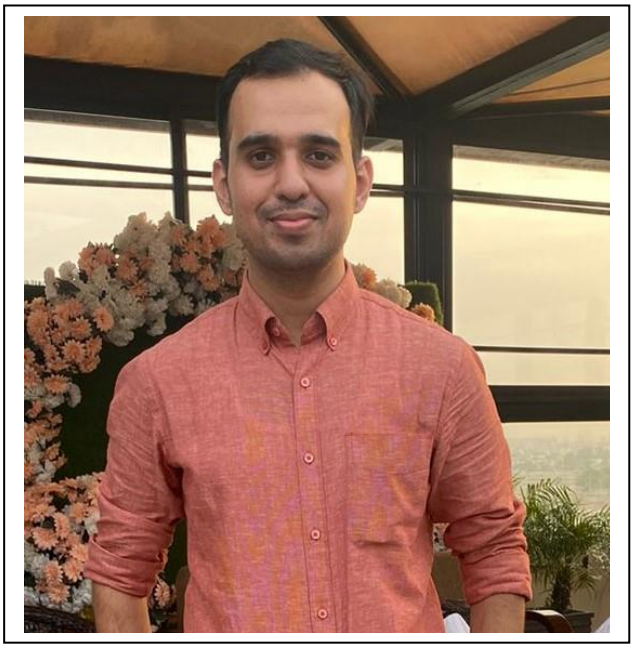
UNIVERSITY OF GOTHENBURG

DEPARTMENT OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES

# PREDICTING CHEMICAL-GENE INTERACTION



**Ali Chishti**

Degree project for Master of Science in Ecotoxicology (45 ects)
ES2520, Master Thesis in Ecotoxicology, 45 ects
Second cycle
Semester/year: Spring 2019
Supervisor: Dr. Pedro Inostroza, Department of Biological and Environmental Sciences
Examiner: Prof. Dr. Thomas Backhaus, Department of Biological and Environmental Sciences

## Acknowledgement

I would like to thank my grandfather Mr. Zahoor Ahmad Chishti and my parents Mr. Masood Chishti and Mrs. Atika Masood Chishti for their undying support. Special thanks to my Supervisor Dr Pedro Inostroza for taking me under his wing and guiding me every step of the way. I would also like to thank Åsa Arrhenius and Lennart Bornmalm for always being willing to help throughout my thesis. I would also like to thank my special friends Mishal Ahsan, Masab Azhar and Naila Nasir for being my support system.

Ali Chishti,
University of Gothenburg, May 2023.

**List of Abbreviations**

| | |
|---|---|
| **STITCH** | **Search Tool for Interactions of Chemical** |
| **DNA** | Deoxyribonucleic Acid |
| **RNA** | Ribonucleic acid |
| **EDC** | endocrine-disrupting chemicals |
| **NCBI** | National Centre for Biotechnology Information |
| **IUPAC** | The International Union of Pure and Applied Chemistry |
| **PFOS** | Perfluoro octane sulfonic acid |
| **SMILES** | Simplified molecular-input line-entry system |
| **PAHs** | polycyclic aromatic hydrocarbons |
| **USEPA** | United States Environmental Protection Agency |
| **BLAST** | Basic Local Alignment Search Tool |
| **AMPA** | α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid |
| **MCPA** | 4-chloro-o-tolyloxyacetic acid |

# Table of Figures

# List of Tables

# Table of Content

# Abstract

Pesticide use in agriculture has become a growing concern as it can have detrimental effects on the environment. The excessive use of these chemicals often leads to them seeping into nearby water bodies, causing harm to aquatic organisms. Recent studies have shown that these compounds can even alter the genetic makeup of these organisms, which can have far-reaching consequences. In this study, pesticide data from two agriculture sites, Skivarpsån and Skåne M42, were analysed to predict the effects of pesticides on a selected aquatic organism. The data for this study was collected from various sources, including the National Centre for Biotechnology Information, the STITCH database, and the PubChem database. The STITCH chemical-protein interaction prediction model provided a careful analysis of the chemical compounds and their interaction with genes. The main result of this study shows that between the two sites, M42 is highly polluted with a greater number of chemicals detected than the Skivarpsån site. Furthermore, this study indicates that higher protein-gene interactions are expected in M42 compared to the Skivarpsån site. The stream with significant pesticide mixture pollution will have more interactions between small molecules and proteins than streams with low pesticide loads. The function enrichment on both sides shows that the chemicals that have the highest effect on the organisms are insecticides affecting their biological aspects. Despite the usefulness of these resources, one of the limitations of the study was that several nodes remained undetermined.


**Keywords:** Pesticides, Small streams, Prediction modelling, Chemical-protein-gene interactions, Aquatic organisms.

# 1 Introduction

Humans have been dependent on chemicals for a long time for various purposes such as food production and drugs. Unfortunately, these chemicals end up in the environment and impact it negatively. Chemicals used in industries, pharmaceuticals, and agriculture are not easily degradable and can take a significant period for their impact to be diminished. They may interact with one another to form a more harmful mixture of toxic chemicals (Backhaus & Faust, 2012). The impact of different chemicals can vary widely (Carpenter, Arcaro, & Spink, 2002). A mixture of two chemicals can interact quite differently when compared to a mixture of different chemicals. The toxicity of these chemicals can be impacted when they react with each other and may end up increasing or decreasing their toxicity (Waiser & Holm, 2005).
A chemical mixture is made up of multiple chemicals, which may produce a wide range of effects on aquatic life. For instance, the use of a formulation herbicide (glyphosate, dicamba, or 2,4-D) together with an increasing application in the field have resulted in serious risks to aquatic and terrestrial life, chronic and acute toxicities in mammals, animals, and birds, extreme toxicities in plants, and environmental degradation [United States Environmental Protection Agency (USEPA), 2005]. Although there is vast scientific evidence of the effect of chemicals on aquatic organisms, there are knowledge gaps regarding effects at the genetic level, the expression of genes, and consequently proteins.
Preliminary studies have shown that chemical pollution may jeopardize genetic diversity leading to genetic erosion, constraining gene flow, and ultimately affecting the adaptability of aquatic organisms (Inostroza et al., 2018). There is growing evidence that some pesticides will be more toxic in mixtures than expected based on the component chemical toxicities, and traditional risk assessment methodologies may underestimate the environmental danger. For example, a study on salmon discovered that sub-lethal mixtures of organophosphate and carbamate pesticides exhibit significant acetylcholinesterase inhibitory synergy, which is critical for salmon survival (Taube et al., 2011). Another study discovered that using multiple fungicides at the same time or in close vicinity can result in a twelve-fold increase in *Daphnia magna* immobility compared to using single pesticides (Laetz et al., 2009b). Many pesticides are persistent in the environment, and new research has revealed that legacy pesticides can still be present in quantities capable of generating ecotoxic consequences (Anderson et al., 2013).

## 1.1 Chemicals in Aquatic Environment

Due to anthropogenic activities, vast quantities of chemicals end up in the aquatic ecosystem every day, threatening the aquatic biodiversity and endangering water ecosystems. The emission of industrial effluents, pesticides, pharmaceuticals, organic wastewater contaminants, alkylphenol ethoxylates, perfluorinated surfactants, flame retardants, and chlorinated paraffin are all manifestations of anthropogenic activity (Kuzmanović et al., 2016; Klečka, Persoon, & Currie, 2010). Some chemicals, for instance, atrazine, imidazole, clothianidin, and imidacloprid are substances with well-documented effects and can harm an organism and can end up destroying an entire food chain of aquatic life; but, when these chemicals combine, their consequences are unpredictable (Yang et al., 2008., Pilling & Jepson, 1993., Denise Forson & Storfer, 2006., Mahmood et al., 2016).
Pesticides can enter the water in several ways which include surface runoff, wastewater, spray drift, and drainage (Iwakuma et al., 1993). This creates a mixture of insecticides, fungicides, and herbicides and in turn impacts aquatic life adversely (Schäfer et al., 2011). This may indirectly increase the mortality of arthropods such as mayfly nymphs (Schulz and Dabrowski, 2001) among other macroinvertebrates.

Freshwater arthropods are important components of freshwater macroinvertebrates, which are invertebrates larger than 0.5 mm in size. They play a role in a variety of ecosystem processes, including nutrient and carbon cycling (Wallace & Webster, 1996). As a result, they've been employed as water quality indicators for over a century and are an important part of current and historical freshwater monitoring systems (Bonada et al., 2006). Macroinvertebrates like *Gammarus pulex* play a role in maintaining the ecosystem. They are involved in leaf litter decomposition (Rasmussen et al., 2012). *Gammarus pulex* shows moderation in adapting to the presence of pesticides in the environment but has led to long-term impairments like reproduction, growth, and survival (Siddique et al., 2020).

## 1.2 Pesticides

Pesticides are chemicals or biological agents (such as viruses, bacteria, or fungi) that dissuade, incapacitate, kill, or otherwise discourage pests. Insects, plant pathogens, weeds, molluscs, birds, animals, fish, nematodes (roundworms), and bacteria that harm property, cause nuisance, or spread disease, as well as disease vectors, are all examples target pests (EPA, 2021). Pesticides cause acute and chronic toxicity in aquatic and terrestrial organisms. These pollutants also act at the DNA level leading to mutations and bringing changes in epigenetics pathways leading to multigenerational levels (Carnevali et al., 2018).
Pesticides are frequently applied to crops in formulation, and residues can be discovered in foods and drinking water. Pesticide formulations, on the other hand, are widespread in the aquatic environment, particularly on surface waters where aquatic life thrives (Laetz et al., 2009). Toxicity studies involving pesticide combinations have yielded a wide range of reactions, with the intricacy of the interactions depending on differences in the pesticides' chemical characteristics and modes of harmful action (Hernández et al., 2013). Insecticides are designed to kill certain arthropod pests, but they are also poisonous to a wide range of nontarget arthropods, including crustaceans, insects, and mites in freshwater (Rubach et al., 2009; Carsten Von Der Ohe & Liess, 2004).
Gammaridae is an amphipod family, where *Gammarus pulex* is one of the model organisms in European freshwater ecosystems. They can be found in clean lakes, ponds, streams, brooks, springs, and underground waters. Gammarids have been used to research chemical stressors over the last decades and are appropriate organisms for use in laboratory and field eco-toxicological studies (Chaumot et al., 2015; Inostroza et al., 2016; Inostroza et al., 2017; Schäfer, 2019). In research by Fu et al. (2018), prochloraz concentrations increase azoxystrobin uptake by causing hyperactivity, increasing its toxicity to *G. pulex*. In agricultural streams, pesticide exposure promotes adaptation in non-target species, and enhanced tolerance to clothianidin (60 h EC75) was linked to reduced overall fitness and long-term viability of *G. pulex* in culture, according to a study (Siddique et al., 2020). In a study by de Castro-Català et al. (2017b) at low quantities, prochloraz and fluoxetine influenced *G. pulex* moulting, eating behaviour, and locomotive behaviour.

## 1.3 The Effect at the Molecular Level

Chemicals affect organisms at the molecular level. The structure, function, and makeup of physiologically relevant molecules such as DNA, RNA, and proteins are studied at the molecular level. The current knowledge of the negative impacts of chemicals on the environment and their more direct consequences on human health is based on structure-activity connections and molecular mechanisms of action. It is critical to discover variables that influence the toxicological effects of foreign chemicals in biological systems and to evaluate the understanding of chemical toxicity mechanisms (McKinney, 1985).

## 1.4 Interactions between Chemicals, Genes, and Proteins

Chemical protein/gene interaction is efficiently used in the field of drug development. The effect of a drug on the body and its efficacy is determined by how well it interacts with the targeted proteins and how much it alters the protein-protein and protein–chemical interaction network (Hopkins, 2008; Hopkins et al., 2004). This is dependent on the chemical's concentration, the strength with which it affects the target's activity, and the distribution of target proteins throughout tissues (Szklarczyk et al., 2015). Chemicals cause compound-specific alterations in an organism's transcriptome (toxicogenomic fingerprints). This could provide information regarding cellular, pathways, or physiological responses to chemical exposure and harmful effects, which is useful for assessing chemical risks or environmental health. Toxicogenomic fingerprinting provides a possible understanding of cellular and physiological responses to chemical vulnerability and unfavourable effects which are required for environmental hazard-related assessment. In a case like this, a comparison study is required for the interpretation of toxicogenomic experiments (Schüttler et al., 2019). This approach would help to understand how chemicals interact with genes and proteins in non-targeted organisms in the aquatic environment and potentially may help to reveal which metabolic pathways are affected. This perspective is new in the field of ecotoxicology.

Life is made possible by a variety of molecular interactions. Protein-protein interactions, protein-nucleic acid interactions, and protein-small molecule interactions are a few examples of these. Major scientific objectives include explaining these connections and figuring out how they regulate biology (McFedries et al., 2013).

Small molecules in the cell can control protein functions as hormones and ligands through allosteric binding in addition to acting as substrates and products in biochemical reactions that are essential to life. Near a protein, small molecules may bind and change its biological processes, including the activity of enzymes and the interaction of proteins. These interactions offer crucial tools for the allosteric (site which allows activation or inhibition of enzyme activity) control of enzymes and receptors (Li et al., 2012). There have been reports of small compounds that target protein-protein interactions in several biological pathways like the inhibitors of protein that regulate immunity recognition by its receptor, receptor that induces an immune response, protein controlling apoptosis binding to the proapoptotic protein which controls apoptosis, and tumour suppressor gene binding to its protein ubiquitination and degradation, a gene that negatively regulates tumour suppressor gene, are a few earlier instances of protein-protein interactions signalling pathways (Braisted et al., 2003c., Vassilev, 2004., Oltersdorf et al., 2005., Tse et al., 2008., Grasberger et al., 2005., Jin et al., 2014). In research from Silvestre (2020), in response to various xenobiotics, aquatic organisms likewise activate key pathways that have been characterized in mammals. These pathways interact with one another, activating a sophisticated cellular antioxidative apparatus in response to various xenobiotics. However, scientists still only have a limited understanding of how aquatic organisms respond to certain stimuli. The number of investigated species should be increased, and efforts should be made to describe the responses' organ- and age-dependency. The process is challenging, though, because of the enormous quantity and variety of chemicals present in the environment. Understanding these important pathways can aid in understanding the mechanism of action of pollutants, which can then aid in determining the environmental risk to aquatic ecosystems.

The use of chemicals in agriculture though having benefits is creating havoc on the environment. These chemicals when interacting with each other create mixtures which are detrimental to aquatic organisms as these chemical mixtures' leachate into the water bodies. The high agricultural activity in Southern Sweden has led to the high-level detection of pesticides in this area. These chemicals may affect organisms on a molecular level and consequently their gene and protein sequences, this leads to a major effect on the gene pool

of the organism. To understand the effects of these pesticides on the aquatic species, *Gammarus pulex* is used as a standard organism to evaluate the chemical-protein-gene interaction effect by using prediction modelling.

## 1.5 Overall and specific objectives

### 2.5.1 Aim:

To predict chemical-protein-gene interactions in freshwater macroinvertebrates exposed to mixtures of pesticides in stream ecosystems stressed by agricultural activities.

### 2.5.2 Hypothesis:

A higher number of small molecules-protein interactions will occur in small streams highly polluted by pesticide mixtures compare to streams with low pesticide loads.

