Green Chemistry

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c6gc00182c

Environmental genome of industrial products (EGIP): the missing link for human health

Michael Overcash

Human health impacts come from genetic factors and from the environment. Human health impacts, as nature *versus* nurture (genetic *versus* non-shared chemical environmental), have currently been represented in genomics only as the human genome (double helix). The discovery of the genomic structure for the non-shared chemical environmental (the environment) component fills a conceptual gap and thus provides a more complete understanding of total impact on human health. When mapped, the environmental genome of industrial products (EGIP) would serve as a global framework for assessing broad and clustered health effects (human and ecological) at the local and global levels. Thus products for society have coupled benefits and impacts. The environmental genome is pyramidal rather than helical. The shape of the EGIP is, $x = n^y$, where n is a structural building unit (SBU) which is different for each of the 100 000 chemicals, x is the complexity of the industrial ecosystem to make the product (width of genome), and y is the levels or stages from natural resources to product (length of genome). The environmental genome is approximately the same informational size (0.5–8 billion pieces) as the human genome (3 billion chemical pairs). The pyramidal environmental genomic pieces span from natural resources (containing much of the Periodic Table) to the final chemical/material products in commerce of which there are about 100 000 which comprise all products. In summary, "the human genome is held in awe as it created each of us. The environmental genome of industrial products must be respected for it is what we as humans have truly created."

Received 19th January 2016,
Accepted 21st March 2016

DOI: 10.1039/c6gc00182c

www.rsc.org/greenchem

Introduction

Genetic is defined as "determined by the origin, development, or causal antecedents of something" (and so is not limited to life forms) (Webster's[1]). When applied in non-biological contexts, genome connotes a fundamental building block toward a larger purpose (NSTC[2]). The environmental genome (EGIP) is the map of the origin and development of molecular building processes antecedent to manufacturing all the final industrial products in which emissions to the air, water, and terrestrial environments create potential human health impacts. It is based on the fundamental chemical engineering and chemistry used to make the approximately 100 000 chemicals products in commerce today (defined as greater than 1 mt per year, (TSCA, 2015;[3] EINECS, 1990[4])). The human genome project has produced a map (with a double helix structure) of genetic information which has been the major advance in understanding the complex influences on human health.

Industrial and Manufacturing Engineering, Wichita State University, Wichita, KS 67260-0035, USA, retired now Environmental Clarity, Inc. Reston, VA. E-mail: mrovercash@earthlink.net

However, the adverse impacts to and susceptibility of humans to disease is derived from both

(1) genetic factors and

(2) non-shared chemical environmental factors.

Both are central to the health effect attribution of nature *versus* nurture (better known as heritable *versus* environmental) (Luch,[5] Collins, 2004,[6] Lichtenstein, 2000[7]).

What are the non-shared chemical environmental factors (specifically those not involving life style choices, such as smoking, alcohol, diet, infection)? These non-shared environmental/occupational exposures are diffuse, chronic, low chemical levels affecting air, water, and terrestrial environments. We know that all of the products we utilize come from a massive network of manufacturing plants that generate the anthropogenic chemical environment of air, water, and land. There are approximately 100 000 chemicals in commerce globally, from which virtually all products you see around you are made. These emissions also include some ingredients contained in these same products. These emissions are the fundamental agents that are reflected in the non-shared chemical environmental factor of disease (in addition to those factors from life style which are not included in EGIP). So, there is a societal benefit of every product (from food, to printers, to cars, to computers, to ocean ships) which is linked to the supply chain

of individual manufacturing plants distributed globally and linked by transportation networks as a source of the causative environmental burden on humans. Thus product benefits and impacts are coupled.

When we take an even broader view, with a human being as the central focus, one is struck with the large void in this puzzle – which is the missing genome of the environmental impact on diseases in humans, Fig. 1. We have long pondered that the environmental genome of industrial products (EGIP) must also exist, as a complement to the human genome and which reflects the environmental influence on human health. Therefore we have sought a fundamental repeating structure

