Mini Review

Evolutionary Toxicology and Genomics: Utility of Natural Populations' Sensitivity and Resistance Mechanisms to Persistent Organic Pollutants in Ecological Risk Assessment and Public Health

Goran Bozinovic*

Division of Biological Sciences, University of California San Diego, USA

***Corresponding author:** Goran Bozinovic, Boz Life Science Research and Teaching Institute, USA.

Abstract

Susceptibility to toxic chemical exposures is determined by genotype-environment interactions. Genetically variable natural populations live in spatially, temporally, and chemically complex environments, and their exposures to persistent organic pollutants should be carefully considered when quantifying evolution's role in mechanisms of sensitivity and resistance. Although most natural populations' genomes are not sequenced, reasonably robust lower-throughput custom platforms and data imputation methods help infer missing genotypes and improve the analysis of non-model organisms' sequencing data. Such datasets yield insights about population processes and variance of allelic diversity within and between sensitive and resistant populations. The advantage of this approach is that an a priori choice of biomarkers of an adverse effect is not necessary. Population genomics provide insights into the sequence variations that define differences in gene expression and protein polymorphisms underlying mechanisms of sensitivity and resistance to polluted environment. Complementary "omics" approaches with model-organism and natural populations offer a better understanding of biological effects and should be an integral component of comprehensive ecological and human health risk assessment.

Keywords: Genomics; Evolution; Risk assessment; Transcriptomics; Toxicology; Public health; Natural populations

Introduction

The prevalence of today's environmental contaminants is unprecedented. The Toxic Substance Control Act Chemical Substance Inventory lists over 85,000 chemicals in the U.S. market [1]. Adapting to persistent organic pollutants (POPs) exposures is difficult, particularly if adaptive responses require novel genetic variation [2]. An individual's fitness is influenced by the sum of all stressors, which can be additive, antagonistic, and/or synergistic [3-5]. Devising effective long-term risk-assessment and mitigation strategies is complicated further by human-induced selection pressures on natural populations (climate change, habitat conversion…), jeopardizing organisms' abilities to survive and adapt to acute and chronic environmental contaminants. Identifying and quantifying such adaptations in individuals and populations in their natural setting is difficult. Yet, technological advances in next generation sequencing (NGS) and transcriptomics provide an opportunity to infer statistically and biologically relevant information within complex but manageable data sets.

POPs are toxic organic compounds resistant to environmental degradation. Most are anthropogenic, lipid-soluble and can cause adverse impacts on the environment and human health. Many POPs are industrial chemicals - solvents, pesticides, insecticides and fungicides, pharmaceuticals, polychlorinated biphenyls (PCBs), dioxins (PCDD), polyfluorinated dibenzofurans (PCDFs), etc. [6,7]. Due to their general use, POPs have high environmental concentrations; many are hydrophobic and bioaccumulate through the food chain due to their slow rate of metabolism [8-10]. Human and animal exposures are mostly through diet (90%), environmental exposures, and during embryo/fetal development via maternal deposition. Chronic POPs exposures can result in acute and chronic toxicity, demonstrated by developmental defects, chronic illnesses, and increased mortality [6]. Some are categorized as carcinogens [11] while many are endocrine disruptors [12-14] and may affect nervous, immune, digestive, urinary and reproductive systems [6]. Low-level exposure to POPs during critical fetal development stages can have a lifetime effect [12,15]. POPs concentrations in human

serum increase with age and seem to be higher in females than males [16]. Evolutionary responses and adaptations to POPs have been documented since the mid-20th century; initially as pesticide resistance cases among invertebrates and plants [17,18], then among other natural populations, including vertebrates and mammals [19,20]. Although industrial toxic chemicals were provoking rapid evolutionary response among natural populations, their effects were mostly ignored.

To evaluate temporal and spatial exposure risks, government agencies study the most environmentally prevalent POPs by examining bioavailability and dose-response relationships in the laboratory setting [21]. Such assessments are limited within relatively static laboratory settings. For organisms exposed to a mixture of POPs, the adverse effects are assumed to be additive as mixtures of POPs can result in synergistic effects, enhancing the toxicity of each compound [3-5]. However, responses to POPs are complex and should be carefully considered when quantifying evolution's role in mechanisms of sensitivity and resistance to complex mixtures of pollutants within and between natural, genetically variable populations [22,23]. Evolutionary changes due to a chronic POP exposure can be quantified in just a few generations [24,25] and include physiological, genetic and epigenetic transgenerational effects influenced by natural selection: studies utilizing natural populations report phenotypic and molecular consequences, including increased mutation rates, receptor desensitization, phenological traits and epigenetic effects [26,27]. While some exposure effects are adaptive, others reduce population size, causing inbreeding depression or genetic drift [28,29] and decreased genetic variation, compromising the ability to adapt to future stressors. Adaptation to POPs can evolve at a cost, such as increased subsequent sensitivity to oxidative stress [30]. Risk assessment of environmental exposures to complex chemical mixtures should not be limited to often oversimplified and outdated controlled laboratory bioassays focused on robust acute and chronic toxicity endpoints such as survival and reproductive viability of model organisms lacking inherent genetic variation.

