكثر قابلية للتفسير مثل IC50، Ki، Kd، وEC50، مع اعتماد العلاقات القياسية =، <، <=، و~. كما تم اعتماد حد نشاط حيوي قدره ≥10 µM، أي ما يعادل تقريبًا ≥10,000 nM، وذلك للاحتفاظ بالعلاقات التي تمتلك نشاطًا بيولوجيًا قابلًا للتفسير دون تضييق الشبكة بشكل مفرط.

بعد تطبيق هذه المعايير، تم بناء روابط مباشرة من نوع Compound → Target. وقد تم تلخيص الأدلة لكل علاقة اعتمادًا على مؤشرات مثل أفضل قيمة نشاط حيوي، عدد الاختبارات، عدد الوثائق، ونوع النشاط الحيوي.

ومع ذلك، يجب التأكيد أن قبول علاقة مركب–هدف بعد Stage 5 لا يعني إثبات تأثير دوائي داخل الجسم. فهذه العلاقة تظل مبنية على بيانات نشاط حيوي مستخرجة من قاعدة بيانات، وقد تكون ناتجة عن أنظمة اختبار مختلفة وظروف تجريبية متباينة.

3.13.الربط الجيني لأهداف**ChEMBL (Stage6)**

بعد بناء روابط المركب–الهدف في Stage 5، تم تنفيذ Stage 6 — Target-to-Gene Mapping بهدف تحويل أهداف ChEMBL إلى رموز جينية قابلة للتفسير داخل شبكة Cluster 3. وتُعد هذه المرحلة ضرورية لأن الهدف النهائي من الدراسة يتطلب ربط الأهداف بطبقة الجينات التي يمكن استعمالها لاحقًا في تحليل التغطية، ترتيب الجينات، ومناقشة المحاور البيولوجية.

اعتمدت هذه المرحلة على المسار المنهجي: ChEMBL target ID → protein/target annotation → gene symbol. وقد تم قبول الأهداف التي ترتبط بسياق بشري واضح، وتتضمن معلومات كافية تسمح بتحويلها إلى رموز جينية.

كان الهدف الأساسي من هذه المرحلة هو إنشاء طبقة Compound → Gene بدل الاكتفاء بطبقة Compound → Target، لأن الرموز الجينية أكثر ملاءمة لدمج البيانات مع الأمراض، وحساب مؤشرات التغطية، وترتيب الجينات حسب درجة تمثيلها داخل الشبكة.

تم التعامل بحذر مع الأهداف التي قد ترتبط بأكثر من جين أو التي تمثل عائلات بروتينية واسعة، لأن هذه الحالات قد تؤدي إلى تضخيم طبقة الجينات أو إدخال علاقات غير دقيقة. لذلك، تم الاحتفاظ بمعلومات نوع الهدف ودرجة وضوح الربط.

3.14. دمج شبكة النبات–المركب–الجين–المرض (**Stage** 7)

بعد الانتهاء من مرحلة الربط الجيني في Stage 6، تم تنفيذ Stage 7 — Plant–Compound–Gene–Disease Integration بهدف دمج طبقات البيانات المختلفة في شبكة واحدة متعددة المستويات. وتُعد هذه المرحلة نقطة التحول الأساسية في المنهجية، لأنها تجمع بين بيانات المصدر الأولية الخاصة بالنباتات والمركبات والأمراض، وبين الأدلة الجينية الناتجة عن ChEMBL ومرحلة الربط الجيني.

اعتمد الدمج على المفتاح الموحد للمركبات compound matchkey، بحيث يتم ربط المركب الموجود في ملف Cluster 3 الأصلي بالمركب الذي تم حله في ChEMBL ثم ربطه بالأهداف والجينات. وبذلك أصبحت العلاقة قابلة للتمثيل في شكل شبكة .Plant → Compound → Gene → Disease

تم في هذه المرحلة حذف العلاقات المكررة لتجنب تضخيم الشبكة اصطناعيًا. فوجود المركب نفسه في أكثر من صف أو ظهوره بصيغ نصية متقاربة قد يؤدي إلى تكرار الروابط نفسها إذا لم يتم ضبطها. لذلك، تم الاعتماد على مفاتيح مطابقة موحدة وعلى إزالة التكرارات عند بناء الجداول النهائية للروابط.

أنتجت هذه المرحلة جداول ملخصة على عدة مستويات، منها مستوى النباتات، المركبات، الجينات، والأمراض. كما تم إنشاء جدول شبكة مدمج يمثل العلاقات الأساسية بين الكيانات الأربع. وقد استُعملت هذه المخرجات لاحقًا في حساب مؤشرات التغطية والترتيب.

3.15.تحليل التغطية والترتيب (**Stage** 8)

بعد بناء الشبكة المدمجة في Stage 7، تم تنفيذ Stage 8 — Advanced Coverage and Ranking بهدف حساب مؤشرات التغطية وترتيب الكيانات داخل شبكة Cluster 3. وقد ركزت هذه المرحلة على قياس مدى انتشار كل نبات أو مركب أو جين داخل الشبكة، اعتمادًا على عدد الأمراض، الجينات، المركبات، أو النباتات المرتبطة به.

اعتمد تحليل التغطية على حساب مؤشرات متعددة مثل plant disease coverage، plant gene coverage، plant coverage، compound disease coverage، compound gene coverage، compound plant coverage، gene disease coverage، gene compound coverage، و.gene plant coverage

سمحت هذه المؤشرات بفهم الشبكة من أكثر من زاوية. فعلى سبيل المثال، يوضح plant disease coverage مدى انتشار النبات عبر أمراض Cluster 3، بينما يوضح plant gene coverage مدى ارتباط مركبات النبات بطبقة الجينات. أما compound disease coverage فيقيس مدى انتشار المركب عبر الأمراض.

مثّلت Stage 8 مرحلة أساسية في تحويل الشبكة المدمجة إلى نتائج كمية قابلة للعرض في الفصل الرابع. كما وفرت هذه المرحلة الأساس الذي استُخدم لاحقًا في تحليل فئات التغطية وفي إنتاج الجداول والأشكال النهائية.

3.16. حساب المؤشرات التكاملية

بعد حساب مؤشرات التغطية المفردة في Stage 8، تم حساب مجموعة من المؤشرات التكاملية بهدف ترتيب النباتات، المركبات، والجينات داخل شبكة Cluster 3 بطريقة تجمع أكثر من بُعد واحد من أبعاد التغطية. وقد استُعملت هذه المؤشرات لأن الاعتماد على مؤشر منفرد، مثل gene coverage أو disease coverage فقط، قد يعطي قراءة جزئية لا تعكس البنية الكاملة للشبكة.

تم حساب plant integrated coverage score لتمثيل التوازن بين تغطية النبات للأمراض، الجينات، المركبات، ونسبة الروابط ذات المطابقة الدقيقة. أما compound integrated coverage score، فقد صُمم لترتيب المركبات وفق انتشارها عبر الأمراض، الجينات، النباتات، ونسبة المطابقة الدقيقة.

كما تم حساب degree-like hub score للجينات، وهو مؤشر مبني على مدى ارتباط الجين بالمركبات، النباتات، والأمراض داخل الشبكة. ولا يمثل هذا المؤشر مقياس مركزية شبكية متقدمًا مثل betweenness centrality أو eigenvector centrality، بل يمثل درجة تمثيل شبكي شبيهة بالـ degree.

اعتمدت الصيغ التكاملية على أوزان تحليلية محددة، بحيث يُعطى لكل بُعد من أبعاد التغطية وزن معين داخل score النهائي. ومع ذلك، تبقى هذه الأوزان اختيارًا تحليليًا يحتاج إلى اختبار لاحق عبر sensitivity analysis.

3.17. تحليل فئات التغطية (Stage 9B)

بعد حساب مؤشرات التغطية والدرجات التكاملية، تم تنفيذ Stage 9B — Coverage Distribution Analysis بهدف دراسة توزيع الكيانات داخل الشبكة حسب فئات التغطية. جاءت هذه المرحلة لتجنب الاكتفاء بعرض أعلى النباتات أو المركبات أو الجينات فقط، لأن ترتيب Top entities قد يعطي انطباعًا غير كامل عن البنية العامة للشبكة.

اعتمدت هذه المرحلة على تصنيف نسب التغطية إلى خمس فئات رئيسية: Very high coverage: 80–100%، High coverage: 60–80%، Moderate-high coverage: 40–60%، Moderate coverage: 20–40%، و Low coverage: <20%.

سمح هذا التصنيف بتحليل الشبكة من منظور توزيعي، أي فهم عدد الكيانات التي تقع في الفئات العالية أو المتوسطة أو المنخفضة، بدل التركيز على الكيان الأعلى ترتيبًا فقط. وهذا مهم في دراسات network pharmacology، لأن قوة الشبكة لا تُفهم فقط من خلال وجود نبات أو مركب واحد عالي الترتيب، بل من خلال توزيع التغطية عبر طبقات النباتات، المركبات، والجينات.

كما ساعد تحليل فئات التغطية على كشف الحاجة إلى مراجعة بعض الجداول النهائية. فقد أظهرت المراجعة اللاحقة وجود بعض الإدخالات غير النباتية أو المتكررة داخل طبقة النباتات، مما استدعى تنفيذ مراحل التصحيح النهائية Stages 10C–10E.

3.18.إنتاج الجداول والأشكال والتصحيح النهائي للنتائج (Stage 10)

بعد الانتهاء من تحليل التغطية وفئاتها في Stage 9B، تم تنفيذ Stage 10 — Tables and Figures Generation بهدف تحويل المخرجات الرقمية إلى جداول وأشكال قابلة للإدراج في فصل النتائج. شملت هذه المرحلة إنتاج الجداول الخاصة بحجم الشبكة، ترتيب النباتات، ترتيب المركبات، ترتيب الجينات، المركبات واسعة التغطية، وفئات التغطية.

في البداية، مثّلت Stage 10 مرحلة إخراج للنتائج، لكنها تحولت لاحقًا إلى مرحلة مراجعة نهائية للجداول قبل اعتمادها. فقد أظهرت المراجعة وجود بعض الأخطاء التصنيفية والتطبيعية في طبقة النباتات، مثل ظهور مركبات كيميائية ضمن جداول النباتات، ووجود مستخلصات أو صيغ عشبية مركبة، إضافة إلى تكرارات ناتجة عن اختلافات كتابية أو رموز غير مرئية.لذلك، أُضيفت مراحل تصحيحية لاحقة ضمن نهاية Stage 10، وهي Stage 10C وStage 10D وStage 10E. كان الهدف من هذه المراحل ضمان أن الجداول والأشكال النهائية لا تعتمد على إدخالات مصنفة خطأ، وأن الأسماء النباتية المتكررة أو المترادفة تم توحيدها قبل إدراج النتائج في الفصل الرابع.

