# **Biological insights results**

## **Shared inflammatory hubs between COVID-19 and neurodegeneration**

Queries:

// ────────────────────────────────────────────────────

// 1. Define parameter lists for COVID-related and NDD-related terms

// ────────────────────────────────────────────────────

:param {

covid: [

'covid 19','sars-cov-2','sars-cov-2 infection',

'severe acute respiratory syndrome coronavirus 2'

],

ndd: [

'neurodegeneration',

"alzheimer's disease",'alzheimers disease',

"parkinson's disease",'parkinson disease',

'als','amyotrophic lateral sclerosis',

'huntington disease','frontotemporal dementia'

]

};

/// ─────────────────────────────────────────────────────

/// 2. Identify shared hubs within 2 hops of both COVID and NDD

/// Compute node degree split by data source (CBM, GPT, GPT-fulltext)

/// ─────────────────────────────────────────────────────

MATCH (c) WHERE toLower(c.name) IN $covid

MATCH (d) WHERE toLower(d.name) IN $ndd

MATCH (c)-[\*..2]-(x)

MATCH (d)-[\*..2]-(x)

WITH DISTINCT x

MATCH (x)-[r]-()

WITH x,

count(r) AS deg\_total,

sum(CASE WHEN r.source = 'CBM' THEN 1 ELSE 0 END) AS deg\_cbm,

sum(CASE WHEN r.source = 'GPT' THEN 1 ELSE 0 END) AS deg\_gpt,

sum(CASE WHEN r.source = 'GPT-fulltext' THEN 1 ELSE 0 END) AS deg\_gpt\_fulltext

RETURN x.name AS entity, labels(x) AS labels,

deg\_total, deg\_cbm, deg\_gpt, deg\_gpt\_fulltext

ORDER BY deg\_total DESC

LIMIT 100;

/// ─────────────────────────────────────────────────────

/// 3. Focus on cytokine/chemokine hubs within 2 hops of both COVID and NDD

/// Restrict to well-known cytokine names and compute degree split by source

/// ─────────────────────────────────────────────────────

MATCH (c) WHERE toLower(c.name) IN $covid

MATCH (d) WHERE toLower(d.name) IN $ndd

MATCH (c)-[\*..2]-(x)

MATCH (d)-[\*..2]-(x)

WHERE toLower(x.name) =~ '(?i)^(il-?[0-9]+|tnf|ifn|cxcl-?[0-9]+|ccl-?[0-9]+|il1b|il1a|il6|il18)$'

WITH DISTINCT x

MATCH (x)-[r]-()

WITH x,

count(r) AS deg\_total,

sum(CASE WHEN r.source = 'CBM' THEN 1 ELSE 0 END) AS deg\_cbm,

sum(CASE WHEN r.source = 'GPT' THEN 1 ELSE 0 END) AS deg\_gpt,

sum(CASE WHEN r.source = 'GPT-fulltext' THEN 1 ELSE 0 END) AS deg\_gpt\_fulltext

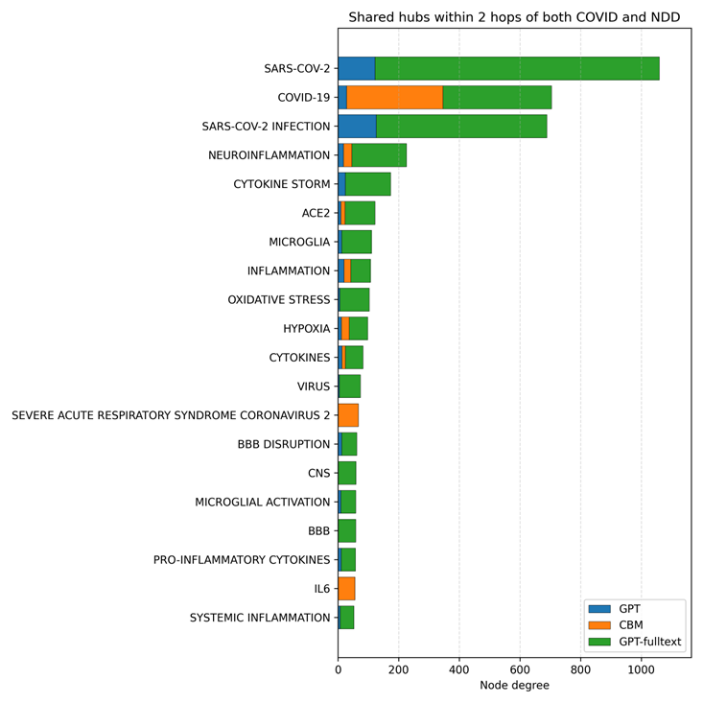
RETURN x.name AS cytokine, labels(x) AS labels,

