Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cross-species transcriptomic analysis identifies candidate genes associated with diquat poisoning

Qiqi Liu
Fuyang City People's Hospital
Tingzhan Lv
Fuyang City People's Hospital
Yong Li
Fuyang City People's Hospital
Min Yan
Fuyang City People's Hospital
Ruijie Li
Fuyang City People's Hospital
Kun Huang
Fuyang City People's Hospital
Hongbo Liu
Fuyang City People's Hospital
Yihong Yang
15856806657@163.com

Fuyang City People's Hospital

Article

Keywords:

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11093547/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Cross-species transcriptomic analysis identifies candidate genes associated with diquat poisoning

Qiqi Liu[a],[1], Tingzhan Lv[a],[1], Yong Li[a], Min Yan[a], Ruijie Li[a], Kun Huang[a], Hongbo Liu[a],* and Yihong Yang[a],*

[a] Department of Emergency Medicine, Fuyang People's Hospital Affiliated to Anhui Medical University, Fuyang, Anhui 236000, China

[1] These authors contributed equally to this work.

* Corresponding authors: Hongbo Liu (xh7293@126.com) and Yihong Yang (15856806657@163.com), Department of Emergency Medicine, Fuyang People's Hospital Affiliated to Anhui Medical University, 501 Sanqing Road, Yingzhou District, Fuyang, Anhui 236000, China.

Abstract

Diquat poisoning can cause central nervous system injury, but the molecular responses shared by patients and experimental models remain unclear. We integrated serum transcriptomes from 17 participants with nine zebrafish transcriptomes using differential expression, ortholog mapping, and directional concordance. Of 532 differentially expressed zebrafish genes, 237 mapped to human orthologs. *ANXA2*, *PRDX1*, and *GNA15* showed concordant increases in exposed zebrafish and in two comparisons of poisoned patients with versus without neurological complications. In separate zebrafish experiments, embryos were allocated at 1 h post-fertilization (hpf), exposed to 0, 100, or 175 mg/L diquat from 72 to 96 hpf, and sampled at 96 hpf. Exposure increased candidate and inflammatory transcripts, reduced proliferating cell nuclear antigen labeling, increased malondialdehyde, and decreased superoxide dismutase activity and reduced glutathione. Serum ANXA2, PRDX1, interleukin-6, and tumor necrosis factor-α showed positive effect estimates, but none of eight comparisons remained significant after correction for multiple testing. These findings identify three candidates for investigating shared diquat-associated stress responses. Their roles in neural injury and clinical relevance require functional studies and independent clinical validation.

Keywords: diquat; central nervous system injury; cross-species transcriptomics; zebrafish; oxidative stress; neuroinflammation

Introduction

Diquat (DQ) is a redox-active bipyridylium herbicide that can cause acute multiorgan failure after ingestion. Severe poisoning has been associated with encephalopathy, brainstem lesions, demyelination, and blood–brain barrier disruption [1–7]. Although these neurological complications are clinically important, the molecular responses associated with them remain poorly characterized. Circulating molecular profiles offer a way to investigate these responses, but they also reflect systemic illness, exposure, sampling time, and treatment.

Redox cycling is central to DQ toxicity. Repeated reduction and reoxidation increase oxygen consumption and reactive oxygen species production, deplete reducing equivalents, and impair antioxidant defenses [8–12]. Mitochondrial injury, regulated cell death, and microglial inflammation have also been implicated in experimental DQ toxicity [13,14]. Zebrafish provide a vertebrate model in which molecular changes can be examined alongside biochemical, developmental, and behavioral responses to exposure [15–18]. Comparing zebrafish and patient transcriptomes may therefore help identify genes that respond consistently across experimental and clinical settings.

A clinical multi-omics study linked acute DQ poisoning and central nervous system injury (CNSI) to histidine–methylmalonate imbalance and neuroinflammation [19]. A remaining question is which DQ-responsive genes in an experimental model have human orthologs that also differ between poisoned patients with and without CNSI. Agreement across these comparisons can help select candidates for experimental investigation, although it does not by itself identify their tissue of origin or establish a causal role in injury.

Here, we integrated human serum and zebrafish transcriptomes to identify candidate genes with concordant expression changes. We then examined their expression in separate zebrafish experiments using a 72–96 h post-fertilization (hpf) exposure window, together with proliferating cell nuclear antigen (PCNA) labeling and biochemical and inflammatory endpoints. Serum protein assays and exploratory molecular docking provided complementary evidence. The study was designed to connect cross-species expression patterns with experimental responses and to identify candidates for functional investigation.

Results

Cross species analysis identifies three concordantly expressed candidate genes

The analysis included serum transcriptomes from 17 participants and nine zebrafish transcriptomes (Fig. 1a). The human cohort comprised six healthy controls and 11 poisoned patients: four without CNSI during hospitalization (NCNSI), four who developed CNSI after admission (CNSI1), and three with CNSI at admission (CNSI2). Candidate selection used the CNSI1-versus-NCNSI and CNSI2-versus-NCNSI comparisons. In zebrafish exposed to 175 mg/L DQ, 532 genes met the differential-expression thresholds relative to controls, and 237 mapped to human orthologs. *ANXA2*, *PRDX1*, and *GNA15* changed in the same direction across the zebrafish comparison and both human comparisons (Fig. 1b,c).

Within-species standardized expression of all three candidates was higher in exposed zebrafish and in the human CNSI groups than in their respective reference groups (Fig. 1d). Estimated $\log_2$ fold changes were positive in all three comparisons, although their magnitudes and false discovery rates differed (Fig. 1e). Transcript abundance distributions are shown in Supplementary Fig. S1. The two human comparisons shared the NCNSI reference group and therefore provided related, rather than independent, evidence of association.

Functional annotations implicate stress and immune responses

The Gene Ontology biological-process results included responses to stress, programmed cell death, innate immunity, cytokines, stress-activated MAPK signaling, and superoxide (Fig. 2a). Among the displayed KEGG and Reactome terms, interferon signaling and interferon-gamma signaling met the false discovery rate (FDR) threshold of 0.05. Other terms, including phagosome, neutrophil degranulation, extracellular matrix organization, ferroptosis, integrin signaling, and cytoskeletal processes, did not meet this threshold and were retained as descriptive annotations (Fig. 2b).

STRING association networks connected ANXA2 with S100A10, S100A4, PLAT, ANXA2R, and AHNAK; GNA15 with phospholipase C β-family proteins and RXFP4; and PRDX1 with TXN, TXN2, SRXN1, PRDX5, and ABL1 (Fig. 2c). The PRDX1 network mainly comprised thiol-redox proteins, whereas the ANXA2 network included proteins involved in membrane organization and inflammation. These associations suggest biological contexts for the candidates without demonstrating physical interactions during DQ exposure. Organism-specific database coverage, network degree rankings, and the evidence convergence matrix are presented in Supplementary Fig. S1.

Zebrafish experiments support the expression findings and reveal redox disturbances

In the targeted zebrafish experiments, embryos were allocated at 1 hpf, exposed to 0, 100, or 175 mg/L DQ from 72 to 96 hpf, and sampled at 96 hpf (Fig. 3a). The exposure window and nominal concentrations matched those used for RNA sequencing. Both exposed groups had fewer PCNA-positive cells per unit area than controls (Fig. 3b,c; Dunnett-adjusted $P < 0.001$). The reduction was descriptively greater at 100 than at 175 mg/L. This pattern did not establish a monotonic concentration response, and the analysis did not establish a difference between the two exposed groups.

DQ exposure increased malondialdehyde (MDA), decreased superoxide dismutase (SOD) activity, and reduced glutathione (GSH), with the largest changes observed at 175 mg/L (Fig. 3d–f; Dunnett-adjusted $P < 0.001$ versus control). Quantitative PCR showed increased expression of the inflammatory genes *il1b*, *tnfa*, and *il6* (Fig. 3g), together with *anxa2a*, *prdx1*, and *gna15.2* (Fig. 3h). Thus, the candidate genes changed in the same direction as in the transcriptomic analysis, alongside biochemical evidence of disturbed redox homeostasis.

Serum protein comparisons remain inconclusive after multiple testing correction

Mean serum ANXA2, PRDX1, IL-6, and TNF-α values were higher in CNSI1 and CNSI2 than in NCNSI (Fig. 4a–d). Individual distributions overlapped in some comparisons, with substantial variability, particularly for ANXA2 in CNSI2. Participant-level normalized measurements are provided in Supplementary Table S3.