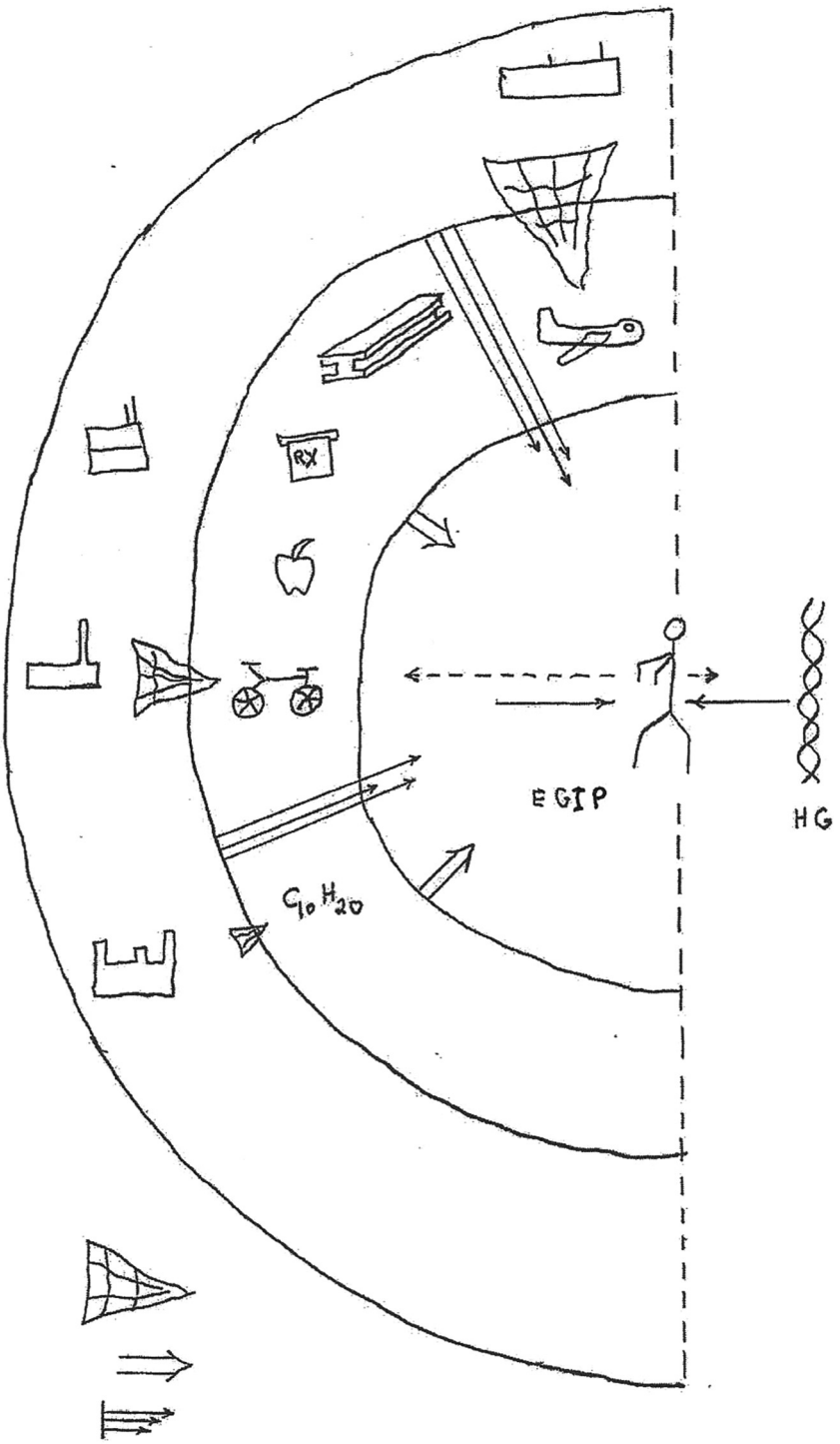

Fig. 1 Discovery sketch (2001) of the existence of an environmental genome of industrial products to mirror human genome as effects to human health.

governing the environmental genome, as a means to organize and then map this entire complex dimension of human health.

At the core of the understanding of non-shared chemical environmental factors is defining where chemicals are entering the environment. These emission locations are where chemicals are manufactured, used as inputs, formed as by-products or unwanted products, and use locations and are the focus of the EGIP. This source information precedes all the follow-on complexity of epigenetic markers, population risks, genetic repairs, *etc.*, because it is these complex manufacturing and product use/disposal stages that start a major part of the process of human impact. This EGIP theory is built on the substantial contributions of many thoughtful scientists and engineers measuring experimentally environmental effects on the global society (Briggs, 2003;[8] Weaver, *et al.*, 1998,[9] Hill, 1965[10]) and those designing manufacturing facilities (Peters, *et al.*, 2003;[11] Ulrich and Vasudevan, 2004[12]).

As a related concept, the exposome encompasses the sequential events beginning with emissions to the multimedia environment, transport to humans including fate in the environment, exposure, and health impacts.[13] In this context, EGIP focuses only on the important beginning emissions events by putting in place the nearly full set of emissions from the production of the approximately 100 000 chemical products that make up virtually all products in the global society (an area of the exposome largely not now available). Thus EGIP would be the critical foundation that supports the larger sequence in the exposome by building a strong geographic emissions framework, based on industrial product manufacturing facilities that could define cluster or dispersed sources of chemical emissions. In that context EGIP is a possible link to such projects as objective one of Helix,[14] the chemical portion arising from indirect (not lifestyle) factors of the Exposomics Project,[15] or objective one of the NIOSH research agenda for exposomics.[16]

However, as with all knowledge in an early stage of development (as was the case for the human genome), EGIP impacts on broader understanding of human health can only be imagined at this time.

