YALE JOURNAL OF BIOLOGY AND MEDICINE 94 (2021), pp.183-191.

INTERVIEW

YALE JOURNAL OF BIOLOGY AND MEDICINE
YJBM
CELEBRATING 90 YEARS 1928–2018

Role of Environmental Genetics in Preventive Medicine

An Interview with Daniel W. Nebert, MS, MD

Brian Thompson*

Department of Environmental Health Sciences, Yale School of Public Health, Yale University, New Haven, CT, USA

Daniel W. Nebert is Professor Emeritus at University of Cincinnati, College of Medicine, Department of Environmental Health, and Division of Human Genetics at Cincinnati Children's Hospital. He received his BA degree (chemistry; biology) from Wesleyan University (CT), and MS (biophysics) and MD degrees from University of Oregon Medical School in Portland (now named Oregon Health & Science University). He then served as a pediatric intern and resident at University of California, Los Angeles Health Sciences Center, before completing a postdoctoral fellowship as Research Associate in the National Cancer Institute (NIH, Bethesda, MD). He then became an Independent Investigator, Section Head, and ultimately Chief of the Laboratory of Developmental Pharmacology at the National Institute of Child Health & Human Development (NICHD). He was one of the founding members of a free medical clinic in southeast Washington, DC.

In December 1989 he became Professor in the Department of Environmental Health, University of Cincinnati Medical Center, where he founded the Center for Environmental Genetics (CEG); this was the first National Institute of Environmental Health Sciences (NIEHS) Center of Excellence on that topic. He also became Principal Investigator of the Gene-Environment Interactions Training Program (GEITP), which continues to include an educational email list of more than 250 colleagues worldwide. Dr. Nebert has published more than 650 papers, which have been cited more than 69,500 times; in March 2016 he was ranked among the top 640 "Most-Cited Scientists/Authors Worldwide since 1900 [1]." For his outstanding work, Dr. Nebert has received numerous awards – among them the 1986 Bernard B. Brodie Award on Drug Metabolism (from the American Society for Pharmacology & Experimental Therapeutics), the 1999 University of Cincinnati George Rieveschl Jr. Award for Distinguished Scientific Researcher, the 2005 Distinguished Lifetime Toxicology Scholar Award (from the Society of Toxicology), and the 2016 R.T. Williams Distinguished Scientific Achievement Award (from the International Society for the Study of Xenobiotics). In 1994 he was elected Fellow of the American Association for the Advancement of Science.

The American College of Preventive Medicine (ACPM) defines preventive medicine as "a medical specialty that focuses on the health of individuals and communities. The goal of preventive medicine is to promote health and well-being and to prevent disease, disability

*To whom all correspondence should be addressed: Brian Thompson, Email: brian.thompson@yale.edu.

Abbreviations: PAH, polycyclic aromatic hydrocarbon; ESR, electron spin resonance; EPR, electron paramagnetic resonance; AHR, aryl hydrocarbon receptor; AHH, aryl hydrocarbon hydroxylase; CYP1A1, cytochrome P450 1A1; GD, gestational day; *CYP*, cytochrome P450.

Keywords: Toxicology, Environmental Genetics, Preventive Medicine, Biochemistry, CYP

and death." In your opinion, how does your career in environmental genetics contribute to preventive medicine?

I'd say that a major portion of preventive medicine embraces the study of drugs that promote health and prevent disease. Another subset of preventive medicine involves understanding the mode- and mechanism-of-action of foreign chemicals (xenobiotics), with the ultimate purpose of promoting health and well-being and preventing morbidity and mortality from drugs and xenobiotics. These two categories represent pharmacology and toxicology, respectively. Classical studies in these fields have typically involved "observing an effect" of *drug A* or *xenobiotic B* in a laboratory animal or cells in culture – following administration of an experimental agent, compared with an untreated or placebo-treated control.

"Environmental genetics" is "the study of interindividual genetic differences in response to environmental agents – which include drugs, hazardous substances in our environment, and even radiation." Thus, a study in "environmental genetics" kicks everything "up a notch." In other words, when an identical amount of *drug A* or *xenobiotic B* is compared in lab animals or cell cultures that are carrying a specific mutant gene versus the wild-type (containing the "normal," unmutated gene), any effect specific to the mutant is more likely to be distinguished from nonspecific effects of the drug or xenobiotic (i.e. "off-targets"). As Nobel Laureate Har Gobind Khorana said in a talk I attended in 1972, "All things being equal, it's better to have a mutant than not to have a mutant!"

Can you describe how you first became interested in environmental genetics, and what sustained your motivation throughout your long career?

Actually, I experienced a very memorable "defining moment." During the 11th-grade of high school, I recall reading an article in *Scientific American* about tryptophan-growth-dependent bacteria that had been deprived of tryptophan; when tryptophan was then added to the culture medium – within minutes there was an ~80,000-fold burst in the specific enzyme activity needed for tryptophan to be utilized as an energy source.