All eight $\log_2$ mean ratios were positive, ranging from 0.458 for ANXA2 in CNSI1 to 0.937 for IL-6 in CNSI2 (Fig. 4e; Supplementary Table S2). The reported unadjusted 95% bias-corrected and accelerated (BCa) bootstrap intervals excluded zero, whereas exact unadjusted P values ranged from 0.0286 to 0.1714. None of the comparisons was significant after Benjamini–Hochberg correction (adjusted P = 0.0914–0.1714). The bootstrap intervals and permutation tests summarize uncertainty using different procedures and are not interchangeable measures of significance. Exploratory receiver operating characteristic (ROC) curves for the pooled CNSI group versus NCNSI are shown in Supplementary Fig. S1. These within-cohort analyses do not establish diagnostic performance in an independent population.

Exploratory docking identifies poses for further investigation

AutoDock Vina generated DQ poses within selected regions of ANXA2 and PRDX1 (Fig. 5). Residues neighboring the displayed ANXA2 pose included Asn136, Leu139, Leu173, Arg178, Val221, Cys261, Ile262, and Lys265; the PRDX1 pose was near Val51, Arg128, Leu147, and Pro148. The interaction maps included π-related and hydrophobic contacts, as well as unfavorable positive–positive contacts. These calculations provide structural hypotheses for testing and do not demonstrate direct DQ binding or quantify experimental affinity.

Discussion

This study identified *ANXA2*, *PRDX1*, and *GNA15* as candidates with concordant expression changes in DQ-exposed zebrafish and poisoned patients with CNSI. Separate zebrafish experiments reproduced the direction of these changes and showed reduced PCNA labeling, increased lipid peroxidation, impaired antioxidant defenses, and increased inflammatory transcripts. Serum protein estimates were also positive, although none remained significant after multiple testing correction. The strongest evidence therefore concerns consistent expression direction across the studied systems and accompanying experimental stress responses; the clinical protein findings remain preliminary.

The known functions of ANXA2 and PRDX1 suggest different ways in which these candidates could relate to DQ-associated stress. ANXA2 participates in membrane organization, inflammation, and host defense [20]. It can also scaffold a PRDX2–STAT3 redox relay [21], although that observation concerns PRDX2 and does not demonstrate an ANXA2–PRDX1 pathway. PRDX1 has location-dependent activities: intracellular PRDX1 removes peroxides, whereas extracellular PRDX1 can stimulate TLR4-dependent cytokine production [22]. GNA15 was associated mainly with phospholipase C β-family proteins in the STRING network. Its role in DQ-associated injury is less clear. These findings support testing the candidates individually before assigning them to a common causal pathway.

The zebrafish biochemical findings are consistent with established DQ redox chemistry [8–11], and the inflammatory transcript changes agree with experimental evidence of DQ-associated inflammation [14]. Increased *anxa2a*, *prdx1*, and *gna15.2* expression may accompany cellular injury, reflect a compensatory response, or both. Because the assays used whole larvae, they cannot distinguish neural responses from changes in other tissues. Agreement between serum and whole-larva transcriptomes supports a shared response to poisoning-associated stress, but tissue-resolved experiments are needed to determine its relevance to neural injury.

The PCNA findings also require a specific interpretation. PCNA participates in DNA replication and repair, so reduced labeling is compatible with altered proliferation or repair activity but does not directly measure neuronal death. The larger reduction at 100 than at 175 mg/L could reflect differences in developmental state, proliferation, sampling variability, or survivor composition. The present data do not distinguish among these possibilities. Because eight of 20 prepared larvae were quantified in each group, the selection of analyzed specimens is also relevant to interpretation.

Exposure duration is important when relating these findings to other DQ studies. The present experiments used a 24 h exposure from 72 to 96 hpf, whereas longer developmental exposures were used in the earlier zebrafish work [19]. The previously reported LC_{50} of approximately 178.2 mg/L therefore cannot be treated as a lethality threshold for the current protocol. The nominal concentrations characterize acute responses to substantial toxic stress; they do not establish equivalence to human internal doses or typical environmental exposure. Using the same exposure window for RNA sequencing and qPCR nevertheless improves the comparability of the expression findings.

The clinical results are limited by the small groups and their shared reference participants. CNSI1 and CNSI2 represent different individuals classified by the timing of CNSI relative to admission, rather than successive stages sampled in the same patients. Under the specified two-sided permutation procedure, a three-versus-four comparison has only 35 possible label allocations and a minimum attainable P value of 2/35 (approximately 0.0571). The BCa intervals were estimated separately and need not agree with these tests. Neither positive effect estimates nor intervals excluding zero overcome the uncertainty caused by the small sample. The exploratory ROC curves use the same participants as candidate discovery and require independent validation.

Several additional limitations affect interpretation. Greater systemic illness was reported in the CNSI groups [19], so poisoning severity, organ dysfunction, sampling time, and treatment may contribute to the observed differences. Direction-based candidate selection can favor large discovery estimates, and one-to-many ortholog mapping may affect gene-set membership. The two human comparisons share four NCNSI patients, while the zebrafish biochemical and qPCR experiments have only three biological pools per group. These features limit precision and the assessment of statistical assumptions. Larger clinical cohorts and additional exposure experiments are needed to test the stability of the candidate selection and the reproducibility of the experimental findings.

The next experimental step is to determine whether changing candidate expression alters relevant injury endpoints. Gene perturbation combined with brain- or cell-type-resolved measurements would help distinguish functional involvement from a general stress response. Protein measurements would also be informative, particularly for GNA15, which was not assessed in serum. The docking results can guide structural hypotheses, but direct binding experiments are needed before assigning ANXA2 or PRDX1 as DQ-binding targets.

Together, the cross-species analysis and zebrafish experiments identify *ANXA2*, *PRDX1*, and *GNA15* as candidates for investigating DQ-associated stress responses. The results provide a basis for focused functional studies, while the contribution of these genes to neural injury and their value in clinical assessment remain open questions.

Methods

Study design and sample sources

The study combined human and zebrafish transcriptomic analysis, functional annotation, targeted zebrafish experiments, and serum protein measurements. Human RNA-seq profiles and stored serum came from the 17-participant cohort reported by Liu et al. [19]. Candidate selection used transcriptomic data; metabolomic and

proteomic features from that report were not included. ANXA2, PRDX1, IL-6, and TNF-α were measured for the present study in serum from the same participants. These measurements therefore provide follow-up within the discovery cohort.

The zebrafish RNA-seq dataset was generated within the broader project and is deposited under NCBI BioProject PRJNA1358919. Targeted PCNA imaging, biochemical assays, and qPCR used separate exposure batches and biological material distinct from the sequencing samples and those used in previous publications. The imaging cohort was separate from the pooled samples used for destructive assays. All current zebrafish experiments used allocation at 1 hpf, exposure from 72 to 96 hpf, and collection at 96 hpf. Supplementary Table S4 summarizes the data and sample sources.

Human cohort and serum sampling

The human cohort comprised healthy controls recruited during health examinations (JKDZ, n = 6), poisoned patients without CNSI during hospitalization (NCNSI, n = 4), patients who developed CNSI after admission (CNSI1, n = 4), and patients with CNSI at admission (CNSI2, n = 3). The cohort report describes the exclusion of patients with pre-existing neurological disease, chronic liver or kidney failure, malignancy, or incomplete clinical data, and the inclusion of age- and sex-comparable healthy controls [19]. DQ poisoning was diagnosed from the ingestion history and compatible clinical findings, with serum DQ measured by LC–MS/MS. CNSI classification was based on neurological findings and/or acute neuroimaging abnormalities. Each group contained different participants.

Peripheral venous blood was collected within 24 h of admission or during a health examination, as described in the cohort protocol [19]. After clotting, samples were centrifuged at 3,000 × g for 10 min at 4 °C. Serum was inspected for hemolysis, aliquoted to limit repeated freeze–thaw cycles, snap-frozen, and stored at −80 °C. The human study was approved by the Ethics Committee of Fuyang People's Hospital (No. 2025-207) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from participants or their legal representatives.

