HHS Public Access

Author manuscript

Environ Int. Author manuscript; available in PMC 2023 February 03.

Published in final edited form as:

Environ Int. 2022 December ; 170: 107613. doi:10.1016/j.envint.2022.107613.

Epidemiology meets toxicogenomics: Mining toxicologic evidence in support of an untargeted analysis of pesticides exposure and Parkinson's disease

Kimberly C. Paul[a,*], **Beate Ritz**[a,b]

[a]Department of Neurology, David Geffen School of Medicine, Los Angeles, CA, USA

[b]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, CA, USA

Abstract

Background: Pesticides have been widely used in agriculture for more than half a century. However, with thousands currently in use, most have not been adequately assessed for influence Parkinson's disease (PD).

Objectives: Here we aimed to assess biologic plausibility of 70 pesticides implicated with PD through an agnostic pesticide-wide association study using a data mining approach linking toxicology and toxicogenomics databases.

Methods: We linked the 70 targeted pesticides to quantitative high-throughput screening assay findings from the Toxicology in the 21st Century (Tox21) program and pesticide-related genetic/disease information with the Comparative Toxicogenomics Database (CTD). We used the CTD to determine networks of genes each pesticide has been linked to and assess enrichment of relevant gene ontology (GO) annotations. With Tox21, we evaluated pesticide induced activity on a series of 43 nuclear receptor and stress response assays and two cytotoxicity assays.

Results: Overall, 59 % of the 70 pesticides had chemical-gene networks including at least one PD gene/gene product. In total, 41 % of the pesticides had chemical-gene networks enriched for ≥ 1 high-priority PD GO terms. For instance, 23 pesticides had chemical-gene networks enriched for response to oxidative stress, 21 for regulation of neuron death, and twelve for autophagy, including copper sulfate, endosulfan and chlorpyrifos. Of the pesticides tested against the Tox21 assays, 79 % showed activity on ≥ 1 assay and 11 were toxic to the two human cell lines. The set

*Corresponding author at: 73-320B CHS; CAMPUS-177220; UCLA, Los Angeles, CA 90095, USA. kimberlp@ucla.edu (K.C. Paul).

CRediT authorship contribution statement

Kimberly C. Paul: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing – original draft. **Beate Ritz:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: 'Beate Ritz reports a relationship with Korein and Tillery Law Firm that includes: paid expert testimony. Beate Ritz reports a relationship with Andrus Wagstaff Law Firm that includes: paid expert testimony.'

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2022.107613.

of PD-associated pesticides showed more activity than expected on assays testing for xenobiotic homeostasis, mitochondrial membrane permeability, and genotoxic stress.

Conclusions: Overall, cross-database queries allowed us to connect a targeted set of pesticides implicated in PD via epidemiology to specific biologic targets relevant to PD etiology. This knowledge can be used to help prioritize targets for future experimental studies and improve our understanding of the role of pesticides in PD etiology.

Keywords

Parkinson's disease; Pesticides; Toxicogenomics; Tox21; Chlorpyrifos; Copper sulfate

1. Introduction

Pesticides have been widely used in agriculture across the world for more than half a century. Rapid growth in the chemical industry over this period contributed to the introduction and commercial use of a wide variety of products. Currently in the state of California, for instance, 13,540 pesticide products and 1,074 different active ingredients are registered for use (California Department of Pesticide Regulation, 2021). Pesticides are important tools of modern commercial agriculture that help maximize food production. However, as pesticides are specifically designed to kill living organisms (e.g., plants, fungi, insects, rodents), they need to be adequately assessed for potential adverse health effects to humans, particularly those that are applied at industrial scales. In the United States, toxicity testing used to register pesticides is generally based on single-product, single-target assessments, primarily in rodents as model organisms, and largely does not consider chronic exposures, mixtures, or sequential treatments with a variety of pesticides (Topping et al., 2020; Milner and Boyd, 2017). In fact, the current accepted toxicology testing routines are generally regarded as not able to assess the full range of toxic effects that could emerge when pesticides are used at scale, especially for effects due to chronic, low-level exposures (Topping et al., 2020; Milner and Boyd, 2017). Furthermore, to date, most commercially applied pesticides have not been adequately assessed for chronic health effects, including Parkinson's disease (PD), by either epidemiologic or experimental research. Toxicologic information produced by other approaches, such as high-throughput screening via *in vitro* cell models, is also not yet systematically employed to inform further studies sufficiently in terms of prioritizing chemical targets for research or providing biologic plausibility of epidemiologic results.

With two primary considerations in mind, namely the large number of individual pesticides commercially used and lack of adequate data on chronic health effects for most, we have designed a multi-step agnostic approach to screen hundreds of agriculturally used pesticides for association with PD. Our approach combines epidemiologic pesticide screening in a population-based study with experimental toxicologic data and toxicogenomics database mining. First, to epidemiologically screen pesticides for association, we developed a record-based pesticide exposure approach using agricultural pesticide application records and land use information in California. This exposure assessment system allowed us to estimate proximity-based ambient exposure to 722 pesticide active ingredients over a 40-year period in a large population-based PD study (Paul et al., 2022). Armed with this information,

we conducted an untargeted screen of all pesticides with widespread agricultural use in California for association with PD, or what we call a “pesticide exposure-wide association study” (PEWAS)(Paul et al., 2022). Our goal was to comprehensively and agnostically investigate long-term pesticide exposure and PD risk in a *pesticide-specific* manner and prioritize agents for follow-up research.

Using this PEWAS approach, we previously implicated 70 pesticide active ingredients as associated with PD (Paul et al., 2022). Many of the pesticides are newly implicated and have not been examined for potential contribution to PD etiology in the human or toxicologic literature thus far. This PEWAS analysis was the first step, i.e., a broad exposure-outcome screen in humans. As a screen, the purpose was not to suggest that every pesticide implicated through a positive association causes PD. Instead, we aimed to prioritize agents and mixtures of agents for further in-depth investigation from amongst the large variety and number of pesticides applied in the geographic location of our study. To further assess biologic plausibility of our epidemiologic screen-based findings and help highlight specific agents as promising candidates for experimental exploration, we secondly employed a data mining approach to link the pesticides our central California study implicated with PD through our epidemiologic screen to toxicology and toxicogenomics databases. Furthermore, given the strength of scientific evidence linking both paraquat and rotenone to PD, we further included both as “positive controls” and 8 pesticides with sufficient exposure prevalence which showed no associations in our epidemiologic screen as “negative controls”.

Here we provide the results from this linkage that allowed us to integrate the 70 pesticides of interest our study identified with toxicologic assay results and toxicogenomic databases. We also previously had the unique opportunity to couple our PEWAS with an *in vitro* testing pipeline that administered the implicated chemicals to sensitized dopaminergic neurons from a PD patient-derived induce pluripotent stem cell (iPSC) model (Paul et al., 2022). Thus, we were able to characterize the epidemiologic hits in terms their performance on multiple toxicologic assays. The information we derived will not only help with future research prioritization but by interrogating a large set of pesticides in this manner we demonstrate how other health endpoints could also be assessed similarly and rapidly through existing sources of experimental evidence. We provide our full analysis code in the supplement to facilitate this. Specifically, we show how we incorporated quantitative high-throughput screening (qHTS) assay data from the Tox21 (Toxicology in the 21st Century (Hur et al., 2018; Richard et al., 2021) program and linked each targeted pesticide to the Comparative Toxicogenomics Database (CTD), a publicly available database that builds networks of chemical, genetic, and disease-related information from manual curation of vast amounts of peer-reviewed literature (Davis et al., 2021).

We ultimately intended to evaluate the extent that pesticides identified in epidemiologic studies could be connected to biologic pathways and mechanisms known to be involved in PD etiology and improve our understanding of the possible role of specific pesticides in PD etiology. We highlight pesticide-related molecular events common across many PD-implicated pesticides as well as specific to individual pesticides, again with our goal being to help prioritize targets for future low-throughput or resource intensive experimental studies.

2. Methods

2.1. Pesticide selection

We began with three sets of pesticides: a set of 70 targeted pesticides and a set each of positive and negative control pesticides. The 70 targeted pesticides of interest were determined through association testing with PD in our previous pesticide-wide association analysis (Paul et al., 2022). For this PEWAS, we used an agnostic strategy to investigate all pesticides agriculturally applied near the homes and workplaces of 1,653 study participants living in three agricultural counties in Central California (Kern, Fresno, ant Tulare). Based on exposure prevalence in the study population, we were able to test 288 pesticides individually for association in the PEWAS using a population-based study of Parkinson's (Parkinson's Environment and Genes (PEG) study)(Ritz et al., 2016). In total, as presented previously, our analyses implicated 70 pesticides out of the 288 as associated with PD (25 at FDR ≤ 0.01, 28 at 0.01 < FDR ≤ 0.05, and 15 at 0.05 < FDR ≤ 0.10)(Paul et al., 2022). These 70 pesticides provided the base set of targeted toxicants for our computational toxicology analysis presented here making use of the CTD and Tox21 databases. The epidemiologic associations of these pesticides with PD from the initial PEWAS are listed in Supplemental Table 1.

Two positive control pesticides, paraquat and rotenone, were selected given the strength of prior evidence linking both to PD through experimental and epidemiologic studies (Liu et al., 2020). We also included eight negative control pesticides not associated with PD in previous epidemiologic study or our PEWAS study and of sufficient prevalence amongst study participants. Specifically, they fit the following criteria of "no association": they showed null association (individual ORs per SD: 0.95–1.05) at both exposure locations (ambient exposure due to applications near residences or workplaces) and in both study populations. Eight pesticides included in the CTD fit these criteria: mefenoxam (metalaxyl-M), cyprodinil, tebuconazole, tebufenozide, halosulfuron-methyl, spinosad, pyrithiobac-sodium, and potassium bicarbonate.