This observation seemed very fascinating to me, yet so fundamental: (a) gene(s) – whose function it is to use tryptophan as an energy source – had been sitting silently on their chromosome; (b) a "signal" (tryptophan) exterior to the bacterium suddenly appears in the medium; (c) that signal is quickly "sensed;" and (d) the "response" to the signal is enzyme induction that provides energy, thereby helping the organism's survival. I was left wondering, "How many steps are involved in sensing the signal, and how many steps might be involved in the organism's response?"

Dr. Daniel W. Nebert, Professor Emeritus at University of Cincinnati, College of Medicine, Department of Environmental Health, and Division of Human Genetics at Cincinnati Children's Hospital, Cincinnati, OH

Fast-forward 12 years. After college (Wesleyan University, Middletown, CT), a 5-year MS-MD program at University of Oregon Medical School (Portland, OR), pediatrics internship and residency (UCLA and Harbor General, Los Angeles, CA) – I was fortunately selected to carry out postdoctoral research at the National Cancer Institute (Bethesda, MD). [*In 1966, MDs on the "Berry Plan" and after internship and residency had little choice: land a position in research or spend several years in the Vietnam War.*] Moreover, I was lucky to have been accepted into the lab of Harry V. Gelboin, who encouraged me to "find whatever I thought was an exciting research project," with no pressure from him.

In 1956, Allan H. Conney, as postdoctoral fellow in the lab of Elizabeth and Jim Miller (McArdle Institute, University of Wisconsin) had reported that a polycyclic aromatic hydrocarbon (PAH), administered into the peritoneal cavity of a rat, caused induction of a liver enzyme that metabolizes the injected inducer PAH and related PAHs. Bingo! Here was a second intriguing gene-environment interactions publication. However, instead of a bacterium, this experiment was carried out in a mammal; instead of the food source tryptophan as the signal, Conney's experiments involved a known human chemical carcinogen as the environmental signal. Is induction of metabolism of this PAH – by the incoming signal itself – intended by Mother Nature as a means to detoxify the carcinogen? Or is the explanation more complicated?

You are a trained physician but have spent most of your career conducting basic science research. What motivated you to perform bench research and how does basic research at the bench advance preventive medicine?

When I considered an MD degree, I already knew I wanted to combine clinical medicine with basic research, because during college I thoroughly enjoyed physics, chemistry, biology, and genetics. Around 1960, MD-PhD programs had not yet become popular; however, University of Oregon Medical School (Portland) offered a 5-year MS-MD program. Also, at the time, about the most intriguing mentor at UOMS was Professor Howard S. Mason, co-discoverer with Osamu Hayaishi (Tokyo) of "mixed-function oxidases" – which later turned out to be synonymous with cytochrome P450 enzymes.

In the Mason lab during my MS program in medical school, I learned many basic aspects of electron spin and electron paramagnetic resonance (ESR, EPR) and related biophysics techniques. In the early 1960s, Mason's lab was the first to use ESR in showing that the "cytochrome P450 enzyme group" contained an essential ferrous/ferric ion. Serendipitously, a PAH-induced cytochrome P450 associated with aryl hydrocarbon receptor (AHR) up-regulation – that I characterized in 1969 – was the unique high-spin ferric-ion-containing protein that the Mason lab had identified more than six years earlier by ESR. None of this benchwork science, which advances our understanding of preventive medicine, could have been solved by clinical research alone.

In academia, I realized that – by publishing primary papers and invited reviews and speaking at national and international symposia – I could "reach a much larger target audience" than I could by being a practicing clinician. However, always aware that federal research dollars might dry up some day, I kept my California and Ohio licenses to practice medicine up to date. This approach might be viewed as "holding an insurance policy" to support my six children!

You characterized the aryl hydrocarbon hydroxylase (AHH) enzyme assay, which was key to the discovery of a cytochrome P-450 (CYP1A1), an inducible enzyme involved in phase I xenobiotic metabolism. Can you briefly describe the role of CYP1A1 in the cell and how developing a deeper understanding of CYP1A1 has been crucial for preventing human disease, including many cancers?

For a postdoc who had just begun to search for an exciting research project – to study an environmental carcinogen's intracellular induction of its own metabolism – seemed to me to be a project that would easily be fundable in the future and provide a lot of exciting results along the way. First, I established a simple reliable enzyme assay for quantifying aryl hydrocarbon hydroxylase (AHH) induced by PAHs; this led to a 1968 paper [2] that has since been cited more than 1,950 times.

Given my strong interest in genetics, my next step was to search for genetic differences in AHH inducibility among inbred strains of mice. Similar to Conney's rat experiments, "AHH-responsive" mice exhibited AHH induction; however, about one-third of inbred strains that I screened were "AHH-nonresponsive," ie, the highest possible level of PAH that could be administered to a mouse did not cause a rise in AHH. Simple genetic crosses showed that lack of AHH induction is inherited as an autosomal recessive trait (eg, similar to "blue eyes"). Thus, in the backcross and F_2 generations, we had mice in the same litter that were basically identical except for this ("large-effect") single-gene difference! This genetic model was then used in hundreds of experiments of PAH-induced cancer, toxicity, and birth defects – to show that AHH induction was intimately associated with these PAH-caused disorders [3].