Human RNA sequencing and differential expression

Low-input serum RNA was used to prepare strand-specific, poly(A)-selected libraries for paired-end 150-bp sequencing on an Illumina platform. The cohort protocol used column-based RNA extraction, spectrophotometric assessment of concentration and purity, and capillary electrophoresis when feasible [19]. FASTQ quality was assessed using FastQC. Adapters and low-quality bases were removed using fastp [23], filtered reads were aligned to GRCh38 using HISAT2 [24], and gene-level counts were generated using featureCounts [25] with GENCODE annotation [26]. Low-count genes were removed before normalization. Further details are provided in Supplementary Methods under “Human serum RNA sequencing and contrasts”.

The CNSI1-versus-NCNSI and CNSI2-versus-NCNSI comparisons were analyzed using a group-based DESeq2 model [27]. P values were adjusted using the Benjamini–Hochberg procedure [28]. Differential expression was defined as FDR < 0.05 and an absolute $\log_2$ fold change of at least 1.

Zebrafish transcriptomic dataset

The zebrafish RNA-seq dataset comprised control (C), 100 mg/L DQ (TL), and 175 mg/L DQ (TH) groups, with three independently prepared pools of 15 larvae per group. Embryos were allocated at 1 hpf, exposed from 72 to 96 hpf, and collected at 96 hpf. Reads were summarized against the GRCz11 reference annotation, and differential expression was assessed using edgeR [29]. Selection thresholds were FDR < 0.05 and an absolute $\log_2$ fold change of at least 1. TH versus C was used for candidate selection; TL was included in

expression plots to show the lower-concentration response (Supplementary Methods, “Zebrafish RNA sequencing and cross species mapping”).

Ortholog mapping and candidate prioritization

Differentially expressed zebrafish genes were mapped from *Danio rerio* GRCz11 to *Homo sapiens* GRCh38 using Ensembl Compara/BioMart, retaining protein-coding orthology records. One-to-many relationships were retained when checking expression direction, after which duplicated human identifiers were collapsed. Candidates were required to be represented in the zebrafish TH-versus-C comparison and both human comparisons, and to change in the same direction across all three. Gene-set overlap and expression direction were evaluated separately. Supplementary Methods under “Zebrafish RNA sequencing and cross species mapping” describes the mapping procedure.

Functional enrichment and protein association analysis

Gene Ontology Biological Process, KEGG, and Reactome over-representation analyses were performed using clusterProfiler and ReactomePA [30,31]. The background comprised unique human-mapped genes eligible for enrichment analysis. Benjamini–Hochberg-adjusted P values below 0.05 defined significant enrichment. Displayed terms above this threshold were treated as descriptive annotations.

DQ-associated gene records were retrieved from the Comparative Toxicogenomics Database (CTD) using the chemical term “Diquat” [32]. Candidate-centered *Homo sapiens* association networks were obtained from STRING v12.0 using a minimum combined confidence score of 0.700, all active evidence channels, and up to five first-shell interactors per seed [33]. Edge widths represented combined scores, and node outlines indicated functional classes assigned from prespecified Gene Ontology annotations. Supplementary Fig. S1 summarizes CTD evidence by organism, network genes ranked by degree, and evidence across the zebrafish and human comparisons, CTD associations, and NF-κB context. Network degree denotes the number of connections to a gene within the analyzed network.

Zebrafish husbandry and targeted exposure

Wild-type AB and Tg(elavl3:EGFP) zebrafish were obtained from the China Zebrafish Resource Center (Wuhan, China). Adults were maintained at 28 ± 1 °C under a 14 h light/10 h dark cycle and fed freshly hatched brine shrimp twice daily. Normally developing embryos were randomly allocated to treatment dishes at 1 hpf and maintained in E3 medium until exposure. Diquat dibromide hydrate ($C_{12}H_{14}Br_2N_2O$; molecular weight 362.06; CAS 6385-62-2; catalog HY-136372; stated purity 99.95%; MedChemExpress, USA) was dissolved in water and diluted in E3 medium. At 72 hpf, larvae were exposed to 0, 100, or 175 mg/L DQ for 24 h and collected at 96 hpf. Controls received E3 medium without DQ.

The nominal concentrations matched those used in the earlier project [19], but the present exposure lasted 24 h. The previously reported LC_{50} of approximately 178.2 mg/L was obtained under a longer developmental exposure and was not treated as a lethality threshold for this protocol. Nominal aqueous concentrations were not assumed to represent equivalent human serum concentrations. Zebrafish procedures were approved by the Animal Ethics Committee of Tongji University (No. TJAA06324B01).

PCNA immunofluorescence

At 96 hpf, 20 Tg(elavl3:EGFP) larvae per group were collected and fixed overnight in 4% paraformaldehyde. Larvae were dehydrated, the external skin was removed, and specimens were rehydrated before acetone treatment at −20 °C for 6 h. Samples were washed in phosphate-buffered saline containing 3% Triton X-100

and blocked in the same buffer supplemented with 4% bovine serum albumin and 0.02% sodium azide. Specimens were incubated at 4 °C with an anti-PCNA antibody (1:100; Abcam, Cambridge, UK), followed by a fluorophore-conjugated secondary antibody. Nuclei were counterstained with DAPI for 45 min. Images were acquired using a Leica TCS SP8 confocal microscope with consistent acquisition settings. Eight larvae per group were quantified in matched regions, and PCNA-positive cell counts were normalized to the imaged area. Thus, 20 larvae were prepared and eight were analyzed per group (Supplementary Methods, "PCNA imaging and zebrafish endpoint assays").

Biochemical assays

MDA, SOD activity, and reduced GSH were measured at 96 hpf in three independently prepared larval pools per group using commercial colorimetric kits according to the manufacturers' protocols. Results were normalized to total protein and expressed as nmol/mg protein for MDA, U/mg protein for SOD activity, and μg/mg protein for GSH. These assays assess redox-associated biochemical changes but do not identify reactive oxygen species or their cellular sources directly [34].

RNA extraction and quantitative PCR

Total RNA was isolated at 96 hpf from three independent larval pools per group using TRIzol. One microgram of RNA was reverse-transcribed using a commercial cDNA synthesis kit. Quantitative PCR used SYBR Green chemistry with technical duplicates, β-actin as the reference gene, and the $2^{-\Delta\Delta Ct}$ method for relative expression [35]. The β-actin and *tnfa* primers followed the project protocol. Primers for *il1b*, *il6*, *anxa2a*, *prdx1*, and *gna15.2* were designed against Ensembl GRCz11 transcripts, checked against corresponding NCBI records, and synthesized by Sangon Biotech (Shanghai, China). Supplementary Table S1 lists the sequences and available predicted product lengths and melting temperatures. Assay details are provided in Supplementary Methods under "PCNA imaging and zebrafish endpoint assays".

Targeted serum protein measurements

Stored serum from all 17 participants was analyzed for ANXA2, PRDX1, IL-6, and TNF-α using marker-specific human ELISA kits purchased from Sangon Biotech. Assays and calibration followed the manufacturers' protocols. Each participant-level value was normalized to the JKDZ mean for the corresponding marker (Supplementary Table S3). ANXA2 and PRDX1 represented prioritized candidates, while IL-6 and TNF-α provided inflammatory context. GNA15 protein was not measured. The assays used the transcriptomic discovery cohort, and no diagnostic cutoff was derived.

Exploratory molecular docking

The DQ dication was obtained from PubChem (CID 6795), with counterions omitted [36]. Human ANXA2 and PRDX1 structures were retrieved from the RCSB Protein Data Bank (PDB 2HYW and 9B7A, respectively) [37]. Crystallographic water and non-target ligands were removed, polar hydrogen atoms were added, and receptor and ligand files were converted to PDBQT format. AutoDock Vina v1.2.7 was used to screen five candidate regions, with nine poses generated per region [38,39]. The lowest-energy poses were visualized using PyMOL and Discovery Studio. The deposited PRDX1 structure contains rosmarinic acid, but no redocking benchmark, negative-control docking, or experimental affinity measurement was performed. Docking was used to generate structural hypotheses (Supplementary Methods, "Exploratory molecular docking").

Statistical analysis

Zebrafish outcomes are presented as mean ± standard error of the mean (SEM). PCNA quantification included eight larvae per group; biochemical and qPCR assays included three biological pools per group. Differences among C, TL, and TH were assessed using one-way analysis of variance followed by Dunnett's comparisons with control. Individual plotted symbols represent biological observations, and technical duplicates were not treated as independent observations.