The post-industrial age of overwhelmingly complex chemical exposures coupled by recent technological advancements in NGS methodology allows for the utility of diverse non-model organisms within an evolutionary context in public health risk assessment. Scientists started using population genomic approaches to study natural populations lacking robust genomic resources a decade ago. Given that genome sequences of non-model organisms are accumulating at an unprecedented pace, with 234 animal and 319 contig genome assemblies currently in development [31], technological advancements in high-throughput sequencing from non-model organisms generate critically important data to infer complex genotype-environment interactions in a natural setting. Population genomic studies target the underlying variation found in the DNA among individuals and populations. They also utilize genome-wide sampling of sequence variation under the premise that demography and the evolutionary history of populations affect neutral loci similarly, whereas loci under selection will be affected differently [32]. Data of multiple, although individually analyzed, loci from natural populations enables studies of the genomic evolution, acclimatization, and adaptation to strong selective pressure such as pollution exposure in natural environment. Such data sets can yield insights about population processes and variance of allelic diversity within and between populations. The advantage is that an a priori choice of biomarkers of an adverse effect is not necessary, which is important when organisms are exposed to complex chemical mixtures in their natural environment and biomarkers of exposure are not well established [33]. Moreover, adaptation-related projects looking for genomic signatures of selection within and between natural populations can take advantage of established adaptive phenotypes to environmental gradients.

A typical population genomic study platform consists of sequencing strategy design, generation of sequence data, mapping of sequence reads to the assembly, genotyping, and population genetic/molecular evolutionary analyses. A sequencing strategy includes depth of coverage and utility of individuals vs. pooled samples, number of individuals per population and number of populations, gender, and the need for outgroup species. High-throughput genotyping of millions of markers simultaneously is available for model organisms. Whole transcriptome sequencing has advantages but is expensive since multiple individuals need to be sequenced to represent the population adequately. Moreover, the analysis of short sequence reads depends on an incomplete reference genome to which sequence reads are aligned. Such sequencing platforms generate massive data sets, meaning that inferring statistical and biological relevance can be challenging. Although most non-model natural population genomes are not yet sequenced, reasonably robust, lower throughput, custom platforms are available. Expressed sequence tags (ESTs) have been developed and used as genome-wide gene expression data within and among natural fish populations historically exposed to chemical pollution. New sequencing technologies such as Roche (454) FLX, Illumina Genome Analyzer, ABI SOLiD, and HeliScope sequencing [33] can be used to sequence thousands of transcripts quickly and cost-effectively, making transcriptomic studies possible in virtually any species. Transcriptomics and population genomics can influence studies of diverse species and populations that ask important ecological, physiological, and evolutionary questions with respect to pollutant exposure. High-throughput sequencing allows enough sequencing depth to gain an adequate representation of all the expressed transcripts and has been used for whole transcriptome sequencing [34]. The coding sequences of the expressed transcripts can be analyzed for mutations, altered splice sites, and protein polymorphisms.

Whole transcriptome approaches can identify significant changes in gene expression that are biologically important, but

the genomic basis underlying altered patterns of gene expression is mostly unknown. Once many genetic markers became available, population geneticists began scanning genomes for reduced nucleotide diversity, extended linkage disequilibrium, or regions of homozygosity as potential selection signatures. This approach requires that large numbers of loci and genetic markers are statistically analyzed for non-random patterns [35]. At loci under selection, local adaptation and directional selection should reduce genetic variability within populations and increase variation among populations. Loci used in population genomics studies include microsatellites, single nucleotide polymorphisms (SNPs), amplified fragment length polymorphisms (AFLPs), randomly amplified polymorphic DNA (RAPDs), and sequences (e.g., whole-genome sequences). Many genetic markers can be generated without a sequenced genome relatively easily and genomes can be mined without measuring phenotypes, allowing for sampling of individuals without knowing their breeding history. For example, SNP analysis can be performed by mining EST data collections for putative SNPs [36]. If the ESTs were derived from multiple individuals (and multiple populations of interest), putative SNPs can be identified in sequence alignments. Conveniently, any SNPs identified as important through population genomic approaches are already associated with a sequenced gene transcript.