It should be noted that these experiments – spanning 1966 to ~1982 – were simply "enzyme assays," because molecular biology techniques (eg, western blots, northern blots, PCR, etc.) were not yet available. Also, by spectrophotometric data in several model systems (1969-79), I showed that "AHH induction" must represent more than one form of PAH-inducible cytochrome P450 (named "CYP1A1" and "CYP1B1" in 1992).

The "classical" means of studying anything in pharmacology or toxicology had been to introduce the test agent intraperitoneally and then measure enzyme activity in liver (a large easily-accessible organ). As with Conney's experiments, most of our early studies also involved intraperitoneal treatment. Yet, in 1962 the laboratory of Lee Wattenberg (University of Minnesota, Minneapolis, MN) fed rats a PAH (benzo[*a*]pyrene) and found AHH induction occurred throughout the gastrointestinal tract; he therefore proposed that this enzyme might be beneficial to the animal – because it detoxified the ingested PAH.

The overriding consensus, from 1973 onward, however, was that CYP1 metabolism of PAHs was harmful, because many studies (especially in cell culture) showed that CYP1-mediated metabolic activation resulted in DNA adducts and other mutagenic intermediates that caused cancer and/or toxicity *directly* in that target cell. Ultimately, with the help of *Cyp1a1(-/-)* and *Cyp1b1(-/-)* knockout mouse models, the Nebert lab in 2004 found that CYP1 metabolism of PAHs reflected a delicate balance between detoxication and metabolic activation – depending on pharmacokinetics. These unequivocal findings thus nullified many dozens of epidemiological "false-positive" studies suggesting that several single-nu-

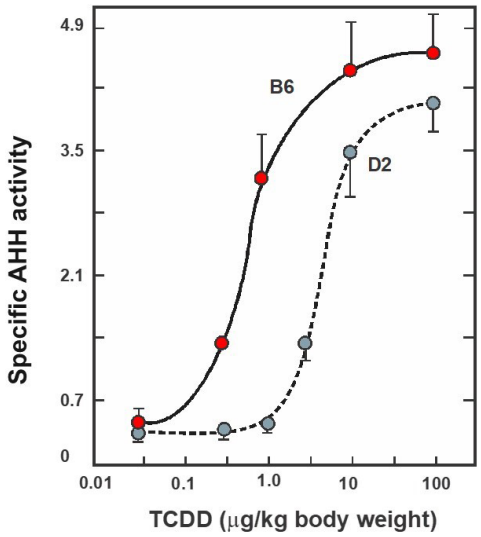

Figure 1. **Dose-response curve**. After intraperitoneal administration of TCDD, liver AHH activity in C57BL/6 (B6) and DBA/2 (D2) was measured 44 h later (modified from [13]).

cleotide variants (SNVs) in or near the human *CYP1A1* gene were specifically associated with increased cancer risk [4]. Innumerable genome-wide association studies (GWAS) on cancer during the past 14 years have confirmed this "absence of *CYP1A1* being a gene significantly associated with malignancy." Our conclusion (as to where and why PAH-induced cancer or toxicity occurs) was found to include: (a) which P450 is predominantly expressed in a particular tissue, (b) PAH dosage, (c) timing of PAH exposure, (d) route-of-administration, and (e) specific oncogenes and tumor-suppressor genes up- and down-regulated during tumorigenesis in the target cell-type [5].

Genetic differences in AHH inducibility resulted in another exciting research direction. In collaboration with Alan Poland (University of Rochester School of Medicine, Rochester, NY), 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD, discovered by Raymond Suskind in 1964 to be a contaminant of Agent Orange) was found to be able to "force" AHH inducibility in "AHH-nonresponsive" mice (Figure 1). This observation – again with the help of studies in backcross and F_2-generation mice from the same litter – led to the proposal and ultimately proof of the existence of aryl hydrocarbon receptor (AHR), the transcription factor that binds to the AHH inducer. In other words, without AHR, the environmental signal is not received in the nucleus and downstream responses to that signal cannot proceed.

Thus, all the "AHH-inducibility" traits associated with susceptibility to PAH-induced cancer, toxicity and birth defects [3] actually reflect function of the "AHR/CYP1 axis" – realized years later. In addition to dozens of AHR ligands proposed [6], our early work – showing that AHR participates in ethanol-induced peritonitis – led to the discovery that AHR has numerous endogenous ligands in the "lipid mediator" pathway (eg, lipoxin A_4, 12(R)-HETE). The 50+ years in my lab uncovered many exciting and fruitful findings relevant to the role of AHR in preventive medicine [7].