Results and discussion

This research began by investigating the most basic principles in manufacturing products to then discover a common environmental genome of industrial products. All industrial enterprises (whether candy bars, sulphuric acid, iPods®, paroxetine, or aircraft) are the result from manufacturing plants that are building either

(a) molecular structure (the chemical/material universe) or

(b) macrostructure (the end-product universe).

It is the building of structure, beginning with natural resources, that implies the existence of a common genome tracing the origin or development of all products. Molecular- and macro-structure building may share a unifying genome, but first attention has been directed at molecular-building (where chemicals and materials are the products). That is, all products consist first of chemicals and materials as building blocks. There are approximately 100 000 chemicals and materials in commerce (TSCA, 2015;[3] EINECS, 1990[4]) and hence these are the basic building blocks for the product universe. Note: while polymers and copolymers are not in these regulatory chemical lists, all of the monomers are included, and hence it is a direct step to also add any desired polymers and copolymers. Importantly, natural substances processed for use are also not in these regulatory schemes and would further increase the size of the EGIP.

In order to discover the structure of the environmental genome, we have studied about fifteen hundred detailed chemical manufacturing processes and plants (about 1.5% of the chemicals-in-commerce universe) with a wide variety of molecular structure, elemental composition, and commercial use (M. Overcash, and E. Griffing, 1998–2015[17]). The domain of these assessments was for the molecular building of a chemical product starting with all the necessary natural resources found in the earth (and represented in the Periodic Table). Molecular building is the overall change in thermodynamic state occurring to produce each chemical in commerce, beginning with resources from nature. This is often referred to as a chemical tree or the supply chain. Each chemical tree is a repeating subpart of the overall environmental genome for chemicals or materials in commerce.

The structure of the environmental genome of industrial products (EGIP) was found to be a pyramidal shape, illustrated in Fig. 2 for three different size supply chains (benzoyl chloride, nylon, and paroxetine). Each final chemical, Fig. 2, is an environmental genome section that has certain repeating common elements. First there is the final chemical product, the top of the pyramidal structure. Second, at the very bottom is the array of elements from the Periodic Table found in various forms in the earth (ores, fossil resources, brines, and air) and hence this is the natural resource base of each genome section. Third, in the middle is a succession of chemical manufacturing plants driven by various simple or complex chemistries and separation unit processes. The expanding environmental genome pyramidal shape is a reflection of the basic chemical reactions used in molecular shape building to reach the final product chemical. The genomic dimensions are comprised of

(1) the integrated synthesis stages of chemical routes needed to form the product; that is, the number of, individual chemical plants making a product as a distance or extent from to natural resources to the final product (Y-axis)

(2) the complexity of the industrial ecosystems by which the molecular structures are built (X-axis).

The global environmental genome shape is the aggregate of all chemicals in commerce, where the collection of each individual manufacturing repeating units (see Fig. 2) are fit by an outer curve or repeating shape of the genome given by,

$$x = n^y \quad (1)$$

where x is the industrial ecosystem complexity (width of genome (along the x-axis) in Fig. 2), y is the levels or stages