For serum proteins, the effect statistic was $\log_2$[mean(CNSI group)/mean(NCNSI)]. Unadjusted two-sided 95% confidence intervals were calculated using the BCa bootstrap with 10,000 within-group resamples and seed 20260817 [40]. Exact independent-label permutation tests enumerated all 70 allocations for CNSI1 versus NCNSI and all 35 allocations for CNSI2 versus NCNSI. Two-sided P values were calculated as twice the smaller upper- or lower-tail probability, including ties and capped at 1. The observed allocation was included in the permutation distribution [41]. P values for the eight marker–group comparisons were adjusted jointly using the Benjamini–Hochberg procedure [28]. Bootstrap intervals and adjusted P values were reported separately because they were obtained using different inferential procedures.

Exploratory ROC curves were constructed for serum ANXA2, PRDX1, IL-6, and TNF-α using CNSI1 and CNSI2 participants combined as the positive group (n = 7) and NCNSI participants as the negative group (n = 4). Healthy controls were not included in this comparison. Areas under the curve describe discrimination within these 11 participants; no independent validation cohort was available.

Manuscript preparation

ChatGPT (OpenAI) assisted with English-language editing, manuscript organization, editorial review, data consistency checks, and figure preparation.

Data availability

Zebrafish RNA-seq data are identified in the NCBI Sequence Read Archive under BioProject PRJNA1358919. Human serum transcriptomic data are from the cohort reported by Liu et al. [19] and are available from the corresponding author on reasonable request, subject to applicable ethical and privacy restrictions. Supplementary Table S3 contains de-identified participant-level normalized serum protein measurements, and Supplementary Table S2 provides effect estimates, BCa intervals, and exact and adjusted P values. Primer sequences and data and sample sources are provided in Supplementary Tables S1 and S4, respectively.

References

[1] Jones, G. M. & Vale, J. A. Mechanisms of toxicity, clinical features, and management of diquat poisoning: a review. J. Toxicol. Clin. Toxicol. 38, 123–128 (2000). https://doi.org/10.1081/CLT-100100926

[2] Magalhães, N., Carvalho, F. & Dinis-Oliveira, R. J. Human and experimental toxicology of diquat poisoning: toxicokinetics, mechanisms of toxicity, clinical features, and treatment. Hum. Exp. Toxicol. 37, 1131–1160 (2018). https://doi.org/10.1177/0960327118765330

[3] Zhou, J. N. & Lu, Y. Q. Lethal diquat poisoning manifests as acute central nervous system injury and circulatory failure: a retrospective cohort study of 50 cases. eClinicalMedicine 52, 101609 (2022). https://doi.org/10.1016/j.eclinm.2022.101609

[4] Ren, Y., Guo, F. & Wang, L. Imaging findings and toxicological mechanisms of nervous system injury caused by diquat. Mol. Neurobiol. 61, 9272–9283 (2024). https://doi.org/10.1007/s12035-024-04172-x

[5] Wang, X. et al. Toxicological effects of diquat on the central nervous system and associated treatment challenges. Neurotoxicology 111, 103327 (2025). https://doi.org/10.1016/j.neuro.2025.103327

[6] Zhou, J. et al. Diquat exposure causes brainstem demyelination by upregulating the mitochondrial calcium uniporter. J. Hazard. Mater. 492, 138063 (2025). https://doi.org/10.1016/j.jhazmat.2025.138063

[7] Li, S. et al. Thrombin-mediated increased permeability of blood-brain barrier in diquat poisoning: mechanistic insights and potential therapeutic targets. Ecotoxicol. Environ. Saf. 315, 120087 (2026). https://doi.org/10.1016/j.ecoenv.2026.120087

[8] Sandy, M. S., Moldeus, P., Ross, D. & Smith, M. T. Cytotoxicity of the redox cycling compound diquat in isolated hepatocytes: involvement of hydrogen peroxide and transition metals. Arch. Biochem. Biophys. 259, 29–37 (1987). https://doi.org/10.1016/0003-9861(87)90466-8

[9] Rikans, L. E., Cai, Y., Kosanke, S. D. & Venkataraman, P. S. Redox cycling and hepatotoxicity of diquat in aging male Fischer 344 rats. Drug Metab. Dispos. 21, 605–610 (1993). https://doi.org/10.1016/S0090-9556(25)07462-8

[10] Drechsel, D. A. & Patel, M. Differential contribution of the mitochondrial respiratory chain complexes to reactive oxygen species production by redox cycling agents implicated in parkinsonism. Toxicol. Sci. 112, 427–434 (2009). https://doi.org/10.1093/toxsci/kfp223

[11] Fussell, K. C. et al. Redox cycling and increased oxygen utilization contribute to diquat-induced oxidative stress and cytotoxicity in Chinese hamster ovary cells overexpressing NADPH-cytochrome P450 reductase. Free Radic. Biol. Med. 50, 874–882 (2011). https://doi.org/10.1016/j.freeradbiomed.2010.12.035

[12] Slaughter, M. R., Thakkar, H. & O'Brien, P. J. Effect of diquat on the antioxidant system and cell growth in human neuroblastoma cells. Toxicol. Appl. Pharmacol. 178, 63–70 (2002). https://doi.org/10.1006/taap.2001.9322

[13] Lai, K. et al. Sensing of mitochondrial DNA by ZBP1 promotes RIPK3-mediated necroptosis and ferroptosis in response to diquat poisoning. Cell Death Differ. 31, 635–650 (2024). https://doi.org/10.1038/s41418-024-01279-5

[14] Wang, P. et al. Diquat exacerbates oxidative stress and neuroinflammation by blocking the autophagic flux of microglia in the hippocampus. Ecotoxicol. Environ. Saf. 286, 117188 (2024). https://doi.org/10.1016/j.ecoenv.2024.117188

[15] Cassar, S. et al. Use of zebrafish in drug discovery toxicology. Chem. Res. Toxicol. 33, 95–118 (2020). https://doi.org/10.1021/acs.chemrestox.9b00335

[16] de Esch, C., Slieker, R., Wolterbeek, A., Woutersen, R. & de Groot, D. Zebrafish as potential model for developmental neurotoxicity testing: a mini review. Neurotoxicol. Teratol. 34, 545–553 (2012). https://doi.org/10.1016/j.ntt.2012.08.006

[17] Kalueff, A. V., Stewart, A. M. & Gerlai, R. Zebrafish as an emerging model for studying complex brain disorders. Trends Pharmacol. Sci. 35, 63–75 (2014). https://doi.org/10.1016/j.tips.2013.12.002

[18] Wang, X. H., Souders, C. L. II, Zhao, Y. H. & Martyniuk, C. J. Mitochondrial bioenergetics and locomotor activity are altered in zebrafish (Danio rerio) after exposure to the bipyridylium herbicide diquat. Toxicol. Lett. 283, 13–20 (2018). https://doi.org/10.1016/j.toxlet.2017.10.022

[19] Liu, Q. et al. Diquat-induced central nervous system injury: serum multi-omics and zebrafish evidence for histidine–methylmalonate imbalance and neuroinflammation. Chem. Biol. Interact. 432, 112046 (2026). https://doi.org/10.1016/j.cbi.2026.112046

[20] Dallacasagrande, V. & Hajjar, K. A. Annexin A2 in inflammation and host defense. Cells 9, 1499 (2020). https://doi.org/10.3390/cells9061499

[21] Talwar, D., Messens, J. & Dick, T. P. A role for annexin A2 in scaffolding the peroxiredoxin 2–STAT3 redox relay complex. Nat. Commun. 11, 4512 (2020). https://doi.org/10.1038/s41467-020-18324-9

[22] Riddell, J. R., Wang, X. Y., Minderman, H. & Gollnick, S. O. Peroxiredoxin 1 stimulates secretion of proinflammatory cytokines by binding to TLR4. J. Immunol. 184, 1022–1030 (2010). https://doi.org/10.4049/jimmunol.0901945

[23] Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890 (2018). https://doi.org/10.1093/bioinformatics/bty560

[24] Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907–915 (2019). https://doi.org/10.1038/s41587-019-0201-4

[25] Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930 (2014). https://doi.org/10.1093/bioinformatics/btt656

[26] Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. 47, D766–D773 (2019). https://doi.org/10.1093/nar/gky955

[27] Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8

[28] Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300 (1995). https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[29] Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). https://doi.org/10.1093/bioinformatics/btp616

[30] Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287 (2012). https://doi.org/10.1089/omi.2011.0118

[31] Yu, G. & He, Q. Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Mol. BioSyst. 12, 477–479 (2016). https://doi.org/10.1039/C5MB00663E