deg\_total, deg\_cbm, deg\_gpt, deg\_gpt\_fulltext

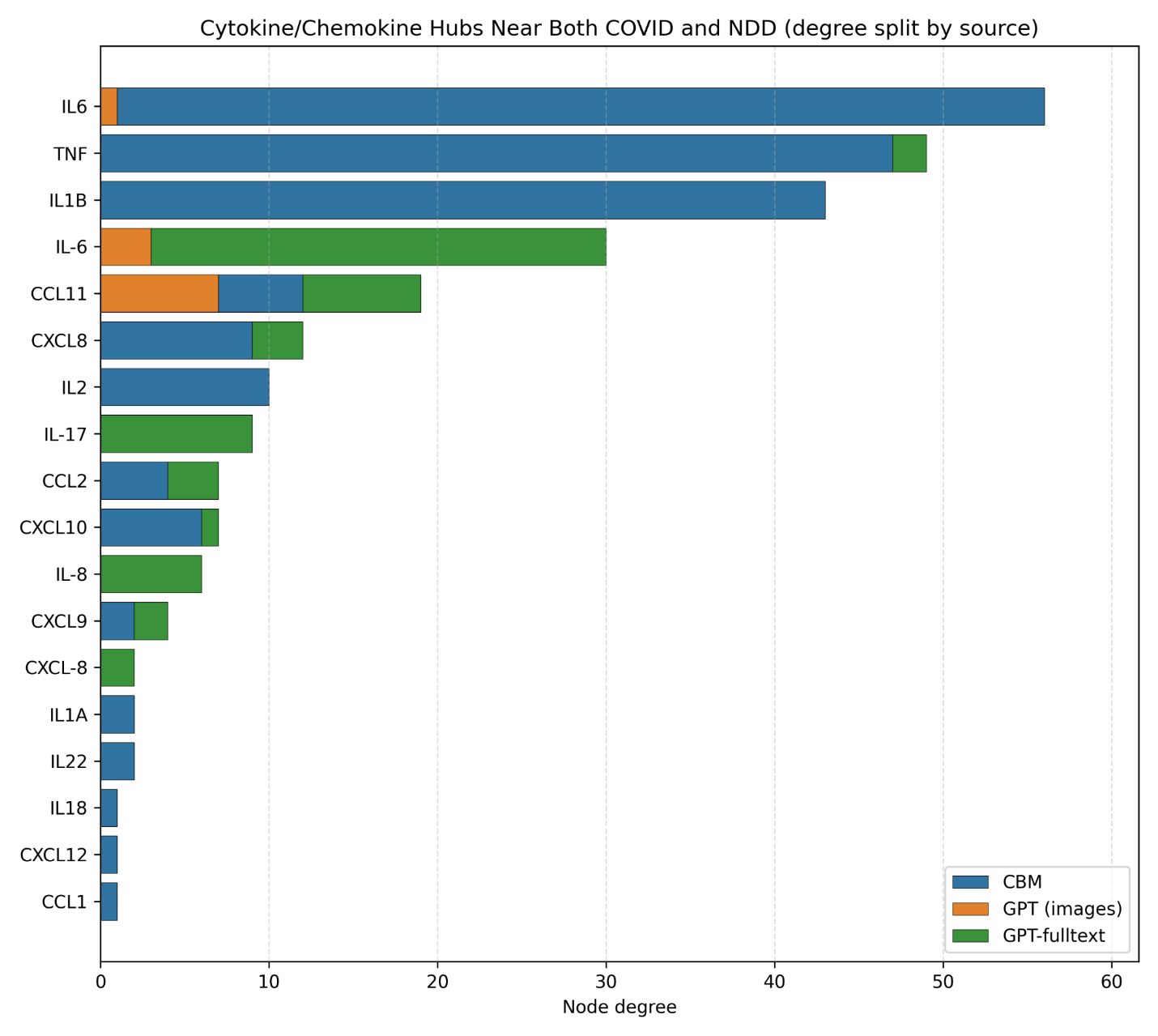
ORDER BY deg\_total DESC

LIMIT 50;

We analyzed nodes that were connected within two steps to both COVID-19 and NDD reference nodes, thereby identifying shared hubs. This analysis confirmed a convergence on inflammatory mediators (Fig. 1, Fig. 2). The curated dataset (CBM) strongly emphasized the canonical cytokine triad IL-6, TNF, IL-1β, which are well-known drivers of systemic and neuroinflammation. GPT extractions (from figures and full text) not only captured this triad but also broadened the inflammatory network to include chemokines CXCL10, CXCL12, and CCL2, and in the case of full-text mining, IL-17, pointing to adaptive immune involvement. This indicates that manual curation reference nodes the KG in canonical inflammatory axes, while GPT methods introduce under-sampled chemokine and antibody-related processes, enriching the representation of post-COVID neuroinflammatory signatures.