في Stage 10C — Corrective Results QC، تم فحص جداول النباتات وتصنيف الإدخالات غير النباتية. وقد تم استبعاد rutin وhesperidin من طبقة النباتات لأنهما مركبان كيميائيان، كما تم استبعاد red_onion_husk_extract لأنه مستخلص، وhangekobokuto لأنه صيغة عشبية مركبة وليست نباتًا مفردًا. كما عالجت هذه المرحلة تكرارات نصية مثل curcuma وbacopa_monnieri الناتجة عن اختلافات في التطبيع النصي.

بعد ذلك، تم تنفيذ Stage 10D — Taxonomic Harmonization بهدف مراجعة بعض الأسماء التي قد تمثل أسماء شائعة أو مرادفات نباتية. وفي هذه المرحلة، تم التعامل مع passionflower باعتباره اسمًا شائعًا مرتبطًا بـ Passiflora incarnata، وتم توثيق قرار دمجه ضمن الاسم الموحد passiflora_incarnata_l. ثم جاءت Stage 10E — Final Passiflora Harmonization لمعالجة التكرار النهائي بين passiflora_incarnat_l وpassiflora_incarnata_l، حيث تم اعتماد الصيغة الأخيرة كاسم نهائي موحد.

بعد هذه المراحل التصحيحية، أصبحت النتائج النهائية المعتمدة تضم 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا. كما تم تثبيت وجود 17 نباتًا بتغطية جينية مرتفعة، و68 مركبًا بتغطية مرضية عالية، و8 مركبات ذات تغطية مرضية كاملة. وتم اعتماد هذه الأرقام في فصل النتائج، المناقشة، الاستنتاجات، والملخصات النهائية.

Table 3.19. إجراءات التصحيح المطبقة في **Stages 10C–10E**

الأثر المنهجي	العناصر المصححة	الإجراء
تنقية طبقة النباتات	rutin, hesperidin, red_onion_husk_extract, hangekobokuto	حذف إدخالات غير نباتية
منع تضخيم الترتيب	curcuma, bacopa_monnieri	دمج تكرارات نصية
إزالة التكرار التصنيفي	passionflower و passiflora_incarnat_l → passiflora_incarnata_l	توحيد Passiflora
اعتماد نسخة نتائج نهائية موحدة	62 نباتًا؛ 299 مركبًا؛ 388 جينًا؛ 15 مرضًا	تثبيت الأرقام النهائية

Table 3.20. ملفات **Stage 10E** النهائية المستعملة في فصل النتائج

الاستخدام في الفصل الرابع	الملف
ملخص الحجم النهائي للشبكة	Table_4_1_final_network_size_summary.csv
القائمة الكاملة للنباتات الموحدة	stage10E_full_final_harmonized_plants.csv
ترتيب أعلى 15 نباتًا	Table_4_2_top15_plants_FINAL.csv
النباتات ذات التغطية الجينية العالية	Table_4_6_high_gene_coverage_plants_FINAL.csv
المركبات ذات التغطية المرضية العالية	Table_4_7_high_disease_coverage_compounds_FINAL.csv
المركبات كاملة التغطية المرضية	Table_4_8_full_disease_coverage_compounds_FINAL.csv
الشكل النهائي لأعلى النباتات حسب الدرجة التكاملية	Figure_4_1_top15_plants_integrated_score_FINAL.png
الشكل النهائي للتغطية الجينية للنباتات	Figure_4_4_top_plants_gene_coverage_FINAL.png

3.21. البرمجيات والأدوات المستعملة

نُفذت مراحل التحليل باستعمال بيئة حاسوبية قائمة على Python ضمن نظام عمل منظم يسمح بتشغيل السكريبتات، حفظ المخرجات، وتوثيق مراحل التنفيذ. وقد تم اختيار Python لأنه يوفر مكتبات قوية لمعالجة الجداول، الاتصال بواجهات قواعد البيانات، إنتاج ملفات ضبط الجودة، وإنشاء الجداول والأشكال النهائية.

استُعملت مكتبة pandas لقراءة ملفات البيانات، تنظيف الجداول، حساب التكرارات، إنشاء مفاتيح المطابقة، دمج الجداول، وحساب مؤشرات التغطية والترتيب. كما استُعملت مكتبات مثل os وpathlib وjson لإدارة الملفات والمجلدات، حفظ مخرجات QC، وتنظيم مسارات العمل عبر مراحل المشروع المختلفة.

تم استعمال مكتبة requests للاتصال بواجهات ChEMBL API أثناء مراحل مطابقة المركبات واستخراج بيانات النشاط الحيوي. وقد سمح ذلك بتحويل أسماء المركبات إلى معرفات قياسية، ثم استرجاع علاقات المركب–الهدف اللازمة لبناء طبقة الأدلة الحيوية داخل الشبكة.

أما إنتاج الأشكال البيانية فتم باستعمال matplotlib، حيث استُخدمت لإنشاء الرسوم الخاصة بترتيب النباتات، المركبات، الجينات، وتوزيع فئات التغطية. كما تم تنفيذ العمل داخل بيئة WSL/Linux على نظام Windows، مع توثيق WSL كأداة تنفيذ تقنية لا كاختصار علمي أساسي في المذكرة.

Table 3.22. البرمجيات والأدوات المستعملة

الدور في الدراسة	الأداة / المكتبة
تنفيذ خط سير التحليل	Python
تنظيف الجداول، الدمج، وحساب المؤشرات	Pandas
الاتصال بواجهات ChEMBL API	Requests
حفظ ملفات QC والتقارير المرحلية	Json
إدارة المسارات والمجلدات	pathlib / os
إنتاج الأشكال البيانية	Matplotlib
مطابقة المركبات واستخراج النشاط الحيوي	ChEMBL API
بيئة تنفيذ وتنظيم ملفات	WSL/Linux

3.23. إدارة الملفات وقابلية إعادة التنفيذ

اعتمدت الدراسة على تنظيم مرحلي للملفات والمجلدات، بحيث تُحفظ مخرجات كل مرحلة في مسار مستقل مرتبط بوظيفتها داخل خط سير التحليل. كان الهدف من هذا التنظيم هو ضمان قابلية تتبع النتائج، وتسهيل الرجوع إلى أي مرحلة عند الحاجة، وتجنب الخلط بين البيانات الخام، البيانات المنظفة، مخرجات ChEMBL، ملفات الربط الجيني، وجداول النتائج النهائية.

```
00_raw/
01_intake/
02_clean/
03_registry/
04_chembl/
05_qc/
06_mapping/
08_network/
09_outputs/
logs/
checkpoints/
config/
```

احتوى مجلد raw_00 على ملف الإدخال الأصلي، بينما خُصصت المجلدات اللاحقة لمخرجات الفحص الأولي، التنظيف، سجل المركبات، بيانات ChEMBL، الربط الجيني، والشبكة المدمجة. أما مجلد outputs_09 فقد استُعمل لحفظ الجداول والأشكال النهائية، بما في ذلك مخرجات Stage 9B وStages 10C–10E التي أنتجت النسخة المصححة من النتائج.

تُعد قابلية إعادة التنفيذ عنصرًا أساسيًا في هذه الدراسة. فقد تم تصميم السكريبتات بحيث يمكن إعادة تشغيل مرحلة محددة عند الحاجة دون إعادة كامل المشروع. على سبيل المثال، بعد اكتشاف أخطاء في جداول النباتات، لم يكن من الضروري إعادة تنفيذ كل مراحل ChEMBL، بل تم تنفيذ مراحل تصحيحية لاحقة ركزت على الجداول النهائية.

3.24.نظام ضبط الجودة QC

اعتمدت الدراسة على نظام ضبط جودة مرحلي QC — Quality Control لمراقبة مخرجات كل مرحلة من مراحل خط سير التحليل. وكان الهدف من هذا النظام هو التأكد من أن البيانات لا تنتقل من مرحلة إلى أخرى دون توثيق، وأن كل عملية تنظيف، فلترة، مطابقة، أو تصحيح يمكن تتبعها لاحقًا.

تم إنشاء ملفات QC بصيغة JSON في مراحل متعددة من العمل. احتوت هذه الملفات على معلومات مثل run_time_utc، input_file، output_files، row counts، unique entity counts، filtering criteria، removed entries، manual review flags، و.final accepted counts

كان نظام QC مهمًا بشكل خاص في المراحل النهائية. ففي Stage 10C، وثّق ملف QC الإدخالات التي حُذفت من جداول النباتات، ومنها rutin، hesperidin، red_onion_husk_extract، وhangekobokuto. وفي Stage 10D وStage 10E، وثقت ملفات QC قرارات التوحيد النباتي، خصوصًا دمج passionflower وpassiflora_incarnat_l ضمن passiflora_incarnata_l.

بعد هذه التصحيحات، وثّق ملف QC النهائي أن الشبكة المعتمدة تضم 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا، إضافة إلى 17 نباتًا بتغطية جينية مرتفعة، 68 مركبًا بتغطية مرضية عالية، و8 مركبات بتغطية مرضية كاملة. وبذلك، مثّل نظام QC عنصرًا جوهريًا في رفع موثوقية الدراسة.

3.25. حدود المنهجية

رغم أن المنهجية المعتمدة في هذه الدراسة صُممت لتكون مرحلية، قابلة للتتبع، ومدعومة بملفات QC، إلا أنها تبقى محدودة بعدة عوامل يجب توضيحها منذ فصل المواد والطرق. ويُعد عرض هذه الحدود داخل المنهجية ضروريًا حتى لا تُفهم النتائج اللاحقة على أنها أدلة علاجية أو تجريبية نهائية.

أول حدود المنهجية يرتبط بطبيعة Cluster 3 نفسه. فقد تم التعامل معه كإطار تحليلي وظيفي يجمع أمراضًا مرتبطة بمحاور خلل النواقل العصبية والميتوكوندريا، وليس كتصنيف سريري رسمي أو مرض واحد. لذلك، فإن أي قياس للتغطية عبر أمراض هذا الـ Cluster يعكس تمثيلًا شبكيًا داخل مجموعة أمراض مختارة.