2.2. Comparative toxicogenomics database

The CTD is a publicly available database aimed at advancing our understanding of how environmental exposures affect human health (Davis et al., 2021). The database provides contextualized content relating chemical exposures with human health, through manual curation of peer-reviewed scientific literature to interrelate chemical, gene, phenotype, disease, and exposure information (Davis et al., 2021). The CTD integrates new research monthly to ensure the database is comprehensive and current. As of May 2022, the CTD included over 2.5 million manually curated chemical-gene interactions, along with chemical-phenotype, chemical-disease, and gene-disease interactions, relaying information from 140,475 curated references about 17,070 chemicals (Davis et al., 2021). Thus, the CTD provides a mechanism to integrate information from vast amounts of previous literature related to chemicals of interest.

The CTD describes relationships (e.g., chemical-disease, chemicalgene, etc.) as direct or inferred (or both). For chemical-disease relationships, direct relationships are established

by CTD curation and indicate direct evidence from peer reviewed literature linking the chemical and disease. Inferred relationships are established through curated chemical-gene interactions, indicating that the chemical interacts with a gene or gene product that is directly associated with the disease based on OMIM (gene-disease). For example, chemical A will have an inferred association with disease B, if chemical A interacts with gene C and gene C has a direct association with disease B (Davis et al., 2008). To assess the level of evidence for inferred relationships, the CTD provides an inference score, which is a statistic for prioritizing inferences based on the topology of networks (King et al., 2012). The higher the inference score, the stronger the pesticide's inference network (e.g., all genes linked to both the pesticide and PD) is linked to PD and the more likely the inference network has atypical connectivity that is not attributable to chance (King et al., 2012). More detail on the inference score has been published (King et al., 2012). We used the most recent CTD data release, last accessed May 2022 through the CTDquerier R package to download CTD data to the R/Bioconductor framework (Hernandez-Ferrer and Gonzalez, 2018).

2.3. Quantitative High-Throughput screening assay Data, Tox21

Tox21 is a U.S. federal collaboration between the Environmental Protection Agency (EPA), the National Toxicology Program (NTP) at NIEHS, the National Center for Advancing Translational Sciences (NCATS) at NIH, and the Food and Drug Administration (FDA), to experimentally assess toxicity of thousands of chemical compounds that have the potential to disrupt processes in the human body using *in vitro* biologic assays (Hur et al., 2018; Richard et al., 2021). We used Tox21 assay data to assess toxicologic activity of the PD-associated pesticides of interest. To date, the Tox21 program has screened 13,128 chemicals (tox21_10k_library_info) using a quantitative high-throughput screening (qHTS) format and various cell-based assays to serve as *in vitro* models (Richard et al., 2021; Tice et al., 2013; Shukla et al., 2010; Huang, 2016). Each assay was screened against the 10 k library as triplicate 15-dose titrations (i.e., 3 replicates for each dose), generating over 50 million data points to date. Details of the Tox21 qHTS assay experimental process have been published (Hur et al., 2018; Huang, 2016; Huang et al., 2016). The toxicity profiling here is based on a nuclear receptor (NR) panel and stress response (SR) panel of forty-three assays. Detailed descriptions, including specific information on targets, technology (e.g., luminescence, fluorescence), cell lines, tissues, protocol details of the NR and SR assays are provided in Supplemental Table 2. These assays measured effects on a range of pathways assessing toxicity to targets including bile acid homeostasis, endoplasmic reticulum stress, genotoxic stress, glucocorticoid homeostasis, heat shock stress, hypoxia stress, inflammation stress, lipid homeostasis, mineral homeostasis, miscellaneous stress, mitochondrial membrane permeability, oxidative stress, retinoid homeostasis, sex hormone homeostasis, steroidogenesis, sterol homeostasis, thyroid homeostasis, and xenobiotic homeostasis.

In addition, we integrated results from two multiplexed, real-time cytotoxicity assays run on two human cell lines: HEK293, a human embryonic kidney cell line, and HepG2, a human hepatocellular carcinoma cell line. Twenty of the 43 NR/SR assays were also run using one of these two cell lines (Supplemental Table 2). For the cytotoxicity assays, the Tox21 10 k library was experimentally interrogated at 6 time points (Huang et al., 2016;

NIEHS/NTP. Tox21 Activity Profiler) and classified across replicates at each time point as active, inconclusive, or inactive for the cytotoxicity endpoint. More detail has been published (Hsieh et al., 2017).

NR/SR activity data was obtained from the Tox21 Activity Profiler online tool, retrieved May 2022. Cytotoxicity data was downloaded from the Tox21 Public Data Repository, May 2022 (Tox21 Public Data Repository, 2022).

2.4. Analytic workflow

We provide a complete code notebook (R markdown html report) in the supplement, which details data retrieval steps and all analyses with output of results presented in this manuscript. Our accompanying MethodsX manuscript provides further details, including all data files used for analysis.

We linked the three sets of pesticides (70 targeted pesticides, 8 negative and 2 positive controls) to the CTD using the CTDquerier R package (Hernandez-Ferrer and Gonzalez, 2018), providing the pesticides by common name as input (see code notebook section 1). To ensure correct linkage, we manually assessed each input and CTD match for equivalency. We allowed for imperfect matches when the matched CTD chemical was equivalent to the California Pesticide Use Report (PUR) chemical. For example, during the linkage, one warning read “No perfect match for term ‘propyzamide’, taking term ‘pronamide’ as new term”. Pronamide is a synonym of prorzamide per PubChem, we therefore allowed this match to go forward.

Of the 70 PEWAS identified and targeted pesticides of interest, we were able to match a total of 59 to the CTD. All positive and negative control pesticides matched successfully. The CTD terms and synonym identifiers for the matched pesticides from the targeted pesticides of interest are shown in Supplemental Table 3. Unmatched pesticides generally fell into a parent category with a matched pesticide. For example, the California Pesticide Use Report database has separate pesticide chemical IDs for “Dicamba” and “Dicamba, Other Related”. However, only dicamba was listed in the CTD. The following petroleum solvents, “petroleum distillates, aromatic”, “petroleum distillates”, “petroleum hydrocarbons”, also all matched to “petroleum”. More information about the matching, including unmatched pesticides and the list of all accepted matches, is provide in the supplemental methods.

We first queried the CTD to find whether the selected pesticides had previously been linked with PD (MeSH:D010300) in the literature, assessing both direct and inferred relations. For each pesticide, we also queried for the network of chemical-gene/gene product interactions to provide information on all genes that have been linked to the pesticides. We then specifically assessed genes/gene products included in the KEGG Parkinson disease pathway (hsa05012). Finally, we queried for enriched chemical-gene ontology (GO) term relationships, including molecular function, biological processes, and cellular components. These GO term relationships are provided by the CTD and can be directly extracted with CTDquerier (see code notebook). To assess enrichment, the CTD uses the network of genes/gene products linked to each chemical and GO annotations associated with each gene. Enrichment significance was determined by the hypergeometric distribution, calculating the

probability that the fraction of genes interacting with a chemical annotated to the specific GO term is significantly higher than the fraction of all human genes annotated to that GO term (CTD). P-values were adjusted using the Bonferroni method.

We specifically assessed PD-relevant GO terms in a targeted query, limiting to 4,500 GO terms that have been linked to PD through the Parkinson's UK GO Annotation Initiative (ParkinsonsUK-UCL). The Parkinson's GO annotation initiative specifically highlighted the following processes as having important roles in PD: WNT signaling pathway, regulation of autophagy, mitophagy, endoplasmic reticulum unfolded protein response, synaptic vesicle transport, oxidative stress, and regulation of neuronal death. The initiative also delineated 48 high priority genes of importance in PD.

For Tox21, we first determined the CAS registry number for each pesticide, extracted from the PAN database (PAN North America), to link the selected pesticides to Tox21 by CAS. In total, 45 of the 70 targeted pesticides, one positive control (rotenone), and five negative controls were included in the Tox21 10 k library (see code notebook, section 2).

We assessed the toxicologic activity of these pesticides first using NR/SR qHTS assay activity data (Kavlock et al., 2009). We uploaded the lists of pesticides (targeted pesticides, negative controls, positive controls) to the Tox21 Activity Profiler, which is part of the Tox21 Toolbox (NIEHS/NTP. Tox21 Toolbox; NIEHS/NTP. Tox21 Activity Profiler). The Activity Profiler is a data analysis tool for accessing and visualizing the Tox21 qHTS 10 K library. The profiler links the input list of pesticide CAS registry numbers to the qHTS assay output and provides pesticide induced activity on 43 NR/SR Tox21 assays. Activity is displayed using the weighted area-under-curve (wAUC), showing the overall effect of the chemical with higher values indicating higher activity [range 0 (inactive) to 1 (active)]. From the 45 targeted pesticides in the Tox21 10 k library, 43 have been processed using the qHTS against the SR/NR assay panels. 1,2-Dichloropropane and Xylenes are included in the 10 k library but not included in the Tox21 Activity Profiler. Activity for antagonistic-type calls due to cytotoxicity was shown as inconclusive. We also excluded activity due to auto-fluorescence and activity with no reporter gene activity readout support.

We further determined assays enriched in the targeted PD-associated pesticide set relative to both (California Department of Pesticide Regulation, 2021) (1) all Tox21 tested chemicals and (Topping et al., 2020) (2) limiting to only other pesticides in the Tox21 library, determined through extracting all pesticide product CAS registry numbers from the CompTox dashboard (EPA) (see code notebook section 2). Specifically, we assessed whether the set of PD-associated pesticides showed more activity on certain assays than expected based on either all chemicals or all pesticides only that had been tested with the assay. Enrichment is based on overrepresentation analysis and a Fisher's exact test and hypergeometric distribution.

Cytotoxicity assay data from the two multiplexed, real-time assays run on two human cell lines, HEK293 and HepG2, was downloaded directly from the Tox21 public data repository (Tox21 Public Data Repository, 2022). After excluding low purity data (purity grade from Tox21 not at C or above), we aggregated the cytotoxicity outcomes determined for each time

points (6 time points, 0–40 h) and each of the two assays to get a range from completely inactive (0, inactive at all time points and on both assays) to completely active (12, active at all 6 time points on both assays (6×2)).