**Fig. 1. Shared hubs connecting COVID-19 and neurodegeneration.**Top 20 nodes within two hops of both COVID and NDD reference nodes ranked by degree, split by source (GPT, CBM, GPT-fulltext). Canonical cytokines dominate, while GPT enriches chemokine and antibody-related hubs.



**Fig. 2. Cytokine/chemokine hubs near both COVID-19 and NDD.**A subset restricted to cytokines/chemokines. CBM emphasizes IL-6, TNF, and IL-1β, while GPT adds CXCL10, CXCL12, CCL2, and IL-17, expanding the inflammatory repertoire.

## **COVID-19–induced blood–brain barrier disruption and glial activation**

// ─────────────────────────────────────────────────────

// 1. Define parameter lists for COVID, BBB-related, and glial terms

// ─────────────────────────────────────────────────────

Queries:

:param {

covid: [

'covid 19','sars-cov-2','sars-cov-2 infection',

'severe acute respiratory syndrome coronavirus 2'

],

bbb: [

'blood brain barrier','bbb','bbb disruption','blood-brain barrier',

'bbb\_breakdown','bbb\_leakage','blood brain barrier disruption'

],

glia: [

'microglial activation','microgliosis',

'astrocyte activation','astrocytosis'

]

};

/// ─────────────────────────────────────────────────────

/// 2. Identify mediators on shortest paths between COVID anchors and BBB nodes

/// Count edges along paths, split by source (CBM vs GPT)

/// ─────────────────────────────────────────────────────

MATCH (b) WHERE toLower(b.name) CONTAINS 'blood brain barrier'

WITH collect(DISTINCT b) AS bbb\_nodes

// Connect BBB nodes with COVID anchors

MATCH (c) WHERE toLower(c.name) IN $covid

UNWIND bbb\_nodes AS b

CALL {

WITH c,b

MATCH p = shortestPath( (c)-[\*..8]-(b) )

RETURN p

}

WITH p

UNWIND nodes(p)[1..-2] AS m

WITH p,m

WHERE NOT toLower(m.name) IN $covid

AND NOT toLower(m.name) CONTAINS 'blood brain barrier'

UNWIND relationships(p) AS r

WITH m,r

WHERE startNode(r)=m OR endNode(r)=m

RETURN

m.name AS mediator,

labels(m) AS labels,

count(\*) AS edges\_on\_paths,

sum(CASE WHEN coalesce(r.source,'NA')='CBM' THEN 1 ELSE 0 END) AS cbm\_edges\_on\_paths,

sum(CASE WHEN coalesce(r.source,'NA')='GPT' THEN 1 ELSE 0 END) AS gpt\_edges\_on\_paths

ORDER BY edges\_on\_paths DESC, mediator

LIMIT 30;

/// ─────────────────────────────────────────────────────

/// 3. Find neighbors of glial activation nodes (microglia, astrocytes)

/// Exclude COVID, BBB, and glial anchors themselves

/// Compute local degree and edge counts split by source

/// ─────────────────────────────────────────────────────

MATCH (g) WHERE toLower(g.name) IN $glia

MATCH (g)-[r1]-(n1)

OPTIONAL MATCH (n1)-[r2]-(n2)