ثانيًا، تعتمد المنهجية على قواعد بيانات عامة، خاصة ChEMBL. ورغم أنها توفر بيانات منظمة حول المركبات والنشاط الحيوي والأهداف، إلا أنها تعكس فقط ما تم اختباره ونشره سابقًا. لذلك، قد تكون بعض المركبات أو الجينات ممثلة أكثر من غيرها بسبب كثافة الدراسات المنشورة، وليس بالضرورة بسبب أهميتها البيولوجية الخاصة داخل Cluster 3.

ثالثًا، تعتمد نتائج الشبكة على جودة أسماء المركبات والنباتات في البيانات الأولية. ورغم أن الدراسة شملت مراحل تنظيف وتصحيح متقدمة، بما في ذلك Stages 10C–10E، إلا أن التوحيد التصنيفي النباتي الكامل لم يُنفذ باستخدام قواعد خارجية متخصصة مثل POWO أو World Flora Online. لذلك، يبقى احتمال وجود بعض المرادفات النباتية غير المكتشفة محدودًا لكنه وارد.

كما أن المؤشرات التكاملية تعتمد على أوزان تحليلية محددة، وأن فلترة النشاط الحيوي اعتمدت حد ≥10 µM، وأن الدراسة لم تشمل ADME/ADMET أو molecular docking أو molecular dynamics أو pathway enrichment أو Gene–Disease overlap analysis. لذلك، يجب تفسير المنهجية باعتبارها إطارًا حاسوبيًا منظّمًا لتوليد الفرضيات البحثية، لا لإثبات الفعالية العلاجية.

3.26. حوصلة الفصل الثالث

قدّم هذا الفصل المنهجية المعتمدة في بناء وتحليل شبكة Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction وفق مقاربة network pharmacology. وقد تم تصميم المنهجية لتتعامل مع Cluster 3 باعتباره مجموعة أمراض وظيفية متعددة، لا مرضًا منفردًا، وذلك من خلال بناء شبكة متعددة الطبقات تربط بين النباتات، المركبات، الجينات، والأمراض.

بدأ خط سير العمل بفحص البيانات الخام وتنظيفها، ثم بناء سجل مركبات موحد، ومطابقة المركبات مع قاعدة ChEMBL، واستخراج بيانات النشاط الحيوي، وفلترة علاقات المركب–الهدف وفق معايير محددة. بعد ذلك، تم تحويل الأهداف إلى رموز جينية، ثم دمج طبقات البيانات في شبكة واحدة من نوع Plant → Compound → Gene → Disease.

كما شملت المنهجية حساب مؤشرات التغطية والترتيب على مستوى النباتات، المركبات، والجينات، إضافة إلى تحليل فئات التغطية لفهم توزيع الكيانات داخل الشبكة بدل الاكتفاء بعرض الكيانات الأعلى ترتيبًا فقط.

وتُعد الإضافة المنهجية الأهم في النسخة النهائية من هذا الفصل هي إدراج مراحل التصحيح النهائية Stages 10C–10E. فقد سمحت هذه المراحل بمراجعة جداول النتائج قبل اعتمادها، حذف الإدخالات غير النباتية من طبقة النباتات، ودمج

الأسماء المتكررة أو المترادفة، خاصة الصيغ المرتبطة بـ Passiflora incarnata. وبذلك، أصبحت النتائج النهائية متوافقة مع شبكة مصححة تضم 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا، مع تثبيت 17 نباتًا عالي التغطية الجينية، و68 مركبًا عالي التغطية المرضية، و8 مركبات كاملة التغطية المرضية.

يضمن هذا التحديث أن تكون المنهجية متطابقة مع النتائج النهائية المعروضة في الفصل الرابع، وأن لا تظهر التصحيحات كإجراء لاحق غير موثق. كما يعزز ذلك قابلية تتبع الدراسة ومراجعتها، لأن كل رقم نهائي في فصل النتائج يمكن ربطه بمرحلة منهجية محددة وملف QC مناسب.

بناءً على ذلك، يشكل هذا الفصل الأساس المنهجي للدراسة، ويوضح أن النتائج النهائية لم تكن مجرد مخرجات مباشرة من قواعد البيانات، بل ناتجة عن خط سير تحليلي منظّم، مرحلي، وقابل للتصحيح. ومع ذلك، تبقى هذه المنهجية حاسوبية واستكشافية، وتهدف إلى توليد فرضيات بحثية قابلة للتحقق لاحقًا عبر تحليلات إضافية مثل pathway enrichment، ADME/ADMET، molecular docking، molecular dynamics، والتحقق التجريبي.

خـــلاصة

4.1. تمهيد الفصل

يختتم هذا الفصل الدراسة من خلال تقديم قراءة نهائية لحدودها، واستنتاجاتها، وتوصياتها المستقبلية بعد اعتماد النسخة المصححة من النتائج الناتجة عن Stage 10E - Final Passiflora Harmonization. وتُعد هذه المرحلة حاسمة في البناء النهائي للمذكرة، لأنها نقلت النتائج من نسخة أولية تضمنت بعض أخطاء التصنيف والتكرار إلى نسخة أكثر اتساقًا من الناحية المنهجية، خاصة في طبقة النباتات.

بعد هذا التصحيح، أصبحت الشبكة النهائية تضم 62 كيانًا نباتيًا موحّدًا، 299 مركبًا، 388 جينًا، و15 مرضًا ضمن Cluster 3 - Neurotransmitter & Mitochondrial Dysfunction. كما أظهرت النتائج وجود 17 نباتًا بتغطية جينية مرتفعة، و68 مركبًا بتغطية مرضية عالية، إضافة إلى 8 مركبات ذات تغطية مرضية كاملة. وتمثل هذه الأرقام النسخة النهائية المعتمدة في الاستنتاجات العامة لهذه الدراسة.

تقوم هذه الدراسة على تحليل شبكة متعددة الطبقات من نوع Plant -> Compound -> Gene -> Disease. وقد سمح هذا النموذج بفحص العلاقات بين النباتات، المركبات، الجينات، والأمراض داخل مجموعة مرضية وظيفية متعددة، بدل التعامل مع كل مرض أو نبات أو مركب بشكل منفصل. ومن هذا المنظور، لا تُقرأ النتائج باعتبارها إثباتًا علاجيًا مباشرًا، بل باعتبارها مخرجات حاسوبية منظّمة قادرة على توليد فرضيات بحثية قابلة للتحقق لاحقًا.

يكتسب هذا الفصل أهميته من كونه يضع حدود النتائج في سياقها الصحيح. فظهور نبات مثل Ginkgo biloba أو مركب مثل quercetin ضمن المراتب العليا لا يمثل بالضرورة إضافة جديدة في ذاته، نظرًا لكثافة الأدبيات حولهما. في المقابل، فإن بروز Passiflora incarnata بعد التصحيح النهائي يمنح الدراسة قيمة تفسيرية خاصة، لأنه يرتبط بمحور النواقل العصبية وخصوصًا المسار GABAergic، وهو محور يتقاطع مباشرة مع طبيعة Cluster 3.

لذلك، يهدف هذا الفصل إلى الإجابة عن سؤالين أساسيين: ماذا يمكن أن نستنتج بثقة من هذه الشبكة المصححة؟ وما الذي لا يمكن استنتاجه دون تحليلات لاحقة مثل ADME/ADMET، وmolecular docking، وmolecular dynamics، وpathway enrichment، والتحقق التجريبي؟

بناءً على ذلك، لا يُقدّم هذا الفصل خاتمة وصفية فقط، بل يقدّم تقييمًا نقديًا لمخرجات الدراسة، يميز بين قوة النتائج كإطار لترتيب الأولويات البحثية، وحدودها كدراسة حاسوبية غير تجريبية. وتمثل هذه القراءة النهائية أساسًا لتحديد المسار المستقبلي الأنسب لتطوير العمل، خاصة نحو اختبار مركبات أكثر ارتباطًا بمحور النواقل العصبية مثل vitexin، isovitexin، apigenin ، وchrysin المرتبطة بـ Passiflora incarnata.

4.2. الحدود العامة للدراسة بعد التصحيح

رغم أن التصحيح النهائي في Stage 10E - Final Passiflora Harmonization حسّن جودة النتائج ورفع موثوقية طبقة النباتات، فإن الدراسة تبقى محدودة بطبيعتها الحاسوبية والاستكشافية. فقد أصبح العدد النهائي المعتمد للشبكة 62 كيانًا نباتيًا موحّدًا، 299 مركبًا، 388 جينًا، و15 مرضًا، مع 17 نباتًا عالي التغطية الجينية، و68 مركبًا عالي التغطية المرضية، و8 مركبات كاملة التغطية المرضية. ومع ذلك، فإن هذه الأرقام تمثل بنية شبكية ناتجة عن قواعد بيانات ومعايير فلترة محددة، ولا تمثل تحققًا تجريبيًا أو سريريًا.

4.2.1. حدود تعريف Cluster 3

يمثل Cluster 3 - Neurotransmitter & Mitochondrial Dysfunction إطارًا تحليليًا وظيفيًا، وليس تصنيفًا سريريًا رسميًا. لذلك، فإن الأمراض المدرجة ضمن هذا الـ Cluster لا تُعامل باعتبارها مرضًا واحدًا أو مجموعة متجانسة تمامًا، بل باعتبارها أمراضًا قد تتقاطع في محاور بيولوجية مشتركة مثل خلل النواقل العصبية، اضطراب الميتوكوندريا،

الإجهاد التأكسدي، الالتهاب العصبي، وproteostasis. وعليه، فإن التداخل الذي أظهرته الشبكة بين النباتات، المركبات، الجينات، والأمراض يجب أن يُفسر كتداخل شبكي وظيفي، وليس كدليل على وحدة مرضية واحدة أو آلية سببية مشتركة مثبتة.

4.2.2. حدود التصحيح النباتي

أدى Stage 10E إلى تحسين طبقة النباتات من خلال حذف الإدخالات غير النباتية ودمج الصيغ المتكررة المرتبطة بـ Passiflora incarnata. وقد ساعد ذلك على تصحيح أخطاء مهمة مثل إدراج مركبات أو مستخلصات ضمن جداول النباتات. ومع ذلك، لا يمثل هذا التصحيح توحيدًا تصنيفيًا نباتيًا كاملًا باستخدام قواعد نباتية متخصصة مثل POWO أو World Flora Online أو GBIF.