[32] Davis, A. P. et al. Comparative Toxicogenomics Database (CTD): update 2023. Nucleic Acids Res. 51, D1257–D1262 (2023). https://doi.org/10.1093/nar/gkac833

[33] Szklarczyk, D. et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 51, D638–D646 (2023). https://doi.org/10.1093/nar/gkac1000

[34] Halliwell, B. & Whiteman, M. Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? Br. J. Pharmacol. 142, 231–255 (2004). https://doi.org/10.1038/sj.bjp.0705776

[35] Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the \(2^{-\Delta\Delta C_T}\) method. Methods 25, 402–408 (2001). https://doi.org/10.1006/meth.2001.1262

[36] Kim, S. et al. PubChem 2023 update. Nucleic Acids Res. 51, D1373–D1380 (2023). https://doi.org/10.1093/nar/gkac956

[37] Burley, S. K. et al. RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. Nucleic Acids Res. 49, D437–D451 (2021). https://doi.org/10.1093/nar/gkaa1038

[38] Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock Vina 1.2.0: new docking methods, expanded force field, and Python bindings. J. Chem. Inf. Model. 61, 3891–3898 (2021). https://doi.org/10.1021/acs.jcim.1c00203

[39] Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461 (2010). https://doi.org/10.1002/jcc.21334

[40] Efron, B. Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82, 171–185 (1987). https://doi.org/10.1080/01621459.1987.10478410

[41] Phipson, B. & Smyth, G. K. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. 9, Article 39 (2010). https://doi.org/10.2202/1544-6115.1585

Funding

This research was funded by the Key Research and Development Program of Fuyang City, Anhui Province - Clinical Medical Research Transformation Project (No. FK20254639).

Author contributions

Qiqi Liu: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft. **Tingzhan Lv:** Methodology, Investigation, Data curation, Formal analysis, Writing – original draft. **Yong Li:** Investigation, Data curation, Writing – review & editing. **Min Yan:** Investigation, Resources, Writing – review & editing. **Ruijie Li:** Investigation, Data curation, Writing – review & editing. **Kun Huang:** Investigation, Visualization, Writing – review & editing. **Hongbo Liu:** Conceptualization, Resources, Supervision, Writing – review & editing. **Yihong Yang:** Conceptualization, Project administration, Supervision, Writing – review & editing.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Information includes expanded methods, primer sequences (Supplementary Table S1), serum effect estimates and statistical comparisons (Supplementary Table S2), participant-level normalized serum measurements (Supplementary Table S3), data and sample sources (Supplementary Table S4), and supporting transcriptomic, database, and exploratory ROC analyses (Supplementary Fig. S1).

Figure legends

Figure 1. Cross-species identification of DQ-associated candidate genes.

(a) Workflow integrating serum transcriptomes from 17 participants and nine zebrafish transcriptomes, followed by candidate selection and experimental follow-up. (b) Filtering of the zebrafish TH-versus-C comparison: 532 differentially expressed genes and 237 genes mapped to human orthologs. (c) UpSet plot of the three comparison-specific gene sets. (d) Within-species, gene-wise z-score heatmaps for *ANXA2*, *PRDX1*, and *GNA15*. Human groups are healthy controls (JKDZ, J), poisoned patients without CNSI (NCNSI, N), patients who developed CNSI after admission (CNSI1, I1), and patients with CNSI at admission (CNSI2, I2). Zebrafish groups are control (C), 100 mg/L DQ (TL), and 175 mg/L DQ (TH). Columns represent participants or biological pools. (e) Log_2 fold changes; color indicates the comparison and point size represents $-log_{10}$(FDR). The dashed line indicates zero. Both human comparisons share the NCNSI reference group. DQ, diquat; CNSI, central nervous system injury; FDR, false discovery rate.

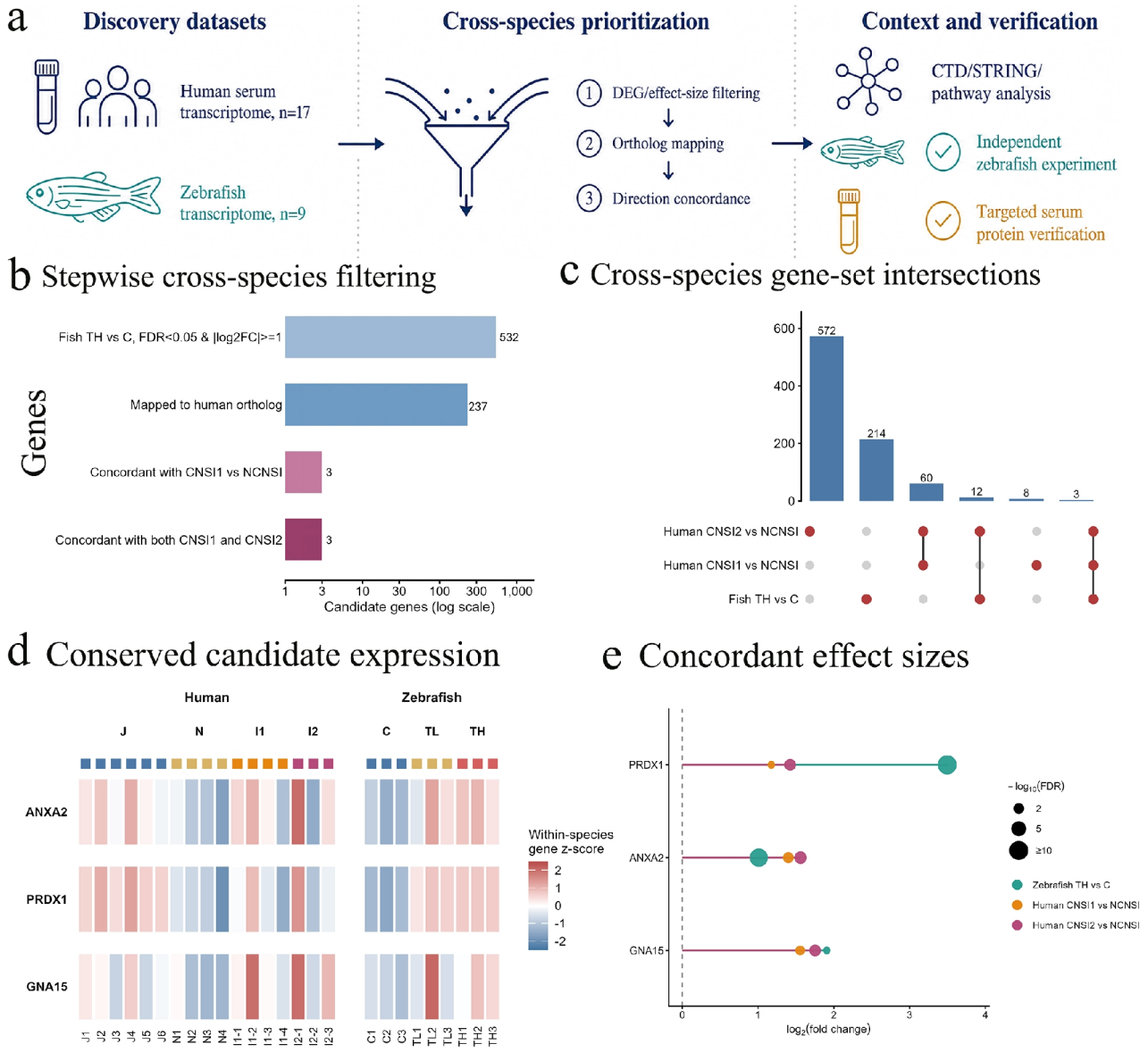


Figure 2. Functional annotations and candidate-centered association networks.

(a) Gene Ontology biological-process annotations. The horizontal axis shows gene ratio, point size indicates gene count, and color represents −log₁₀(Benjamini–Hochberg-adjusted P). (b) KEGG and Reactome results. Circles indicate KEGG terms and squares indicate Reactome terms; red denotes FDR < 0.05 and gray denotes terms above this threshold. The dashed line marks adjusted P = 0.05. (c) STRING association networks centered on ANXA2, GNA15, and PRDX1. Colored nodes represent seed proteins and gray nodes represent first-shell interactors. Node outlines indicate functional classes, and edge widths represent combined scores from 0.70 to 1.00. These database associations do not establish physical interactions during diquat exposure. FDR, false discovery rate.