WITH g, collect({nbr:n1, rel:r1}) + collect({nbr:n2, rel:r2}) AS pairs

UNWIND pairs AS p

WITH g, p.nbr AS x, p.rel AS r

WHERE x IS NOT NULL AND g <> x

AND NOT toLower(x.name) IN $covid

AND NOT toLower(x.name) IN $bbb

AND NOT toLower(x.name) IN $glia

WITH g.name AS glia\_anchor,

x.name AS neighbor,

labels(x) AS labels,

count(r) AS deg\_local,

sum(CASE WHEN coalesce(r.source,'NA')='CBM' THEN 1 ELSE 0 END) AS edges\_cbm,

sum(CASE WHEN coalesce(r.source,'NA')='GPT' THEN 1 ELSE 0 END) AS edges\_gpt

RETURN glia\_anchor, neighbor, labels, deg\_local, edges\_cbm, edges\_gpt

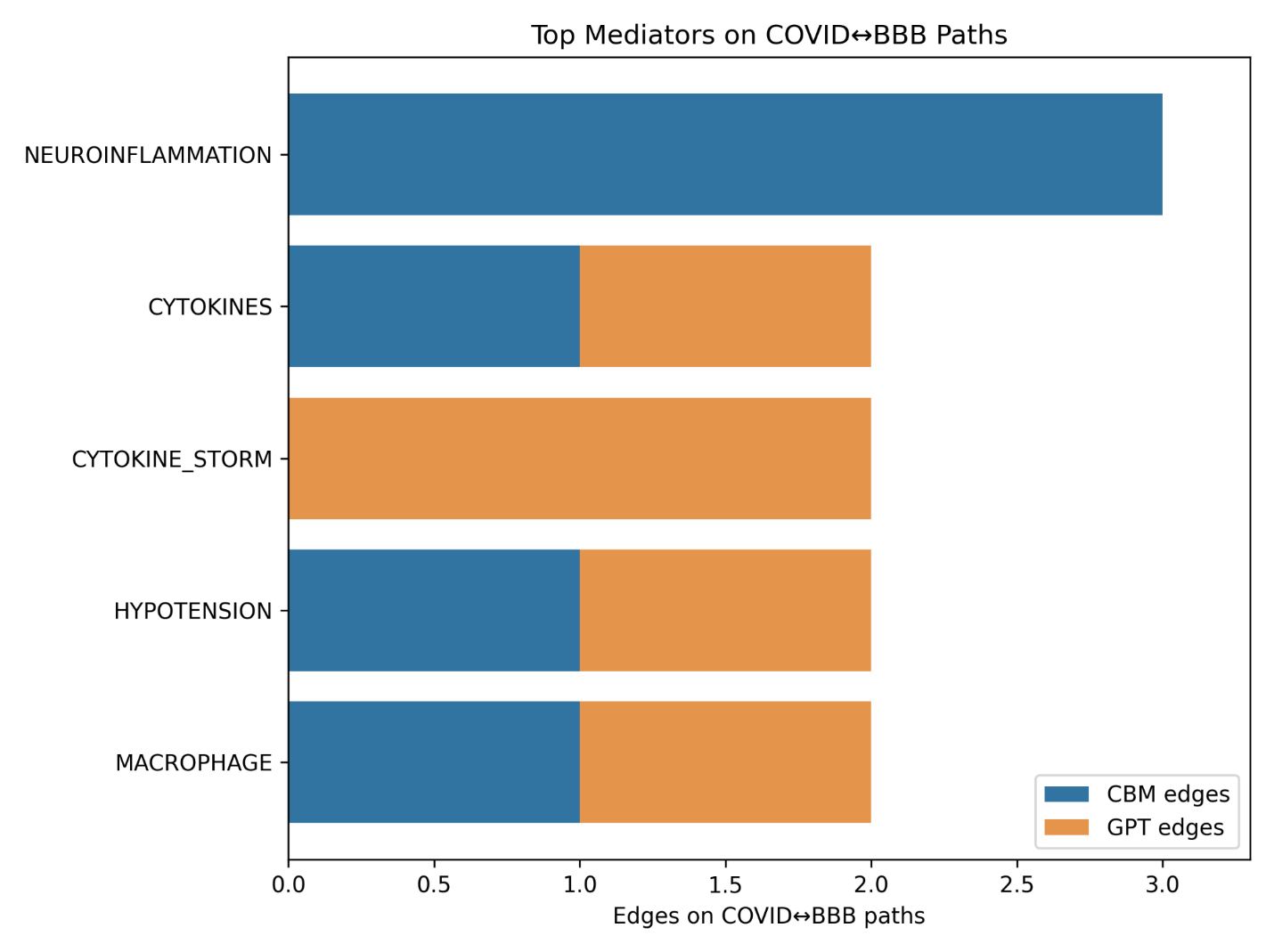
ORDER BY glia\_anchor, deg\_local DESC, neighbor

LIMIT 300;