4.2.3. حدود قواعد البيانات

اعتمدت الدراسة أساسًا على قواعد بيانات عامة، خاصة ChEMBL، لاستخراج علاقات المركب-الهدف وربطها بالجينات. ورغم أن ChEMBL تُعد قاعدة قوية ومنظمة، إلا أنها تعكس فقط المركبات والأهداف التي دُرست ونُشرت عنها بيانات نشاط حيوي. لذلك، قد تظهر بعض المركبات أو الجينات بتغطية مرتفعة لأنها مدروسة بكثافة في الأدبيات، وليس بالضرورة لأنها الأكثر أهمية بيولوجيًا داخل Cluster 3.

4.2.4. حدود أوزان الدرجات التكاملية

استُخدمت في الدراسة مؤشرات تكاملية مثل plant integrated score وcompound integrated score وdegree-like hub score. هذه المؤشرات مفيدة لترتيب الكيانات داخل الشبكة، لكنها تعتمد على أوزان تحليلية محددة مسبقًا. لذلك، فإن ترتيب النباتات أو المركبات قد يتغير إذا تغيرت الأوزان، ولهذا تبقى النتائج الحالية مرتبطة بنموذج الأوزان المعتمد، وتحتاج مستقبلًا إلى sensitivity analysis.

4.2.5. حدود معيار النشاط الحيوي <=10 **uM**

اعتمدت الدراسة حدًا قدره bioactivity threshold <=10 uM لقبول روابط النشاط الحيوي بين المركبات والأهداف. وهذا الحد يوازن بين الصرامة والحفاظ على حجم شبكة كافٍ للتحليل، لكنه ليس الحد الوحيد الممكن؛ إذ يمكن استعمال حدود أكثر صرامة مثل <=5 uM أو <=1 uM أو pChEMBL =>5 أو pChEMBL =>6.

4.2.6. حدود غياب **ADME/ADMET**

لم تتضمن الدراسة الحالية تحليل ADME/ADMET، ولذلك لا يمكن الحكم على خصائص مهمة للمركبات مثل الامتصاص، التوافر الحيوي، عبور الحاجز الدموي الدماغي، السمية، تداخلات CYP، أو الاستقرار الاستقلابي. وهذه النقطة مهمة جدًا في Cluster 3، لأن محاوره مرتبطة بالجهاز العصبي.

4.2.7. حدود غياب **docking** و**molecular dynamics**

لم يتم تنفيذ molecular docking أو molecular dynamics في هذه الدراسة. لذلك، فإن روابط المركبات بالأهداف لا تعني أن المركبات ترتبط بنيويًا وبشكل مستقر بتلك الأهداف. غياب docking يعني أن الدراسة لم تختبر binding pose أو binding affinity أو protein-ligand interactions أو active site compatibility، بينما غياب molecular dynamics يعني أنها لم تختبر استقرار المعقدات الجزيئية عبر الزمن.

4.2.8. حدود غياب pathway enrichment

رغم أن الدراسة ناقشت محاور مثل خلل الميتوكوندريا، اضطراب النواقل العصبية، الإجهاد التأكسدي، الالتهاب العصبي، وproteostasis، إلا أن هذه المحاور لم تُثبت بعد عبر تحليل pathway enrichment مستقل. لذلك، يبقى الربط البيولوجي في هذه الدراسة تفسيرًا مدعومًا بالأدبيات وببنية الشبكة، لكنه يحتاج إلى تحليل enrichment باستخدام KEGG أو Reactome أو Gene Ontology أو WikiPathways.

4.2.9. حدود غياب التحقق التجريبي

لم تشمل الدراسة تجارب خلوية أو حيوانية أو سريرية. لذلك، لا يمكن اعتبار النتائج دليلًا على فعالية علاجية مباشرة لأي نبات أو مركب أو جين. وبناءً على ذلك، فإن نتائج الدراسة يجب أن تُقرأ باعتبارها hypothesis-generating network results، أي نتائج تولد فرضيات بحثية قابلة للاختبار لاحقًا.

بشكل عام، حسّن التصحيح النهائي موثوقية الدراسة، لكنه لم يلغِ حدودها الأساسية كدراسة حاسوبية قائمة على قواعد بيانات. لذلك، فإن قوة العمل تكمن في إنتاج شبكة مصححة ومنظمة تساعد على ترتيب الأولويات البحثية، بينما تبقى مراحل pathway enrichment، ADME/ADMET، docking، molecular dynamics ، والتحقق التجريبي ضرورية قبل أي استنتاج دوائي أو علاجي قوي.

4.3. الاستنتاجات الرئيسية

4.3.1. بناء شبكة متعددة الطبقات قابلة للتتبع

تمكنت الدراسة من بناء شبكة متعددة الطبقات من نوع Plant -> Compound -> Gene -> Disease تضم في نسختها النهائية 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا. ويُعد هذا البناء أحد أهم مخرجات الدراسة، لأنه سمح بدمج مصادر نباتية، مركبات كيميائية، أهداف جينية، ومجموعة أمراض وظيفية ضمن إطار تحليلي واحد.

4.3.2. التصحيح النهائي حسّن موثوقية النتائج

أظهر التصحيح النهائي أن جودة النتائج لا تعتمد فقط على حجم الشبكة، بل على نقاء كل طبقة من طبقاتها. فقد أدى حذف المركبات والمستخلصات والصيغ العشبية المركبة من طبقة النباتات إلى اعتماد قائمة نباتية أكثر دقة، كما سمح توحيد صيغ Passiflora incarnata بتجنب تضخيم ترتيب النبات أو تكراره تحت أسماء مختلفة.

4.3.3. Cluster 3 يمثل شبكة موزعة لا كيانًا واحدًا

أظهرت النتائج أن Cluster 3 لا يعتمد على نبات واحد أو مركب واحد أو جين واحد، بل يقوم على بنية موزعة متعددة الطبقات. فقد ظهرت عدة نباتات ذات تمثيل شبكي مرتفع، و17 نباتًا بتغطية جينية مرتفعة، و68 مركبًا بتغطية مرضية عالية.

4.3.4. Ginkgo biloba وquercetin يؤكدان اتساق المنهجية لا الجدة العلمية

أظهر التحليل ظهور كيانات معروفة مثل Ginkgo biloba على مستوى النباتات وquercetin على مستوى المركبات. ويُعد ذلك متسقًا مع كثافة الأدبيات حولهما في سياق الاضطرابات العصبية والإجهاد التأكسدي والالتهاب، لكنه لا يمثل الجدة الأهم في الدراسة.

4.3.5. Passiflora incarnata يمثل إضافة تفسيرية مهمة

يمثل ظهور Passiflora incarnata ضمن النباتات ذات التمثيل الشبكي المرتفع إحدى النتائج الأكثر أهمية من الناحية التفسيرية. فبعد التصحيح النهائي ودمج الصيغ المختلفة لهذا النبات، بقي ضمن الطبقة العليا من النباتات، وهو ما يجعله مرشحًا شبكيًا مهمًا داخل Cluster 3، خصوصًا بالنظر إلى ارتباطه المحتمل بمحور GABAergic modulation.

3.6.4. طبقة المركبات أظهرت انتشارًا مرضيًا واسعًا

أظهرت الدراسة وجود 68 مركبًا بتغطية مرضية عالية، و8 مركبات ذات تغطية مرضية كاملة داخل أمراض Cluster 3 الخمسة عشر. ويشير ذلك إلى أن طبقة المركبات تمثل محورًا مهمًا في الربط بين الأمراض داخل الشبكة، مع ضرورة تفسير هذا الانتشار بوصفه network coverage فقط.

3.7. 4 الجينات المحورية تحتاج تفسيرًا حذرًا

أظهرت طبقة الجينات وجود جينات ذات degree-like hub score مرتفع. غير أن هذه الجينات تمثل عقدًا عالية التمثيل داخل الشبكة، وليست بالضرورة أهدافًا علاجية مثبتة. كما أن بعض الجينات أو العائلات الجينية كثيرة الدراسة قد تظهر بقوة بسبب تحيز قواعد البيانات.

3.8. 4 الدراسة تنتج فرضيات بحثية لا أدلة علاجية

الاستنتاج العام الأهم هو أن هذه الدراسة تقدم إطارًا قويًا لتوليد الفرضيات البحثية داخل Cluster 3، لكنها لا تقدم إثباتًا علاجيًا مباشرًا. فالنتائج الحالية تُحدد كيانات ذات أولوية بحثية وتبرز محاور تستحق التحقق، لكنها تحتاج إلى مراحل لاحقة تشمل ADME/ADMET، docking، molecular dynamics، enrichment analysis ، والتحقق التجريبي.

4.4. القيمة العلمية للدراسة بعد التصحيح

تتمثل القيمة العلمية لهذه الدراسة في أنها لم تكتفِ ببناء شبكة حاسوبية أولية، بل قامت بمراجعة نتائجها وتصحيحها قبل اعتماد الاستنتاجات النهائية. فقد أظهر Stage 10E أن بعض الأخطاء الصغيرة في التصنيف أو التسمية يمكن أن تؤثر مباشرة في قراءة النتائج، خاصة عندما تُدرج مركبات أو مستخلصات ضمن طبقة النباتات.

أول قيمة علمية للدراسة هي القيمة المنهجية. فقد تم اعتماد مسار تحليلي مرحلي قابل للتتبع، يبدأ من تنظيف البيانات ومطابقة المركبات، ثم استخراج النشاط الحيوي، والربط الجيني، ودمج الشبكة، وصولًا إلى تحليل التغطية. كما ساعدت ملفات QC ومرحلة التصحيح النهائية على كشف الأخطاء وتصحيحها، بدل تمريرها إلى الاستنتاجات.

القيمة الثانية هي القيمة الشبكية. فقد نجحت الدراسة في بناء شبكة متعددة الطبقات تضم 62 نباتًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا. هذه البنية تسمح بفهم Cluster 3 كشبكة موزعة بين النباتات والمركبات والجينات والأمراض، بدل التعامل معه كقائمة منفصلة من النباتات أو المركبات.

القيمة الثالثة هي القيمة التفسيرية. فقد أظهرت النتائج أن بعض المرشحات مثل Ginkgo biloba وquercetin كانت متوقعة بسبب كثافة الأدبيات حولهما، بينما أضاف ظهور Passiflora incarnata بعد التصحيح بعدًا تفسيريًا أكثر خصوصية. فارتباط Passiflora بمحور النواقل العصبية، خاصة GABAergic modulation، يجعل ظهوره منسجمًا مع هوية Cluster 3 أكثر من بعض المرشحات العامة واسعة الشهرة.