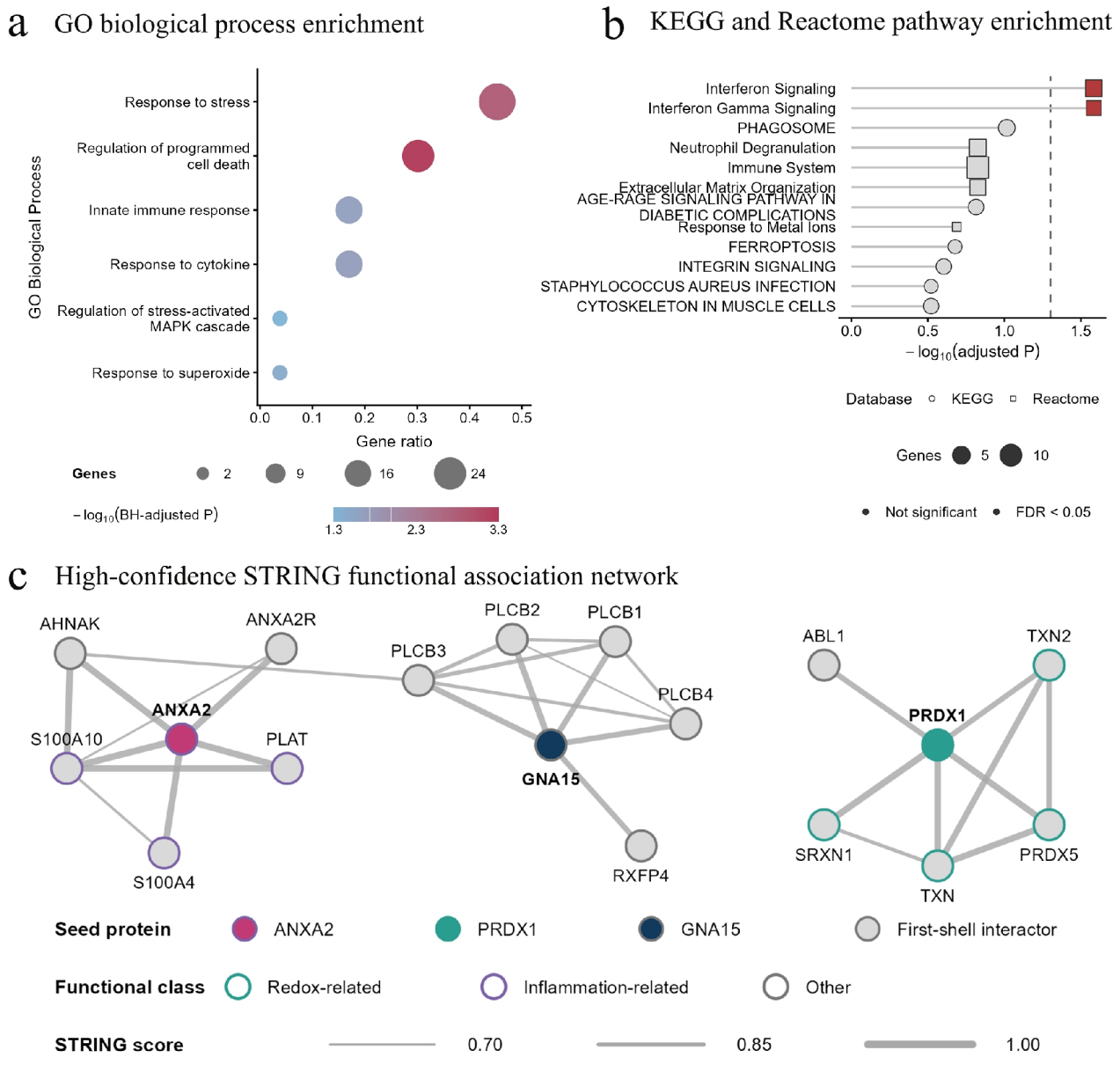


Figure 3. Cellular, biochemical, and transcriptional responses to diquat in zebrafish.

(a) Timeline showing allocation at 1 h post-fertilization (hpf), exposure from 72 to 96 hpf, and collection at 96 hpf. Groups received 0 (C), 100 (TL), or 175 mg/L diquat (TH). (b) Representative Tg(elavl3:EGFP), PCNA,

DAPI, and merged images; scale bars are indicated. (c) PCNA-positive cell density in eight larvae analyzed from 20 prepared larvae per group. (d–f) MDA, SOD activity, and reduced GSH, normalized to total protein. (g) Relative expression of *il1b*, *tnfa*, and *il6*. (h) Relative expression of *anxa2a*, *prdx1*, and *gna15.2*. Biochemical and qPCR assays used three biological pools per group. Bars show mean ± SEM; symbols represent biological observations. ***P < 0.001 versus control by one-way ANOVA followed by Dunnett's test. Imaging used a separate cohort from the pooled assays. PCNA, proliferating cell nuclear antigen; MDA, malondialdehyde; SOD, superoxide dismutase; GSH, glutathione; SEM, standard error of the mean.

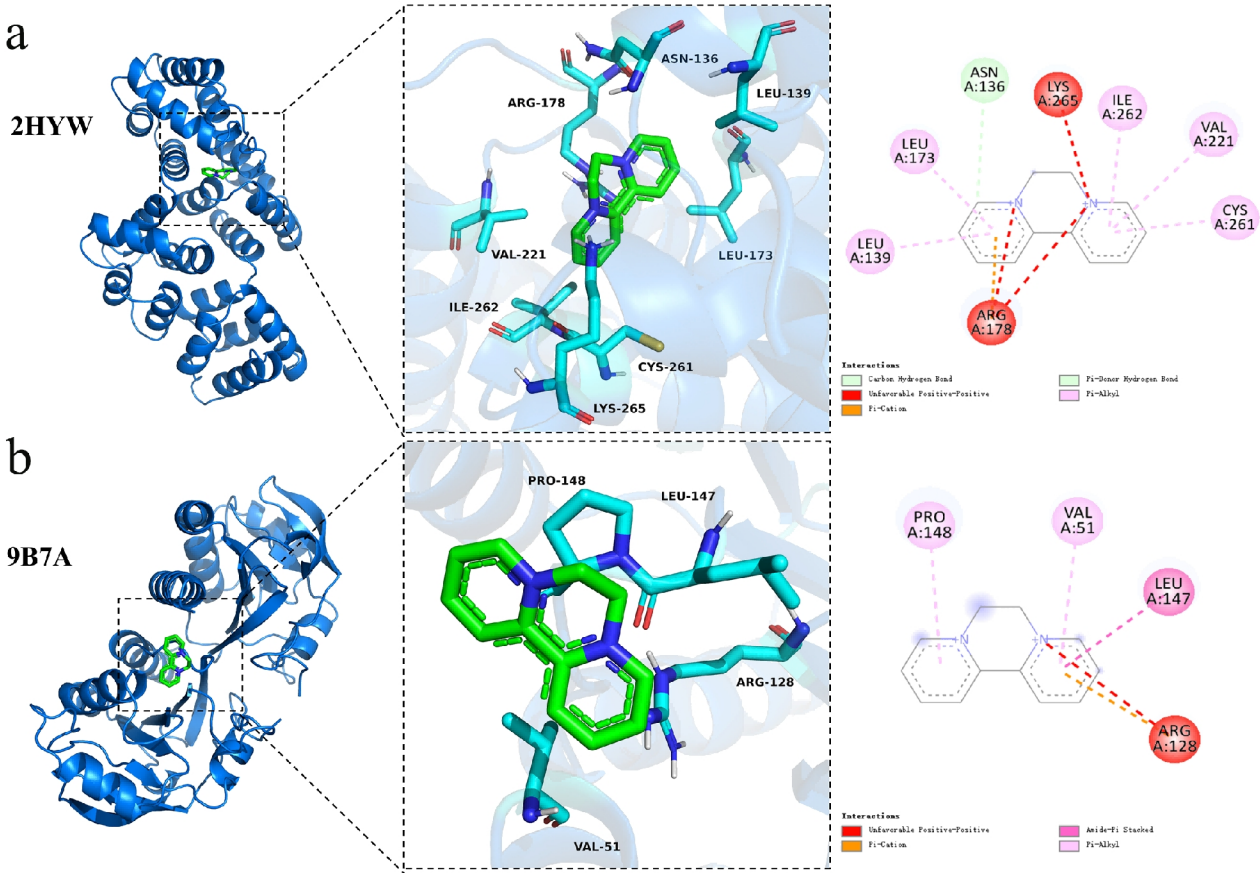


Figure 4. Serum protein measurements and effect estimates in the clinical groups.