The World Health Organization and almost every regulatory agency in the world has expanded upon this pioneering work on AHR and TCDD, using the so-called Toxic Equivalence Factors (TEFs) to estimate the relative contribution of multiple different dioxins, dibenzofurans and polychlorinated biphenyls (PCBs); this, in turn, generated Toxic Equivalency Quotients (TEQs) to facilitate risk assessment and government regulations on air-water-ground pollution, factory emissions, clean-up levels, etc. However, this field is very complicated – due to striking interspecies and intraspecies differences in AHR binding affinities (eg, AHR affinity in rat is 10,000-fold greater than that of guinea pig; AHR affinity varies >30-fold among mouse strains; AHR affinity in mouse and rat can be 100 times greater than that in human; but for certain dioxins and dibenzofurans, rat AHR affinity is several orders of magnitude poorer than that of human AHR, etc.). Yet billions of dollars are spent in cleaning up PAH-contaminated sites, generally based on risk assessments using rat TEF values. But don't get me started on the topic of "government waste, resulting from policies and regulations based on poor science!" In conclusion, "species differences in AHR affinity" and other environmental response genes – is an enormously important area in the gene-environment interactions field, but it hasn't gotten nearly the attention it deserves.

You asked, "what's the role of CYP1A1 in the cell." Anup Dey in the Nebert lab showed that AHH (CYP1A1) inducibility is expressed in the fertilized ovum. The most fundamental "role" (in my mind) is therefore to ensure cell division of the fertilized zygote. It seems most likely that a striking abundance of constitutive CYP1A1 mRNA – detectable in the fertilized ovum at mouse gestational day (GD) 0.5 – might be important for maintaining sufficient amounts of CYP1A1 enzyme during the transition from maternal to zygotic control in the GD 0.5 to GD 1.5 embryo. Availability of catalytically active CYP1A1 would ensure that any (endogenous or exogenous) AHR ligand be rapidly degraded, so that any chance of an AHR–RB1 block in E2F-mediated gene transcription (see Fig. 17 in [8]) would be eliminated during the very critical moments of early mammalian embryogenesis such as the first cell division. Moreover, if an endogenous AHR ligand is an essential substrate for the CYP1A1 enzyme after the 2-cell embryo stage, one would expect that *Ahr* and *Cyp1a1* gene expression might go hand-in-hand during early embryogenesis.

I understand you also discovered a second gene, expressed in embryonic stem cells, which thus participates in a multitude of critical-life processes as well as diseases. Can you briefly describe this additional "research project" and how it relates to preventive medicine?

Yes, this could be an entertaining cocktail hour story! In the early 1970s, Ben Taylor (The Jackson Laboratory, Bar Harbor, ME) described a fascinating simple-gene difference; he showed that "resistance to cadmium (Cd^{2+})-caused testicular necrosis" segregated as an autosomal recessive trait. In the mid-1990s, largely because of 25 years of advances in genomics since Ben Taylor's initial report, I decided we should pursue this project further. In the late 1990s, at a restaurant overlooking the bay and ocean, I chatted with a famous heavy-metal toxicologist about my plans; his response was, "Don't waste your time, we already know that Taylor's observation was an artefact." This was all the encouragement I needed – to forge ahead on this intriguing Cd-toxicity project!

First, using the power of Cd-sensitive and Cd-resistant inbred mouse strains, we confirmed the original study. By means of various genetic tricks plus having access to a much larger genomics database, we then proceeded to narrow Taylor's "cadmium resistance" (*Cdm*) locus from more than 24 centiMorgans (cM) to 0.64 cM – which represented ~4.96 megabases (Mb). With additional "gene-isolation genomics tricks," the chromosomal region associated with the Cd-induced toxicity phenotype was narrowed to a 400-kilobase (kb) haploblock (DNA segment) in which one of five genes leaped out as an excellent candidate: the only homologous gene in the genome database in 2003 to our candidate was a putative "**z**inc- and **i**ron transporter **p**rotein" called "ZIP8" in plant and yeast. This was exciting: could *Cdm* code for a divalent cation transporter?

In mouse fetal fibroblast cultures, we were thrilled to find that ZIP8 cDNA expression was associated with large increases in Cd^{2+}influx, intracellular Cd^{2+} accumulation, and intracellular toxicity. Curiously, ZIP8 mRNA was found to be prominent in testicular vascular endothelial cells of Cd-sensitive, but not Cd-resistant, strains of mice! We generated a bacterial artificial chromosome (BAC) carrying the Cd-sensitive gene, which we inserted into a Cd-resistant mouse. Cd treatment reversed the Cd-resistance trait seen in nontransgenic littermates, to Cd sensitivity in our BAC-transgenic mouse. Reversal of the testicular necrosis phenotype thus unequivocally confirmed that our candidate gene (*Slc39a8*) was Ben Taylor's *Cdm* locus. Further studies revealed that Mn^{2+}, Zn^{2+}, Fe^{2+}, Co^{2+} and Se^{4+} are the endofenous substrates for ZIP8, encoded by *Slc39a8*. [*As is true in all of toxicology – foreign agents (such as Cd^{2+}) enter the cell by imitating an endogenous molecule, and then "hitchhiking" into the cell.*]