Finally, we provide a summary profile for each pesticide based on the original PEWAS association, cytotoxicity in the two human cell lines (embryonic kidney and liver carcinoma), the CTD inference score between the pesticides and PD, and the Tox21 ToxScore computed by Tox21 Activity Profiler based on the NR and SR activity panels. We prioritized the targeted pesticides into five categories: high priority (pesticides with both a CTD inference score and Tox21 ToxScore > 0), medium–high priority (pesticides with a CTD PD inference score but missing Tox21 testing or pesticides with a Tox21 ToxScore but missing CTD data), medium priority (pesticides with a CTD PD inference score or Tox21 ToxScore > 0 and no inference or activity for the other score), low priority (pesticides with CTD data and Tox21 testing, with both no inference and no activity), and unknown (pesticides which could not be linked to either the CTD or Tox21).

3. Results

3.1. Comparative toxicogenomics database integration

We foremost considered chemical-disease relationships for each of the CTD linked pesticides and idiopathic PD (59 targeted pesticides, 2 positive controls, and 8 negative controls). Of the targeted pesticides, nine pesticides had direct pesticide-PD evidence, meaning peer reviewed literature had directly linked the pesticide to PD: chlorpyrifos, permethrin, endosulfan, ethylene dibromide, methomyl, dicamba, omite (equivalent of propargite), phorate, and trifluralin (Supplemental Table 4). While both of the positive control pesticides (paraquat and rotenone) and none of the negative control pesticides had direct pesticide-PD evidence.

Although there was no direct epidemiologic evidence linking the other targeted pesticides to PD apart from our PEWAS, 35 of the pesticides were indirectly linked to PD through inferred relationships, signifying direct interaction between the pesticide and at least one PD gene.

The inference scores for each pesticide and PD are displayed in Fig. 1A, along with the number of references supporting the inference network. Both positive controls show, as one would expect, the highest inference scores (paraquat = 36.89; rotenone = 35.06). Of the targeted pesticides, chlorpyrifos has the highest inference score (Fig. 1A; inference score = 31.97). This indicates that the inference network for chlorpyrifos, e.g., all OMIM determined genes linked to PD that chlorpyrifos directly interacts with based on previous literature, is the most atypical and least likely to be due to chance. The chlorpyrifos-gene inference network involves 44 OMIM-defined PD linked genes (OMIM #168600) with 14 high priority genes (ParkinsonsUK-UCL); including *SNCA*, *LRRK2*, *MAPT*, *PARK7*, *PINK1*, and *PRKN* (Fig. 1B; Supplemental Table 4). For instance, chlorpyrifos has four links in the CTD to *SNCA* based on previous research reporting that chlorpyrifos increases phosphorylation of α-synuclein (SNCA protein)(Singh et al., 2018) and expression of SNCA mRNA (Slotkin and Seidler, 2011; Su and Niu, 2015). Chlorpyrifos is also directly

linked to PD in the CTD, but the previous epidemiologic evidence stems from the first wave of our own PEG study, meaning there is no independent confirmation from other human studies for our chlorpyrifos PEWAS result (Narayan et al., 2013; Gatto et al., 2009 Dec; Manthripragada et al., 2010). As another example of a pesticide with a lower inference score, methomyl has only been related to decreased expression of one PD gene, *HSPA1A* (Skandrani et al., 2006), and therefore received a much lower inference score of 3.54. According to NCBI, the *HSPA1A* protein product stabilizes proteins against aggregation and mediates the folding of newly translated proteins, processes highly relevant for PD α-synuclein pathology.

Overall, the mean inference score for the 35 targeted pesticides with inferred relationships was 8.08, with support from 541 references linking the pesticides to PD genes/gene products. The mean inference score for the positive control pesticides was 35.97 and for the negative control pesticides it was 2.69. After chlorpyrifos, inference scores were highest for diquat (inference score = 24.59; interacts with 12 PD genes), parathion (21.92; 9 PD genes), permethrin (21.56; 24 PD genes), and simazine (18.74; 7 PD genes). We provide the complete set of PD inference networks, i.e., all PD genes/gene products linked to the pesticides, the inference scores, and reference counts for each pesticide in Supplemental Table 4.

We next assessed the complete chemical-gene network for each pesticide, extracting all chemical-gene interactions from the CTD for the pesticides i.e., all genes/gene products that the pesticides interact with, not just PD genes. Overall, among the targeted pesticides, 57 of the 59 matched pesticides (96.%) were directly linked to at least one gene or gene product. Both positive controls (100 %) and 4 of the negative control pesticides (50 %) have reported interactions with at least one gene/gene product.

The complete chemical-gene network from the CTD for each pesticide in the PEWAS targeted list can be found in Supplemental Table 5. In total there were 36,382 chemical-gene interactions linking the pesticides to 15,471 different genes. However, most interactions were only supported by a single reference. When limiting to replicated interactions (Reference.Count > 1), there were 1260 chemical-gene interactions linking 37 pesticides to 747 genes. For example, “Mevinphos results in increased expression of NOS1 protein” was a chemical-gene interaction supported by seven references. The most supported individual interaction was “Chlorpyrifos results in decreased activity of ACHE protein”, which to date, has support from 67 references in 18 different organisms (Supplemental Table 5). This of course is not surprising as the organophosphate pesticides are designed to inhibit acetylcholinesterase, the ACHE protein. Fig. 2A shows the genes most involved in pesticide interactions as well as how many unique pesticides have been reported to interact with each gene. Each gene in Fig. 2A has been reported to interact with ≥ 9 of the pesticides of interest. Overall, based on previous literature, the genes/gene products interacting with the most pesticides were *ACHE*, which interacts with 31 of the PD-associated pesticides, *ESR1*, interacting with 28 pesticides, and *CAT*, interacting with 27 pesticides.

Fig. 2B is limited to interactions involving genes in the PD KEGG pathway (hsa05012) and the 59 targeted pesticides. Among the PD KEGG pathway genes, 16 targeted pesticides

of interest have shown interaction with *CASP3*, 15 with *TP53*, 13 with *SOD1*, 12 with *NFE2L2*, and 12 with *DDIT3*.

Fig. 2C details the reference counts for each pesticide- gene interaction by PD KEGG pathway gene. The reference counts are aggregated across all interaction types within a gene. For instance, *CASP3* is reported to interact with 16 of the targeted pesticides of interest. Those 16 pesticides can be seen in Fig. 2C, with the heatmap indicating the reference count supporting different interactions within the gene, ranging from 1 reference to 52 (chlorpyrifos-*CASP3*). Both positive control pesticides have reported interactions across a wide range of PD KEGG pathway genes, while two of the eight negative controls have reported interactions with two PD KEGG pathway genes (Fig. 2C). On average, among pesticides with reported chemical-gene interactions, the targeted pesticides interacted with 24.48 PD KEGG pathway genes, the positive controls with 136, and the negative controls with 2.5 PD KEGG pathway genes. The specific interactions for the targeted pesticides are detailed in Supplemental Table 5. Of note, if a single reference was used by the CTD to support multiple different interactions within a gene, this reference would be counted more than once in the aggregation. The individual references can be found via the CTD online database.

For the targeted pesticides, Fig. 3 maps the genes/gene products from the pesticide-gene interaction table to the KEGG Parkinson disease pathway (hsa05012). This figure highlights specific mechanisms in the Parkinson's disease pathway that have been most strongly linked to the pesticides of interest based on toxicogenomics. Twenty-five of the PEWAS-implicated pesticides have been linked to at least one gene/gene product in the Parkinson disease pathway (Supplemental Table 6). For instance, 18 of the pesticides have been implicated in the mitochondrial pathway, with 11 specifically interacting with genes/gene products related to mitochondrial complex I activity (e.g. NADH dehydrogenase related genes, including *NDUFC2*, *ND2*, *ND3*, etc.). Other mechanisms implicated across many of the targeted PD-associated pesticides include 20 pesticides influencing genes/gene products involved in reactive oxygen species handling (e.g. *TP53*, *SOD*, Nrf2/*NFE2L2*), 18 pesticides influencing apoptosis through mitochondria-dependent pathways (e.g. *CASP3* or *CASP9*), 14 pesticides influencing protein processing in the endoplasmic reticulum (e.g. BIP, *XBP1*), and 12 influencing endoplasmic reticulum-stress induced apoptosis (e.g. CHOP). Overall, nearly every gene/gene product in the KEGG Parkinson pathway has reportedly been influenced by at least one targeted pesticide per the CTD (e.g., 67 out of 69 gene product components detailed in Fig. 3 were linked to at least one pesticide).

To quantitatively assess enrichment of the pesticide-gene networks for biologic pathways of interest in PD etiology, we extracted the chemical-GO term relationships from the CTD. In total, 52 (88 %) of the targeted pesticides, all positive controls, and four (50 %) negative controls had chemical-gene networks significantly enriched for at least one of the 4500 PD GO terms. Fig. 4A shows enrichment for the 12 high-priority processes in PD according to the Parkinson's UK Annotation Initiative (see methods). These relationships link the pesticides indirectly to PD through highly relevant etiologic processes. For instance, 23 of the targeted pesticides were linked to response to oxidative stress (GO:0006979), while 21 were linked to regulation of neuron death (GO:1901214). Thirteen of the targeted

pesticides had chemical-gene networks significantly enriched for autophagy (GO:0006914), including copper sulfate which had the most statistically significant enrichment. Fig. 4B shows the network of all autophagy genes that copper sulfate has been linked with leading to enrichment. For instance, copper has been found to affect the expression of *LRRK2* (Min et al., 2009), which is linked to autophagy and lysosomal activity in neurons in PD. The full table showing the chemical-GO term relationships for the 4500 PD GO terms, along with the gene inference network linking the chemical to the GO term for the targeted pesticides can be found in Supplemental Table 7. This table provides detailed information that can be used to further prioritize agents for mechanistic research. For instance, 7 pesticides are linked through their chemical-gene networks to ubiquitination, another protein degradation pathway. Copper sulfate's gene network, for example, is also enriched for "protein K48-linked ubiquitination". PD involves protein misfolding/aggregation (Lewy bodies) and impairment in misfolded protein degradation. Polyubiquitination on the K48 and K29 lysines has been related to degradation by the proteasome, whereas polyubiquitination of other lysines (e.g. K63) does not seem to be related to protein degradation (Grice and Nathan, 2016). Thus, in prioritizing the 70 pesticides of interest here, these 7 pesticides and particularly copper sulfate would be interesting targets for protein degradation research in PD for both ubiquitination and autophagy. Supplemental Table 7 provides a means to prioritize the pesticide for other specific mechanistic studies of interests based on both gene ontology and gene target.