Normalized serum (a) ANXA2, (b) PRDX1, (c) IL-6, and (d) TNF-α. Points represent participants; boxes show medians and interquartile ranges, and whiskers span the observed range. Group sizes were JKDZ, n = 6; NCNSI, n = 4; CNSI1, n = 4; and CNSI2, n = 3. (e) Log_2 mean ratios relative to NCNSI, with unadjusted 95% BCa bootstrap intervals. Blue circles represent CNSI1 and red squares represent CNSI2; the dashed line indicates zero. Exact two-sided permutation P values were adjusted jointly across eight comparisons using the Benjamini–Hochberg procedure; none was below 0.05 after adjustment. Measurements and statistical estimates are provided in Supplementary Tables S3 and S2, respectively. Serum assays used the transcriptomic discovery cohort. JKDZ, healthy controls; NCNSI, poisoned patients without central nervous system injury (CNSI); CNSI1, patients who developed CNSI after admission; CNSI2, patients with CNSI at admission; BCa, bias-corrected and

accelerated.

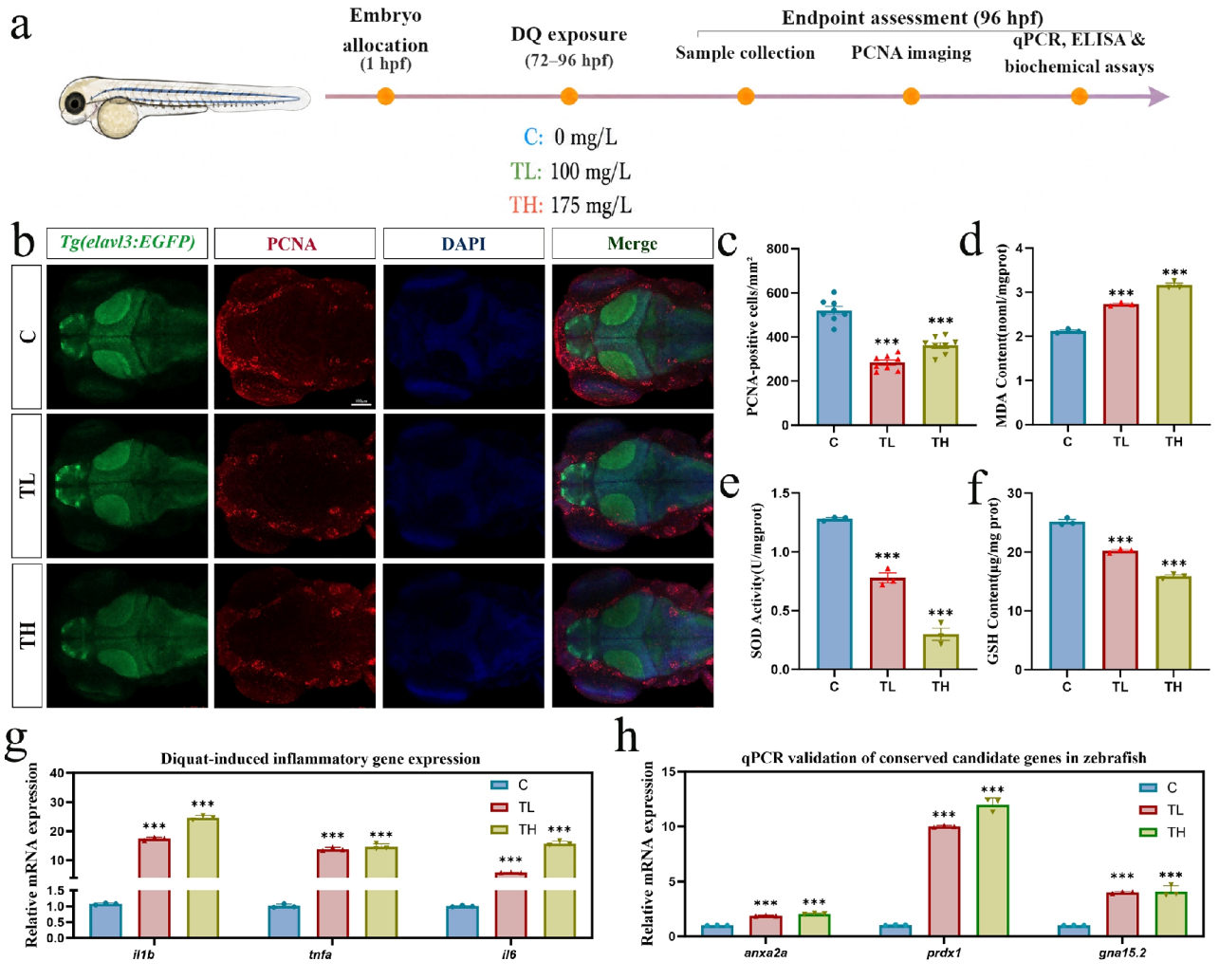


Figure 5. Exploratory docking poses of diquat in ANXA2 and PRDX1.

Poses of the diquat dication in (a) ANXA2 (PDB 2HYW) and (b) PRDX1 (PDB 9B7A), generated using AutoDock Vina v1.2.7. Panels show the protein structure, the selected pose and neighboring residues, and the corresponding two-dimensional interaction map. Diquat is shown in green and neighboring residues in cyan. Nine poses were generated per evaluated region, and a lowest-energy pose was selected for display. The maps include unfavorable positive–positive contacts. These calculations do not establish direct binding, affinity, or a causal

mechanism.