### 2.5.3 Objectives:

To build a pesticide database from selected streams impacted by pesticide pollution.
To predict pesticide-protein-gene interactions using prediction modelling. To assess pesticide-gene interaction (nodes) according to the degree of pesticide loads in the environment.

# 2 Methodology

## 2.1 Pesticides in the aquatic environment

Pesticide concentrations were retrieved from the Jordbruksvatten Database (http://jordbruksvatten.slu.se/). This database is part of Datavärdskap Jordbruksmark which is a task of SLU (Swedish University of Agricultural Sciences) from the Swedish Environmental Protection Agency. Jordbruksvatten Database contains data about the type of land area, the plant's nutrients, field observations and plant protection agents like pesticides in surface water, from national environmental monitoring of plant protection products in surface water from two rivers and four minor catchments with a high concentration of agriculture.

Measured environmental concentrations and the mode of action for the all the chemical were retrieved for Skivarpsån and M42. Both sampling sites are located in Skåne, Southern Sweden. Skåne is characterized by land for agricultural use. It is referred to as Sweden's breadbasket and boasts some of the best agricultural soil. About half of Sweden's food production is represented by it (Ul Hassan et al., 2021). Compared to other regions in Sweden, Skåne has a high percentage of agricultural land at 40%. Skivarpsån, a 25-kilometre-long stream in Scania's deep south, has a catchment area of around 102 square kilometres. This catchment's percentage of agricultural land is about 89%. A smaller body of water called M42 passes across farmland. The stream's catchment area is only 8.24 km2, and about 93–94% of it is made up of agricultural land. These locations were picked as pesticide-effect areas. Both Skivarpsån and M42 are locations that SLU regularly monitors for pesticides on an annual basis (Swedish University of Agricultural Sciences, Uppsala) (Håkansson, 2017).

## 2.2 *In silico* chemical-protein/gene interactions

To conduct chemical-protein-gene interaction, information is required about the chemical structure of the measured pesticide in Skåne. The uses of SMILES (Chemical identifiers) are required. SMILES are unique for every chemical and are generated by a Canonical algorithm, so the term canonical SMILES is used. Compared to other chemical identifiers like IUPAC, SMILES are easily read by humans and fixed if any error occurs, unlike IUPAC which can only be interpreted by an online database system. The canonical SMILES were retrieved from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) for each of the chemicals measured in Skivarpsån and Skåne M42. SMILES were used as input files in STICHT.

STICHT is an interaction information database for chemicals. Information regarding more than 0.5 million chemicals of different types is available on STICHT (http://stitch.embl.de/). It also contains information regarding 2200 drugs. It also interlinks these chemicals and drugs to 1.5 million genes and their interconnections that are available in the STRING database and about 2031 genomes (Szklarczyk et al., 2016b). STICHT predicts a network interaction, which is comprised of interactions between proteins and chemicals. Chemicals and proteins are represented by nodes, while biological interactions between them are represented by edges (protein-protein interaction). Edges are the interaction between proteins to achieve a specified function. (Kuhn et al., 2007).

To carry out the prediction interaction, chemical data and a genome is required, since *Gammarus pulex* our target organism genome is not available, the surrogate species *Drosophila melanogaster* is used. *Gammarus pulex* is a crustacean and crustaceans have similarities with the genetic makeup of arthropods. STICH provides several predictions for the input chemical SMILES. These different structures of chemicals are ranked according to the Tanimoto score. The Tanimoto score is a similarity metric, it compares two chemicals and creates an

intersection, and provides us with a similarity ratio (Chung et al., 2019). The substance with the highest ratio then proceeds to assess a prediction action model.

The first information that is obtained from the database is the chemical protein interaction model. The information that is observed from the interaction model entails the molecular action interactions among a variety of proteins and the chemical (Figure 1). The analysis of the prediction action provides information on nodes; proteins serve as network nodes. The network node is comprised of shells, where the first shell corresponds to proteins which are directly involved with the input protein and the second shell relates to the proteins linked to the first shell proteins or your input protein makes up the second shell of interactors. The total number of proteins produced by a single protein-coding gene is indicated by each node. It also provides information on the size and colour of the nodes. The size of the nodes gives information on the prediction of the protein structure. If the size of the node is small, then the 3D-Protein structure is unknown and if the size of the node is large then the 3D-Protien structure is known or can be predicted. The second piece of information is the colour of the nodes. If the nodes are coloured, such as purple, green, or red, they are connected to the first shell proteins and are query proteins (protein sequence) and if the nodes are white, they are directly connected to the input chemical and are second shell interactors. First shell proteins are directly connected to the input chemical. They are involved in the initial interaction between chemical and the protein and the second shell interactor proteins are indirectly connected to the input chemical through the first shell interactor protein. These proteins are less likely to be involved in the first interaction but a make a combination interaction and lead to affecting the signalling pathways. The action effect of the nodes is also discussed, node impacts are classified as positive (arrow), negative (bar), or unspecified (circle). These effects show how these substances affect the function of the organism in the area, either negatively or positively. Some consequences are vague or unspecified, and sometimes the effect is not revealed.

The next information given is on edges which is the association of the proteins with each other. Proteins work together to perform a specific function, so the association of protein in a prediction interaction gives information on specific action happening. It also explains the action types and action effects among the proteins and the chemical).

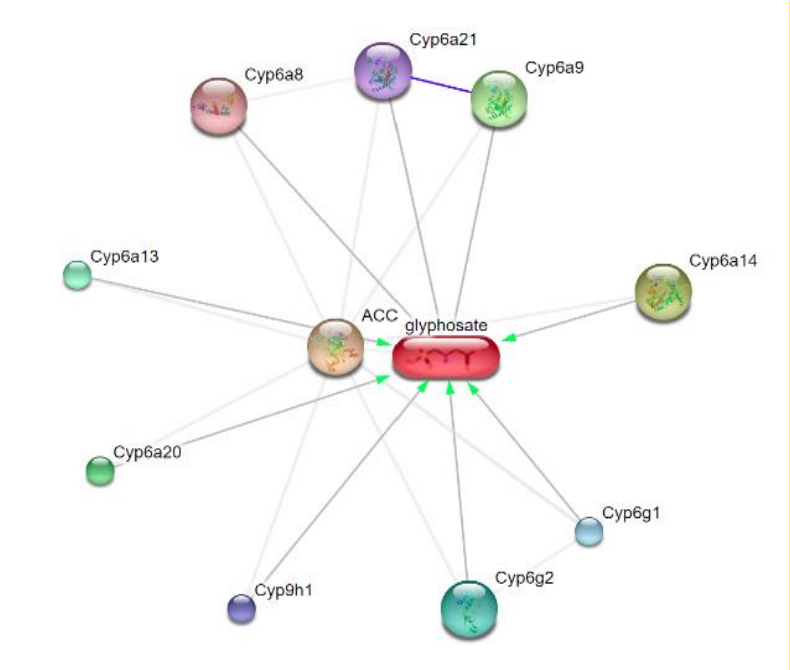



*Figure 1 Glyphosate Chemical-Protein Interaction, the nodes identifying the 3D protein structure are the large node and the node not knowing the protein structure are small nodes. The interaction between two nodes represented by a line is known as an edge.*

STITCH is capable to predict the total number of nodes involved and the total number of edges. Additionally, a functional enrichment is conducted, and it focuses on gene sets, which are groups of genes that share common biological functions, chromosomal locations, or regulation (Subramanian et al., 2005). The functional enrichment in the network may be a biological process, molecular function, etc. The function enrichment network provides us with information regarding if any biological process or molecular function or cellular process is affected.

STICH provides a node identifier for each chemical-protein interaction. These identifiers were used to retrieve the protein sequence data from NCBI. The protein sequence was then used to retrieve the gene sequence from the GenBank at NCBI.

## 2.3 From proteins to genes

### 3.3.1 National Centre for Biotechnology Information:

A protein-gene library was created by using the National Centre for Biotechnology Information. The data for the homologous gene/protein was obtained from the NCBI database and then it was used to generate the library. The fly database, a database of Drosophila genes and genome, is then updated with the node identifiers (http://flybase.org/). The Uniport database (https://www.uniprot.org), which is a database of protein sequences and functional information, is accessed if node identification is present. If it is, the external cross-reference and links section is searched. The Uniport database gives the protein sequences and the gene sequences by utilizing the node identifiers from the fly database. Thus, the obtained protein and gene sequences generate a protein-gene library, and the gene and protein sequences were compiled in an Excel file. The data was then exported to an Excel file where the duplicate genes and protein sequences were removed using the condition formatting function. After the file was cleaned up then it was converted into a Fasta file for the blast analysis. Blast is Basic Local Alignment Search Tool which compares the protein sequence with the sequence database and computes the statistical relevance.

## 2.4 Graph Plotting:

For graph plotting, all the data that have been gathered in the exploration of the research have been analysed and plotted in Microsoft Excel.

# 3 Results and Discussion:

## 3.1 Pesticide characterization of small streams affected by agricultural activities.

The study conducted an analysis of 59 chemicals retrieved from two streams in Skåne, Sweden - M42 and Skivarpsån. The concentrations of these chemicals were found to be generally higher in the M42 site as compared to the Skivarpsån site, indicating a higher degree of contamination in M42 (Table 1 Appendix). The study revealed the presence of three classes of pesticides - herbicides, fungicides, and insecticides at both sites (Figure 2).

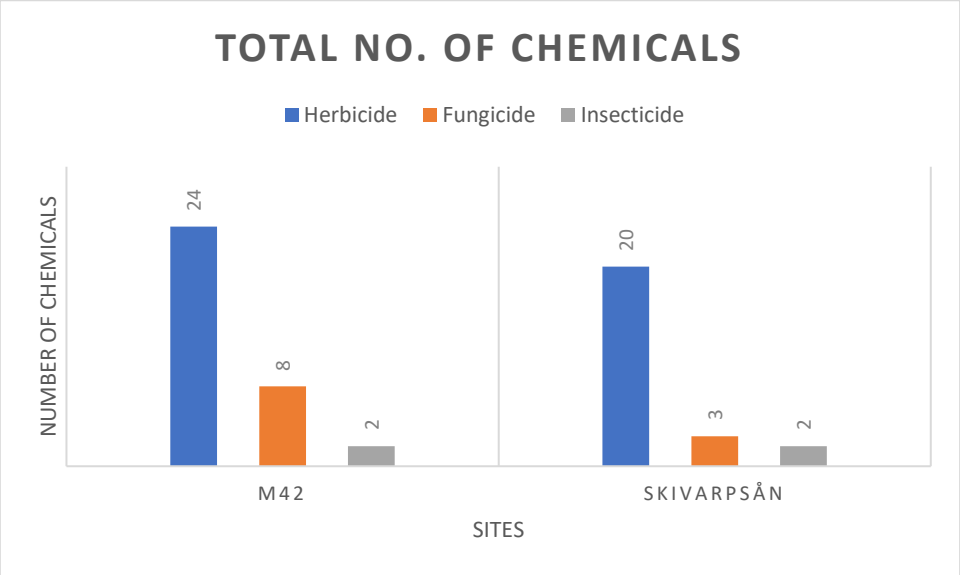



*Figure 2 Total Number of Chemicals*

Out of the 34 chemicals detected in the M42 site, 24 were herbicides, 8 were fungicides, and 2 was an insecticide (Figure 2). Similarly, the Skivarpsån site had 25 chemicals, of which 20 were herbicides, 3 were fungicides, and 2 were insecticides (Figure 2). It is noteworthy that herbicides were detected at both sites at high levels, with 22 common chemicals detected at both sites out of all the pesticides.

The study identified Glyphosate, α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA), Metamitron, and Fluopyram as the chemicals with the highest concentrations at the M42 site (Figure 3). Glyphosate and AMPA are broad-spectrum herbicides that inhibit the activity of an enzyme involved in the synthesis of amino acids. Metamitron is a herbicide, while Fluopyram is a fungicide. Similarly, at the Skivarpsån site, the chemicals with the highest concentrations were 4-chloro-o-tolyloxyacetic acid (MCPA), AMPA, Fluopyram, and Chloridzon (Figure 4). MCPA and Chloridzon are herbicides, while Fluopyram is a fungicide.

Glyphosate and its degradation product AMPA were found to be present at both sites, indicating the widespread use and persistence of these chemicals in the environment. The study highlights the need for effective measures to minimize the contamination of water bodies.

The insecticides thiacloprid, imidacloprid and pirimicarb found at the concentration 1 consolidated ng/l respectively as mentioned in the figures 3 & 4 will lead to reduced life span of Drosophila melanogaster. In a study conducted by Riva et al. (2018), exposure of pirimicarb at 1.05mM to bees will lead to rapid decrease in food intake.