كما أن الدراسة أبرزت أهمية عدم الاكتفاء بالمركبات العامة مثل quercetin، بل توجيه الاهتمام إلى مركبات أكثر ارتباطًا بالنباتات ذات الدلالة التفسيرية، مثل مركبات Passiflora incarnata: vitexin، isovitexin، apigenin، وchrysin. وتمثل هذه المركبات نقطة انطلاق مناسبة لمراحل لاحقة، خاصة ADME/ADMET وdocking.

القيمة الرابعة هي القيمة النقدية. فقد حافظت الدراسة على تمييز واضح بين network coverage و therapeutic efficacy. وهذا التمييز يحمي النتائج من المبالغة، ويجعلها مناسبة لتوليد الفرضيات بدل تقديم ادعاءات علاجية غير مثبتة. كما أن الدراسة ناقشت بوضوح حدود قواعد البيانات، تحيز النشر، حدود الأوزان، وغياب التحقق التجريبي.

القيمة الخامسة هي القيمة التطويرية. فالدراسة لا تنتهي عند ترتيب النباتات والمركبات والجينات، بل تقترح مسارًا واضحًا للمتابعة، يشمل تحليل الحساسية، Gene-Disease overlap، Jaccard similarity، pathway enrichment ، ADME/ADMET، docking، molecular dynamics ، والتحقق التجريبي. وهذا يجعلها قاعدة قابلة للتطوير نحو عمل أكثر تقدمًا في CADD و network pharmacology.

بناءً على ذلك، يمكن اعتبار القيمة العلمية للدراسة بعد التصحيح في أنها تجمع بين التحليل الشبكي، ضبط الجودة، والنقد المنهجي. فهي لا تقدم فقط قائمة مرشحات، بل تقدم إطارًا قابلًا للتتبع والتطوير لفهم 3 Cluster، مع إبراز Passiflora incarnata كإضافة تفسيرية مهمة تستحق التحقق في المراحل اللاحقة.

4.5. التوصيات المستقبلية

4.5.1. تنفيذ **sensitivity analysis** للأوزان

يُوصى بإجراء تحليل حساسية لاختبار مدى استقرار ترتيب النباتات والمركبات والجينات عند تغيير أوزان integrated scores. يمكن اختبار أوزان متساوية، أو زيادة وزن gene coverage، أو زيادة وزن disease coverage، أو زيادة وزن exact edge ratio، أو تقليل تأثير المركبات العامة كثيرة الانتشار.

4.5.2. تنفيذ **Gene-Disease overlap analysis**

ينبغي تنفيذ تحليل تداخل الجينات بين أمراض 3 Cluster عبر إنشاء Gene-Disease overlap matrix. يسمح هذا التحليل بتحديد عدد الجينات المشتركة بين كل زوج من الأمراض، وبالتالي اختبار درجة التداخل الوظيفي داخل المجموعة المرضية.

4.5.3. حساب **Jaccard similarity** بين الأمراض

إضافة إلى حساب عدد الجينات المشتركة، يُوصى بحساب Jaccard similarity بين أمراض 3 Cluster، وفق العلاقة: Jaccard similarity = عدد الجينات المشتركة / عدد الجينات الكلية الفريدة بين المرضين. ويمكن عرض النتائج في شكل heatmap.

4.5.4. تنفيذ **pathway enrichment**

يُوصى بتحليل قائمة الجينات النهائية باستعمال قواعد مثل KEGG، Reactome، Gene Ontology ، وWikiPathways. يسمح هذا التحليل بالتحقق مما إذا كانت المحاور التي نوقشت نظريًا تظهر فعليًا كمسارات غنية إحصائيًا.

4.5.5. تنفيذ **ADME/ADMET** للمركبات ذات الأولوية

يجب إخضاع المركبات ذات التغطية العالية، خاصة المركبات كاملة التغطية والمركبات المرتبطة بـ Passiflora incarnata، لتحليل ADME/ADMET يشمل GI absorption، bioavailability، BBB permeability، CYP interactions، toxicity alerts، drug-likeness ، و synthetic accessibility.

4.5.6. إعطاء أولوية لمركبات **Passiflora incarnata**

لا ينبغي أن تقتصر المراحل اللاحقة على مركبات عامة ومدروسة بكثافة مثل quercetin فقط. يُوصى بإعطاء أولوية لمركبات مثل vitexin، isovitexin، orientin، isoorientin، apigenin ، وchrysin، لأنها قد تكون أكثر ارتباطًا بمحور النواقل العصبية وGABAergic modulation.

4.5.7. تنفيذ **molecular docking**

بعد ADME/ADMET، يمكن اختيار عدد محدود من المركبات والأهداف لتنفيذ molecular docking. يجب أن يعتمد اختيار المركبات على network coverage، exact mapping، ADME acceptability، BBB ، permeability، وbiological relevance.

4.5.8. تنفيذ **molecular dynamics**

بالنسبة لأفضل نتائج docking، يُوصى بتنفيذ molecular dynamics simulation لتقييم استقرار معقدات المركب-البروتين عبر الزمن، مع متابعة RMSD، RMSF، Radius of gyration، Hydrogen bonds ، وMM/PBSA أو MM/GBSA binding energy.

4.5.9. التحقق التجريبي

في المرحلة النهائية، يجب اختبار المركبات أو النباتات المرشحة عبر نماذج تجريبية مناسبة لمحاور Cluster 3، مثل GABAergic signaling assays، ATP production، inflammatory cytokines، neuronal viability ، وoxidative stress markers، mitochondrial membrane potential، apoptosis markers.

4.5.10. تحسين التوحيد التصنيفي للنباتات

رغم أن Stage 10E صحح الأخطاء الكبرى في طبقة النباتات، إلا أن الدراسات المستقبلية يجب أن تنفذ توحيدًا تصنيفيًا أعمق باستعمال قواعد نباتية متخصصة مثل POWO، World Flora Online ، وGBIF.

بشكل عام، ينبغي أن ينتقل العمل مستقبلًا من مستوى network coverage إلى مستوى التحقق المتدرج. ويبدأ ذلك بتحليل حساسية للأوزان، ثم تحليل تداخل الجينات والمسارات، يليه ADME/ADMET، ثم docking و molecular dynamics، وأخيرًا validation تجريبي. وتُعد Passiflora incarnata ومركباتها من أكثر الاتجاهات الواعدة التي تستحق المتابعة، لأنها تضيف بعدًا أكثر ارتباطًا بمحور النواقل العصبية مقارنة بالمرشحات العامة واسعة الشهرة.

4.6. الخلاصة العامة النهائية

قدمت هذه الدراسة تحليلًا شبكيًا متعدد الطبقات لـ Cluster 3 - Neurotransmitter & Mitochondrial Dysfunction باعتباره مجموعة أمراض وظيفية متعددة، وليس مرضًا منفردًا أو تصنيفًا سريريًا رسميًا. وقد اعتمدت الدراسة على منهجية network pharmacology لبناء شبكة تربط بين النباتات، المركبات الكيميائية، الجينات، والأمراض وفق نموذج Plant -> Compound -> Gene -> Disease.

بعد تنفيذ التصحيح النهائي في Stage 10E - Final Passiflora Harmonization، أصبحت الشبكة النهائية تضم 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا. كما أظهرت النتائج وجود 17 نباتًا بتغطية جينية مرتفعة، و68 مركبًا بتغطية مرضية عالية، إضافة إلى 8 مركبات كاملة التغطية المرضية. وقد مكّن هذا التصحيح من اعتماد نتائج أكثر اتساقًا بعد حذف الإدخالات غير النباتية ودمج الأسماء المتكررة أو المترادفة.

أظهرت الدراسة أن Cluster 3 لا يعتمد على نبات واحد أو مركب واحد أو جين واحد، بل يعكس بنية شبكية موزعة متعددة الطبقات. وقد برزت نباتات ومركبات معروفة مثل Ginkgo biloba وquercetin، وهو ما يدعم اتساق المنهجية مع الأدبيات. غير أن القيمة التفسيرية الأهم تمثلت في بروز Passiflora incarnata بعد التصحيح النهائي، نظرًا لارتباطه المحتمل بمحور النواقل العصبية، خصوصًا GABAergic modulation، وهو محور يتماشى مع هوية Cluster 3.

كما بيّنت الدراسة أن مركبات Passiflora incarnata، مثل vitexin، isovitexin، apigenin ، وchrysin، تمثل اتجاهًا بحثيًا واعدًا للمراحل اللاحقة، لأنها قد تكون أكثر خصوصية لمحور النواقل العصبية مقارنة ببعض المركبات العامة كثيرة الدراسة. ومع ذلك، فإن هذه المركبات لا تُعد مرشحات علاجية مثبتة في هذه المرحلة، بل مرشحات شبكية تحتاج إلى تحقق دوائي وتجريبي لاحق.

تؤكد الدراسة أيضًا أن مؤشرات network coverage وintegrated scores مفيدة لترتيب الأولويات البحثية، لكنها لا تكفي لإثبات الفعالية العلاجية أو تحديد أهداف علاجية نهائية. فالنتائج ما تزال محدودة بطبيعة قواعد البيانات، احتمال تحيز النشر، اختيار الأوزان، حد النشاط الحيوي، وغياب تحليلات مثل ADME/ADMET، docking ، molecular dynamics، pathway enrichment ، والتحقق التجريبي.

بناءً على ذلك، يمكن اعتبار هذه الدراسة خطوة أولى منظمة نحو فهم Cluster 3 من منظور حاسوبي شبكي. وتكمن قوتها في بناء شبكة قابلة للتتبع، تصحيح الأخطاء قبل اعتماد النتائج، وتمييز المرشحات العامة عن المرشحات الأكثر ارتباطًا بخصوصية الكتلة. أما المرحلة المستقبلية، فينبغي أن تنتقل من تحليل التغطية الشبكية إلى التحقق المتدرج، بدءًا من تحليل الحساسية وتداخل الجينات والمسارات، ثم تقييم ADME/ADMET، ثم docking وmolecular dynamics، وصولًا إلى validation تجريبي.