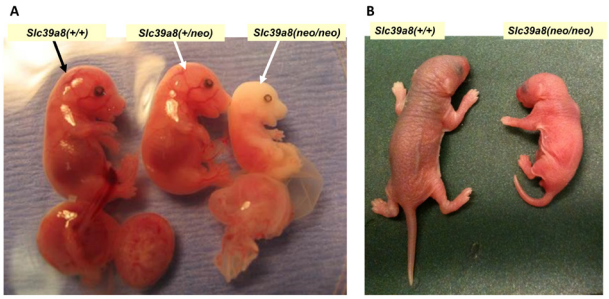

Figure 2. ***Slc39a8(+/+)* wild-type (left) and *Slc39a8(neo/neo)* (right) newborns, shortly before death of the mutant**. Compared with wild-type that is pink in color and normal in size, *Slc39a8(neo/neo)* littermates are extremely pale, show stunted growth, and have deformed skulls and limbs (from [14]).

Other labs then found that ZIP8 is expressed in gastrula and visceral endoderm; this finding was consistent with our discovery that a *Slc39a8(–/–)* global knockout is very early embryolethal. Intriguingly, during our attempts to create the global knockout, we inadvertently created the "knockdown" *Slc39a8(neo/neo)* mouse; this "hypomorph" expresses ZIP8 mRNA and protein levels ~15% of that found in all wild-type tissues examined, and the mouse is viable until the end of gestation. This experimental model thus provides a "window of time" for studying ZIP8 functions in placenta, yolk sac, and fetal tissues during *in utero* growth.

We found that the *Slc39a8(neo)* allele is associated with striking anemia, severely hypoplastic spleen and substantially reduced sizes of liver, kidney, lung, and brain (Figure 2). Bioinformatics analysis of the transcriptome concluded that the deficient ZIP8-mediated divalent cation transport in *Slc39a8(neo/neo)* mice – predominantly in yolk sac – affects zinc-finger-protein transcription factors. Our transcriptomics data thus were strongly consistent with the *in utero* phenotypes of dysmorphogenesis, dysregulated hematopoietic stem cell fate, and anemia that are seen in *Slc39a8(neo/neo)* mice.

During the past decade of clinical research, many GWAS have identified a handful of human *SLC39A8* (low activity) variants correlated with numerous traits: heart disease and blood pressure regulation, schizophrenia, osteoarthritis, Crohn's disease, and retinal iron accumulation. Some *SLC39A8* variants are also associated with congenital deformed skull, cerebellar atrophy, severe seizures, short limbs, and hearing loss defects – traits similar to those seen in the *Slc39a8(neo/neo)* newborn (Figure 2). Perhaps most importantly, (human and mouse) ZIP8 deficiency was found to impair Mn^{2+}-dependent enzyme functions, which dramatically diminishes posttranslation-

al glycosylation; this global effect might help explain the extreme pleiotropy of clinical disorders associated with *SLC39A8* variants [9].

As a pioneer in the field of environmental health, you spearheaded the shift from 20th-century "classical toxicology" toward 21st-century environmental molecular genetics and genomics. What in your mind drove this shift?

When I accepted the position of Professor, Department of Environmental Health (DEH), at University of Cincinnati College of Medicine (COM) in 1989, I was appalled that the DEH faculty members in the Division of Toxicology lacked any understanding of basic concepts in genetics, developmental biology, or recombinant-DNA research. It thus came as no surprise that these DEH members had little (if any) ability to interact with colleagues in the other COM basic-science departments. In contrast, I saw no fundamental difference in research being carried out by these other basic-science departments and what I viewed as environmental genetics research.

Thus, we immediately changed the name of the DEH group to Division of Environmental Genetics and Molecular Toxicology (EGMT). In addition, I interviewed colleagues from all over University of Cincinnati COM and Cincinnati Children's Hospital Research Center – to find common avenues of interest and goals; this led in 1992 to the birth of the first National Institute of Environmental Health Sciences (NIEHS)-funded Center for Environmental Genetics (CEG). I included more than a dozen colleagues outside DEH as CEG members who had more research interests in common with one another than those found inside DEH when I had arrived. Success of the CEG exploded, and our Center is currently in its 29th year of NIEHS funding support. Also, virtually all of the ~25 NIEHS Centers of Excellence now have an "environmental genetics" component. In 1992, however, the idea of gene-environment interactions was regarded as heretical!