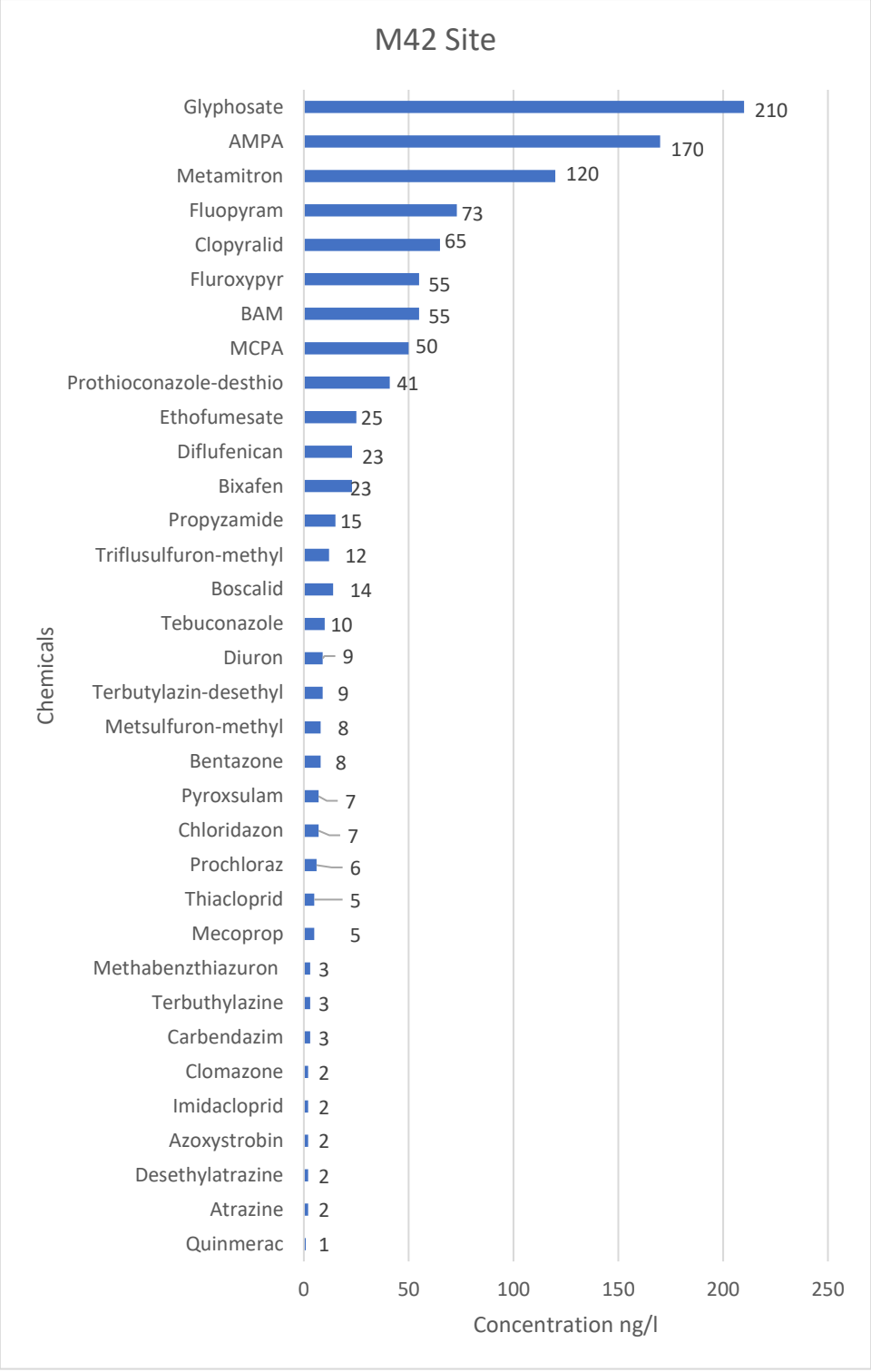



*Figure 3 Chemical Concentration at M42 Site.*

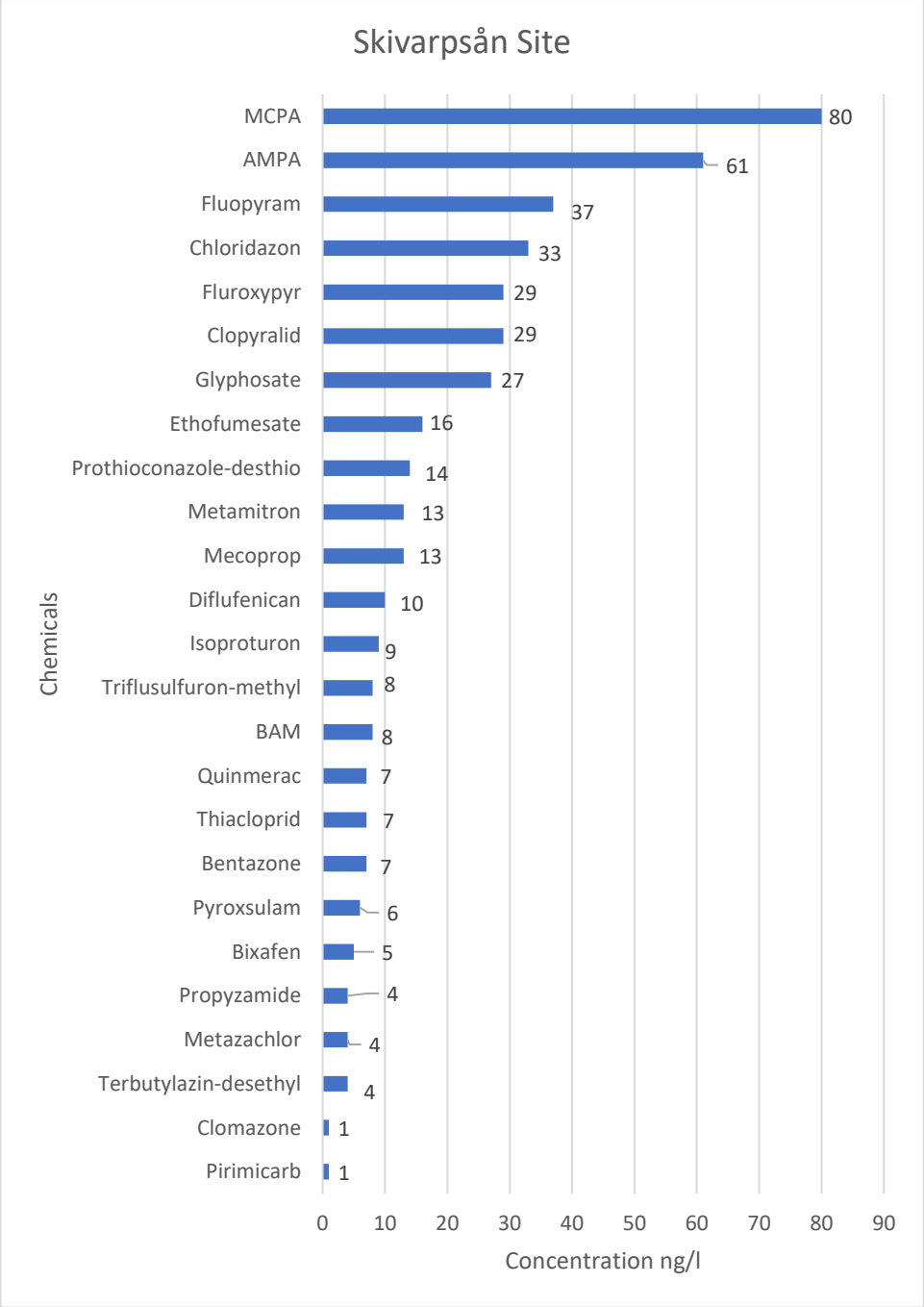



*Figure 4 Chemical Concentration at Skivarpsån.*

## 3.2 Mode of Action:

The mode of action is a process through which a chemical reaction affects an organism. This can either disrupt or enhance the specific action of the organism. In this study, 59 chemicals were analyzed from two different sites. The mode of action for each chemicals were retrieved from the SLU database.

Out of the 34 chemicals analysed at the M42 site, 12 chemicals had the same modes of action and were identified. The most common mode of action detected was photosynthetic inhibition, which was observed in 10 of the chemicals. Additionally, five chemicals had a synthetic auxin-producing mode of action, which replicates the effects of the growth hormone

indole-3-acetic acid in plants (Figure 5). Synthetic auxin is a critical plant hormone that affects growth, tropism, and development (Moulton et al., 2020).

At the Skivarpsån site, all 25 chemicals analysed had a known mode of action. 12 modes of action were detected, with the most common being photosynthetic inhibition and synthetic auxin, each observed in five of the chemicals (Figure 6).

The mode of action frequency for both sites is shown in Figures 5 and 6, indicating that M42 and Skivarpsån have the same number of modes of action detected, 12, and the same highest mode of action, photosynthetic inhibition. However, Skivarpsån has two dominant modes of action, photosynthetic inhibition and synthetic auxin, while M42 has one dominant mode of action, photosynthetic inhibition. Notably, the highest mode of action is not necessarily associated with the chemicals present in the highest concentration. Only one chemical, 2-methyl-4-chlorophenoxyacetic acid (MCPA), coincided with the highest mode of action detected, a synthetic auxin, at the Skivarpsån site.

The mode of action frequency appears to be more related to the types of chemicals used, specifically herbicides which when used in combination with insecticides will have a cojoined effect on the arthropods. This can lead to drastic effects on the organisms as chemical mixtures may lead to reduced life span of the arthropods. The effect on the growth hormone and the process of photosynthesis will lead to the retardation of the aquatic plants which will lead to starvation of the arthropods.

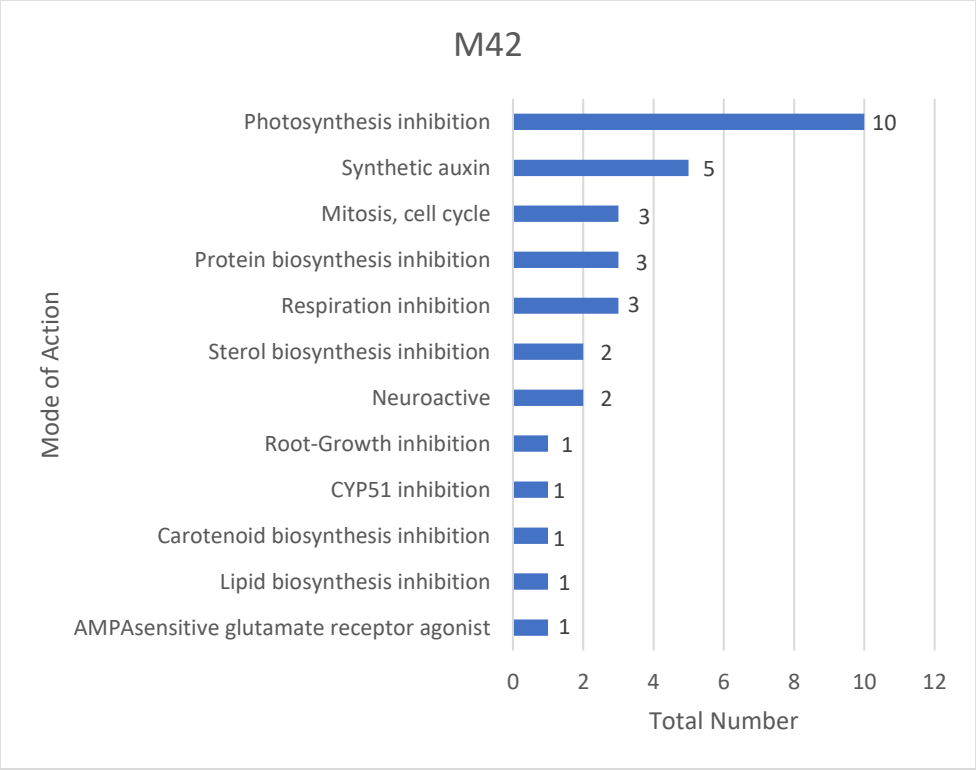



*Figure 5 Chemicals Mode of Action at M42 site.*

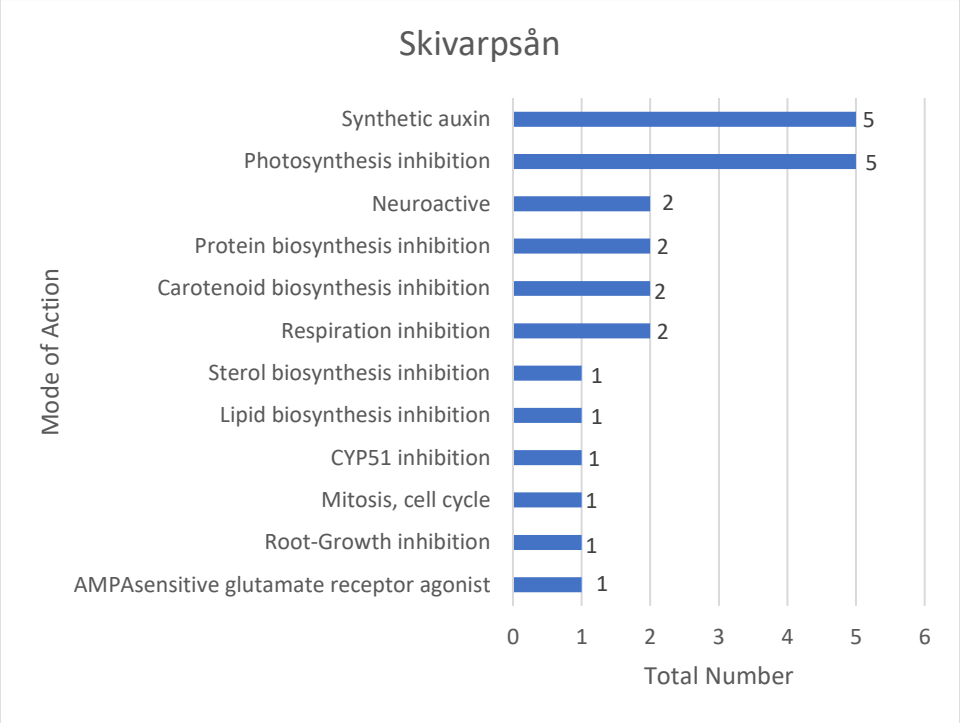



*Figure 6 Chemical Mode of Action at Skivarpsån site.*

## 3.3 Prediction modelling

### 4.3.1 Nodes:

Nodes in this context represent proteins that are predicted to interact by a single protein-coding gene locus. In this study, the interaction between chemicals and genes was analysed at two different sites, M42 and Skivarpsån. At the M42 site, 14 out of 34 chemicals showed interaction predictions, while at the Skivarpsån site, 11 out of 25 chemicals showed interaction predictions (Figure 7).

The M42 site had a higher number of predicted interacting nodes, with 76 nodes predicted from 14 chemicals (Figure 8), as compared to the Skivarpsån site, which had 39 interacting nodes predicted from 11 chemicals (Figure 9). This suggests that chemical-protein interactions are higher at the M42 site as compared to the Skivarpsån site.

Imidacloprid, mecoprop, atrazine, and metsulfuron-methyl were found to heavily affect the protein sequences at the M42 site, while mecoprop, glyphosate, and thiacloprid were found to heavily affect protein sequences at the Skivarpsån site. Notably, mecoprop, glyphosate, and thiacloprid were common chemicals found at both sites and had the highest number of nodes predicted at each site. Glyphosate, in particular, had the highest concentration among these three chemicals.

These chemical-protein interactions can potentially impact organisms, both positively and negatively. Understanding the nature of these interactions can provide insights into the effects of chemical exposure on the organism and may inform strategies to mitigate negative effects.