وفي الخلاصة، توفر هذه الدراسة أساسًا علميًا ومنهجيًا يمكن تطويره لاحقًا في إطار CADD و network pharmacology، خاصة إذا تم التركيز على المرشحات الأكثر اتساقًا مع محور النواقل العصبية مثل Passiflora incarnata ومركباتها، مع الحفاظ على القراءة النقدية التي تميز بين الفرضيات الشبكية والأدلة العلاجية المثبتة.

4.7. حوصلة الفصل الرابع

يمثل هذا الفصل الخاتمة النقدية للدراسة، حيث تم الانتقال من عرض النتائج ومناقشتها إلى تحديد حدودها واستنتاجاتها وآفاق تطويرها. وقد أظهرت الدراسة، بعد التصحيح النهائي في Stage 10E، أن شبكة Cluster 3 النهائية أصبحت أكثر اتساقًا من الناحية المنهجية، بعد اعتماد 62 كيانًا نباتيًا موحدًا، 299 مركبًا، 388 جينًا، و15 مرضًا، مع تحديد 17 نباتًا عالي التغطية الجينية، و68 مركبًا عالي التغطية المرضية، و8 مركبات كاملة التغطية المرضية.

تؤكد هذه النتائج أن قيمة الدراسة لا تكمن في إثبات فعالية علاجية مباشرة، بل في بناء إطار شبكي منظم يسمح بتوليد فرضيات بحثية قابلة للتطوير. كما أبرزت الدراسة أهمية التمييز بين المرشحات العامة والمدروسة بكثافة، مثل Ginkgo biloba وquercetin، وبين المرشحات ذات القيمة التفسيرية الأكثر خصوصية، مثل Passiflora incarnata ومركباته المرتبطة بمحور النواقل العصبية.

وبناءً على ذلك، فإن المسار المنهجي المنطقي بعد هذه الدراسة يتمثل في الانتقال من network coverage إلى validation pipeline، بدءًا بتحليل الحساسية للأوزان، ثم تحليل تداخل الجينات بين الأمراض، ثم pathway enrichment، يليه ADME/ADMET، docking، molecular dynamics ، وأخيرًا التحقق التجريبي. وبهذا، تمثل هذه الدراسة قاعدة أولية قوية يمكن تطويرها لاحقًا ضمن إطار CADD و.network pharmacology

Table 6.1. Final synthesis of study conclusions and future directions

المحور	الخلاصة النهائية	الاتجاه المستقبلي
الشبكة النهائية	شبكة متعددة الطبقات تضم نباتات، مركبات، جينات، وأمراض	توسيع التحليل بمقارنات بين clusters
النباتات	62 كيانًا نباتيًا بعد التصحيح	توحيد تصنيفي أعمق عبر POWO أو WFO
Passiflora incarnate	مرشح شبكي ذو قيمة تفسيرية لمحور النواقل العصبية	ADME/ADMET وdocking لمركباته
المركبات	299 مركبًا، منها 8 كاملة التغطية المرضية	فلترة دوائية وبنيوية لاحقة
الجينات	388 جينًا عالي التمثيل الشبكي	pathway enrichment و Gene-Disease overlap
حدود الدراسة	حاسوبية وتعتمد على قواعد بيانات	تحقق تجريبي وخطوات validation
القيمة العلمية	توليد فرضيات بحثية منظمة	تطوير نحو CADD متقدم

يعرض Table 6.1 خلاصة مركزة للاستنتاجات النهائية والاتجاهات المستقبلية المقترحة بعد التصحيح النهائي للشبكة.

Figure 6.1. Recommended future validation pathway

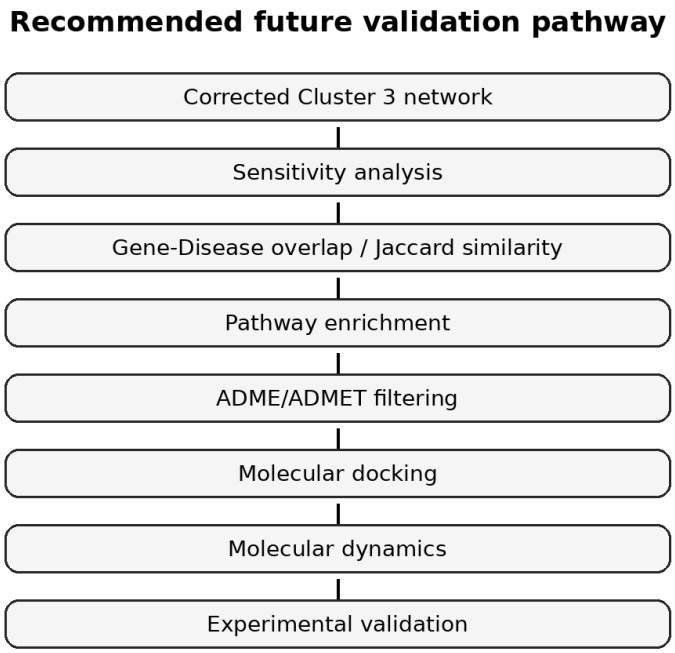


Figure 6.1. Proposed validation pathway for Cluster 3 candidates after network pharmacology analysis.

قائمة الاختصارات، المراجع، والملحقات

Cluster 3 — Neurotransmitter & Mitochondrial Dysfunction

نسخة نهائية محدثة بعد Stage 10E — Final Passiflora Harmonization

قائمة الاختصارات

تعرض هذه القائمة الاختصارات العلمية والتقنية الأساسية المستعملة في المذكرة. تم حذف الاختصارات غير العلمية من هذه القائمة، مثل WSL، والاكتفاء بذكرها ضمن أدوات التنفيذ في فصل المنهجية.

الاختصار	Full term	التوضيح
AD	Alzheimer's Disease	مرض ألزهايمر
ADME	Absorption, Distribution, Metabolism, Excretion	الامتصاص، التوزيع، الاستقلاب، الإطراح
ADMET	Absorption, Distribution, Metabolism, Excretion, Toxicity	الامتصاص، التوزيع، الاستقلاب، الإطراح، السمية
BBB	Blood–Brain Barrier	الحاجز الدموي الدماغي
CADD	Computer-Aided Drug Design	التصميم الدوائي بمساعدة الحاسوب
ChEMBL	Chemical database of bioactive molecules	قاعدة بيانات للمركبات النشطة حيويًا
CNS	Central Nervous System	الجهاز العصبي المركزي
DNA	Deoxyribonucleic Acid	الحمض النووي الريبوزي منقوص الأكسجين
EC50	Half-maximal Effective Concentration	التركيز الفعال لنصف الاستجابة القصوى
GABA	Gamma-Aminobutyric Acid	حمض غاما-أمينوبيوتيريك
GBIF	Global Biodiversity Information Facility	مرفق معلومات التنوع الحيوي العالمي
GO	Gene Ontology	أنطولوجيا الجينات
IC50	Half-maximal Inhibitory Concentration	التركيز المثبط لنصف الاستجابة القصوى
Ki	Inhibition Constant	ثابت التثبيط
Kd	Dissociation Constant	ثابت التفكك
KEGG	Kyoto Encyclopedia of Genes and Genomes	موسوعة كيوتو للجينات والجينومات
MD	Molecular Dynamics	الديناميكا الجزيئية
MM/GBSA	Molecular Mechanics Generalized Born Surface Area	طريقة تقدير طاقة الارتباط
MM/PBSA	Molecular Mechanics Poisson–Boltzmann Surface Area	طريقة تقدير طاقة الارتباط
pChEMBL	Negative logarithmic bioactivity value in ChEMBL	قيمة النشاط الحيوي اللوغاريتمية في ChEMBL
POWO	Plants of the World Online	قاعدة نباتات العالم على الإنترنت
QC	Quality Control	ضبط الجودة
RMSD	Root Mean Square Deviation	الانحراف الجذري المتوسط
RMSF	Root Mean Square Fluctuation	التذبذب الجذري المتوسط
ROS	Reactive Oxygen Species	أنواع الأكسجين التفاعلية
WFO	World Flora Online	الفلورا العالمية على الإنترنت

قائمة المراجع

تغطي قائمة المراجع المحاور المنهجية والبيولوجية والدوائية الأساسية للدراسة، بما في ذلك network pharmacology، قواعد البيانات الحيوية، خلل الميتوكوندريا، الالتهاب العصبي، Passiflora incarnata، Ginkgo biloba، quercetin، ومرحلة CADD المستقبلية.

[1] A. L. Hopkins, "Network pharmacology," Nature Biotechnology, vol. 25, no. 10, pp. 1110–1111, 2007, doi: 10.1038/nbt1007-1110.

[2] A. L. Hopkins, "Network pharmacology: the next paradigm in drug discovery," Nature Chemical Biology, vol. 4, no. 11, pp. 682–690, 2008, doi: 10.1038/nchembio.118.

[3] A.-L. Barabási, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," Nature Reviews Genetics, vol. 12, no. 1, pp. 56–68, 2011, doi: 10.1038/nrg2918.

[4] J. Gertsch, "Botanical drugs, synergy, and network pharmacology: forth and back to intelligent mixtures," Planta Medica, vol. 77, no. 11, pp. 1086–1098, 2011, doi: 10.1055/s-0030-1270904.

[5] S. Berger and R. Iyengar, "Network analyses in systems pharmacology," Bioinformatics, vol. 25, no. 19, pp. 2466–2472, 2009, doi: 10.1093/bioinformatics/btp465.

[6] B. Zdrazil et al., "The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods," Nucleic Acids Research, vol. 52, no. D1, pp. D1180–D1192, 2024, doi: 10.1093/nar/gkad1004.

[7] The UniProt Consortium, "UniProt: the Universal Protein Knowledgebase in 2025," Nucleic Acids Research, vol. 53, no. D1, pp. D609–D617, 2025, doi: 10.1093/nar/gkae1010.

[8] D. Ochoa et al., "The next-generation Open Targets Platform: reimagined, redesigned, rebuilt," Nucleic Acids Research, vol. 51, no. D1, pp. D1353–D1359, 2023, doi: 10.1093/nar/gkac1046.

[9] S. Kim et al., "PubChem 2023 update," Nucleic Acids Research, vol. 51, no. D1, pp. D1373–D1380, 2023, doi: 10.1093/nar/gkac956.

[10] D. Gfeller et al., "SwissTargetPrediction: a web server for target prediction of bioactive small molecules," Nucleic Acids Research, vol. 42, no. W1, pp. W32–W38, 2014, doi: 10.1093/nar/gku293.