3.2. Tox21 qHTS assay performance

We used Tox21 assay performance to directly assess *in vitro* based activity related to the pesticides of interest. Activity (wAUC) on the nuclear receptor (NR) panel and stress response (SR) panels, accessed through Tox21 Activity Profiler, is displayed in Fig. 5 and Supplemental Table 8. Of the 70 targeted pesticides of interest, 43 were tested against 43 different assays. Among the positive control pesticides, only rotenone was included in the 10 k library, and for the negative controls, five of the eight were tested. Nine targeted pesticides and two negative control pesticides did not show conclusive activity on any of these assays, with six showing no activity (five targeted, 1,3-dichloropropene, dimethoate, ethephon, methamidophos, methomyl, and one negative control, pyrithiobac-sodium) and five only showing inconclusive results or no activity (four targeted, 2,4-D dimethylamine salt, 2,4-D isopropyl ester, disulfoton, and sodium chlorate, and one negative control, metalaxyl-m). The other 34 targeted pesticides showed activity across several processes of interest, such as mitochondrial membrane permeability, xenobiotic homeostasis, and genotoxic stress (Fig. 5).

After excluding inconclusive activity, primarily resulting from activity due to cytotoxicity, the most highly active targeted pesticides across assays were naled (activity on 12 different assays), endosulfan (11 assays), bromoxynil octanoate (10 assays), azinphos-methyl (9 assays), carbaryl (8 assays), and endothall (8 assays). Propargite, folpet, and endosulfan showed the strongest activity across multiple assays (wAUC $\geq$ 0.75 on 6 assays for propargite, 6 for folpet, and 3 for endosulfan). Some assays targeted processes of interest in PD etiology, including mitochondrial membrane permeability (pesticides resulting in activity include folpet, propargite, endosulfan, dinoseb, bromoxynil octanoate, dicloran, dicofol,

parathion, permethrin, chlorathal-dimethyl, diuron, carbaryl, and trifluralin), oxidative stress (positive with folpet, propargite, endosulfan, naled, carbaryl, malathion, endothall, and azinphosmethyl), and genotoxic stress (positive with folpet, endosulfan, naled, chlorpyrifos, bromoxynil octanoate, carbaryl, oryzalin, endothall, and dicofol). The heatmap (Fig. 5) displays the average activity on the assay across multiple tests, the full table of all activities can be found in Supplemental Table 9.

Among the positive control pesticides, paraquat was not included in the Tox21 library and thus not tested, but rotenone showed strong activity (wAUC ≥ 0.75) on 7 assays. Interestingly, among the negative controls two pesticides (cyprodinil and tebuconazole) also showed activity on several assays, while the other negative controls (3/5) were primarily inactive or inconclusive.

Overall, the targeted set of PD-associated pesticides in the 10 k library (43 pesticides) was enriched for activity on several different assays, as more of these pesticides showed activity on the assay than expected based on comparing to both the activity of (1) all chemicals assessed in the 10 k library and (2) pesticides only. The most enriched assays were tox21-car-agonist-p1/xenobiotic homeostasis ($p = 9.4e-05$ compared to all 10 k chemicals and $p = 0.001$ compared to pesticides only), tox21-mitotox-antagonist-p1/mitochondrial membrane permeability ($p = 0.008$ with all 10 k chemicals; $p = 0.008$ pesticides only), tox21-rar-antagonist-p2/retinoid homeostasis ($p = 0.03$ all 10 k chemicals; $p = 0.02$ pesticides only), and tox21-h2ax-cho-agonist-p2/genotoxic stress ($p = 0.04$ all 10 k chemicals; $p = 0.01$ pesticides only). The enrichment results are shown in Supplemental Table 10.

Eleven of the targeted pesticides were also cytotoxic to either the HEK293 line or HepG2 based on at least one of the two Tox21 cytotoxicity assays (Supplemental Tables 12 and 13). After excluding results with a low purity, oryzalin, methomyl, and folpet were toxic to the HepG2 hepatocellular carcinoma cells, while oryzalin, dinoseb, folpet, fluazifop-butyl, chlorpyrifos, azinphos-methyl, propargite, endothall, trifluralin, and bromoxynil octanoate were toxic to the HEK293 embryonic kidney cells over at least one of the time intervals (0–40h after exposure). Furthermore, several other pesticides were inconclusive, that is they showed a mid-range effect that was not inactive or conclusively active, including methidathion, chlorthal-dimethly, endothall, azinphos-methyl, ethephon, dinoseb, prometyrn, phorate, chlorpyrifos, propargite, endothall, tribufos (HepG2) and chlorthal-dimethly, tribufos, carbaryl, iprodione, dinoseb, prometryn, diuron, dicloran, propargite, ethephon, dimethoate, methamidophos, propyzamide (HEK293).

Of the positive controls, rotenone was actively toxic to both cell lines, while none of the negative control pesticides were toxic to the HepG2 line and cyprodinil and tebuconazole showed some cytotoxicity to the HEK293 line.

3.3. CTD and Tox21

Fig. 6 displays gene targets implicated via both Tox21 and the CTD for the most active targeted pesticides. The 25 genes shown are gene targets from the Tox21 assays. Many of the pesticides were linked to the same target products in both Tox21 and the CTD. For instance, of the targeted pesticides, endosulfan induced oxidative stress activity (tox21-are-

bla-agonist-p1) via action on NFE2L2, had a chemical-gene network enriched for “response to oxidative stress” GO term (adj p-value = 2.98E-57), and was linked specifically to *NFE2L2* in the CTD (“Endosulfan results in increased expression of NFE2L2 mRNA”). Malathion also resulted in activity on the oxidative stress Tox21 assay and was linked to increased activity of NFE2L2 via CTD. Tribufos (aka s,s,s-tributyl phosphorotrithioate) was linked to NR1I3 (xenobiotic homeostasis) and chlorpyrifos, carbaryl, diuron, iprodione, malathion, and parathion were all linked to AHR (xenobiotic homeostasis) both via Tox21 and CTD. Surprisingly, however, several pesticides showed a wide range of activity on the Tox21 assays, but they have not been linked to the target previously in the literature according to the CTD. These pesticides include naled, which showed activity on 12 different Tox21 assays, but has not been linked to any of the gene products in the CTD, and bromoxynil octanoate, which showed activity on 10 Tox21 assays but again was not related to any of the genes in the CTD (Supplemental Table 5). Chlorthal-dimethyl (aka DCPA), dinoseb, mepiquat, and methidathion also showed activity on a range of Tox21 assays but have not previously been linked to any gene targets. Propargite and folpet both showed very strong activity (high wAUC) on varying assays (e.g., mitochondrial membrane permeability, oxidative stress, sex hormone homeostasis, genotoxic stress), but were not widely linked (via reference count) to any targets in CTD.

Supplemental Fig. 1 shows this information for all 43 targeted pesticides tested against the Tox21 assays and Supplemental Fig. 2 shows the six control pesticides (rotenone as positive control and five negative control pesticides) tested against the Tox21 assays.

Fig. 7 shows the summary profile for each PEWAS targeted pesticide based on the original PEWAS association, cytotoxicity in the two human cell lines, the CTD inference score between the pesticides and PD, and the Tox21 ToxScore. Several pesticides scored on all three measures (Tox21 cytotoxicity, Tox21 ToxScore, CTD inference), including chlorpyrifos, dinoseb, and folpet. Whereas several other pesticides were only implicated through toxicity or toxicogenomics, or neither. For instance, propargite had the highest Tox21 ToxScore, showing activity on assays highly relevant to PD etiology, such as impairing mitochondrial membrane permeability, but has not been linked to any PD genes in the CTD.

4. Discussion

Pesticides are among the environmental risk factors most consistently associated with PD, as demonstrated in many epidemiologic investigations (Kamel, 2013; Goldman et al., 2017). However, given the needs of commercial agriculture for variety in agents that are tailored to emerging and reemerging pests and the large number of pesticides currently in use, i.e., 1074 active ingredients alone registered in California, epidemiologic studies and experimental investigations have not been able to fully evaluate long-term, low dose exposure related health effects for most agents. Complex experimental work on model organisms for PD has also necessarily been limited to addressing effects of the most common and likely suspects in use (Liu et al., 2020). Consequently, most of the pesticide active ingredients used in agriculture have to date not been assessed for their potential influence on PD etiology.

Here, we show that an agnostic strategy to screen pesticides for a health effect of interest (PD) can be further informed with existing toxicology and toxicogenomic evidence to prioritize agents for further in-depth experimentation or targeted epidemiologic studies. Our first step, a population-based pesticide-wide association study investigating hundreds of specific pesticide agents for association with PD, implicated 70 targets as ostensibly associated with PD. Here, we further investigated these 70 pesticides of interest with toxicologic and toxicogenomic databases allowing us to assess biologic links to PD. This agnostic approach combined pesticide association screening in a population-based study with readily available experimental data, including toxicogenomics data from the CTD and high-throughput *in vitro* results from Tox21.