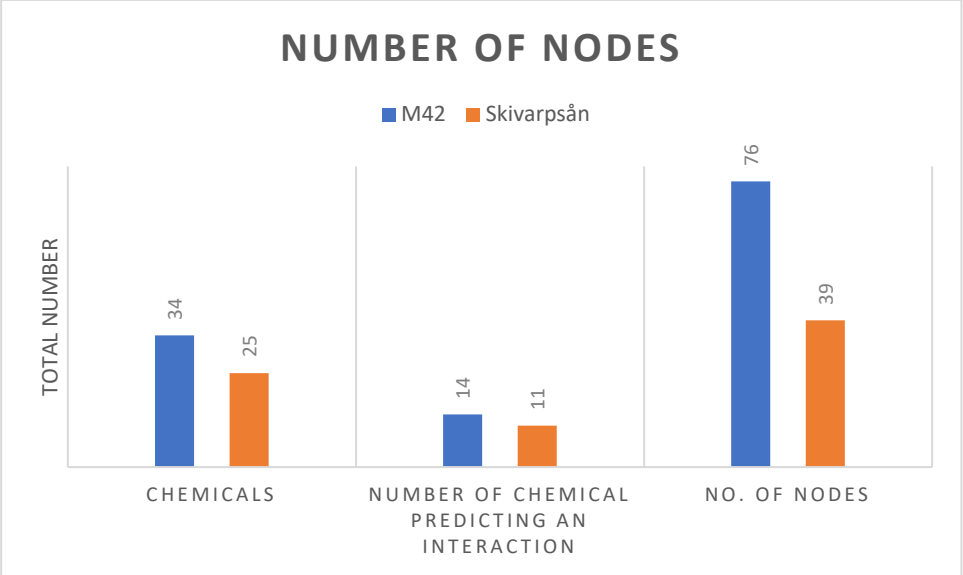



*Figure 7 Total Number of Nodes.*

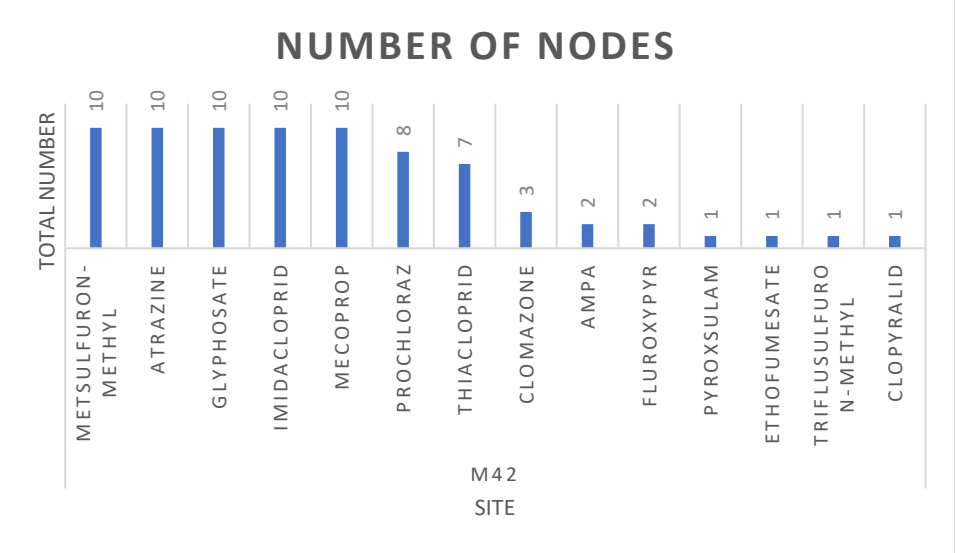



*Figure 8 Number of nodes detected in the sites M42.*

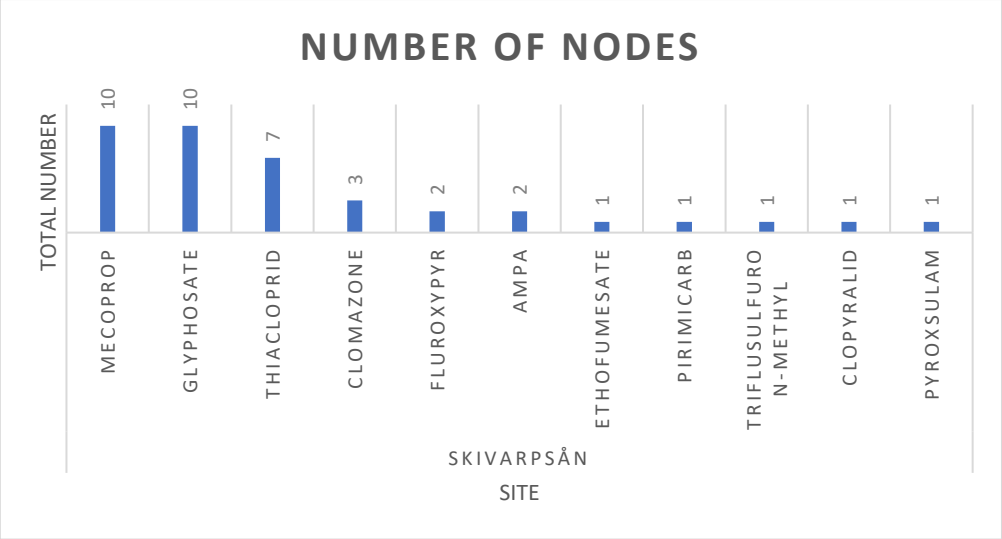



*Figure 9 Number of nodes detected in the sites Skivarpsån.*

### 4.3.2 Size of Nodes

The prediction of chemical-protein interactions involves the use of nodes to represent proteins. These nodes can be of two types based on their sizes - large nodes and small nodes. The size of these nodes is indicative of the 3D structure of the protein they represent. Large nodes have known or predictable 3D protein structures, while small nodes have unknown 3D protein structures.

In this study conducted at the M42 and Skivarpsån sites, a total of 76 interacting nodes were predicted at the M42 site, with 55 of them being large nodes and 21 being small nodes. Similarly, at the Skivarpsån site, 39 nodes were predicted, with 26 large nodes and 13 small nodes (Figure 10).

Most of the nodes at both sites had known 3D protein structures, making their sequences easier to analyze. This information is crucial as it can aid in predicting the effects of chemicals on protein structures and thus provide insights into the potential impact on organisms.

Overall, the use of large and small nodes in predicting chemical-protein interactions allows for a more nuanced understanding of the potential impact of chemical exposure on proteins and organisms.

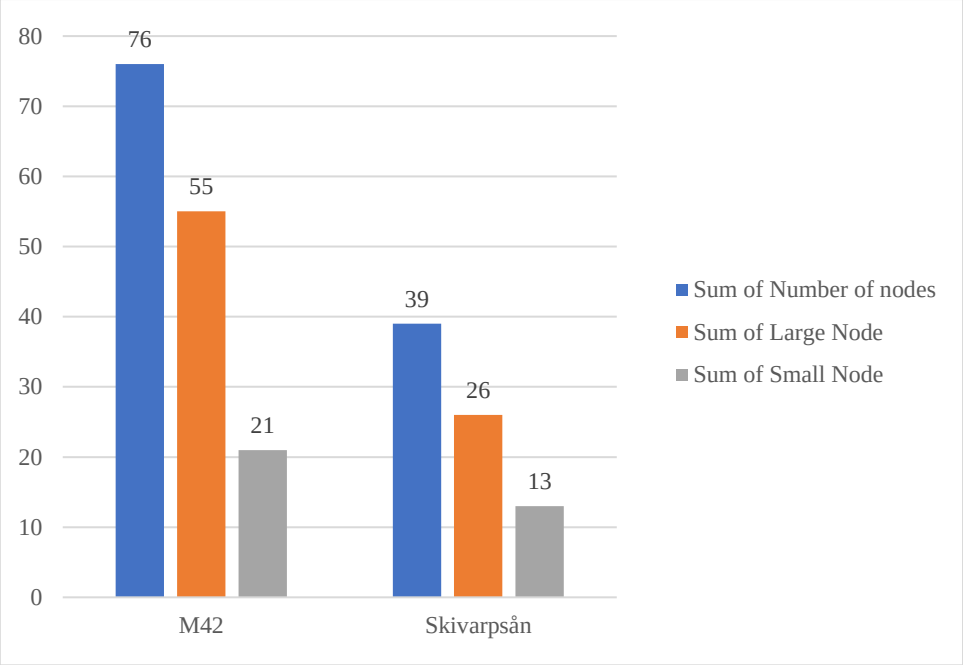



*Figure 10 Sum of the Number of Nodes, Sum of large Nodes, and Sum of Small Nodes in both sites M42 and Skivarpsån.*

### 4.3.3 Nodes Effects:

Chemical-protein interactions can have a significant impact on the organism, and the effect of the interaction can either be positive or negative. In some cases, the chemical can promote the organism's energy synthesis, while in other cases, it can disrupt the ongoing process of the organism. Therefore, understanding the effect of chemical-protein interactions is essential to predict the potential risks or benefits of the chemical.

At the M42 site, out of 76 nodes, only 20 nodes have a positive effect, while 17 nodes have a negative effect, and the rest of the nodes' effects are either undetermined or not informed (Figure 11). Similarly, at the Skivarpsån site, only 9 out of 39 nodes have a positive effect, while 10 nodes have a negative effect, and the effects of the remaining nodes are unknown (Figure 12).

The positive or negative effects of chemical-protein interactions can have a profound impact on the organism's protein sequences. For instance, Atrazine interaction with the node Cyp303a1 in *Drosophila melanogaster* has a positive effect, which is involved in the metabolism of insect hormones and the breakdown of synthetic pesticides. On the other hand, Atrazine interaction with the node *chic* in *Drosophila melanogaster* is negative, which binds to actin and changes the cytoskeleton structure.

The M42 site and the Skivarpsån site both have more nodes with positive effects than negative effects (Figures 11 & 12). However, a majority of the nodes' effects remain unknown, which indicates a significant gap in the data. Further research is required to fill this

gap and better understand the potential impact of chemical-protein interactions on the organism's protein sequences.

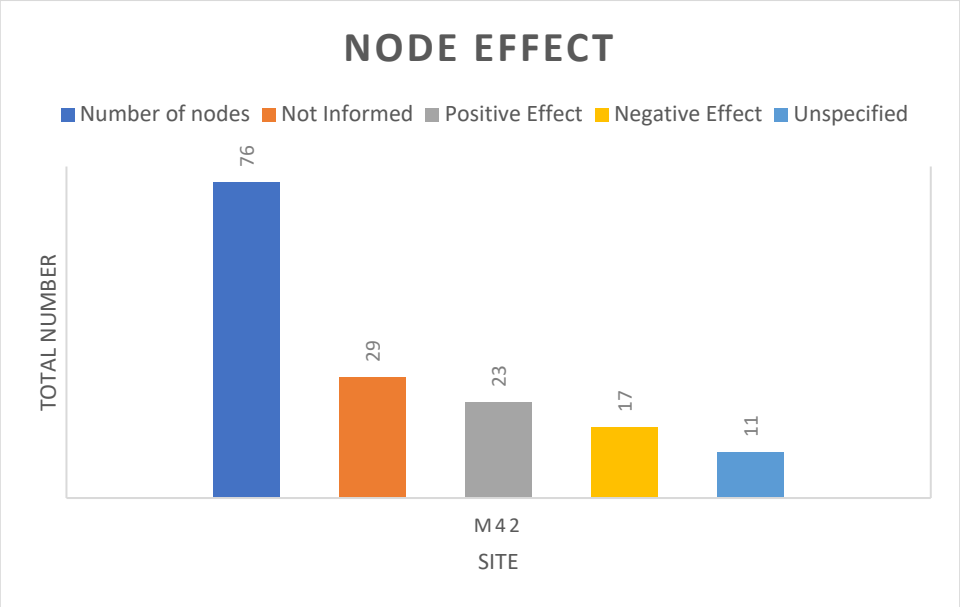



*Figure 11 Sum the Number of Nodes showing Positive affect, Negative affect, Unspecified, and Not informed.*

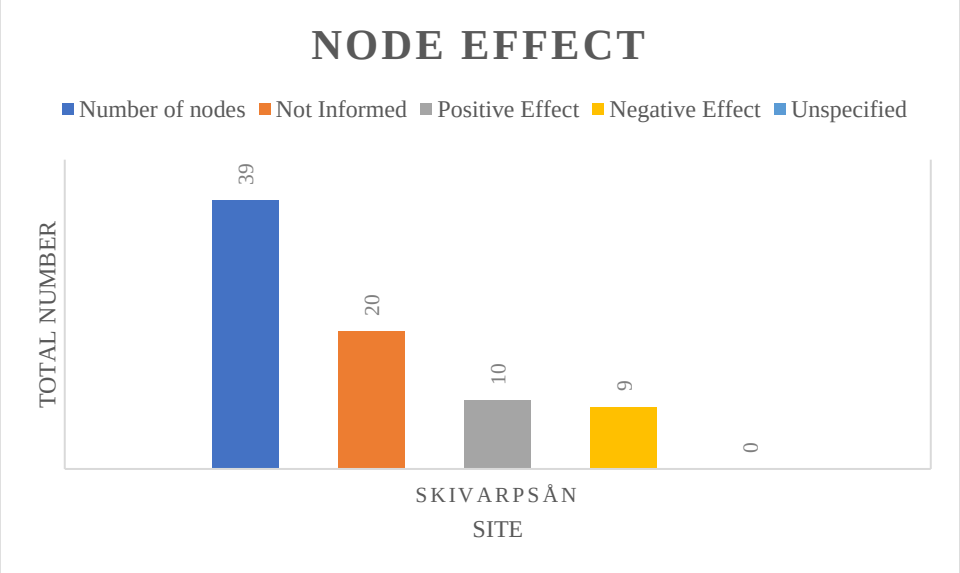



*Figure 12 Sum the Number of Nodes showing Positive affect, Negative affect, Unspecified, and Not informed.*

## 3.4 EDGES:

Edge means protein-protein interactions which play a crucial role in various biological processes, including signal transduction, enzymatic reactions, and gene regulation. The interactions between proteins can occur through direct physical binding or indirect interactions. These interactions can lead to shared functions between the proteins, which can be affected by external factors such as chemical exposure.

The analysis of edges in the M42 and Skivarpsån sites reveals important information about the chemical-protein interactions that occur in these environments. A total of 125 edges were detected at both sites, indicating that several protein-protein interactions take place in these environments.

The M42 site shows a higher number of edges than the Skivarpsån site, with a total of 94 edges detected. Imidacloprid, a chemical commonly used as a pesticide, was found to have the highest number of edges at the M42 site, with 33 edges. This indicates that Imidacloprid has a significant effect on the protein sequences and function at the M42 site.

On the other hand, the Skivarpsån site showed a total of 31 edges, with Thiacloprid having the highest number of edges at 15 (Figure 13). Thiacloprid is also a widely used pesticide and its high number of edges suggests that it has a significant impact on the protein sequences and function at the Skivarpsån site.

Overall, the analysis of edges in these environments highlights the significant impact of chemical exposure on protein function and emphasizes the importance of understanding chemical-protein interactions in environmental studies.

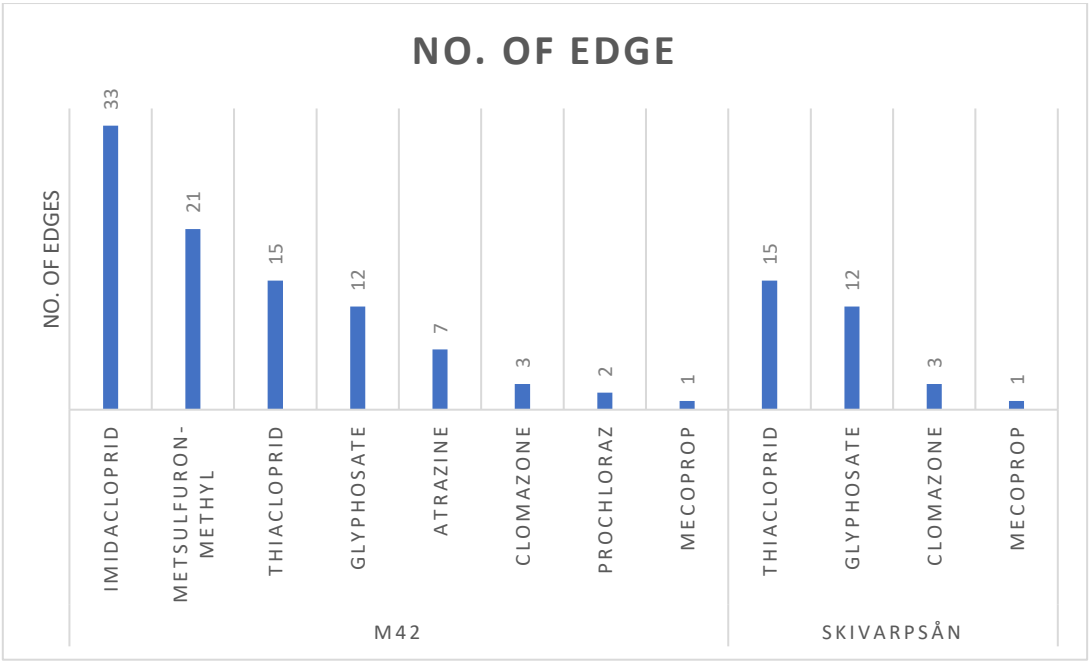



*Figure 13 No. of Edge at M42 and Skivarpsån sites.*

## 3.5 Functional enrichments in a network:

It is utilized for the analysis of gene expression data. It focuses on gene sets, which are collections of genes with related chromosomal locations, biological roles, or regulatory mechanisms. It offers us numerous typical biological paths. When we look at the biological process shown in Figures 15 and 15, we can see numerous routes that are similar to one another. For instance, the biological process “Protein Targeting Mitochondrion” (Figure 14) is labelled with the number 5, meaning that this process is detected five times. The graph also displays numerous more biological processes that are annotated with the number of genes that influence those pathways. As we move forward, curated the process from STITCH and created a database, and produced these graphs from where we can get the information about the Biological Processes, Intero Protein Domain and Features (IPDF), Molecular Functions, Cellular Components, and KEGG (Kyoto Encyclopedia of Genes and Genome) pathways.