Custom SNP genotyping platforms provide efficient genotyping of thousands of individuals, resulting in biologically important sequence variations from multiple populations. The advantage of AFLPs and RAPDs is that, again, no prior sequence information is necessary. The drawback is that AFLP and RAPD analyses depend on high-quality DNA and provide only dominant markers, so heterozygotes cannot be directly measured, and generated markers are often anonymous or in a nonfunctional area of the genome. Genome-wide association studies (GWAS) are challenging with non-model organisms because they lack an abundance of readily available, high quality, sequenced genomes, and genome-wide genotyping data. To improve the use of NGS data obtained from non-model organisms, new imputation methods such as Link Imputer [37] help infer missing genotypes and facilitate the analysis of non-model organisms sequencing data.

Susceptibility to toxic substances depends on the organism's genotype and interactions with the polluted environment. Natural populations inhabit spatially and temporally uncontrolled environments where pollutants are present as complex mixtures, so the challenges of "omics" technologies with natural populations are the challenges of basic biological research: which data is biologically meaningful? Considering cost, challenging sample material, lack of biological replicates, and complexity of genotype-environment interactions, choosing an optimal method is difficult. Yet, instead of a single gene, protein, or polymorphism, one can analyze thousands of genes, proteins, and polymorphisms utilizing biologically realistic, genetically variable individuals, extrapolating a population response rather than a strain-specific one. Complementary

"omics" approaches offer a better understanding of biological effects at multiple levels: model-organisms are used to study many important questions in biology, population genomics provide insights into the sequence variations that govern differences in gene expression and protein polymorphisms, while transcriptomics provides insights into altered protein levels. A focused approach, using known POP inducers and inhibitors present in such mixtures and analyzing responses of cellular and tissue targets, can provide insights into mechanisms of toxicity seen in natural populations and help devise a more comprehensive risk assessment strategy. Thus, just as the integration of laboratory and field studies improves our knowledge of genes and proteins, so is the integration of laboratory and field studies critical for "omic" approaches and public health risk assessment modeling.

References

1. (2019) About the TSCA Chemical Substance Inventory. United States Environmental Protection.

2. Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. Trends in Ecology & Evolution 23(1): 38-44.

3. Wassenberg DM, Di Giulio RT (2004) Synergistic Embryotoxicity of Polycyclic Aromatic Hydrocarbon Aryl Hydrocarbon Receptor Agonists with Cytochrome P4501A Inhibitors in Fundulus heteroclitus. Environ Health Perspect 112(17): 1658-1664.

4. Billiard SM, Timme Laragy AR, Wassenberg DM, Cockman C, Di Giulio RT (2006) The Role of the Aryl Hydrocarbon Receptor Pathway in Mediating Synergistic Developmental Toxicity of Polycyclic Aromatic Hydrocarbons to Zebrafish. Toxicol Sci 92(2): 526-536.

5. Bozinovic G, Sit TL, Di Giulio R, Wills LF, Oleksiak MF (2013) Genomic and physiological responses to strong selective pressure during late organogenesis: few gene expression changes found despite striking morphological differences. BMC Genomics 14(1): 779.

6. Ritter LA, Solomon KR, Forget JF, Stemeroff M, leary CO (2001) PERSISTENT ORGANIC POLLUTANTS DDT-Aldrin-Dieldrin-Endrin-Chlordane Heptachlor-Hexachlorobenzene Mirex-Toxaphene Polychlorinated Biphenyls Dioxins and Furans. 1-149.

7. United Nations Environment Programme (2016) Listing of POPs in the Stockholm Convention.

8. Kelly BC, Ikonomou MG, Blair JD, Morin AE, Gobas FAPC (2007) Food web-specific biomagnification of persistent organic pollutants. Science 31(5835):236-239.

9. Wania F, MacKay D (1996) Peer Reviewed: Tracking the Distribution of Persistent Organic Pollutants. Environ Sci Technol 30(9): 390A-396A.

10. Lohmann R, Breivik K, Dachs J, Muir D (2007) Global fate of POPs: Current and future research directions. Environmental Pollution 150(1):150-165.

11. WHO (2019) IARC – International Agency for Research on Cancer.

12. Damstra T (2002) Potential effects of certain persistent organic pollutants and endocrine disrupting chemicals on the health of children. J Toxicol Clin Toxicol 40(4): 457-465.