We queried the CTD first for prior epidemiologic support that linked each of the targeted pesticides directly to PD and found such evidence for nine pesticides implicated in our screen. This epidemiologic evidence, however, came from only two studies, our California PEG studies (chlorpyrifos, endosulfan, methomyl, omite/propargite) and the Agricultural Health Study (AHS) (permethrin, ethylene dibromide, dicamba, phorate, and trifluralin (Kamel et al., 2007; Furlong et al., 2015; Shrestha et al., 2020). This demonstrates the need to develop and implement feasible exposure assessment approaches, such as record-based exposure assessments that we employed in the PEG studies, for future epidemiologic investigations of specific pesticide health effects. Connecting specific pesticides to PD is among the most pressing needs in the field (Kamel, 2013), yet very few epidemiologic studies to date are equipped to investigate most of the various pesticides currently used in the agricultural and other industries.

The AHS is a longitudinal study of pesticide applicators and their spouses residing in Iowa and North Carolina (n = 38,274 pesticide applicators and 27,836 of their spouses), among whom 373 male applicators (0.97 %) and 118 female spouses (0.42 %) developed PD during follow-up (Kamel et al., 2007; Furlong et al., 2015; Shrestha et al., 2020). This study has multiple publications, and most recently reported on 50 pesticides assessed through self-reports at baseline and over follow-up. The AHS investigated 14 of the 70 pesticides implicated via our PUR-GIS approach and reported positive associations with PD for four: permethrin, dicamba, phorate, and trifluralin (Skandrani et al., 2006; Min et al., 2009; Grice and Nathan, 2016). The AHS also reported an increased risk for PD with 2,4-D but did not investigate the ester (2,4-D, isopropyl ester), which was associated with PD in our study. The CTD therefore did not return a direct evidence match as it did not equate 2,4-D with 2,4-D, isopropyl ester. Furthermore, AHS also implicated terbufos and 2,4,5-T as associated with PD (Shrestha et al., 2020). Neither of these pesticides were applied in sufficient quantity in PEG study areas and, thus, could not be investigated as part of our approach. Interestingly, results from a Dutch case-control study (n = 352 PD patients and 607 hospital-based controls) which investigated 157 specific pesticides (including 21 of the 70 pesticides implicated through our PEWAS approach) have not been included in the CTD manual curation. Perhaps, because exposure was based on residential proximity to presumed pesticide applications (Brouwer et al., 2017). Nevertheless, the Dutch study reported 21 pesticides in total were associated with an increased risk of PD, including fluazifop-butyl, MCPA, 1,3-dichloropropene (1,3-D), oxamyl, and aldicarb which were also associated with PD in our PEWAS analysis. Overall, however, most of the 70 PD pesticides implicated in

PEG had limited prior epidemiologic support, with no previous research or direct evidence of association listed in the CTD.

Thus, it is especially important that toxicogenomic analyses revealed pesticide-gene interaction networks for most of the targeted pesticides, pointing to effects on gene expression, protein phosphorylation, or protein activity for a variety of PD genes/gene products. Prior evidence allowed us to link many of the pesticides directly with genes or gene products like *SNCA*, *LRRK2*, *MAPT*, *PARK7*, *PINK1*, and *PRKN*. Over half of the PEWAS-implicated pesticides (34 out of 58 uniquely matched in CTD) had an inferred relationship with PD through direct interaction with at least one PD gene. In fact, when we investigated the pesticide-gene ontology (GO) relationships from the CTD through enrichment analysis of the gene networks, focusing on processes with high-priority for PD etiology according to the Parkinson's UK Annotation Initiative, 24 pesticides had chemical-gene networks enriched for at least one of the high-priority terms, such as autophagy, mitophagy, and protein folding/endoplasmic reticulum stress. While these pesticides need to be further investigated in PD-specific experimental models, the toxicogenomic output we generated helps to rank agents as likely targets for specific mechanistic experiments. For instance, we found that the chemical-gene network for copper sulfate is strongly enriched for multiple protein degradation pathways, including autophagy and protein K48-linked ubiquitination. While organic copper (Cu^{+}) from food sources is generally processed by the liver and transported/sequestered safely, inorganic copper (Cu^{2+}), such as in copper sulfate pesticides, can bypass the liver and enter the blood supply as 'free copper' where it is able to penetrate the blood–brain barrier (Brewer, 2009). As a result, copper sulfate would be an interesting target for PD mechanistic models assessing neuronal protein degradation impairment in PD. Furthermore, seven pesticides (copper sulfate, cacodylic acid, sodium arsenate/DSMA, chlorpyrifos, endosulfan, monomethylarsonic acid (MSMA), and permethrin), have been shown to interact with *LRKK2*, mostly via their influence on mRNA expression, and thus may be good targets for LRKK2-kinase and protein degradation focused research. Notably, this list includes four heavy metal-based pesticides, including copper sulfate and three arsenic-based pesticides, two of which are now banned in the United States (cacodylic acid and sodium arsenate/DSMA).

Beyond links through toxicogenomics, many of the targeted pesticides of interest elicited activity on a range of Tox21 assays, including cytotoxicity, xenobiotic homeostasis, genotoxic stress, and sex hormone homeostasis. While only 43 of the PEWAS PD-associated pesticides were included in the Tox21 10 k library and tested, 34 (79 %) showed activity on at least one assay. Rotenone also showed a range of strong activity across multiple assays. While two of the eight negative controls also showed activity on a range a Tox21 assays. This it is not surprising that even these negative controls would show some level of activity, as pesticides by design influence living systems. Furthermore, these assays are not specific to PD or neurodegenerative systems, and are not tested in neuronal cell lines including dopaminergic cells, which are selectively influenced in PD.

Still, eleven targeted pesticides were cytotoxic to either the HepG2 hepatocellular carcinoma cells or the HEK293 embryonic kidney cells. Interestingly, several of these pesticides were also found to be toxic to dopaminergic neurons in the previously mentioned testing of our

PEWAS pesticide hits in a PD-patient derived iPSC stem cell line (Paul et al., 2022). These include trifluralin, folpet, endothall, and propargite. However, these same experiments also identified several other pesticides as toxic to midbrain dopaminergic neurons which were not conclusively toxic to the other Tox21 cell lines, including naled, dicofol, and endosulfan. Still, cytotoxicity is an extreme endpoint. It may be just as if not more important to closely consider the pesticides that showed activity on assays targeting processes of particular interest as part of PD etiology. For instance, thirteen (30 %) that impaired mitochondrial membrane permeability and eight (19 %) that induced oxidative stress. While the Tox21 assays are not specific to neurodegeneration or PD, the results do provide some support for interaction with pertinent biologic mechanisms and provide a proof-of-concept that these pesticides may affect key systems in PD. Interestingly, many of the pesticides showed activity on pathways and targets that were not reported in the CTD. For instance, naled, which was toxic to the dopaminergic neurons in prior testing, showed a wide range of activity across 10 Tox21 assays, inducing genotoxic and oxidative stress and disrupting lipid and sex hormone homeostasis. However, there is very limited previous research related to this pesticide, with only nine references and 5 links to gene products in the CTD. Similarly, bromoxynil octanoate, chlorthal-dimethyl (aka DCPA), dinoseb, fluazifop-butyl, and methidathion also showed activity on a range of Tox21 assays but the data provided in CTD is limited.

Our PEWAS approach also implicated trifluralin, which is clearly a compelling target, as it has now been associated with PD in two independent epidemiologic studies with over 1300 PD patients (Paul et al., 2022; Shrestha et al., 2020). Trifluralin also emerged as toxic to the dopaminergic neurons in the follow-up iPSC cell testing and mechanistically was determined to disrupt mitochondrial respiration in the neurons (Paul et al., 2022). While several lines of evidence now converge for trifluralin and PD, trifluralin only has 28 references linking it to health impacts or genes in the CTD. One-third of the references are from the Agricultural Health Study, the rest linking it to 18 genes, none of which are thought to be related to PD (for comparison, chlorpyrifos has 594 references in the CTD, endosulfan has 188). Thus, although trifluralin is a widely used pesticide and now has been linked to several diseases including PD, glioma, and other neoplasms (Furlong et al., 2015; Shrestha et al., 2020; Hoppin et al., 2002; Kang et al., 2008; Carreón et al., 2005), it is still not well studied experimentally. Naled, bromoxynil octanoate, propargite, and several other pesticides share similar narratives. Our agnostic approach that linked epidemiology to existing toxicogenomic databases, however, was able to highlight specific agents that have gone relatively unnoticed and have not been studied yet for their neurotoxic potential after chronic exposure.

This also highlights a limitation of relying on data mining approaches and a caveat for interpreting the CTD chemical-disease-gene networks. While the CTD provides a valuable tool to synthesize findings from nearly 150,000 manuscripts, it is still limited to manual curation of published literature. It can be used to target agents or pathways based on *known* information. This means agents which have not been prioritized for research will not rise to the surface through CTD data mining. Among the targeted PD pesticides of interest here, the CTD inference score for PD was highest for chlorpyrifos and lowest for dicofol. However, there is a large disparity in the amount of research that has been undertaken

for chlorpyrifos (again 594 references in the CTD) versus dicofol (34 references). Thus, it is certainly possible that dicofol influences more than just one PD gene/gene product (increased expression of CYP2E1 protein (Chan et al., 2009)), but research has yet to be conducted to reveal such an influence. Importantly, dicofol was not only toxic to the dopaminergic neurons (Paul et al., 2022), but also showed high levels of activity on multiple Tox21 assays, including strongly impairing mitochondrial membrane permeability, a known mechanism related to dopaminergic neuron loss. It is likely therefore, that the CTD networks only provide the tip of the iceberg for exposure-gene links, suggesting a need for more comprehensive screening approaches to help understand pesticide health effects.