### 4.5.1 Biological Processes:

The biological process is required for an organism to survive and helps an organism to interact with the environment. In Figures 14 & 15 it is observed that some of the biological

processes are affected by pesticides. A total of 24 biological processes are detected all over the M42 site, which is being induced due to these chemical protein/gene interactions. For example, the oxidation-reduction process has been identified 17 times. This biological process is induced due to the interaction with a insecticides. The cation transport membrane process has been identified 13 times. This process is induced due to the interaction with insecticides, in this case, imidacloprid and thiacloprid at the M42 site (Figure 14).

At the Skiverpsån site compared to M42 a smaller number of biological processes are observed. Almost all the biological processes identified at the Skivarpsån site are common with the M42 site and also affected by the insecticides thiacloprid, imidacloprid and pirimicarb. The oxidation-reduction process, signalling pathways, cell communication and single-organism signalling are highly affected at the Skivarpsån site (Figure 15).

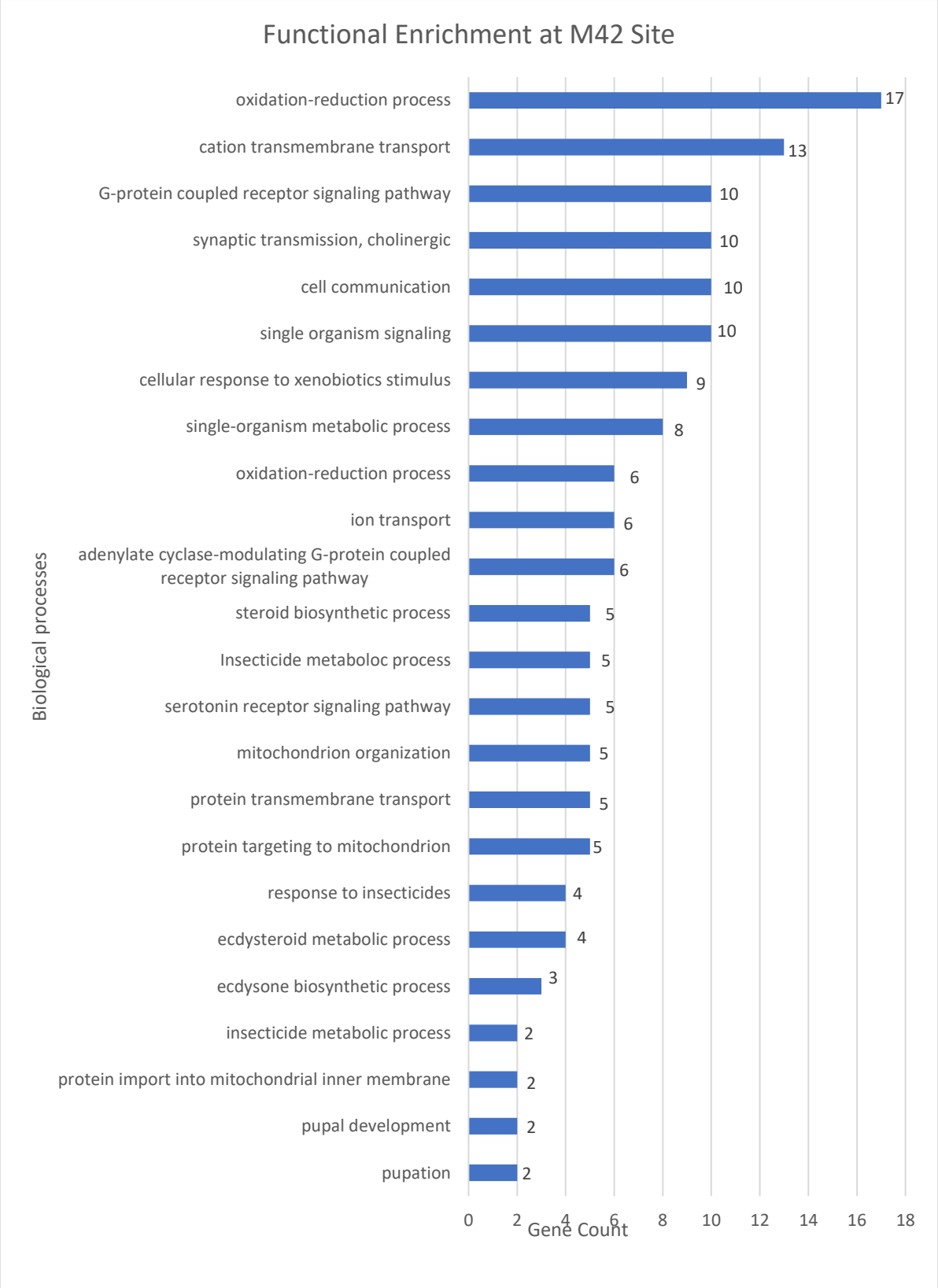



*Figure 14 Biological processes predicted by the influence of chemical-gene-protein interaction M42 site.*

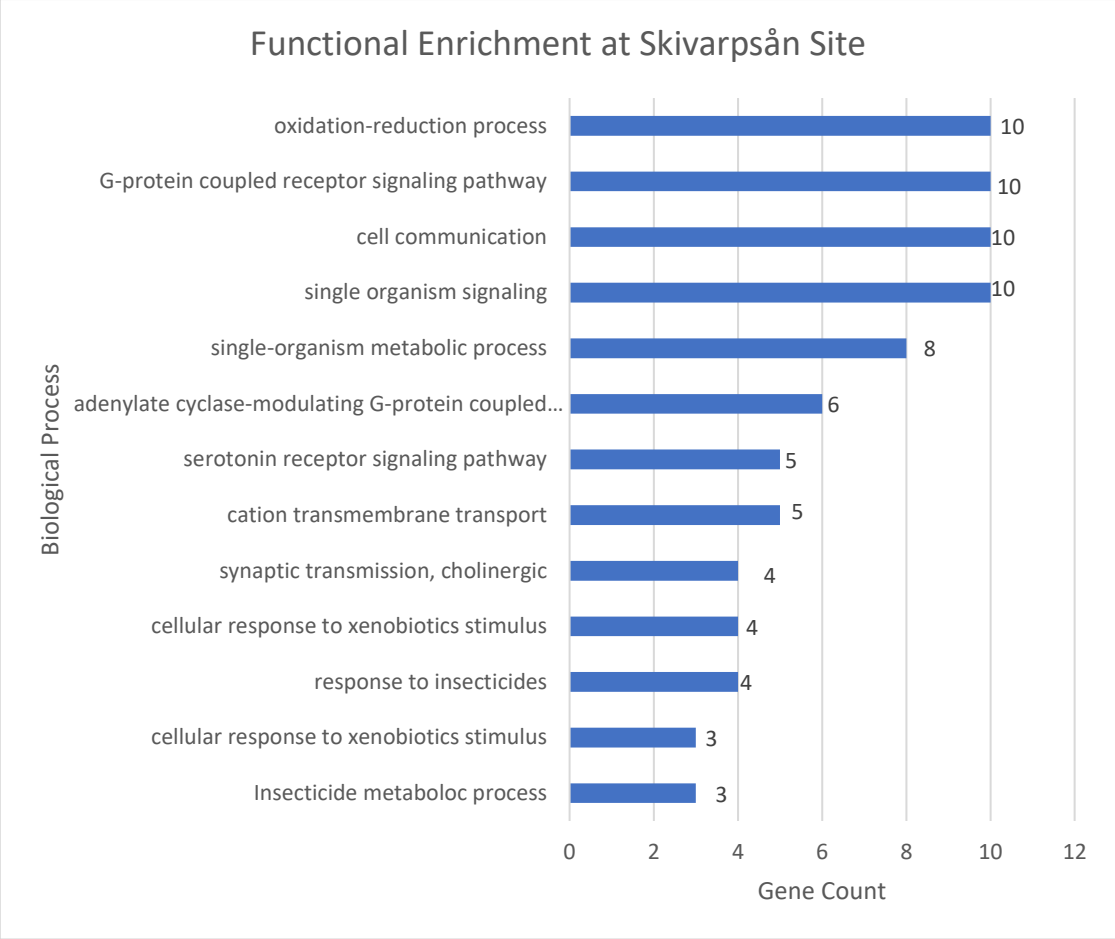



*Figure 15 Biological processes predicted by the influence of chemical-gene-protein interaction at the Skivarpsån site.*

### 4.5.2 Intero Protein Domain and Features:

InterPro is a valuable resource for the functional analysis of proteins, as it groups proteins into families, predicts domains, and identifies key locations using predictive algorithms known as signatures. These signatures are provided by several databases that make up the InterPro collaboration (Blum et al., 2020). InterPro has been specifically useful in understanding how insecticides such as imidacloprid and thiacloprid affect the processes in protein domains.

The Neurotransmitter-gated ion channel, Nicotinic acetylcholine receptor, Nicotinic acetylcholine-gated receptor, transmembrane domain, Neurotransmitter-gated ion-channel transmembrane domain, and Neurotransmitter-gated ion-channel conserved site are among the most commonly affected protein domains at the M42 site, each being identified 15 times (Figure 16). These domains play important roles in the transmission of nerve impulses, and the interference of their normal function by insecticides can have serious concerns for the nervous system.

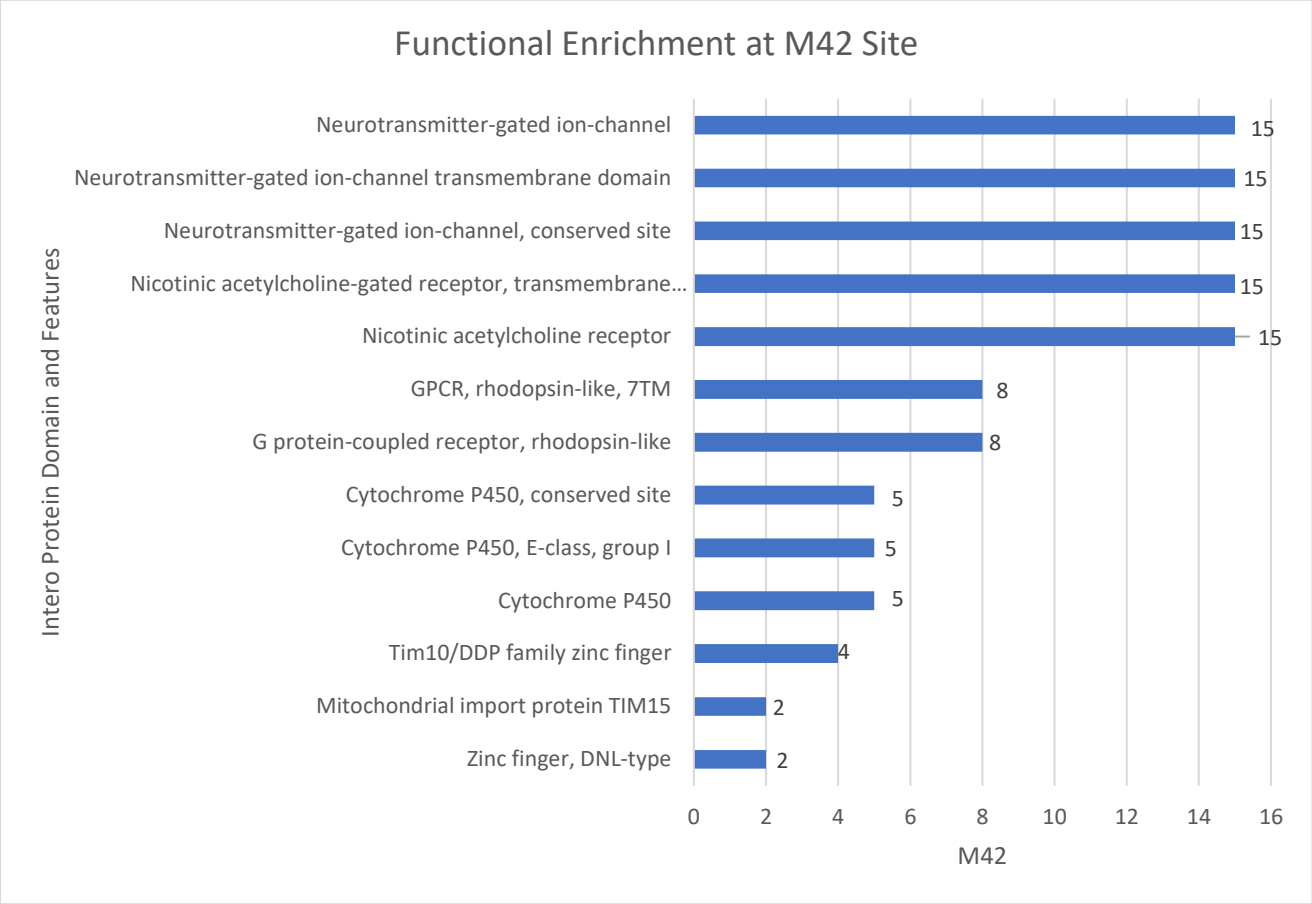



*Figure 16 INTERO Protein Domain and Features predicted by chemical-gene-protein interaction at M42 site.*

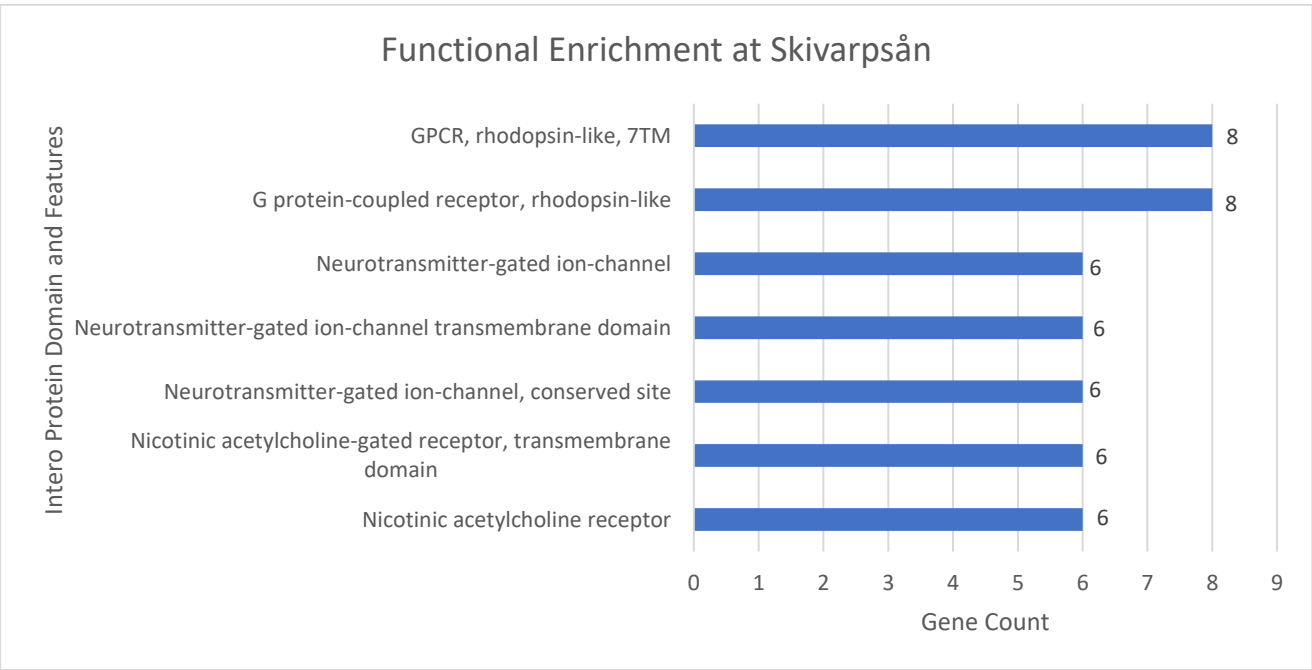



*Figure 17 INTERO Protein Domain and Features predicted by chemical-gene-protein interaction Skivarpsån site.*

At the Skivarpsån site, all the protein domains and features that are being identified are overlapping with the proteins being affected at the M42 sites. The G protein-coupled receptors, Rhodopsin-like and 7TM are highly identified at the Skivarpsån site (Figure 17). This protein domains are highly affected by the neonicotinoids thiacloprid, imidacloprid and pirimicarb.