In my mind, I envision environmental genetics and gene-environment interactions research as one and the same. With each research question, an "exogenous signal" is presented to an organism (cell, tissue, organ, animal, plant, or clinical patient), and a "response" is elicited, based upon that organism's genome. Each "exogenous signal" can arise from a neighboring cell or tissue in the same organism, or from that in a different plant or animal, or from xenobiotics in the environment. Each "response" might reflect a single gene's activity, but the vast majority of "responses" are now known to involve numerous genes that are each part of one or another gene-signaling pathway. Notice that we have now transitioned from "1980s genetics mindthink" to "genomics thinking of the 1990s up to the present." In terms of a genetic response, activity of one or a few genes will usually present as a Mendelian trait (dominant, recessive, codominant), whereas contributions from dozens (or hundreds or thousands) of genes will appear as a polygenic multifactorial trait, and the genetic response will represent a gradient. Finally, every "response" ultimately is designed by Mother Nature to benefit evolutionary survival of that species (to find food, avoid predators, and/or generate healthy offspring).

As a pioneer in the field of pharmacogenetics, you seem to have promoted the idea of complexity in the genome and its response to any drug or other environmental agent – as far back as the 1970s. Can you explain why?

The more I had read, and the more meetings I had attended – in the basic fields of developmental biology, and bacterial, yeast, fungal, plant, and worm, and fly genetics – the more I became convinced that the clinical response to a drug (or any toxicant or mixture of xenobiotics) could not possibly be any less simple than what happens in all of non-human biology. Consequently, since the late 1970s I've written a number of invited reviews proposing this complexity, which contrasted with "simple genetics models" being proposed by most of my colleagues. With each review over the past four decades, these concepts have continued to evolve, in order to keep up with the latest understanding in the fields of genomics and population genetics (cf. our latest review in [10]).

As mentioned earlier, the field of pharmacology is absolutely central to "preventive medicine." To dissect the mechanisms of interindividual response to drugs is another way to describe "personalized medicine." Again, how could environmental genetics and gene-environment interactions research not be fundamentally intertwined with preventive medicine?

I think that many trainees fail to appreciate that our current standardized gene nomenclature system is relatively new. Can you describe how the lack of a standardized gene nomenclature system obfuscated findings from independent research groups; can you detail your efforts in helping to create a standardized gene nomenclature system – including the need for the HUGO Gene Nomenclature Committee?

In the 1970s, with purification of many cytochrome P450 proteins and generation of antibodies, more than a dozen laboratories – working independently with rat, rabbit or mouse – gave each of their purified proteins or antibodies their own pet names. Before the mid-1980s, I could see that this haphazard approach would soon lead to chaos. And this confusion would be greatest among graduate students and postdoctoral fellows just entering

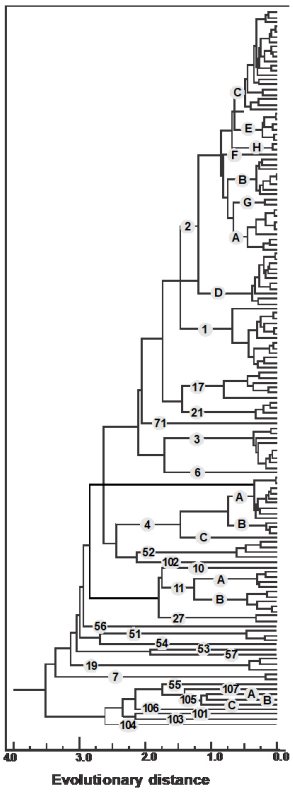

Figure 3. **Phylogenetic tree of the *CYP* gene superfamily, based on evolutionary divergence, as of December 1990**. Analysis was performed by the **u**nweighted **p**air-**g**roup **m**ethod with **a**rithmetic mean (**UPGMA**) algorithm (taken from an invited talk in Lisbon, Portugal, in early 1991).

the field.

In the early 1980s, clones of P450 partial cDNAs began to be sequenced, from which translated amino-acid sequences could be deduced. Intriguingly, I noticed that if one aligned the protein sequences of P450 proteins from *Pseudomonas*, yeast, and eight vertebrates, including human – a highly-conserved peptide of eight amino acids was found [11]. To me, this high degree of similarity in the P450 protein consensus sequence (between bacteria and human) could only be explained by evolution from a common ancestral gene.

At one of our annual P450 symposia, I therefore called a meeting for all of us "guilty" of providing pet names to our P450 proteins or antibodies, and suggested we all be coauthors on a "nomenclature paper." After a great deal of reluctance, this was agreed upon. I insisted we name the members of this P450 superfamily (or group of genes) – based on evolutionary divergence. The "root," or "symbol" ***CYP*** (for **cy**tochrome **P**450) for each gene arising from a common ancestor was eventually agreed upon. The original "cut-off" for members within one "family" was >40% amino-acid similarity; P450 protein sequences having <40% similarity would represent *CYP* genes of different families. The original cut-off for members within one "subfamily" was >60% similarity. Of course, many overlapping complications arose, and each new gene had to be manually curated and decisions made. The earliest method of illustrating phylogenetic trees was the **u**nweighted **p**air-**g**roup **m**ethod with **a**rithmetic mean (**UPGMA**), a simple bottom-to-top hierarchical clustering method, which we applied to the *CYP* gene superfamily (Figure 3). Today – in this expanding field of evolutionary divergence analysis – many new sophisticated algorithms have appeared.