[11] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," Scientific Reports, vol. 7, Art. no. 42717, 2017, doi: 10.1038/srep42717.

[12] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings," Advanced Drug Delivery Reviews, vol. 46, no. 1–3, pp. 3–26, 2001, doi: 10.1016/S0169-409X(00)00129-0.

[13] O. Trott and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," Journal of Computational Chemistry, vol. 31, no. 2, pp. 455–461, 2010, doi: 10.1002/jcc.21334.

[14] M. J. Abraham et al., "GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX, vol. 1–2, pp. 19–25, 2015, doi: 10.1016/j.softx.2015.06.001.

[15] J. D. Durrant and J. A. McCammon, "Molecular dynamics simulations and drug discovery," BMC Biology, vol. 9, Art. no. 71, 2011, doi: 10.1186/1741-7007-9-71.

[16] J. Nunnari and A. Suomalainen, "Mitochondria: in sickness and in health," Cell, vol. 148, no. 6, pp. 1145–1159, 2012, doi: 10.1016/j.cell.2012.02.035.

[17] B. Halliwell, "Oxidative stress and neurodegeneration: where are we now?," Journal of Neurochemistry, vol. 97, no. 6, pp. 1634–1658, 2006, doi: 10.1111/j.1471-4159.2006.03907.x.

[18] H. Sies, "Oxidative stress: a concept in redox biology and medicine," Redox Biology, vol. 4, pp. 180–183, 2015, doi: 10.1016/j.redox.2015.01.002.

[19] M. T. Heneka et al., "Neuroinflammation in Alzheimer's disease," The Lancet Neurology, vol. 14, no. 4, pp. 388–405, 2015, doi: 10.1016/S1474-4422(15)70016-5.

[20] P. H. Reddy and M. F. Beal, "Amyloid beta, mitochondrial dysfunction and synaptic damage: implications for cognitive decline in aging and Alzheimer's disease," Trends in Molecular Medicine, vol. 14, no. 2, pp. 45–53, 2008, doi: 10.1016/j.molmed.2007.12.002.

[21] K. Janda et al., "Passiflora incarnata in neuropsychiatric disorders: a systematic review," Nutrients, vol. 12, no. 12, Art. no. 3894, 2020, doi: 10.3390/nu12123894.

[22] S. M. Elsas et al., "Passiflora incarnata L. extracts elicit GABA currents in hippocampal neurons in vitro, and show anxiogenic and anticonvulsant effects in vivo, varying with extraction method," Phytomedicine, vol. 17, no. 12, pp. 940–949, 2010, doi: 10.1016/j.phymed.2010.03.002.

[23] M. K. Harit et al., "Randomized, double-blind, placebo-controlled clinical study of Passiflora incarnata in participants with stress and sleep problems," Nutrients, vol. 16, no. 7, Art. no. 1054, 2024, doi: 10.3390/nu16071054.

[24] M. Ożarowski et al., "Extracts and flavonoids of Passiflora species as promising anti-inflammatory and antioxidant substances," Molecules, vol. 26, no. 2, Art. no. 321, 2021, doi: 10.3390/molecules26020321.

[25] G. H. Kim et al., "Improvement in neurogenesis and memory function by administration of Passiflora incarnata L. extract applied to sleep disorder in rodent models," Journal of Chemical Neuroanatomy, vol. 98, pp. 27–40, 2019, doi: 10.1016/j.jchemneu.2019.04.003.

[26] S. Weinmann, S. Roll, C. Schwarzbach, C. Vauth, and S. N. Willich, "Effects of Ginkgo biloba in dementia: systematic review and meta-analysis," BMC Geriatrics, vol. 10, Art. no. 14, 2010, doi: 10.1186/1471-2318-10-14.

[27] Y. Luo, "Alzheimer's disease, the nematode Caenorhabditis elegans, and Ginkgo biloba leaf extract," Life Sciences, vol. 78, no. 18, pp. 2066–2072, 2006, doi: 10.1016/j.lfs.2005.12.004.

[28] F. Dajas, "Life or death: neuroprotective and anticancer effects of quercetin," Journal of Ethnopharmacology, vol. 143, no. 2, pp. 383–396, 2012, doi: 10.1016/j.jep.2012.07.005.

[29] S. Kumar and A. K. Pandey, "Chemistry and biological activities of flavonoids: an overview," The Scientific World Journal, vol. 2013, Art. no. 162750, 2013, doi: 10.1155/2013/162750.

[30] J. A. Ross and C. M. Kasum, "Dietary flavonoids: bioavailability, metabolic effects, and safety," Annual Review of Nutrition, vol. 22, pp. 19–34, 2002, doi: 10.1146/annurev.nutr.22.111401.144957.

[31] E. Nic Lughadha et al., "Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates," Phytotaxa, vol. 272, no. 1, pp. 82–88, 2016, doi: 10.11646/phytotaxa.272.1.5.

[32] GBIF Secretariat, "GBIF Backbone Taxonomy," Checklist dataset, Global Biodiversity Information Facility, 2023, doi: 10.15468/39omei.

الملحقات

Appendix A. Summary of QC Reports

أهم المخرجات	الهدف	المرحلة
تحديد الاعمدة، التكرارات، الصفوف المقبولة والمراجعة اليدوية	فحص البيانات الأولية	Stage 0
حذف/وسم المستخلصات، الخلطات، والإدخالات غير الدقيقة	التنظيف المحافظ	Stage 1
استبعاد الفئات الكيميائية العامة والاوصاف غير المحددة	تحسين التنظيف	Stage 1.5
compound registry وpriority tiers	سجل المركبات	Stage 2
مطابقة المركبات مع معرفات ChEMBL	ChEMBL resolution	Stage 3
استخراج علاقات المركب–الهدف	Bioactivity extraction	Stage 4
قبول العلاقات البشرية ذات النشاط الحيوي المناسب	Bioactivity filtering	Stage 5
تحويل الأهداف إلى رموز جينية	Target-to-gene mapping	Stage 6
بناء شبكة Plant–Compound–Gene–Disease	Network integration	Stage 7
فئات التغطية للنباتات والمركبات والجينات	Coverage distribution	Stage 9B
حذف non-plant entries من جداول النباتات	Corrective QC	Stage 10C
مراجعة passionflower	Taxonomic harmonization	Stage 10D
تثبيت النسخة النهائية المصححة	Final Passiflora harmonization	Stage 10E

Appendix B. Stage 10E Final Correction Note

تم تنفيذ Stage 10E كمرحلة تصحيح نهائية قبل اعتماد نتائج الفصل الرابع. هدفت هذه المرحلة إلى دمج الصيغ المختلفة المرتبطة بـ Passiflora incarnata وتثبيت النسخة النهائية المصححة للشبكة.

passiflora_incarnat_l → passiflora_incarnata_l

passionflower → passiflora_incarnata_l

كما تم اعتماد مخرجات Stage 10C التي أزالت الإدخالات غير النباتية من جداول النباتات: rutin, hesperidin, red_onion_husk_extract, hangekobokuto.

المؤشر	القيمة النهائية
النباتات بعد harmonization	62
المركبات	299
الجينات	388
الأمراض	15
النباتات ذات gene coverage ≥50%	17
المركبات ذات disease coverage ≥60%	68
المركبات ذات disease coverage =100%	8

Appendix C. Final Harmonized Plant Table

يعرض هذا الملحق القائمة النهائية للنباتات بعد Stage 10E. تم اختصار الأعمدة إلى المؤشرات الأكثر أهمية لتسهيل القراءة داخل المذكرة.

Plant	Diseases	Compounds	Genes	Disease cov. %	Gene cov. %	Integrated score
ginkgo_biloba	6	76	274	40.000	70.619	0.550
withania_somnifera	6	49	233	40.000	60.052	0.488
rhodiola_rosea	4	64	279	26.667	71.907	0.486
Curcuma	2	63	286	13.333	73.711	0.440
passiflora_incarnata_l	2	62	277	13.333	71.392	0.408
bacopa_monnieri	4	26	211	26.667	54.381	0.399
euterpe_oleracea	1	50	267	6.667	68.814	0.378
pueraria_lobata	1	49	269	6.667	69.330	0.377
crocus_sativus	4	26	169	26.667	43.557	0.364
mimosa_pudica_l	1	39	251	6.667	64.691	0.359
myristica_fragrans_houtt	1	46	235	6.667	60.567	0.353
garcinia_mangostana	1	38	240	6.667	61.856	0.351
valeriana_officinalis_l	1	47	228	6.667	58.763	0.349
lavandula_angustifolia	3	22	180	20.000	46.392	0.346
hericium_erinaceus	2	28	193	13.333	49.742	0.331
matricaria_chamomilla	1	37	208	6.667	53.608	0.326
alisma_orientale	1	37	201	6.667	51.804	0.319
glycyrrhiza_glabra	1	32	204	6.667	52.577	0.315
curcuma_longa	2	27	175	13.333	45.103	0.312
panax_quinquefolius	1	34	193	6.667	49.742	0.310
Kava	2	20	172	13.333	44.330	0.307
korean_red_ginseng	1	36	182	6.667	46.907	0.306
Coriander	1	30	190	6.667	48.969	0.305
rosmarinus_officinalis	2	46	158	13.333	40.722	0.304
andrographis_paniculata	1	28	199	6.667	51.289	0.304
uncaria_rhynchophylla	1	32	189	6.667	48.711	0.297
camellia_sinensis	1	11	218	6.667	56.186	0.295
hypericum_perforatum	1	21	192	6.667	49.485	0.295
panax_ginseng	2	32	136	13.333	35.052	0.281
Valerian	3	23	107	20.000	27.577	0.273
lemo_nbalm	1	19	160	6.667	41.237	0.259
lemon_balm	1	19	160	6.667	41.237	0.259
Chamomile	2	8	121	13.333	31.186	0.257
polygalae_radix	1	12	150	6.667	38.660	0.254
astragalus_mongholicus	1	13	143	6.667	36.856	0.253
st_john_s_wort_hypericum_perforatum	1	12	154	6.667	39.691	0.253
feverfew_tanacetum_parthenium	1	16	155	6.667	39.948	0.249
epigallocatechin_gallate	1	11	139	6.667	35.825	0.245
Jujube	1	20	125	6.667	32.216	0.241
trigonella_foenum-graecum	1	7	130	6.667	33.505	0.234

melissa_officinalis_l	1	12	124	6.667	31.959	0.224
rhodomyrtus_tomentosa	1	4	116	6.667	29.897	0.221
tianma_gastrodia_elata	1	29	97	6.667	25.000	0.221
Skullcap	1	12	110	6.667	28.351	0.219
Mangosteen	1	8	118	6.667	30.412	0.212
Berberine	1	5	108	6.667	27.835	0.206
lingzhi_ganoderma_lucidum	1	8	84	6.667	21.649	0.201
lemon_verbena_aloysia_citriodora	1	5	77	6.667	19.845	0.185
Caffeine	1	7	57	6.667	14.691	0.180
mucuna_pruriens_(2)	1	9	63	6.667	16.237	0.166
mucuna_pruriens	1	7	61	6.667	15.722	0.162
schisandra_chinensis	1	7	27	6.667	6.959	0.151
radix_paeoniae	1	12	53	6.667	13.660	0.147
Ashwagandha	1	4	18	6.667	4.639	0.135
Cinnamon	1	7	31	6.667	7.990	0.131
cistanche_deserticola	1	6	25	6.667	6.443	0.108
lepidium_meyenii	1	4	23	6.667	5.928	0.089
cannabis_sativa_l	2	6	14	13.333	3.608	0.088
garcinia_mangostana_linn	1	4	11	6.667	2.835	0.086
butterbur_petasites_hybridus	1	13	21	6.667	5.412	0.073
boswellia_serrata	1	5	13	6.667	3.351	0.056
reishi_ganoderma_lucidum	1	1	1	6.667	0.258	0.029