The use of InterPro simultaneously with other databases has allowed researchers to advance a vast knowledge of protein functions and their interactions with environmental factors. This information is necessary to create strategies to reduce the possible risks linked with exposure to toxic chemicals.

### 4.5.3 Molecular Function:

All living organisms rely on a collection of complex molecular functions to maintain their normal biological processes. Disruptions to these molecular functions can lead to various physiological problems and can affect the overall health of the organism.

As depicted in Figure 18, the interaction of pesticides with the genome can lead to the disruption of various molecular functions at the M42 site. Specifically, the iron ion binding, heme binding and oxidation-reduction activity functions are highly affected by insecticides such as thiacloprid, imidacloprid and pirimicarb.

Iron ion binding is an important function for many biological processes, including oxygen transport, DNA synthesis, and energy metabolism. The interference of iron ion binding can lead to severe consequences and impaired cognitive function. Oxidation-reduction activity is another necessary molecular function that is affected by pesticides. This function is necessary for a wide range of physiological processes, including energy production, detoxification, and regulation of cell signalling pathways. The disruption of oxidation-reduction activity can lead to the accumulation of toxic compounds, impaired cellular respiration, and oxidative stress, which can ultimately lead to cell death and tissue damage (Reeder, 2010).

At the Skivarpsån site, the molecular functions being identified overlap with the molecular functions affected at the M42 sites. The most identified molecular functions are mostly binding receptors like heterocyclic receptor binding, organic cyclic compound binding and metal binding. The most identified molecular function is G-protein coupled receptor activity (Figure 19). This activity is highly influenced by insecticides, thiacloprid and imidacloprid.

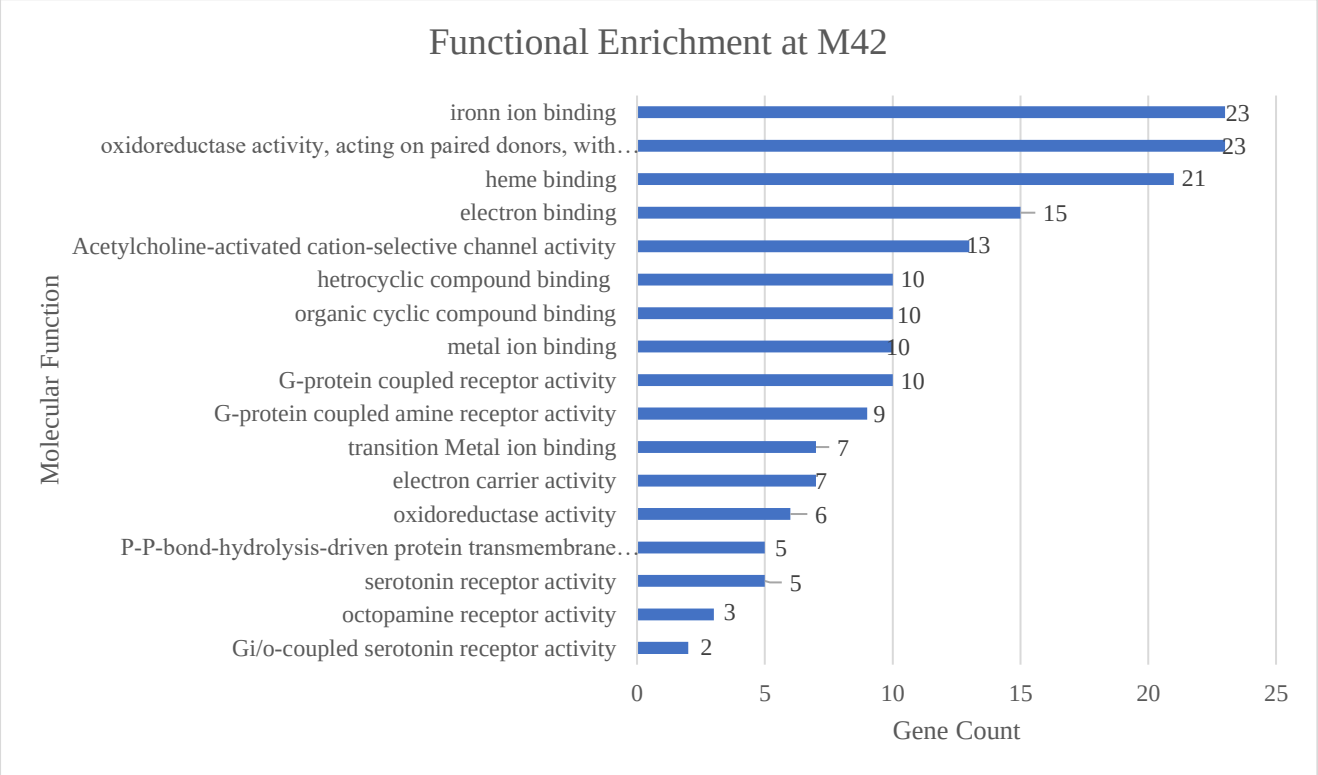



*Figure 18 Molecular Function predicted by chemical-gene-protein interaction at M42 site.*

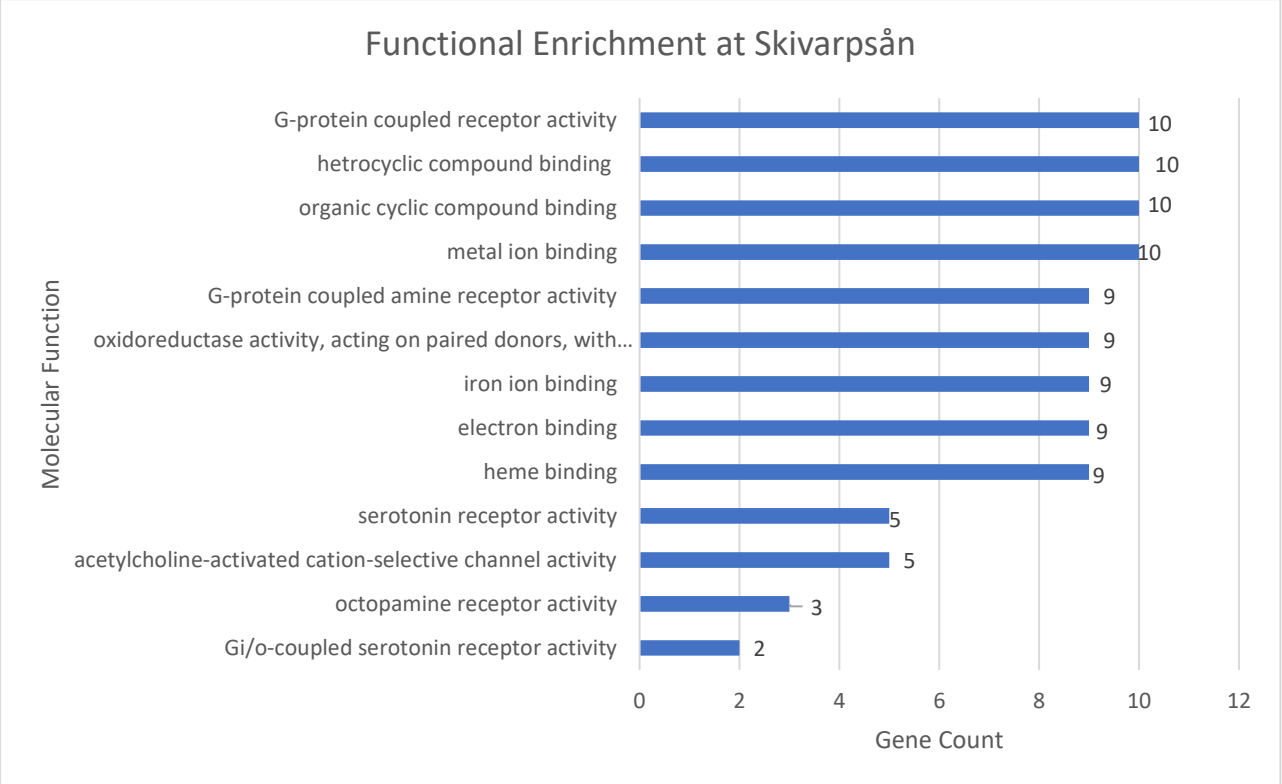



*Figure 19 Molecular Function predicted by chemical-gene-protein interaction at Skivarpsån site.*

### 4.5.4 Cellular Component:

Cells are the fundamental unit of life and are composed of various biomolecules and structures that carry out essential functions required for survival. One of the most important components of a cell is its plasma membrane, which serves as a physical barrier separating the intracellular environment from the extracellular environment. The plasma membrane also plays a crucial role in maintaining the integrity of the cell by controlling the transport of materials in and out of the cell (Alberts, 2004).

Chemical-gene interaction analysis is an incredible tool that allows us to analyze how chemicals interact with cellular components at a molecular level. The data obtained from this analysis can provide us with insights into the potentially harmful effects of chemicals on living organisms. In the context of pesticides, it can help us understand how these chemicals can affect the health of both humans and the environment.

According to the data presented in Figure 20 at the M42 site, the integral component of the plasma membrane is the most identified cellular component by insecticides such as imidacloprid, thiacloprid, and pirimicarb. The chemical gene/protein interaction analysis shows that the plasma membrane is identified 20 times overall on both sites. This indicates that these chemicals have a significant impact on the structure and function of the plasma membrane, which can have far-reaching consequences for the cell and the organism as a whole.

The synapse, post-synapse, and acetylcholine-gated channel complex are also mostly identified by the insecticides imidacloprid and thiacloprid, with 13 chemical gene/protein interactions observed for each of these cellular components. Synapses are the junctions between nerve cells that allow for the transmission of signals between neurons. The post-synapse is the area of the synapse that receives the signal, while the acetylcholine-gated channel complex is a type of ion channel that plays a role in the transmission of nerve impulses (Alberts, 2004).

At the Skivarpsån site, the cellular components that are being identified are overlapping. The highest gene count is for integral components of the plasma membrane which is 12.

The fact that these cellular components are mostly identified by imidacloprid and thiacloprid suggests that these chemicals can have a significant impact on the nervous system of organisms exposed to them. This can lead to a range of neurological symptoms and disorders, depending on the extent and duration of exposure.

The chemical-gene interaction analysis provides valuable insights into how pesticides such as imidacloprid, thiacloprid, and mecoprop can affect cellular components and their function. The data presented in (Figures 20 & 21) indicates the potentially harmful effects of these chemicals on the plasma membrane, synapses, post-synapses, and acetylcholine-gated channel complex.

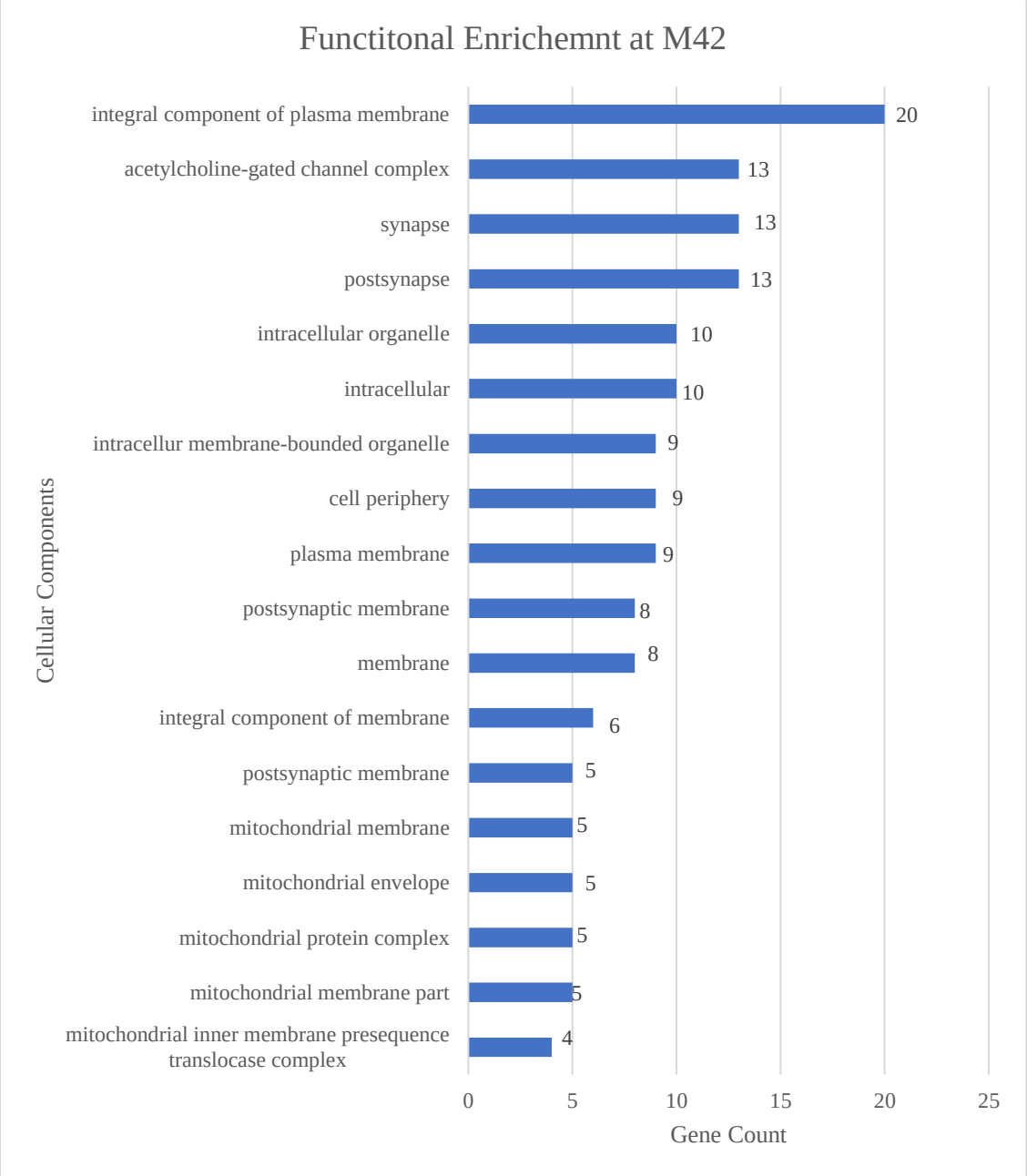



*Figure 20 Cellular Component predicted by chemical-gene-protein interaction at M42 site.*