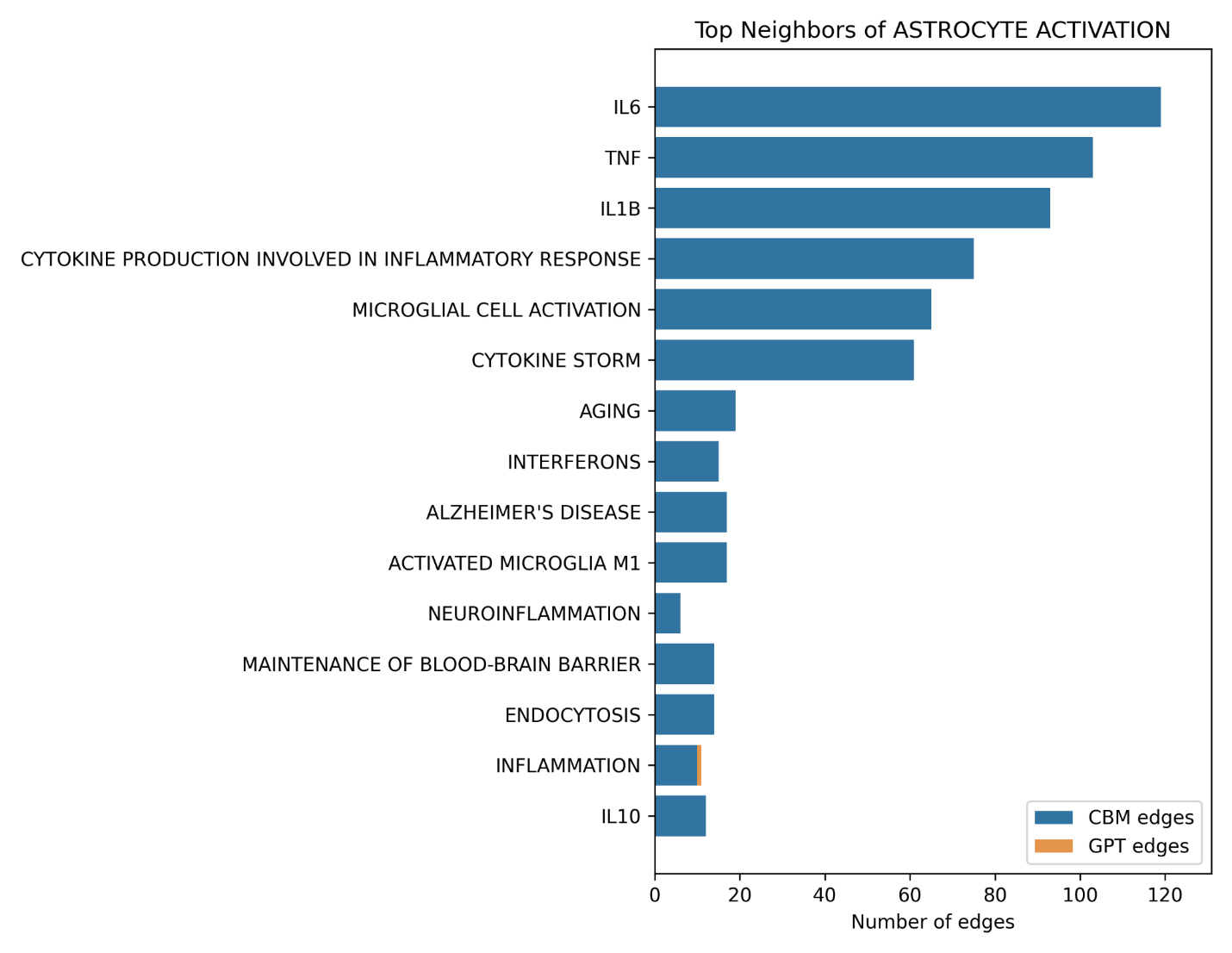
Analysis of shortest paths between COVID-19 entities and blood–brain barrier (BBB) nodes revealed two distinct classes of mediators (Fig. 3). CBM annotations emphasized canonical neuroinflammatory processes, with NEUROINFLAMMATION emerging as the dominant bridge. In contrast, GPT broadened the mechanistic landscape by introducing additional systemic mediators such as cytokine storm, hypotension, and macrophage activity, which provide plausible but less conventional routes for BBB compromise. This illustrates how GPT complements expert annotation by surfacing under-represented systemic factors.

Neighborhood analysis of astrocyte activation further supported this pattern (Fig. 4). The most prominent neighbors were IL-6, TNF, and IL-1β, along with related canonical inflammatory processes (e.g., cytokine production, microglial activation), all driven primarily by CBM curation. GPT contributions were weaker in degree but expanded the network with additional nodes such as oxidative stress, autoantibodies, and diverse chemokines, highlighting the model’s ability to suggest alternative, potentially novel drivers of glial activation.

Together, these findings show that while CBM annotations reliably anchor the knowledge graph in well-established neuroinflammatory mechanisms, GPT extractions enrich the landscape by adding broader systemic and exploratory links, providing complementary strengths for mechanistic hypothesis generation.

****

**Fig. 3. Top mediators on COVID↔BBB shortest paths.**Stacked bar plot showing mediators linking COVID and BBB disruption. CBM strongly emphasizes neuroinflammation; GPT adds systemic factors (cytokine storm, hypotension, macrophages).



**Fig. 4. Top neighbors of astrocyte activation.**Top 15 neighboring nodes of astrocyte activation. CBM emphasizes IL-6, TNF, and IL-1β, while GPT contributes oxidative stress, autoantibodies, and chemokines.