Tox21 to that end provides a proof-of-principal for high throughput, automated screening platforms which can be used as an additional tool to agnostically assess large amounts of chemicals. NIEHS/NTP, EPA, and NCATS have gone to great lengths to develop the Tox21 10 k library, robotic qHTS, and a system of predictive toxicology. Currently Tox21 uses a 1536-well microtiter plate format, allowing testing of 1408 samples and concurrent negative and positive controls on each plate. Each plate tests a different concentration. Processing 100 plates per day, the qHTS robots can test up to a million samples per week (NIEHS/NTP. Tox21 Research Phases FAQs). While there are still challenges and limitations to interpreting results from these model systems and assays, our results are encouraging as we were able to inform epidemiologic results with biologic understanding of disease pathways of importance in PD. We recommend further developing this system to screen every pesticide formulation upon registration with the US EPA at the expense to the pesticide manufacturer as a step towards ‘pesticidovigilance’ (Milner and Boyd, 2017). That is, long-term safety monitoring of pesticide products, post-regulation, to assess health and environmental effects when used at large scale, similar to pharmacovigilance (Milner and Boyd, 2017). If not yet fit for regulatory consideration, these high-throughput assays provide at least some information on potential downstream health effects and provide an agnostic screening approach that can help prioritize agents for low throughput and more resource intensive toxicologic testing which currently feeds risk assessment meant to improve public health.

5. Conclusions

Overall, our comprehensive, pesticide-wide association study and toxicologic/toxicogenomic integration has implicated a variety of individual pesticides in PD risk. Cross-database queries we conducted and described here helped us connect pesticides to specific biologic and mechanistic targets relevant to PD etiology. This knowledge can be used to prioritize experimental pesticide research and improve our understanding of the role of pesticides in PD etiology. With epidemiologic evidence linking common, currently used pesticides to PD and support from both toxicogenomic and high-throughput toxicologic assays, it is time to design additional investigations based on PD-specific models. Furthermore, our agnostic screening approach for pesticide exposure effects indicates that research in general should target a broader range of pesticides as possibly related to PD than those traditionally studied in the past. The approach we described here can be employed for other pesticides and health outcomes and is promising as a tool to help prioritize research targets such as specific agents and mechanisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The Michael J. Fox Foundation for Parkinson's Research (MJFF-001018), National Institute of Environmental Health Science (Grant No 2R01ES010544, R56ES026600), Department of Defense Parkinson's Disease Research Program (W81XWH1910696).

Data availability

The data used are public data sources, free to download from Tox21 and CTD. We have provided a complete codebook and explanation on how to obtain data.

References

NIEHS/NTP. Tox21 Activity Profiler. Available from: <https://sandbox.ntp.niehs.nih.gov/tox21-activity-browser/>.

Brewer GJ, 2009. The risks of copper toxicity contributing to cognitive decline in the aging population and to alzheimer's disease. J. Am. Coll. Nutr 28 (3).

Brouwer M, Kromhout H, Vermeulen R, Duyzer J, Kramer H, Hazeu G, et al. , 2017. Assessment of residential environmental exposure to pesticides from agricultural fields in the Netherlands. J. Exposure Sci. Environ. Epidemiol

California Department of Pesticide Regulation. Actively Registered Active Ingredients (AI) by Common Name [Internet]. [cited 2021 Sep 6]. Available from: <https://www.cdpr.ca.gov/docs/label/actai.htm>.

Carreón T, Butler MA, Ruder AM, Waters MA, Davis-King KE, Calvert GM, et al. , 2005. Gliomas and farm pesticide exposure in women: The Upper Midwest Health Study. Environ Health Perspect. 113 (5).

Chan WH, Liao JW, Chou CP, Chan PK, Wei CF, Ueng TH, 2009. Induction of CYP1A1, 2B, 2E1 and 3A in rat liver by organochlorine pesticide dicofol. Toxicol Lett. 190 (2).

CTD. Chemical Enriched GO Terms. Available from: <http://ctdbase.org/help/chemGODetailHelp.jsp>.

Davis AP, Murphy CG, Rosenstein MC, Wiegers TC, Mattingly CJ, 2008. The Comparative Toxicogenomics Database facilitates identification and understanding of chemical-gene-disease associations: arsenic as a case study. BMC Med. Genomics 1 (1).

Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegers J, Wiegers TC, et al. , 2021. Comparative toxicogenomics database (CTD): update 2021. Nucleic Acids Res. 49(D1).

EPA CompTox. EPA CompTox Dashboard. Available from: <https://comptox.epa.gov/dashboard/search-results?input_type=consumer_product&inputs=CPCat%20Use%20Category%3APesticides>.

Furlong M, Tanner CM, Goldman SM, Bhudhikanok GS, Blair A, Chade A, et al. , 2015. Protective glove use and hygiene habits modify the associations of specific pesticides with Parkinson's disease. Environ Int. 75.

Gatto NM, Cockburn M, Bronstein J, Manthripragada AD, Ritz B, 2009. Well-water consumption and Parkinson's disease in rural California. Environ Health Perspect. 117 (12), 1912–1918. [PubMed: 20049211]

Goldman SM, Musgrove RE, Jewell SA, di Monte DA, 2017. Pesticides and Parkinson's Disease: Current Experimental and Epidemiological Evidence.

Grice GL, Nathan JA, 2016. The recognition of ubiquitinated proteins by the proteasome. Cell. Mol. Life Sci 73.

Hernandez-Ferrer C, Gonzalez JR, 2018. CTDquerier: A bioconductor R package for comparative toxicogenomics database™ data extraction, visualization and enrichment of environmental and toxicological studies. Bioinformatics 34 (18).

Hoppin JA, Umbach DM, London SJ, Alavanja MCR, Sandler DP, 2002. Chemical predictors of wheeze among farmer pesticide applicators in the Agricultural Health Study. Am. J. Respir. Crit. Care Med 165 (5).

Hsieh JH, Huang R, Lin JA, Sedykh A, Zhao J, Tice RR, et al. , 2017. Real-Time cell toxicity profiling of Tox21 10K compounds reveals cytotoxicity dependent toxicity pathway linkage. PLoS One 12 (5).

Huang R, 2016. A quantitative high-throughput screening data analysis pipeline for activity profiling. Methods Mol. Biol 1473.

Huang R, Xia M, Sakamuru S, Zhao J, Shahane SA, Attene-Ramos M, et al. , 2016. Modelling the Tox21 10 K chemical profiles for in vivo toxicity prediction and mechanism characterization. Nat. Commun 7.

Hur J, Danes L, Hsieh JH, McGregor B, Krout D, Auerbach S, 2018. Tox21 Enricher: Web-based Chemical/Biological Functional Annotation Analysis Tool Based on Tox21 Toxicity Screening Platform. Vol. 37, Molecular Informatics.

Kamel F, 2013. Paths from pesticides to Parkinson's. Science 341.

Kamel F, Tanner CM, Umbach DM, Hoppin JA, Alavanja MCR, Blair A, et al. , 2007. Pesticide exposure and self-reported Parkinson's disease in the agricultural health study. Am. J. Epidemiol 165 (4).

Kang D, Park SK, Beane-Freeman L, Lynch CF, Knott CE, Sandler DP, et al. , 2008. Cancer incidence among pesticide applicators exposed to trifluralin in the Agricultural Health Study. Environ Res. 107 (2).

Kavlock RJ, Austin CP, Tice RR, 2009. Toxicity testing in the 21st century: Implications for human health risk assessment. Risk Anal. 29.

King BL, Davis AP, Rosenstein MC, Wiegers TC, Mattingly CJ, 2012. Ranking transitive chemical-disease inferences using local network topology in the comparative toxicogenomics database. PLoS One. 7 (11).

Liu C, Liu Z, Zhang Z, Li Y, Fang R, Li F, et al. , 2020. A scientometric analysis and visualization of research on Parkinson's disease associated with pesticide exposure. Front Public Health.

Manthripragada AD, Costello S, Cockburn MG, Bronstein JM, Ritz B, 2010. Paraoxonase 1, agricultural organophosphate exposure, and Parkinson disease. Epidemiology [Internet]. 21(1):87–94. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3117899&tool=pmcentrez&rendertype=abstract>. [PubMed: 19907334]

Milner AM, Boyd IL, 2017. Toward pesticidovigilance. Science (1979) [Internet]. 357 (6357), 1232–1234. Available from: <http://www.sciencemag.org/lookup/doi/10.1126/science.aan2683>.

Min OS, Li J, Freedman JH, 2009. Physiological and toxicological transcriptome changes in HepG2 cells exposed to copper. Physiol. Genomics 38 (3).

Narayan S, Liew Z, Paul K, Lee PC, Sinsheimer JS, Bronstein JM, et al. , 2013. Household organophosphorus pesticide use and parkinson's disease. Int. J. Epidemiol 42 (5).

NIEHS/NTP. Tox21 Research Phases FAQs. Available from: <https://ntp.niehs.nih.gov/whatwestudy/tox21/phases/index.html>.

PAN North America. Available from: <Pesticide Info. https://www.pesticideinfo.org/>.

ParkinsonsUK-UCL. Parkinson's UK Annotation Initiative. Available from: <https://www.ucl.ac.uk/cardiovascular/research/pre-clinical-and-fundamental-science/functional-gene-annotation/neurological-gene>.

Paul KC, Krolewski RC, Moreno EL, Blank J, Holton KM, Ahfeldt TD, et al. , 2022. Coupling comprehensive pesticide-wide association study to iPSC dopaminergic screening identifies and classifies Parkinson-relevant pesticides. bioRxiv.

Richard AM, Huang R, Waidyanatha S, Shinn P, Collins BJ, Thillainadarajah I, et al. , 2021. The Tox21 10K compound library: collaborative chemistry advancing toxicology. Chem. Res. Toxicol 34.

Ritz BR, Paul KC, Bronstein JM, 2016. of pesticides and men: a california story of genes and environment in Parkinson's disease. Curr. Environ. Health Rep 3 (1).

Shrestha S, Parks CG, Umbach DM, Richards-Barber M, Hofmann JN, Chen H, et al. , 2020. Pesticide use and incident Parkinson's disease in a cohort of farmers and their spouses. Environ Res. 191.