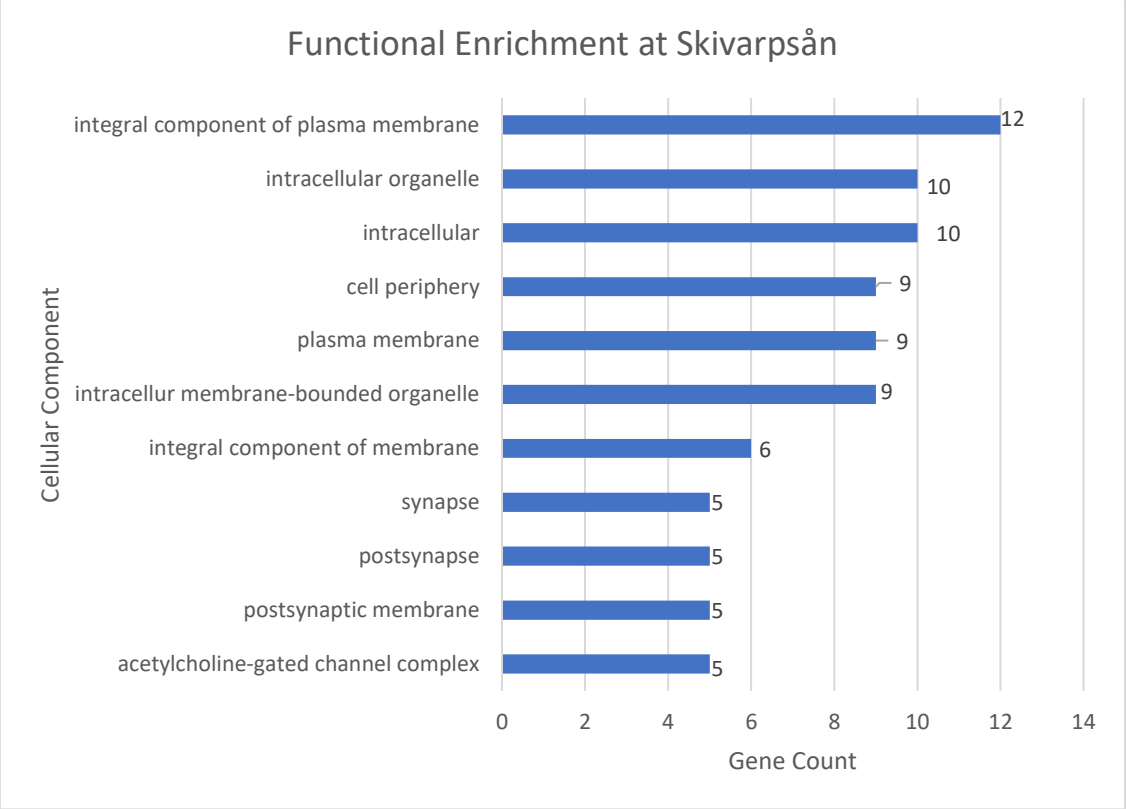



*Figure 21 Cellular Component predicted by chemical-gene-protein interaction at Skivarpsån site.*

### 4.5.5 KEGG Pathway:

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a highly popular database that provides extensive information on biological systems, such as genetic, chemical, and systemic aspects of living things. It contains a vast collection of pathway maps, which allow researchers to map genomic or transcriptome content of genes to KEGG reference pathways, thus facilitating the study of systemic behaviours of cells or organisms. One of the main applications of KEGG is in pathway mapping, where the database is used to infer the systemic behaviour of biological systems. For example, scientists can use KEGG to map the genomic content of genes to KEGG reference pathways to better understand the metabolic pathways and functions of cells or organisms (Kanehisa et al., 2006; Kanehisa et al., 2007; Kanehisa et al., 2012).

Figure 22 shows how the chemical-gene interactions affect three KEGG pathways at the M42 site, namely the biosynthesis of insect hormones, the nicotinate and nicotinamide effects, and neuroactive ligand-receptor interaction. These pathways are crucial for the proper functioning of cells and organisms, and the disruption of these pathways can have significant consequences.

The biosynthesis of insect hormones is being affected by the fungicide prochloraz, which is used to control fungal infections in crops. Insect hormones play a vital role in the regulation of growth and development in insects, and disruption of this pathway can lead to abnormal growth and development of insects.

The nicotinate and nicotinamide metabolism pathways are affected by the insecticides thiacloprid and imidacloprid. This pathway is responsible for the synthesis and breakdown of nicotinamide and nicotinate, which are essential for a wide range of cellular processes, including DNA repair and energy metabolism.

The neuroactive ligand-receptor interaction pathway is being affected by the insecticide imidacloprid. This pathway is responsible for the transmission of signals between nerve cells in the brain and plays a critical role in regulating a wide range of physiological processes, including mood, behaviour, and cognition.

At the Skivarpsån site, figure 23 only the neuroactive ligand-receptor and nicotinate and nicotinamide metabolism pathway have been identified.

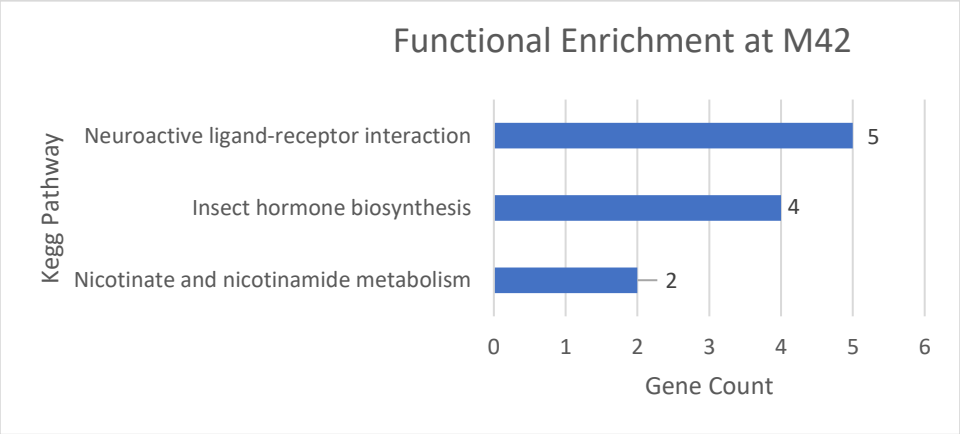



*Figure 22 KEGG Pathway predicted by chemical-gene-protein interaction M42 site.*

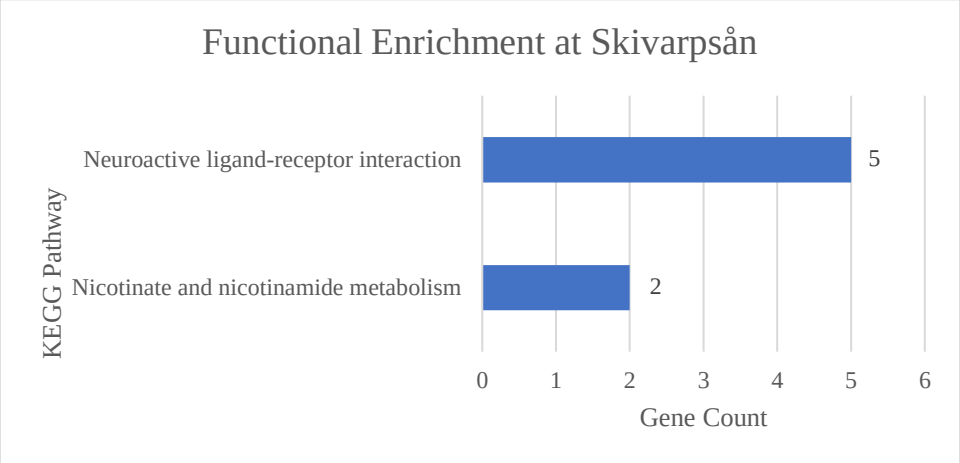



*Figure 23 KEGG Pathway predicted by chemical-gene-protein interaction at Skivarpsån site.*

.

# 4 Conclusion:

The study is carried out on two river sites, the M42 and Skivarpsån, which are highly contaminated with agricultural pesticides. Through the study, it was observed that the site of Skivarpsån is less infected with pesticides. The Skåne M42 has a larger ratio of pesticides due to this a higher number of interacting chemical genes were found on it than on Skivarpsån. The hypothesis of the study is majorly supported by the observations that are found at Skåne M42. The site of Skivarpsån is seen to have low node interactions in comparison to Skåne M42.

Furthermore, it was observed that there were a high number of chemical node interactions which have not been reported so far. It was also seen that among the chemical node interactions that have been reported there still was a very high number of node interactions which has not been found yet. Most of the findings only showed the negative and positive effects of the chemical node interactions and the remaining node interactions were marked as unspecified.

The function enrichment on both sides shows effects of chemical-gene-protein interactions on multiple biological aspects. Neonicotinoids such as thiacloprid, imidacloprid and pirimicarb highly influence theses biological aspects. At the M42 site a number of biological processes were identified which include the cation transport membrane process which is induced by insecticides. The most affected protein domains were neurotransmitter-gated ion channel and nicotinic acetylcholine receptors. The molecular functions such as the oxidation-reduction activities and cellular components such as the integral components of the plasma membrane and synapse were significantly influenced. The neuroactive ligand-receptor interaction pathway showed disruption.

At the Skivarpsån site, almost all the results overlapped with the M42 site. The protein domains like G protein-coupled receptors and Rhodopsin-like domains were highly influenced. Molecular functions, G-protein coupled receptor activity and multiple binding receptors were highly affected. The cellular components, integral components of the plasma membrane were most impacted. The neuroactive ligand-receptor pathway was affected.

Furthermore, the gene protein sequences will be used to identify the organisms the chemicals have affected.

# 5 Referencing:


Anderson, B., Phillips, B., Hunt, J., Siegler, K., Voorhees, J., Smalling, K., Kuivila, K., Hamilton, M., Ranasinghe, J. A., & Tjeerdema, R. (2013). Impacts of pesticides in a Central California estuary. *Environmental Monitoring and Assessment*, *186*(3), 1801–1814. https://doi.org/10.1007/s10661-013-3494-7

Backhaus, T., & Faust, M. (2012). Predictive Environmental Risk Assessment of Chemical Mixtures: A Conceptual Framework. *Environmental Science &Amp; Technology*, *46*(5), 2564–2573. https://doi.org/10.1021/es2034125

Blum M, Chang H, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, RichardsonL, Salazar GA, Williams L, Bork P, Bridge A, Gough J, Haft DH, Letunic I, Marchler-Bauer A, Mi H, Natale DA, Necci M, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Sillitoe I, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Bateman A and Finn RD The InterPro protein families and domains database: 20 years on. Nucleic Acids Research, Nov 2020, (doi: 10.1093/nar/gkaa977)

Bonada, N., Prat, N., Resh, V. H., & Statzner, B. (2006). DEVELOPMENTS IN AQUATIC INSECT BIOMONITORING: A Comparative Analysis of Recent Approaches. *Annual Review of Entomology*, *51*(1), 495–523. https://doi.org/10.1146/annurev.ento.51.110104.151124

Braisted, A. C., Oslob, J. D., Delano, W. L., Hyde, J., McDowell, R. S., Waal, N., Yu, C., Arkin, M. R., & Raimundo, B. C. (2003c). Discovery of a Potent Small Molecule IL-2 Inhibitor through Fragment Assembly. *Journal of the American Chemical Society*, *125*(13), 3714–3715. https://doi.org/10.1021/ja034247i

Bundschuh, M., Zubrod, J. P., Klemm, P., Elsaesser, D., Stang, C., & Schulz, R. (2013). Effects of peak exposure scenarios on Gammarus fossarum using field-relevant pesticide mixtures. *Ecotoxicology and Environmental Safety*, *95*, 137–143. https://doi.org/10.1016/j.ecoenv.2013.05.025

Carnevali, O., Santangeli, S., Forner-Piquer, I., Basili, D., & Maradonna, F. (2018). Endocrine-disrupting chemicals in aquatic environment: what are the risks for fish gametes? *Fish Physiology and Biochemistry*, *44*(6), 1561–1576. https://doi.org/10.1007/s10695-018-0507-z

Carpenter, D. O., Arcaro, K., & Spink, D. C. (2002). Understanding the human health effects of chemical mixtures. *Environmental Health Perspectives*, *110*(suppl 1), 25–42. https://doi.org/10.1289/ehp.02110s125

Carsten von der Ohe, P., & Liess, M. (2004). RELATIVE SENSITIVITY DISTRIBUTION OF AQUATIC INVERTEBRATES TO ORGANIC AND METAL COMPOUNDS. *Environmental Toxicology and Chemistry*, *23*(1), 150. https://doi.org/10.1897/02-577

Céspedes, R., Urrutia, R., & Barra, R. (2018). Acute toxicity of glyphosate, AMPA, and its mixture on freshwater shrimp (Neocaridina denticulata) and early development of the crayfish (Parastacus sp.). Environmental Science and Pollution Research, 25(20), 19696-19706. doi: 10.1007/s11356-018-2153-3

Chaumot, A., Geffard, O., Armengaud, J., & Maltby, L. (2015). Gammarids as Reference Species for Freshwater Monitoring. Elsevier EBooks, 253–280. https://doi.org/10.1016/b978-0-12-800949-9.00011-5

Chung, N. C., Miasojedow, B., Startek, M., & Gambin, A. (2019). Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinformatics*, *20*(S15). https://doi.org/10.1186/s12859-019-3118-5

De Castro-Català, N., Muñoz, I., Riera, J., & Ford, A. (2017b). Evidence of low dose effects of the antidepressant fluoxetine and the fungicide prochloraz on the behavior of the keystone freshwater invertebrate Gammarus pulex. *Environmental Pollution*, *231*, 406–414. https://doi.org/10.1016/j.envpol.2017.07.088

Denise Forson, D., & Storfer, A. (2006). ATRAZINE INCREASES RANAVIRUS SUSCEPTIBILITY IN THE TIGER SALAMANDER,*AMBYSTOMA TIGRINUM*. *Ecological Applications*, *16*(6), 2325–2332. https://doi.org/10.1890/1051-0761(2006)016

EPA. (2021, June 8). Basic Information about Pesticide Ingredients. US EPA.

https://www.epa.gov/ingredients-used-pesticide-products/basic-information-about-pesticide-ingredients

Grasberger, B. L., Lu, T., Schubert, C., Parks, D. J., Carver, T. E., Koblish, H. K., Cummings, M. D., LaFrance, L. V., Milkiewicz, K. L., Calvo, R. R., Maguire, D., Lattanze, J., Franks, C. F., Zhao, S., Ramachandren, K., Bylebyl, G. R., Zhang, M., Manthey, C. L., Petrella, E. C., . . . Bone, R. F. (2005). Discovery and Cocrystal Structure of Benzodiazepinedione HDM2 Antagonists That Activate p53 in Cells. *Journal of Medicinal Chemistry*, *48*(4), 909–912. https://doi.org/10.1021/jm049137g

Håkansson, A. (2017). *Effects of environmentally occurring levels of pesticide pollution on the structure, function, and nutritional quality of periphytic algae in agricultural streams in Skåne, Sweden* [Masters Thesis]. University of Gothenburg.