## **Therapeutic bridge candidates**

Queries:

// ─────────────────────────────────────────────────────

// 1. Define parameter lists for keywords and relationship polarity

// ─────────────────────────────────────────────────────

// COVID-related keyword filters

:param covid\_kw => ['cytokine storm','inflammation','interferon','endothelial',

'hypoxia','bbb','microglia','viral','replication','blood brain barrier'];

// Neuro “bad” keywords (pathological processes)

:param neuro\_kw\_bad => ['neuroinflammation','neurodegeneration','cognitive','memory',

'synaptic','tau','amyloid','alpha-synuclein','injury'];

// Neuro “good” keywords (protective/regenerative processes)

:param neuro\_kw\_good => ['neuroprotection','synaptic plasticity','remyelination'];

// Relationship type classification

:param beneficial\_down => ['DECREASE','INHIBIT','BLOCK','REDUCE',

'ALLEVIATES','COUNTERACTS','PREVENTS','ATTENUATES'];

:param beneficial\_up => ['INCREASE','UPREGULATE','ACTIVATE','PROMOTE']; // for “good” neuro items

:param harmful\_up => ['INCREASE','UPREGULATE','ACTIVATE','PROMOTE']; // for “bad” neuro items

:param harmful\_down => ['DECREASE','INHIBIT','BLOCK','REDUCE']; // for “good” items

/// ─────────────────────────────────────────────────────

/// 2. Identify potential therapeutic targets

/// Candidates: Gene / Protein / Chemical

/// Conditions: beneficial effects on COVID-side nodes,

/// beneficial effects on neuro-side nodes,

/// support from ≥ 2 sources

/// ─────────────────────────────────────────────────────

MATCH (t)

WHERE t:Gene OR t:Protein OR t:Chemical

// --- COVID side aggregation ---

CALL {

WITH t

MATCH (t)-[cov\_r]->(covid\_node)

WHERE any(l IN labels(covid\_node) WHERE l IN ['Biological\_Process','Cell','Anatomical\_Structure','Phenotype'])

AND any(kw IN $covid\_kw WHERE toLower(covid\_node.name) CONTAINS kw)

WITH t, cov\_r, covid\_node,

(type(cov\_r) IN $beneficial\_down) AS cov\_benefit,

(type(cov\_r) IN $harmful\_up) AS cov\_harm

RETURN

t AS tt,

sum(CASE WHEN cov\_benefit THEN 1 ELSE 0 END) AS cov\_ben,

sum(CASE WHEN cov\_harm THEN 1 ELSE 0 END) AS cov\_harm\_cnt,

[x IN collect(DISTINCT CASE WHEN cov\_benefit THEN covid\_node.name END) WHERE x IS NOT NULL] AS covid\_connections,

[x IN collect(DISTINCT CASE WHEN cov\_benefit THEN cov\_r.source END) WHERE x IS NOT NULL] AS cov\_sources

}

WITH t, cov\_ben, cov\_harm\_cnt, covid\_connections, cov\_sources

WHERE cov\_ben > 0 // require at least one beneficial COVID-side hit

// --- NEURO side aggregation ---

CALL {

WITH t

MATCH (t)-[neu\_r]->(neuro\_node)

WHERE any(l IN labels(neuro\_node) WHERE l IN ['Biological\_Process','Disease','Phenotype'])

AND (

( any(kw IN $neuro\_kw\_bad WHERE toLower(neuro\_node.name) CONTAINS kw)

AND (type(neu\_r) IN $beneficial\_down OR type(neu\_r) IN ['ALLEVIATES','COUNTERACTS','PREVENTS','ATTENUATES']) )

OR ( any(kw IN $neuro\_kw\_good WHERE toLower(neuro\_node.name) CONTAINS kw)

AND type(neu\_r) IN $beneficial\_up )

)

WITH t, neu\_r, neuro\_node,

( (any(kw IN $neuro\_kw\_bad WHERE toLower(neuro\_node.name) CONTAINS kw) AND type(neu\_r) IN $harmful\_up)

OR (any(kw IN $neuro\_kw\_good WHERE toLower(neuro\_node.name) CONTAINS kw) AND type(neu\_r) IN $harmful\_down) ) AS neu\_harm

RETURN

t AS tt2,

count(\*) AS neu\_ben, // all matched above are beneficial

sum(CASE WHEN neu\_harm THEN 1 ELSE 0 END) AS neu\_harm\_cnt,

collect(DISTINCT neuro\_node.name) AS neuro\_connections,

[x IN collect(DISTINCT neu\_r.source) WHERE x IS NOT NULL] AS neu\_sources

}

WITH t, covid\_connections, cov\_sources, cov\_ben, cov\_harm\_cnt,

neuro\_connections, neu\_sources, neu\_ben, neu\_harm\_cnt

WHERE cov\_ben > cov\_harm\_cnt

AND neu\_ben > neu\_harm\_cnt

// Deduplicate sources without APOC

WITH t, covid\_connections, neuro\_connections, (cov\_sources + neu\_sources) AS all\_sources

UNWIND [s IN all\_sources WHERE s IS NOT NULL] AS s

WITH t, covid\_connections, neuro\_connections, collect(DISTINCT s) AS sources

WHERE size(sources) >= 2 // set to 1 to debug; keep 2 for manuscript

RETURN labels(t) AS target\_type,

t.name AS potential\_target,

covid\_connections, neuro\_connections, sources,

size(covid\_connections)+size(neuro\_connections) AS support

ORDER BY support DESC

LIMIT 15;