Shukla SJ, Huang R, Austin CP, Xia M, 2010. The future of toxicity testing: a focus on in vitro methods using a quantitative high-throughput screening platform. Drug Discovery Today 15.

Singh N, Lawana V, Luo J, Phong P, Abdalla A, Palanisamy B, et al. , 2018. Organophosphate pesticide chlorpyrifos impairs STAT1 signaling to induce dopaminergic neurotoxicity: Implications for mitochondria mediated oxidative stress signaling events. Neurobiol Dis. 117.

Skandrani D, Gaubin Y, Vincent C, Beau B, Claude Murat J, Soleilhavoup JP, et al. , 2006. Relationship between toxicity of selected insecticides and expression of stress proteins (HSP, GRP) in cultured human cells: Effects of commercial formulations versus pure active molecules. Biochim. Biophys. Acta Gen Subj 1760 (1).

Slotkin TA, Seidler FJ, 2011. Developmental exposure to organophosphates triggers transcriptional changes in genes associated with Parkinson's disease in vitro and in vivo. Brain Res Bull. 86(5–6).

Su C, Niu P, 2015. Low doses of single or combined agrichemicals induces α-synuclein aggregation in nigrostriatal system of mice through inhibition of proteasomal and autophagic pathways. Int. J Clin. Exp. Med 8 (11).

Tice RR, Austin CP, Kavlock RJ, Bucher JR, 2013. Improving the human hazard characterization of chemicals: a Tox21 update. Environ. Health Perspect 121.

NIEHS/NTP. Tox21 Toolbox. Available from: <https://ntp.niehs.nih.gov/whatwestudy/tox21/toolbox/index.html>.

Topping CJ, Aldrich A, Berny P, 2020. Overhaul environmental risk assessment for pesticides. Science 367.

Tox21 Public Data Repository. Available from: <https://tripod.nih.gov//tox21/pubdata/>. 2022.

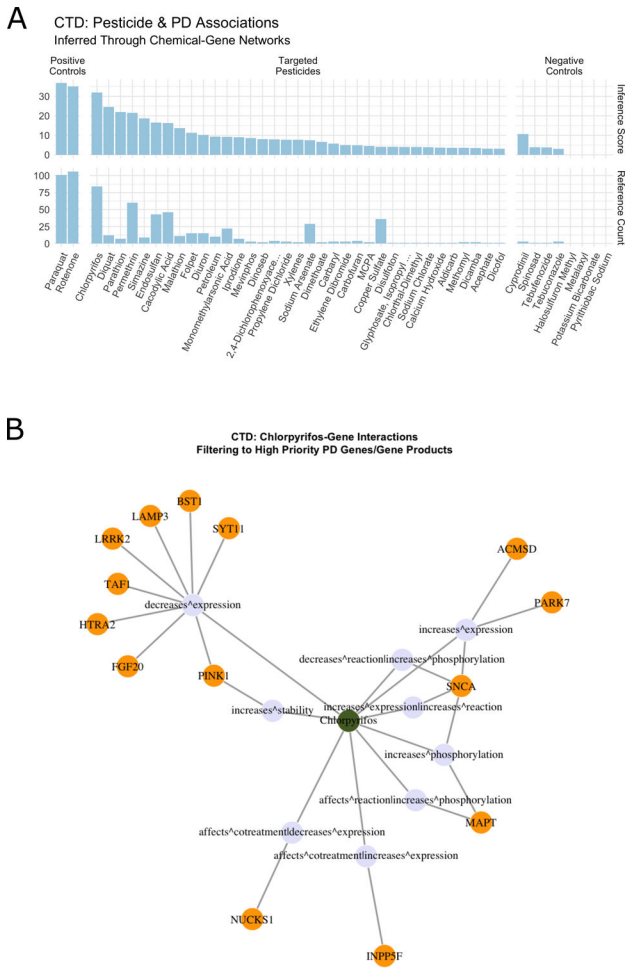

Fig. 1.

A. Inference score and reference count from the CTD linking each pesticide to PD through the pesticide-gene networks. Three sets of pesticides are shows: pesticides selected as positive controls, the targeted pesticide set, and a set of negative control pesticides. **B.** The pesticide-gene network for chlorpyrifos from the CTD, limiting to PD genes/gene products.

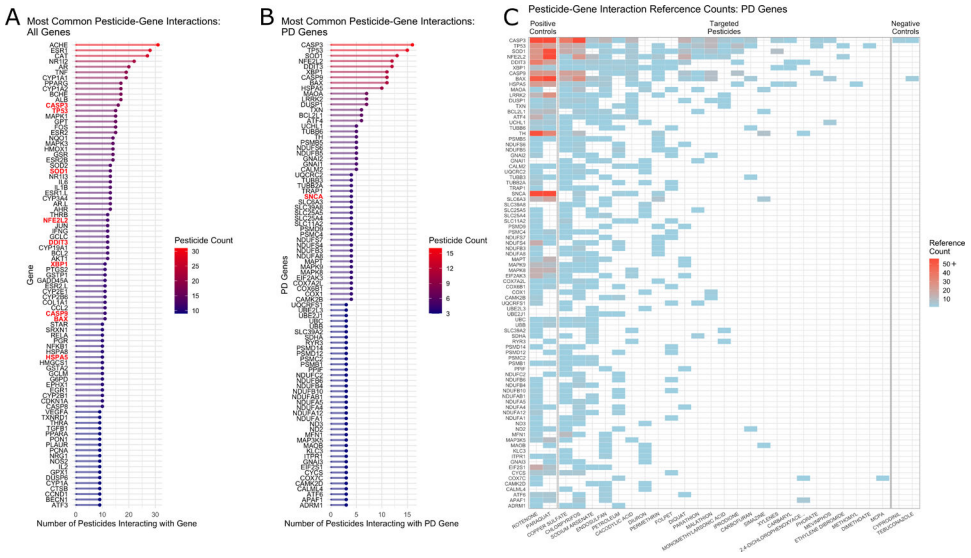

Fig. 2.

A. The most common genes involved in pesticide interactions with the targeted set of pesticides based on the pesticide-gene interactions in the CTD. Figure shows the number of unique pesticides each gene has been reported to interact. **B.** The most common genes involved in pesticide interactions with the targeted set of pesticides based on the pesticide-gene interactions in the CTD, limiting to KEGG PD pathway genes. **C.** Heatmap indicating the reference count supporting the different pesticide-gene interactions for KEGG PD pathway genes. The reference counts are aggregated across all interaction types within a gene. Reference counts, therefore, may not be unique, if the same reference is indicated in the CTD for multiple unique interactions within the gene. References can be accessed through query of the CTD database online tool. We did not display-six negative controls in the heatmap as they were unrelated to all genes.

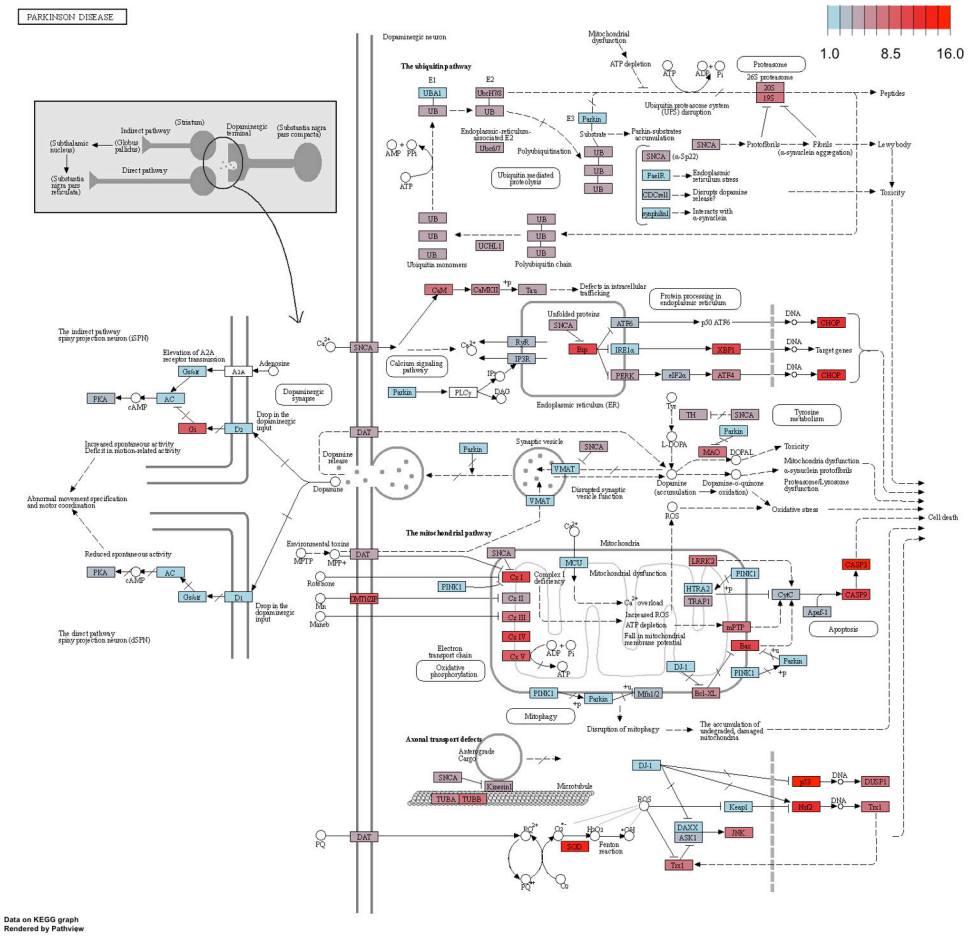

Fig. 3.

Map of the KEGG Parkinson disease pathway (hsa05012), indicating pesticide-gene interactions, with the color indicating the number of pesticides in the targeted pesticide set shown to interact with different genes/gene products in the pathways, based on the CTD.