In 1996 was the last paper publication [12] before the total number of *CYP* genes had become unwieldy for journal updates. David R. Nelson (University Tennessee Health Sciences Center, Memphis) had volunteered, early on, to become curator of the *CYP* gene nomenclature homepage website and he continues to tirelessly perform this task today. With each whole-genome sequence publication of a new species – animal, plant, fungus, protist, eubacteria, archaebacteria, or virus – Nelson, by means of BLAST-searches, gives a name to each new *CYP* gene having a putatively functional protein-coded product, as well as pseudogenes (https://drnelson.uthsc.edu/cytochromeP450.html).

The total number of CYP (putative protein-coding) genes now exceeds 90,000 – across a wide diversity of species, from viruses to humans – and is increasing every week. With the recent introduction of the **t**ranscriptome **s**hotgun **a**ssembly (**TSA**) database, Nelson predicts the number of functional individually named *CYP* genes will soon exceed 250,000! This massive number reminds me of a talk I had presented in Ireland in 1978, in which I proposed "there might be hundreds, even thousands, of P450 genes on the planet;" most in the audience laughed hysterically, because they could not imagine there might possibly be any more than the ten or so P450 proteins known to exist at that time.

Outside the "P450 world," many other scientists were similarly envisioning the need for standardized gene nomenclature. In the 1990s, I was invited to begin interacting with the External Advisory Committee of the Human Genome Organization (HUGO)-associated Human Gene Nomenclature Committee (HGNC). The fundamental use of nomenclature – of "genes within any group having a shared consensus sequence and thus derived from a common ancestor" based on divergent evolution – is now extremely common throughout all living organisms on the planet (https://www.genenames.org/).

You have immense passion for teaching and mentoring, which has included mentorship of ~104 graduate students and postdoctoral fellows, including my current mentor, Dr. Vasilis Vasiliou. Why is it important to continue

mentoring young scientists?

Despite many big-time offers, I wished to carry out research in a small lab. With the development of growing numbers of consortia performing research, it comes as no surprise that the individual investigator with a small lab of five or ten is becoming increasingly rare. This is no different from evolution of human civilization, ie, "hunter-gatherers" 25,000 years ago gave way to "farming communities" 10,000 years ago. However, in scientific research I view this "evolution" as something that should not necessarily be looked upon as beneficial. In my opinion, the individual creative mind can easily be subdued by consortium group-thinking.

Combining this phenomenon with the sheer numbers of new PhD students looking for employment and diminishing funds in academia, we're seeing more scientists flee directly from a PhD or postdoc program to Big Pharma or other industry. As grant funding becomes more difficult than ever, it is therefore extremely important to identify those gifted students who have unique creative-thinking skills, and to try and persuade them to "stay in academia." This is what I have always tried to do.

As you look back on your illustrious career, do you have any advice for trainees? Are there common pitfalls that we should avoid? How can we best utilize our training period?

It has always been my belief that – in selecting a lifelong career – one should choose a particular project, or theme, and expand upon that, rather than jumping around from year-to-year, depending on "buzzwords" and what seems to be the "projet du jour" in study sections. Unfortunately, today I see increasing numbers of young investigators taking the latter route, which I consider as "shallow" research. This "studysectionthink" behavior also extends to choosing a mentor, ie, if your mentor's research is "all over the place," unfortunately, it's more likely the trainees will "learn" to do the same. Good trainees do best in labs with an underlying complementary theme. Often, I've joked that "carrying out creative research is better than working for a living."

In the lab you have chosen, be sure to have frequent journal clubs where all participants must take turns selecting a cutting-edge paper and describing why they like, or don't like, it. If your lab does not have frequent journal clubs, you should ask a nearby lab if you might attend their journal club. Read voraciously. Discuss scientific findings as frequently as possible with everyone around you. Never be surprised by Mother Nature; if you set up an experiment and expect Result A, but instead find Result B, don't be angry or disappointed but rather ask yourself why you've found this Unexpected Result. Keep your eyes open and your mind clicking. My philosophy has always been based on evolution; if I find an unexpected result, I ask myself why Mother Nature might have chosen to "do it this way" – instead of the way my little brain thought was the better way.

Lastly, where is toxicology (and environmental health) headed? What questions remain to be answered, what emerging contaminants are you most concerned about, and what technologies will revolutionize the field?

In my own case, I accidentally chose two phenotypes involving toxicology – "induction by a human PAH carcinogen of an enzyme that metabolizes its own inducer" and "cadmium-induced testicular necrosis." And then (with a dynamic lab and other collaborators) I was fortunate to be able to identify the (mouse and human) genes responsible for each of the two traits. Moreover, for both projects, the genes are expressed in embryonic stem cells and therefore participate in myriad clinical functions in many cell-types, as well as being relevant and linked to multiple clinical diseases. Both projects clearly reflect gene-environment interactions, the constant theme of my lab. Also, going from the phenotype to identifying the gene could not have been accomplished without environmental genetics, ie, "detailed knowledge of genetics and genomics" and knowing the tricks therein, that can be used to locate and identify the gene. Moreover, the knowledge gained from both projects is central to clinical advances in preventive medicine.