/// ─────────────────────────────────────────────────────

/// 3. Identify potential drug repurposing opportunities

/// Candidates: Chemicals (excluding common neurotransmitters)

/// Conditions: beneficial effects on COVID-side nodes,

/// beneficial effects on neuro-side nodes,

/// support from ≥ 2 sources

/// ─────────────────────────────────────────────────────

MATCH (d:Chemical)

WHERE NOT toLower(d.name) IN ['acetylcholine','dopamine','glutamate','gaba','serotonin']

// --- COVID side aggregation ---

CALL {

WITH d

MATCH (d)-[cov\_r]->(covid\_node)

WHERE any(l IN labels(covid\_node) WHERE l IN ['Biological\_Process','Cell','Anatomical\_Structure','Phenotype'])

AND any(kw IN $covid\_kw WHERE toLower(covid\_node.name) CONTAINS kw)

WITH d, cov\_r, covid\_node,

(type(cov\_r) IN $beneficial\_down) AS cov\_benefit,

(type(cov\_r) IN $harmful\_up) AS cov\_harm

RETURN

d AS dd,

sum(CASE WHEN cov\_benefit THEN 1 ELSE 0 END) AS cov\_ben,

sum(CASE WHEN cov\_harm THEN 1 ELSE 0 END) AS cov\_harm\_cnt,

[x IN collect(DISTINCT CASE WHEN cov\_benefit THEN covid\_node.name END) WHERE x IS NOT NULL] AS covid\_targets,

[x IN collect(DISTINCT CASE WHEN cov\_benefit THEN cov\_r.source END) WHERE x IS NOT NULL] AS cov\_sources

}

WITH d, cov\_ben, cov\_harm\_cnt, covid\_targets, cov\_sources

WHERE cov\_ben > 0

// --- NEURO side aggregation ---

CALL {

WITH d

MATCH (d)-[neu\_r]->(neuro\_node)

WHERE any(l IN labels(neuro\_node) WHERE l IN ['Biological\_Process','Disease','Phenotype'])

AND (

( any(kw IN $neuro\_kw\_bad WHERE toLower(neuro\_node.name) CONTAINS kw)

AND (type(neu\_r) IN $beneficial\_down OR type(neu\_r) IN ['ALLEVIATES','COUNTERACTS','PREVENTS','ATTENUATES']) )

OR ( any(kw IN $neuro\_kw\_good WHERE toLower(neuro\_node.name) CONTAINS kw)

AND type(neu\_r) IN $beneficial\_up )

)

WITH d, neu\_r, neuro\_node,

( (any(kw IN $neuro\_kw\_bad WHERE toLower(neuro\_node.name) CONTAINS kw) AND type(neu\_r) IN $harmful\_up)

OR (any(kw IN $neuro\_kw\_good WHERE toLower(neuro\_node.name) CONTAINS kw) AND type(neu\_r) IN $harmful\_down) ) AS neu\_harm

RETURN

d AS dd2,

count(\*) AS neu\_ben,

sum(CASE WHEN neu\_harm THEN 1 ELSE 0 END) AS neu\_harm\_cnt,

collect(DISTINCT neuro\_node.name) AS neuro\_benefits,

[x IN collect(DISTINCT neu\_r.source) WHERE x IS NOT NULL] AS neu\_sources

}

WITH d, covid\_targets, cov\_sources, cov\_ben, cov\_harm\_cnt,

neuro\_benefits, neu\_sources, neu\_ben, neu\_harm\_cnt

WHERE cov\_ben > cov\_harm\_cnt

AND neu\_ben > neu\_harm\_cnt

// Deduplicate sources without APOC

WITH d, covid\_targets, neuro\_benefits, (cov\_sources + neu\_sources) AS all\_sources

UNWIND [s IN all\_sources WHERE s IS NOT NULL] AS s

WITH d, covid\_targets, neuro\_benefits, collect(DISTINCT s) AS sources

WHERE size(sources) >= 2 // set to 1 to debug; keep 2 for manuscript

RETURN d.name AS potential\_drug,

covid\_targets, neuro\_benefits, sources,

size(covid\_targets)+size(neuro\_benefits) AS support

ORDER BY support DESC

LIMIT 15;

Applying a majority-beneficial filter (down-modulating COVID-side inflammatory/viral/BBB processes and down-modulating neuroinflammation or neuronal injury, with ≥2 evidence sources), we identified minocycline as a prominent therapeutic candidate (Table 1). Minocycline decreased or counteracted multiple COVID-related processes including viral replication, microglial activation, and SARS-CoV-2 crossing of the blood–brain barrier, and was simultaneously linked to reduced neuroinflammation, hypoxia-induced neuroinflammation, and neuronal/glial injury. Importantly, this signal was supported across CBM curation and GPT-fulltext extractions, underscoring its robustness.