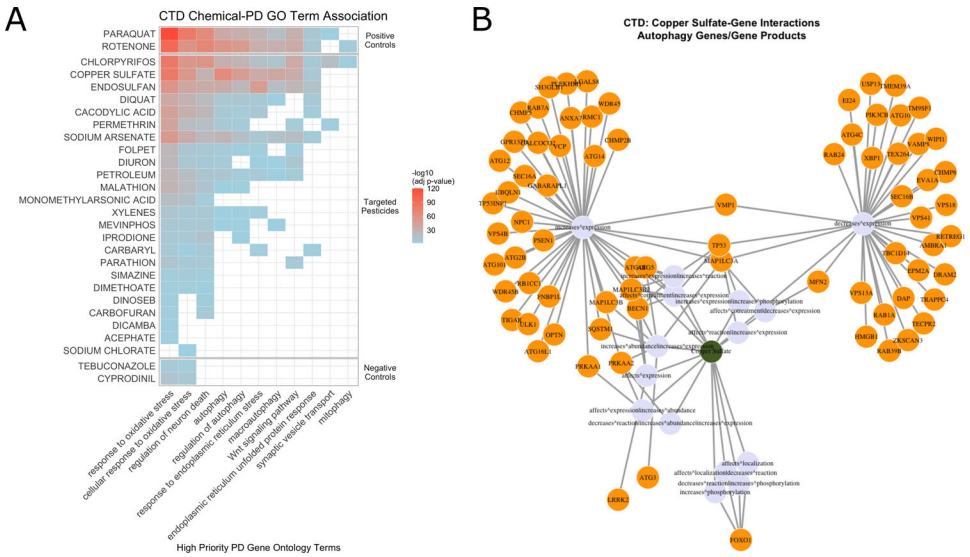

Fig. 4.

A. Heatmap indicating pesticides with chemical-gene networks significantly enriched high-priority PD GO terms, based on the CTD. **B.** The network of all the autophagy genes that copper sulfate has been shown to interact with leading to significant enrichment. We did not display-six negative controls in the heatmap as they were unrelated to all pathways.

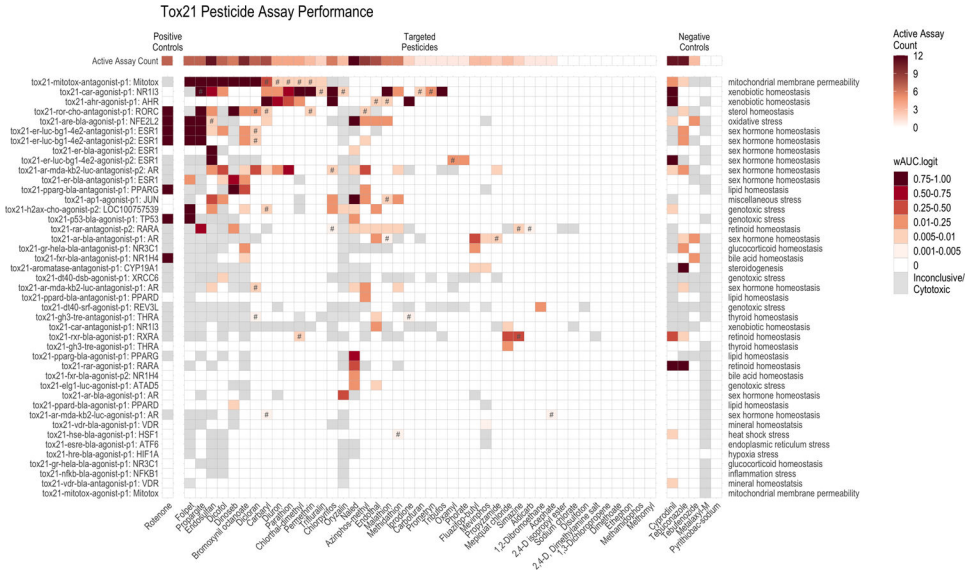

Fig. 5.

Heatmap indicating activity on the Tox21 assays due to the pesticides of interest. The # indicates variability between sources, the heatmap displays the mean activity across screens, however the individual activities across all sources are listed in Supplemental Table 8. Excludes activity from chemicals with suboptimal QC grade, excludes activity (antagonist-type calls) due to cytotoxicity, excludes activity due to auto-fluorescencent, excludes activity with no reporter gene activity readout support. # indicates results varied between source.

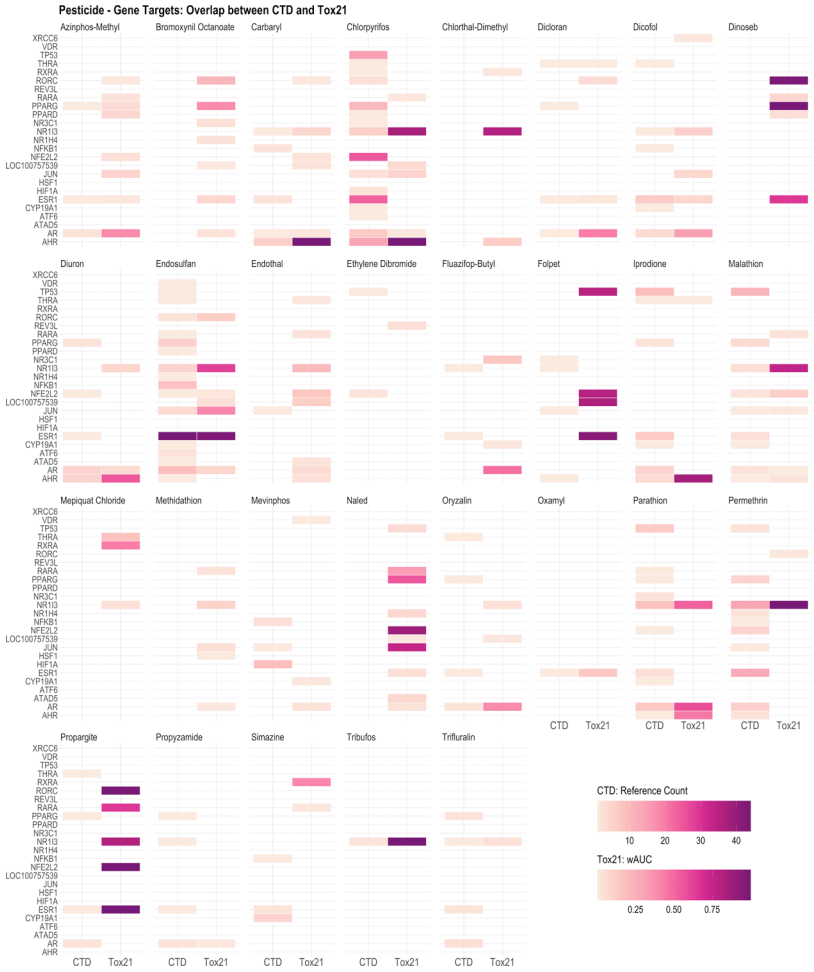

Fig. 6.

Overlap between the Tox21 and CTD, indicating if a pesticide showed activity on the target (Tox21) or interaction with the gene (CTD). If different Tox21 assays tested for activity on the same gene targets (e.g., 5 different assays with ESR1 gene target), the highest activity level is shown.

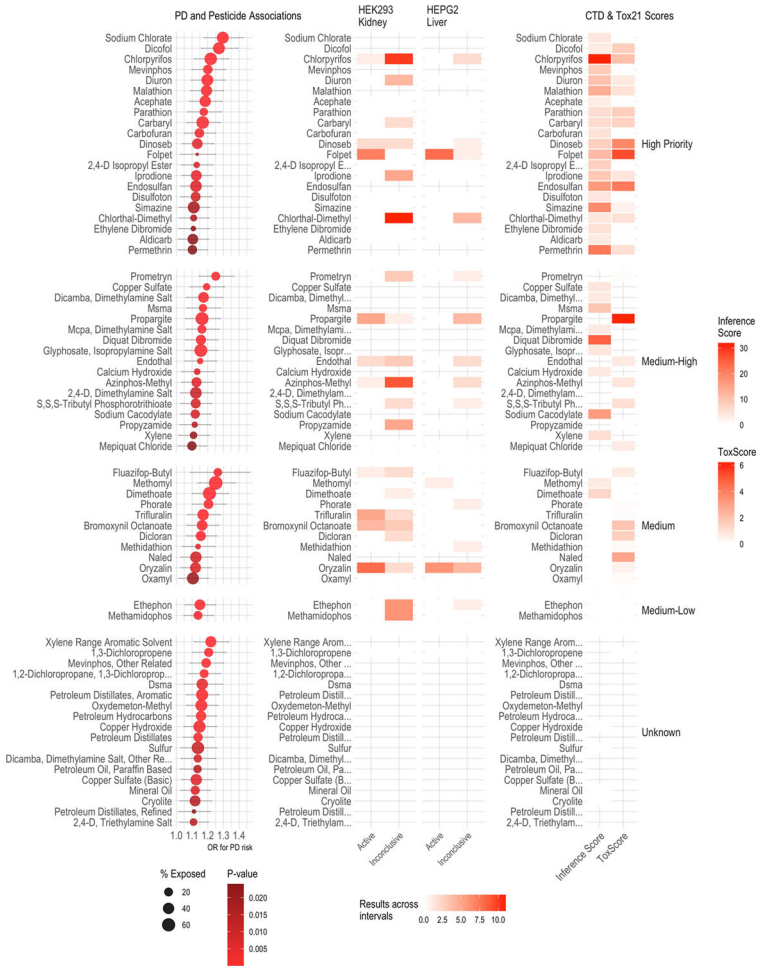

Fig. 7.

Summary profile for the targeted set of pesticides and PD, based on the PEWAS association, Tox21 cytotoxicity, CTD inference score, and Tox21 SR/NR panel ToxScore. The cytotoxicity outcomes (active, inconclusive) are aggregated across the intervals of exposure (6intervalsfrom0 to 40 h), so a higher count indicates more of the respective outcomes was measured at more intervals. For the Tox21 heatmaps, a box color is not shown (e.g., transparent) for pesticides that were not assessed or had low purity data that was excluded.