As far as "research into the future" – toxicology and environmental health are intimately intertwined with genetics and genomics. I would caution young researchers not to be too reductionist with their experimental designs. In the real world, we are not exposed to individual, discrete environmental toxicants but rather complex mixtures. Furthermore, when we administer eg, "100 mg per kg to a lab animal," we must keep in mind that realistic clinical doses are usually in the nanomolar range, not millimolar range.

Technologies of GWAS, epigenome-wide association studies (EWAS), genotype-tissue expression (GTEx) transcriptomics, machine learning, and polygenic risk scores (PRS) – are current exciting fields of study that are all going to benefit from increasingly large cohorts (eg, 500,000 or 2 million individuals). The idea of one "large-effect" gene contributing all or most to one trait (eg, "response to a drug"), is almost always overly simplistic. We now know that most traits involve hundreds or thousands of "small-effect" genes each contributing a miniscule amount to the trait. Most genes are pleiotropic, ie, contributing a small "effect" to several, or many dozens of, disorders. Scientists in all walks of biological research including toxicology should "always try to think outside the box." Every phenotype represents the contri-

bution of gene differences (variations in DNA sequence), epigenetic factors (changes within the chromosome, but no DNA sequence alterations), environmental effects, endogenous influences, and each patient's microbiome. Finally, except for the germline alleles that we are each born with, everything else is dynamic and keeps changing over time.

Acknowledgments: We thank our Reviewers, in particular David L. Eaton (Dean & Vice Provost, The Graduate School, University of Washington, Seattle; now professor emeritus), for a very thoughtful in-depth review.

REFERENCES

1. 1040 Most Highly Cited Researchers (h > 100) according to their Google Scholar Citations public profiles, 4th edition). 2016. Available from: https://web.archive.org/web/20160819045703/http://www.webometrics.info/en/node/58
2. Nebert DW, Gelboin HV. Substrate-inducible microsomal aryl hydroxylase in mammalian cell culture. I. Assay and properties of induced enzyme. J Biol Chem. 1968 Dec 10;243(23):6242–9.
3. Nebert DW. The *Ah* locus: genetic differences in toxicity, cancer, mutation, and birth defects. Crit Rev Toxicol. 1989;20(3):153–74.
4. Nebert DW, Dalton TP. The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. Nat Rev Cancer. 2006 Dec;6(12):947–60.
5. Nebert DW, Shi Z, Gálvez-Peralta M, Uno S, Dragin N. Oral benzo[*a*]pyrene: understanding pharmacokinetics, detoxication, and consequences--*Cyp1* knockout mouse lines as a paradigm. Mol Pharmacol. 2013 Sep;84(3):304–13.
6. Nguyen LP, Bradfield CA. The search for endogenous activators of the aryl hydrocarbon receptor. Chem Res Toxicol. 2008 Jan;21(1):102–16.
7. Nebert DW. Aryl hydrocarbon receptor (AHR): "pioneer member" of the basic-helix/loop/helix per-Arnt-sim (bHLH/PAS) family of "sensors" of foreign and endogenous signals. Prog Lipid Res. 2017 Jul;67:38–57.
8. Nebert DW, Roe AL, Dieter MZ, Solis WA, Yang Y, Dalton TP. Role of the aromatic hydrocarbon receptor and [*Ah*] gene battery in the oxidative stress response, cell cycle control, and apoptosis. Biochem Pharmacol. 2000 Jan;59(1):65–85.
9. Nebert DW, Liu Z. *SLC39A8* gene encoding a metal ion transporter: discovery and bench to bedside. Hum Genomics. 2019 Sep;13 Suppl 1:51.
10. Zhang G, Nebert DW. Personalized medicine: genetic risk prediction of drug response [Addendum in: Pharmacol Ther. 2017 Dec 10]. Pharmacol Ther. 2017 Jul;175:75–90.
11. Nebert DW, Gonzalez FJ. P450 genes: structure, evolution, and regulation. Annu Rev Biochem. 1987;56(1):945–93.
12. Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, et al. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. Pharmacogenetics. 1996 Feb;6(1):1–42.
13. Niwa A, Kumaki K, Nebert DW, Poland AP. Genetic expression of aryl hydrocarbon hydroxylase activity in the mouse. Distinction between the "responsive" homozygote and heterozygote at the *Ah* locus. Arch Biochem Biophys. 1975 Feb;166(2):559–64.
14. Chen J, Gálvez-Peralta M, Zhang X, Deng J, Liu Z, Nebert DW. In utero gene expression in the *Slc39a8(neo/neo)* knockdown mouse. Sci Rep. 2018 Jul;8(1):10703.