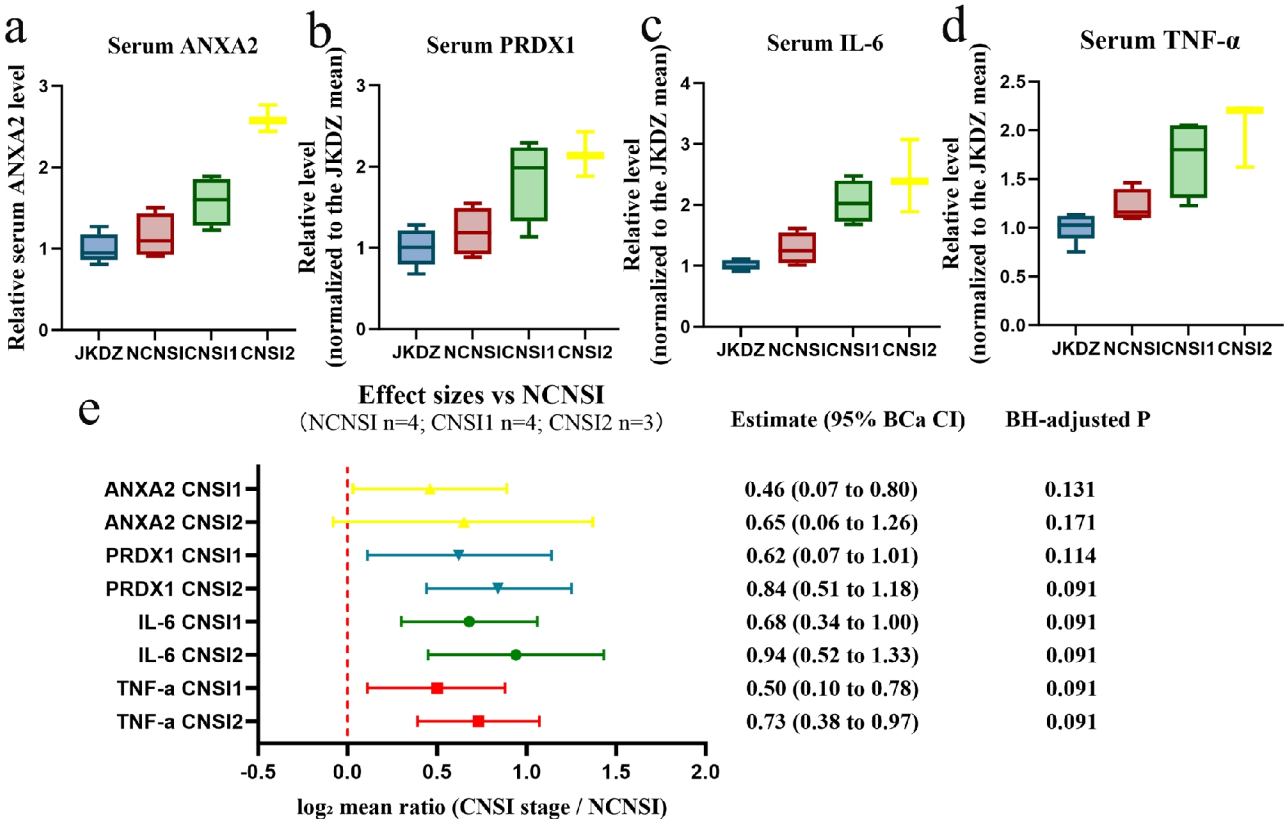
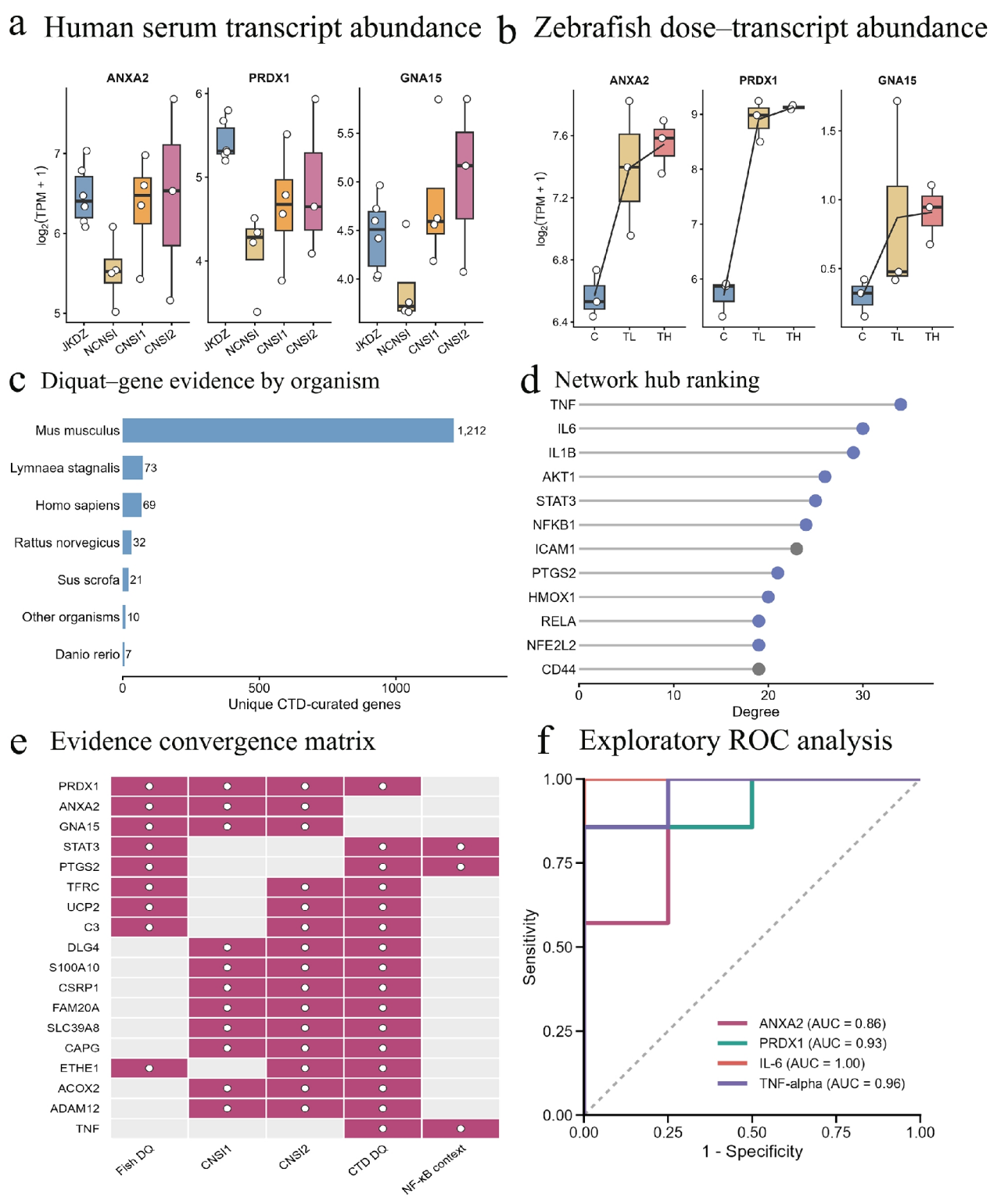


Figure S1. Supplementary cross-species transcript profiling, network analysis, and exploratory discrimination of candidate markers associated with diquat exposure.

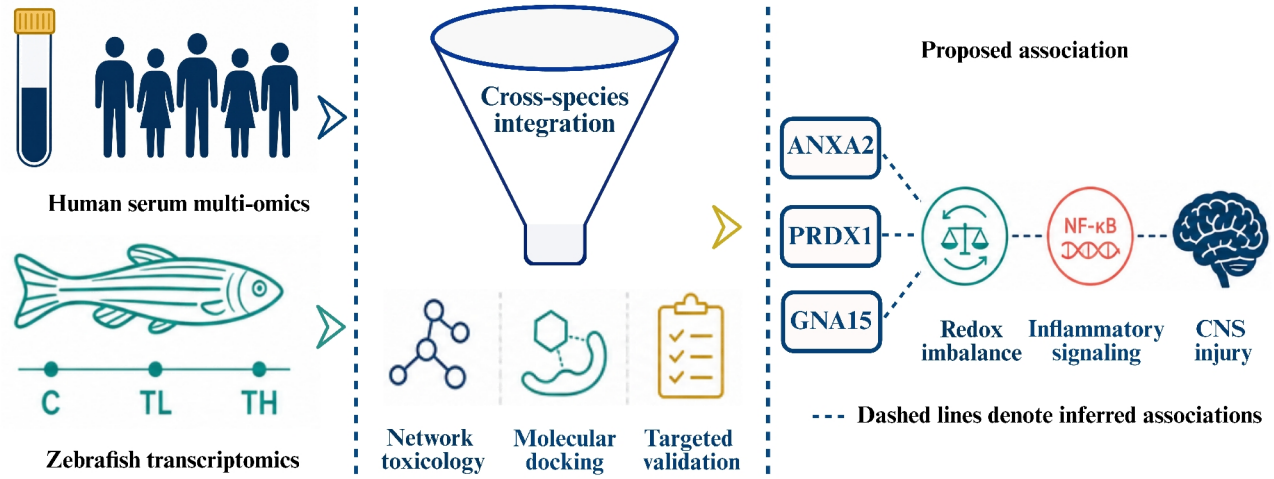
(a) Transcript abundance of ANXA2, PRDX1, and GNA15 in human serum samples from the JKDZ, NCNSI, CNSI1, and CNSI2 groups, expressed as $\log_2(TPM + 1)$. (b) Transcript abundance of the corresponding zebrafish candidates in the control (C), low-dose (TL), and high-dose (TH) exposure groups. Open circles in panels a and b represent individual samples. (c) Numbers of unique diquat-associated genes curated in the Comparative Toxicogenomics Database (CTD), grouped by organism. (d) Degree-based ranking of selected nodes in the analyzed interaction network; degree represents the number of connections to each node. (e) Evidence convergence matrix summarizing support for selected genes across the zebrafish transcriptomic comparison, human CNSI1 and CNSI2 comparisons, CTD diquat associations, and the NF-κB/inflammatory gene set. Colored cells containing white dots indicate that the corresponding evidence criterion was met; gray cells indicate that it was not met in the analysis. (f) Exploratory receiver operating characteristic (ROC) curves for ANXA2, PRDX1, IL-6, and TNF-α, with areas under the curve (AUCs) of 0.86, 0.93, 1.00, and 0.96, respectively. The dashed diagonal represents chance-level discrimination. These exploratory results do not establish diagnostic performance in an independent population. TPM, transcripts per

million.



Graphical Abstract. Cross-species integration to prioritize candidate markers associated with diquat-related neurotoxicity.

Human serum multi-omics and zebrafish transcriptomic data are integrated with network toxicology, molecular docking, and targeted validation to investigate ANXA2, PRDX1, and GNA15 as candidate markers associated with diquat-related neurotoxicity. The proposed model links these candidates to redox imbalance, inflammatory signaling involving NF-κB, and central nervous system (CNS) injury. Dashed lines indicate inferred associations rather than experimentally established causal relationships. C, control; TL, low-dose diquat exposure; TH, high-dose diquat exposure.



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.pdf