Appendix D. Full Disease-Coverage Compounds

يعرض هذا الملحق المركبات الثمانية التي حققت disease coverage = 100% داخل Cluster 3.

Compound	Plants	Diseases	Genes	Disease cov. %	Gene cov. %	Integrated score
Kaempferol	33	15	39	100.000	10.052	0.653
Caffeic acid	36	15	18	100.000	4.639	0.643
Chlorogenic acid	36	15	14	100.000	3.608	0.640
Palmitic acid	22	15	9	100.000	2.320	0.605
Linoleic acid	22	15	8	100.000	2.062	0.605
Protocatechuic acid	22	15	8	100.000	2.062	0.605
Gentisic acid	18	15	4	100.000	1.031	0.593
Ferulic acid	31	15	13	100.000	3.351	0.528

Appendix E. Final Tables and Figures Inventory

الجدول	العنوان
Table 4.1	Final network size summary after correction and harmonization
Table 4.2	Top 15 plants by integrated coverage score
Table 4.3	Top 20 compounds by integrated coverage score
Table 4.4	Top 20 degree-like hub genes
Table 4.5	Coverage tier summary
Table 4.6	High gene-coverage plants
Table 4.7	High disease-coverage compounds
Table 4.8	Full disease-coverage compounds
Table 4.9	Coverage tier pivot
Table 4.10	Summary of corrected main quantitative results
Table 5.1	Summary of key discussion points after Stage 10E correction
Table 6.1	Final synthesis of study conclusions and future directions
الشكل	العنوان
Figure GA.1	Graphical abstract of the Cluster 3 workflow and findings
Figure 4.1	Top 15 plants by integrated coverage score
Figure 4.2	Top 20 compounds by integrated coverage score

Disease coverage of top compounds	Figure 4.3
Gene coverage of top plants	Figure 4.4
Top degree-like hub genes	Figure 4.5
Coverage tier distribution	Figure 4.6
Interpretation framework after final correction	Figure 5.1
Proposed validation pathway	Figure 6.1

Appendix F. Reproducibility Note

تم تنظيم المشروع وفق بنية ملفات مرحلية تسمح بإعادة تتبع النتائج من البيانات الخام إلى الجداول النهائية. وقد حُفظت ملفات الإدخال، مخرجات التنظيف، نتائج ChEMBL، ملفات الربط الجيني، ملفات QC، والجداول النهائية في مجلدات منفصلة.

تُعد ملفات QC وسجلات التنفيذ جزءًا أساسيًا من قابلية إعادة التنفيذ، لأنها تسمح بالتحقق من أعداد الصفوف، الكيانات الفريدة، الإدخالات المستبعدة، وقرارات التصحيح النهائية.

الفهـــــرس

العنوان	الـــرقـــم
اهداء	
الشكر والعرفان	
الملخص	
الفصل الاول: المقدمة العامة، الإشكالية، وأهداف الدراسة	
1.1.تمهيد عام	
1.2. خلفية الدراسة	
1.3 إشكالية الدراسة	
1.4. أسئلة الدراسة	
1.5. فرضية الدراسة	
1.6أهداف الدراسة	
1.6.1 الهدف العام	
1.6.2 الأهداف الخاصة	
1.7 أهمية الدراسة	
1.7.1 الأهمية العلمية	
1.7.2 الأهمية المنهجية	
1.7.3 الأهمية التطبيقية	
1.7.4 الأهمية النقدية	
1.8حدود الدراسة	
1.9 بنية المذكرة	
الفصل الثاني: الإطار النظري	
2.1 تمهيد الفصل	**9**
2.2مفهوم Network Pharmacology	**9**
2.3 من نموذج Single-Target إلى Multi-Target	**9**
2.4 المركبات النباتية في علم الأدوية	**10**
2.5المحاور البيولوجية الأساسية المرتبطة ب 3cluster	**10**
2.5.1 خلل الميتوكوندريا (Mitochondrial Dysfunction)	**10**
2.5.2 خلل النواقل العصبية (Neurotransmitter Dysfunction)	**11**
2.5.3 الإجهاد التأكسدي (Oxidative Stress)	**11**
2.5.4 الالتهاب العصبي (Neuroinflammation)	**11**
2.5.5 اضطراب Proteostasis	**12**
2.6 التداخل بين المحاور البيولوجية	**12**
2.7مفهوم Multi-Disease Networks	**13**

2.8 تحليل الشبكات الحيوية (Biological Network Analysis)	**14**
2.9مفهوم Network Coverage	**14**
2.10 التمييز بين Network Coverage وTherapeutic Effect	**14**
2.11 دور قواعد البيانات الحيوية	**15**
2.12 حدود المقاربات الحاسوبية	**15**
2.13 حوصلة الفصل	**16**
الفصل الثالث : المواد والطرق	
3.1 تمهيد الفصل	**19**
3.2 تصميم الدراسة	**19**
3.3تعريف cluster3 ونطاق الدراسة	**20**
3.4 مصادر البيانات	**21**
3.5 بنية المنهجية العامة	**22**
3.6 فحص البيانات الأولي والإدخال (0stage)	**23**
3.7 التنظيف لمحافظ للبيانات(1stage)	**23**
3.8 التنظيف المتقدم والتحسين المحافظ(1.5stage)	**24**
3.9 بناء سجل المركبات وترتيب الأولويات(2stage)	**24**
3.10 مطابقة المركبات مع ChEMBL (stage3)	**25**
3.11 استخراج بيانات النشاط الحيوي منChEMBL (stage4)	**25**
3.12 فلترة الأدلة الحيوية وبناء روابط المركب–الهدف(5stage)	**25**
3.13 الربط الجيني لأهداف ChEMBL (stage6)	**26**
3.14 دمج شبكة النبات–المركب–الجين–المرض(7stage)	**26**
3.15 تحليل التغطية والترتيب(8stage)	**27**
3.16 حساب المؤشرات التكاملية	**27**
3.17 تحليل فئات التغطية stage9B	**27**
3.18 إنتاج الجداول والأشكال والتصحيح النهائي للنتائج(10stage)	**28**
3.19 البرمجيات والأدوات المستعملة	**29**
3.20 إدارة الملفات وقابلية إعادة التنفيذ	**31**
3.21 نظام ضبط الجودة QC	**31**
3.22حدود المنهجية	**32**
3.23 **3** حوصلة الفصل الثالث	**32**
الفصل الرابع: الحدود، الاستنتاجات، والتوصيات	
4.1 تمهيد الفصل	**34**
4.2 الحدود العامة للدراسة بعد التصحيح	**34**

4.2.1. حدود تعريف 3cluster	**34**
4.2.2 حدود التصحيح النباتي	**35**
4.2.3 حدود قواعد البيانات	**35**
4.2.4 حدود أوزان الدرجات التكاملية	**35**
4.2.5حدود معيار النشاط الحيوي uM <=10	**35**
4.2.6 حدود غياب ADME/ADMET	**35**
4.2.7 حدود غياب docking وmolecular dynamics	**35**
4.2.8 حدود غياب pathway enrichment	**36**
4.2.9 حدود غياب التحقق التجريبي	**36**
4.3 الاستنتاجات الرئيسية	**36**
4.3.1 بناء شبكة متعددة الطبقات قابلة للتتبع	**36**
4.3.2 التصحيح النهائي حسّن موثوقية النتائج	**36**
4.3.3 3 Cluster يمثل شبكة موزعة لا كيانًا واحدًا	**36**
4.3.4 Ginkgo biloba وquercetin يؤكدان اتساق المنهجية لا الجدة العلمية	**36**
4.3.5 Passiflora incarnata يمثل إضافة تفسيرية مهمة	**37**
4.3.6 طبقة المركبات أظهرت انتشارًا مرضيًا واسعًا	**37**
4.3.7 الجينات المحورية تحتاج تفسيرًا حذرًا	**37**
4.3.8الدراسة تنتج فرضيات بحثية لا أدلة علاجية	**37**
4.4 القيمة العلمية للدراسة بعد التصحيح	**37**
4.5 التوصيات المستقبلية	**38**
4.5.1 تنفيذ sensitivity analysis للأوزان	**38**
4.5.2 تنفيذ Gene-Disease overlap analysis	**38**
4.5.3 حساب Jaccard similarity بين الأمراض	**38**
4.5.4 تنفيذ pathway enrichment	**38**
4.5.5 تنفيذ ADME/ADMET للمركبات ذات الأولوية	**38**
4.5.6 إعطاء أولوية لمركبات Passiflora incarnata	**39**
4.5.7 تنفيذ molecular docking	**39**
4.5.8 تنفيذ molecular dynamics	**39**
4.5.9 التحقق التجريبي	**39**
4.5.10 تحسين التوحيد التصنيفي للنباتات	**39**
4.6 الخلاصة العامة النهائية	**39**
4.7 حوصلة الفصل الرابع	**40**
قائمة الاختصارات	

قائمة الصور والبيانات	
قائمة الجداول	
قائمة المراجع والملحقات	