It is noteworthy that the GPT shapes did not extract any edges for the minocycline, whereas CBM and GPT-fulltext showed a consistent pattern. This highlights the incomplete overlap between sources and their complementarity: images introduce new mechanisms (for example, ferroptosis), while text/curation covers specific pharmacological relationships.

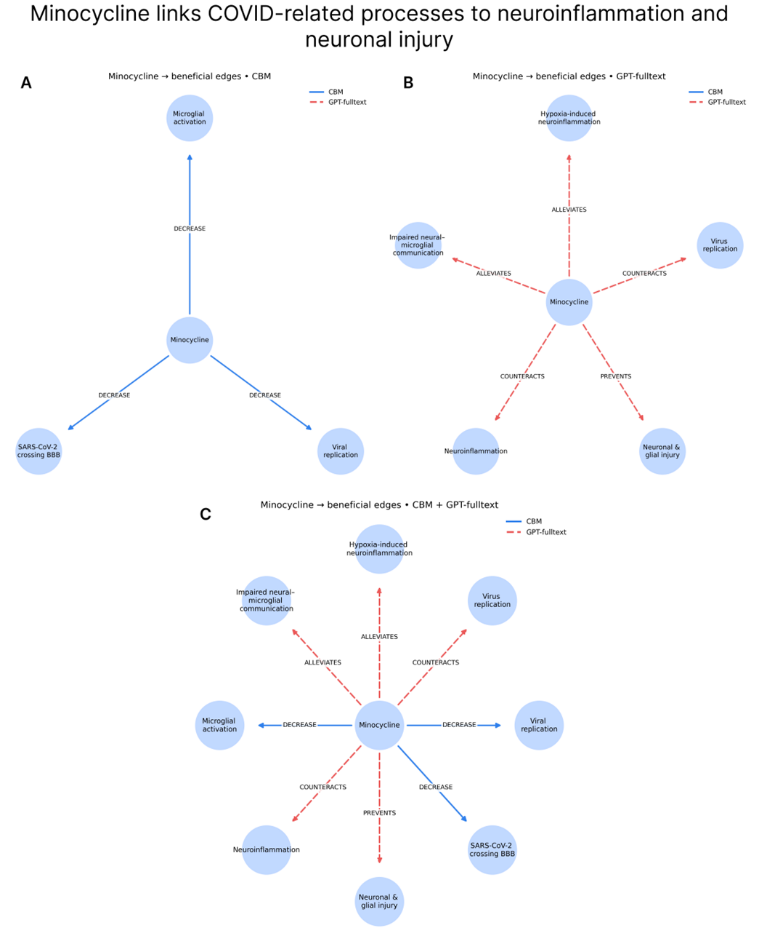
**Table 1. Therapeutic targets and repurposing opportunities.**

| **Type** | **Candidate** | **COVID-side processes** | **Neuro-side benefits** |  | **Sources** | **Support** |
| --- | --- | --- | --- | --- | --- | --- |
| Chemical | Minocycline | Viral replication; Microglial cell activation; SARS-CoV-2 crossing BBB; Impaired neural–microglial communication; Hypoxia-induced neuroinflammation; Virus replication; Neuroinflammation | Neuronal and glial injury; Hypoxia-induced neuroinflammation; Neuroinflammation |  | CBM, GPT-fulltext | 10 |

Therapeutic bridge analysis identified minocycline as the most consistent repurposing candidate, linking COVID-19–related inflammatory/viral/BBB processes with neuroprotective outcomes across modalities.

**Therapeutic bridge analysis**

To explore potential therapeutic leverage points, we focused on minocycline, which emerged as the top chemical candidate from our majority-beneficial filtering (see Table K). Subgraph visualization (Fig. 5) highlights that CBM curation and GPT-fulltext extractions provide complementary perspectives. CBM annotations emphasize established mechanisms — minocycline *decreases* microglial activation, viral replication, and SARS-CoV-2 crossing of the blood–brain barrier — consistent with canonical roles in glial suppression and anti-viral activity. In contrast, GPT-fulltext extractions expand the scope toward outcomes, connecting minocycline with reduced neuroinflammation, hypoxia-induced neuroinflammation, neuronal & glial injury, and *counteraction* of virus replication. The combined network underscores how curation and text-mined evidence converge on a protective role of minocycline at the interface of COVID-induced inflammation and neurodegeneration, while also revealing complementary nodes contributed by different modalities.



**Fig. 5. Minocycline links COVID-related processes to neuroinflammation and neuronal injury.**  
(A) GPT-fulltext: connections to neuroinflammatory outcomes via ALLEVIATES, COUNTERACTS, PREVENTS. (B) CBM: decreases microglial activation, viral replication, and SARS-CoV-2 BBB crossing. (C) Combined view: complementary evidence across sources.