13. Chen H, Wang C, Li H, Ma R, Yu Z, et al. (2019) A review of toxicity induced by persistent organic pollutants (POPs) and endocrine-disrupting chemicals (EDCs) in the nematode *Caenorhabditis elegans*. Journal of Environmental Management 237: 519-525.

14. Gregoraszczuk EL, Ptak A (2013) Endocrine-Disrupting Chemicals: Some Actions of POPs on Female Reproduction. Int J Endocrinol 2013: 828532.

15. Dewan P, Jain V, Gupta P, Banerjee BD (2013) Organochlorine pesticide residues in maternal blood, cord blood, placenta, and breastmilk and their relation to birth size. Chemosphere 90(5): 1704-1710.

16. Vallack HW, Bakker DJ, Brandt I, Broström Lundén E, Brouwer A, et al. (1998) Controlling persistent organic pollutants–what next? Environ Toxicol Pharmacol 6(3): 143-175.

17. KEIDING J, VAN DEURS H (1949) D.D.T.-Resistance in Houseflies in Denmark. Nature 163(4155): 964-965.

18. Metcalf RL, March RB (1949) Studies of the Mode of Action of Parathion and Its Derivatives and Their Toxicity to Insects1. J Econ Entomol 42(5): 721-728.

19. Ishizuka M, Tanikawa T, Tanaka KD, Heewon M, Okajima F, et al. (2008) Pesticide resistance in wild mammals--mechanisms of anticoagulant resistance in wild rodents. J Toxicol Sci 33(3): 283-289.

20. Georghiou GP, Taylor CE (1977) Pesticide resistance as an evolutionary phenomenon [Insects, acarines].

21. Szabo DT, Loccisano AE (2012) POPs and Human Health Risk Assessment. In: Dioxins and Health Hoboken, NJ, USA. pp. 579-618.

22. Bozinovic G, Oleksiak MF (2010) Embryonic gene expression among pollutant resistant and sensitive Fundulus heteroclitus populations. Aquat Toxicol 98(3): 221-229.

23. Oleksiak MF (2008) Changes in gene expression due to chronic exposure to environmental pollutants. Aquat Toxicol 90(3): 161-171.

24. Hendry AP, Gotanda KM, Svensson EI (2017) Human influences on evolution, and the ecological and societal consequences. Philos Trans R Soc B Biol Sci 372(1712): 20160028.

25. Schoener TW (2011) The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. Science 331(6016): 426-429.

26. Crews D, Gore AC (2012) Epigenetic synthesis: a need for a new paradigm for evolution in a contaminated world. F1000 Biol Rep 4: 18.

27. Oziolor EM, Bigorgne E, Aguilar L, Usenko S, Matson CW (2014) Evolved resistance to PCB- and PAH-induced cardiac teratogenesis, and reduced CYP1A activity in Gulf killifish (Fundulus grandis) populations from the Houston Ship Channel, Texas. Aquat Toxicol 150: 210-219.

28. Bickham J ShabegSandhubPaul D, NHebertc, LounesChikhid, RaghbirAthwale (2000) Effects of chemical contaminants on genetic diversity in natural populations: implications for biomonitoring and ecotoxicology. Mutation Research/Reviews in Mutation 463(1): 33-51.

29. Falk JL, Parent C, Agashe D, Bolnick DI (2012) Drift and selection entwined: asymmetric reproductive isolation in an experimental niche shift.

30. Harbeitner RC, Hahn ME, Timme Laragy AR (2013) Differential sensitivity to pro-oxidant exposure in two populations of killifish (Fundulus heteroclitus). Ecotoxicology 22(2): 387-401.

31. NCBI (2018) Genome List Genome NCBI

32. Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. Nat Genet 32(2): 261-266.

33. Bozinovic G, Oleksiak MF (2011) Genomic approaches with natural fish populations from polluted environments. Environ Toxicol Chem 30(2): 283-289.

34. Morozova O, Hirst M, Marra MA (2009) Applications of New Sequencing Technologies for Transcriptome Analysis. Annu Rev Genomics Hum Genet 10(1): 135-151.

35. Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4(12): 981-994.

36. Williams LM, Oleksiak MF (2011) Ecologically and Evolutionarily Important SNPs Identified in Natural Populations. Molecular Biology and Evolution 28(6): 1817-1826.

37. Money D, Migicovsky Z, Gardner K, Myles S (2017) LinkImputeR: user-guided genotype calling and imputation for non-model organisms. BMC Genomics 18(1): 523.