Hernández, A. F., Parrón, T., Tsatsakis, A. M., Requena, M., Alarcón, R., & López-Guarnido, O. (2013). Toxic effects of pesticide mixtures at a molecular level: Their relevance to human health. *Toxicology*, *307*, 136–145. https://doi.org/10.1016/j.tox.2012.06.009

Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology*, *4*(11), 682–690. https://doi.org/10.1038/nchembio.118

Hopkins, A. L., Groom, C. R., & Alex, A. (2004). Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today*, *9*(10), 430–431. https://doi.org/10.1016/s1359-6446(04)03069-7

Inostroza, P. A., Vera-Escalona, I., Wicht, A. J., Krauss, M., Brack, W., & Norf, H. (2016).

Anthropogenic Stressors Shape Genetic Structure: Insights from a Model Freshwater Population

along a Land Use Gradient. Environmental Science &Amp; Technology, 50(20), 11346–11356. https://doi.org/10.1021/acs.est.6b04629

Inostroza, P. A., Massei, R., Wild, R., Krauss, M., & Brack, W. (2017). Chemical activity and distribution of emerging pollutants: Insights from a multi-compartment analysis of a freshwater system. Environmental Pollution, 231, 339–347. https://doi.org/10.1016/j.envpol.2017.08.015

Inostroza, P. A., Vera-Escalona, I., Wild, R., Norf, H., & Brauns, M. (2018a). Tandem Action of Natural and Chemical Stressors in Stream Ecosystems: Insights from a Population Genetic Perspective. *Environmental Science &Amp; Technology*, *52*(14), 7962–7971. https://doi.org/10.1021/acs.est.8b01259

Jin, L., Wang, W., & Fang, G. (2014). Targeting Protein-Protein Interaction by Small Molecules. *Annual Review of Pharmacology and Toxicology*, *54*(1), 435–456. https://doi.org/10.1146/annurev-pharmtox-011613-140028

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., & Hirakawa, M. (2006). From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research*, *34*(90001), D354–D357. https://doi.org/10.1093/nar/gkj102

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S.,

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, *40*(D1), D109–D114. https://doi.org/10.1093/nar/gkr988

Kehrer, J. P., & Klotz, L. (2015). Free radicals and related reactive species as mediators of tissue injury and disease: implications for Health. Critical Reviews in Toxicology, 45(9), 765–798. https://doi.org/10.3109/10408444.2015.1074159

Klečka, G., Persoon, C., & Currie, R. (2010). Chemicals of Emerging Concern in the Great Lakes

Basin: An Analysis of Environmental Exposures. Reviews of Environmental Contamination and

Toxicology, 1–93. https://doi.org/10.1007/978-1-4419-6406-9_1

Kuhn, M., von Mering, C., Campillos, M., Jensen, L. J., & Bork, P. (2007). STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Research*, *36*(Database), D684–D688. https://doi.org/10.1093/nar/gkm795

Kuzmanović, M., López-Doval, J. C., De Castro-Català, N., Guasch, H., Petrović, M., Muñoz, I., Ginebreda, A., & Barceló, D. (2016). Ecotoxicological risk assessment of chemical pollution in four Iberian river basins and its relationship with the aquatic macroinvertebrate community status. *Science of the Total Environment*, *540*, 324–333. https://doi.org/10.1016/j.scitotenv.2015.06.112

Laetz, C. A., Baldwin, D. H., Collier, T. K., Hebert, V., Stark, J. D., & Scholz, N. L. (2009b). The Synergistic Toxicity of Pesticide Mixtures: Implications for Risk Assessment and the Conservation of Endangered Pacific Salmon. *Environmental Health Perspectives*, *117*(3), 348–353. https://doi.org/10.1289/ehp.0800096

Li, X., Wang, X., & Snyder, M. (2012). Systematic investigation of protein-small molecule interactions. *IUBMB Life*, *65*(1), 2–8. https://doi.org/10.1002/iub.1111

Moulton, D. E., Oliveri, H., & Goriely, A. (2020). Multiscale integration of environmental stimuli in plant tropism produces complex behaviors. Proceedings of the National Academy of Sciences of the United States of America, 117(51), 32226–32237. https://doi.org/10.1073/pnas.2016025117

McKinney, J. D. (1985). The molecular basis of chemical toxicity. Environmental Health

Perspectives, 61, 5–10. https://doi.org/10.1289/ehp.85615

Mahmood, I., Imadi, S. R., Shazadi, K., Gul, A., & Hakeem, K. R. (2016). Effects of Pesticides on Environment. *Plant, Soil and Microbes*, 253–269. https://doi.org/10.1007/978-3-319-27455-3_13

McFedries, A., Schwaid, A., & Saghatelian, A. (2013). Methods for the Elucidation of Protein-Small Molecule Interactions. *Chemistry &Amp; Biology*, *20*(5), 667–673. https://doi.org/10.1016/j.chembiol.2013.04.008

Oltersdorf, T., Elmore, S. W., Shoemaker, A. R., Armstrong, R. C., Augeri, D. J., Belli, B. A., Bruncko, M., Deckwerth, T. L., Dinges, J., Hajduk, P. J., Joseph, M. K., Kitada, S., Korsmeyer, S. J., Kunzer, A. R., Letai, A., Li, C., Mitten, M. J., Nettesheim, D. G., Ng, S., . . . Rosenberg, S. H. (2005). An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature*, *435*(7042), 677–681. https://doi.org/10.1038/nature03579

Pilling, E. D., & Jepson, P. C. (1993). Synergism between EBI fungicides and a pyrethroid insecticide in the honeybee (*Apis mellifera*). *Pesticide Science*, *39*(4), 293–297. https://doi.org/10.1002/ps.2780390407

Rasmussen, J. J., Wiberg-Larsen, P., Baattrup-Pedersen, A., Monberg, R. J., & Kronvang, B. (2012). Impacts of pesticides and natural stressors on leaf litter decomposition in agricultural streams. *Science of the Total Environment*, *416*, 148–155. https://doi.org/10.1016/j.scitotenv.2011.11.057

Reeder, B. J. (2010). The Redox Activity of Hemoglobins: From Physiologic Functions to Pathologic Mechanisms. *Antioxidants & Redox Signaling*, *13*(7), 1087–1123. https://doi.org/10.1089/ars.2009.2974

Riva, C., Sokolowski, M. B. C., Normand, J., Santos, J. S. O., & Halm-Lemeille, M. (2018).

Effect of oral exposure to the acaricide pirimicarb, a new varroacide candidate, on Apis mellifera feeding rate. *Pest Management Science*, *74*(8), 1790–1797. https://doi.org/10.1002/ps.4876

Rubach, M. N., Baird, D. J., & Van den Brink, P. J. (2009). A new method for ranking mode-specific sensitivity of freshwater arthropods to insecticides and its relationship to biological traits. *Environmental Toxicology and Chemistry*, *29*(2), 476–487. https://doi.org/10.1002/etc.55

Schulz, R., & Dabrowski, J. M. (2001). Combined effects of predatory fish and sublethal pesticide

contamination on the behavior and mortality of mayfly nymphs. Environmental Toxicology

and Chemistry, 20(11), 2537–2543. https://doi.org/10.1002/etc.5620201120

Schäfer, R. B., Pettigrove, V., Rose, G., Allinson, G., Wightwick, A., von der Ohe, P. C., Shimeta, J.,

Kühne, R., & Kefford, B. J. (2011). Effects of Pesticides Monitored with Three Sampling

Methods in 24 Sites on Macroinvertebrates and Microorganisms. Environmental Science

& Amp; Technology, 45(4), 1665–1672. https://doi.org/10.1021/es103227q

Schäfer, R. B. (2019). Responses of freshwater macroinvertebrates to pesticides: insights from field studies. *Current Opinion in Environmental Science &Amp; Health*, *11*, 1–7. https://doi.org/10.1016/j.coesh.2019.06.001

Schüttler, A., Altenburger, R., Ammar, M., Bader-Blukott, M., Jakobs, G., Knapp, J., Krüger, J., Reiche, K., Wu, G. M., & Busch, W. (2019). Map and model—moving from observation to prediction in toxicogenomics. *GigaScience*, *8*(6). https://doi.org/10.1093/gigascience/giz057

Siddique, A., Liess, M., Shahid, N., & Becker, J. M. (2020). Insecticides in agricultural streams exert pressure for adaptation but impair performance in Gammarus pulex at regulatory acceptable concentrations. *Science of the Total Environment*, *722*, 137750. https://doi.org/10.1016/j.scitotenv.2020.137750

Silvestre, F. (2020). Signaling pathways of oxidative stress in aquatic organisms exposed to xenobiotics. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *333*(6), 436–448. https://doi.org/10.1002/jez.2356

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Szklarczyk, D., Santos, A., von Mering, C., Jensen, L. J., Bork, P., & Kuhn, M. (2015). STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Research*, *44*(D1), D380–D384. https://doi.org/10.1093/nar/gkv1277

Szklarczyk, D., Santos, A. B., Von Mering, C., Jensen, L. J., Bork, P., & Kuhn, M. (2016b). STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Research*, *44*(D1), D380–D384. https://doi.org/10.1093/nar/gkv1277

Taube, F., Krawinkel, M., Susenbeth, A., & Theobald, W. (2011). The booklet “Genetically modified crops” published from the German Research Foundation, does not meet the given claim. *Environmental Sciences Europe*, *23*(1). https://doi.org/10.1186/2190-4715-23-1

Tse, C., Shoemaker, A. R., Adickes, J., Anderson, M. G., Chen, J., Jin, S., Johnson, E. F., Marsh, K. C., Mitten, M. J., Nimmer, P., Roberts, L., Tahir, S. K., Xiao, Y., Yang, X., Zhang, H., Fesik, S., Rosenberg, S. H., & Elmore, S. W. (2008). ABT-263: A Potent and Orally Bioavailable Bcl-2 Family Inhibitor. *Cancer Research*, *68*(9), 3421–3428. https://doi.org/10.1158/0008-5472.can-07-5836

Ul Hassan, M., Noreen, Z., & Ahmed, R. (2021). Regional frequency analysis of annual daily rainfall maxima in Skåne, Sweden. *International Journal of Climatology*, *41*(8), 4307–4320. https://doi.org/10.1002/joc.7074

United States Environmental Protection Agency (USEPA), 2005. Reregistration eligibility decision for 2,4-D. In: USEPA. Report No. EPA 738-R-05-002.

Vassilev, L. T. (2004). Small-Molecule Antagonists of p53-MDM2 Binding: Research Tools and Potential Therapeutics. *Cell Cycle*, *3*(4), 417–419. https://doi.org/10.4161/cc.3.4.801

Waiser, M. J., & Holm, J. (2005). Cumulative and synergistic effects of complex mixtures of herbicides on wetland biodiversity: Implications for beneficial herbicide application practice and environmental standards for Prairie wetlands. In NWRI technical report# AEP-TN05-003.

National Water Research Institute Saskatoon, SK, Canada.Wallace, J. B., & Webster, J. R. (1996). The Role of Macroinvertebrates in Stream Ecosystem Function. *Annual Review of Entomology*, *41*(1), 115–139. https://doi.org/10.1146/annurev.en.41.010196.000555

Yang, E. C., Chuang, Y. C., Chen, Y. L., & Chang, L. H. (2008). Abnormal Foraging Behavior Induced by Sublethal Dosage of Imidacloprid in the Honey Bee (Hymenoptera: Apidae). *Journal of Economic Entomology*, *101*(6), 1743–1748. https://doi.org/10.1603/0022-0493-101.6.1

Zeliger, H. (2011). *Human Toxicology of Chemical Mixtures* (2nd ed.). William Andrew.

# 6 Appendix.

*Table 1 Chemicals detected at the M42 and Skivarpsån Site at different concentrations in ng/L.*

| Chemical name | Type | M42 | Skivarpsån | MoA |
|---|---|---|---|---|
| AMPA | Herbicide | 170 | 61 | glutamate receptor agonist |
| Atrazine | Herbicide | 2 | - | Photosynthesis inhibition |
| Desethylatrazine | Herbicide | 2 | - | Photosynthesis inhibition |
| Azoxystrobin | Fungicide | 2 | - | Respiration inhibition |
| BAM | Herbicide | 55 | 8 | Root-Growth inhibition |
| Bentazone | Herbicide | 8 | 7 | Photosynthesis inhibition |
| Bixafen | Fungicide | 23 | 5 | Respiration inhibition |
| Boscalid | Fungicide | 14 | - | Respiration inhibition |
| Diflufenican | Herbicide | 23 | 10 | Carotenoid biosynthesis inhibition |
| Diuron | Herbicide | 9 | - | Photosynthesis inhibition |
| Ethofumesate | Herbicide | 25 | 16 | Lipid biosynthesis inhibition |
| Fluopyram | Fungicide | 73 | 37 | Respiration inhibition |
| Fluroxypyr | Herbicide | 55 | 29 | Synthetic auxin |
| Glyphosate | Herbicide | 210 | 27 | Protein biosynthesis inhibition |
| Imidacloprid | Insecticide | 2 | - | Neuroactive |

| | | | | |
|---|---|---|---|---|
| Carbendazim | Fungicide | 3 | - | Mitosis, cell cycle |
| Clomazone | Herbicide | 2 | 1 | Photosynthesis inhibition |
| Clopyralid | Herbicide | 65 | 29 | Synthetic auxin |
| Chloridazon | Herbicide | 7 | 33 | Photosynthesis inhibition |
| Quinmerac | Herbicide | 1 | 7 | Synthetic auxin |
| MCPA | Herbicide | 50 | 80 | Synthetic auxin |
| Mecoprop | Herbicide | 5 | 13 | Synthetic auxin |
| Methabenzthiazuron | Herbicide | 3 | - | Photosynthesis inhibition |
| Metamitron | Herbicide | 120 | 13 | Photosynthesis inhibition |
| Metsulfuron-methyl | Herbicide | 8 | - | Mitosis, cell cycle |
| Prochloraz | Fungicide | 6 | - | Sterol biosynthesis inhibition |
| Propyzamide | Herbicide | 15 | - | Mitosis, cell cycle |
| Prothioconazole-desthio | Fungicide | 41 | - | CYP51 inhibition |
| Pyroxsulam | Herbicide | 7 | - | Protein biosynthesis inhibition |
| Tebuconazole | Fungicide | 10 | - | Sterol biosynthesis inhibition |
| Terbuthylazine | Herbicide | 3 | - | Photosynthesis inhibition |
| Terbutylazin-desethyl | Herbicide | 9 | 4 | Photosynthesis inhibition |
| Thiacloprid | Insecticide | 5 | 7 | Neuroactive |

| | | | | |
|---|---|---|---|---|
| Triflusulfuron-methyl | Herbicide | 12 | 8 | Protein biosynthesis inhibition |