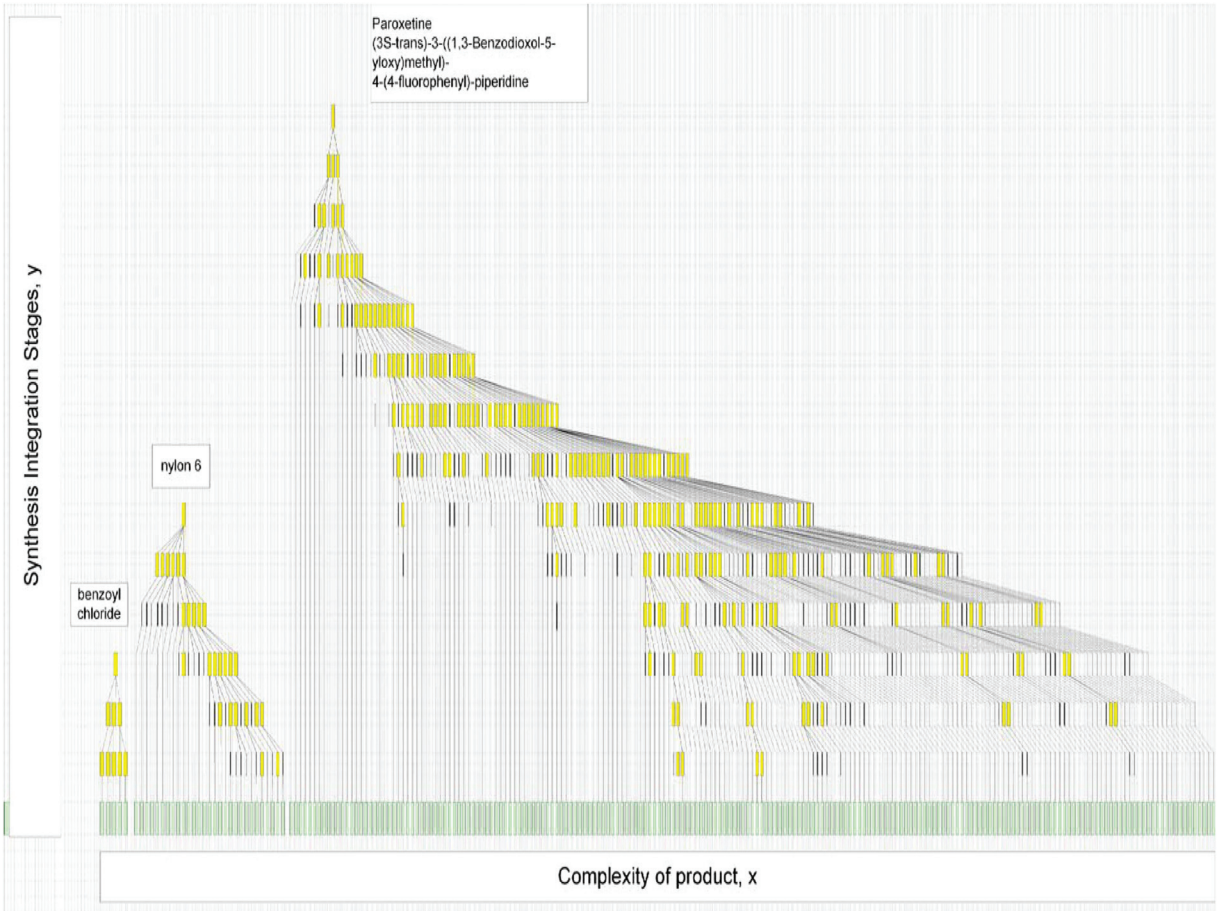

Fig. 2 Repeating environmental genome structures for three chemicals manufactured from natural resources to completed chemical structure.

from natural resources to product (length of genome, along the y axis), and n is a structural building unit (SBU) calculated for each chemical from data such as in Fig. 2 and is different for each of the 100 000 chemicals in commerce. Since the natural resource origins of each environmental genomic structure are elements (in mixtures found as ores, fossil resources, brines, and air), these pyramidal structures are each then connected to the Periodic Table. Some elements, such as carbon are widely used, while others such as neodymium (Nd) are infrequently used. Depiction of the entire environmental genome as all 100 000 chemicals in commerce is too complex visually and so a subset of chemicals is used to illustrate the larger EGIP, Fig. 3. The size of the cylinders above each element would represent the magnitude of element use in the entire EGIP and the repeating structures (the pyramidal shapes) go from basic chemicals up to more complex chemicals, Fig. 3.

The repeating pyramidal shape, Fig. 2, found in the environmental genome of industrial products is thus characterized by the values of y, n, and x as exemplified in Table 1. The SBU of n = 2 would correspond to the full sequence throughout the supply chain of chemical synthesis A + B → C. The range of x (ecosystem complexity) so far mapped has been 1 to 239 and y (cradle-to-gate stages) has been between 1 to 15. Thus we anticipate that the entire chemicals-in-commerce genome (molecular shape-building) has this same repeating pyramidal shape from which the chemical losses to the environment with the potential impact on human health can then be estimated (that is, mapped). Also in the environmental genome information are the chemicals in products, another potential source of health impacts. Since all nodes in these structures (Fig. 2) are full scale manufacturing facilities, each having estimated chemical emissions to the environment (derived from genome structure), the full EGIP is the total global environmental influences on human health derived from all molecular product-building processes. It is further postulated that the size of the structural building unit (SBU = n) reflects the impact on the humans through the environment. That is, more complex sequences of manufacturing plants (having larger values of n) lead to greater emissions. An early test of this hypothesis used ecotoxicity (TRACI 2011[18]) as the environmental health impact and the larger SBU was found to correspond to the larger total ecotoxicity, Table 1. Ecotoxicity is just one measure of health impacts and others such as human health cancer units (kg benzene eq. per kg chemical), human health non-cancer units (kg toluene eq. per kg chemical), and human health particulate units (disability-

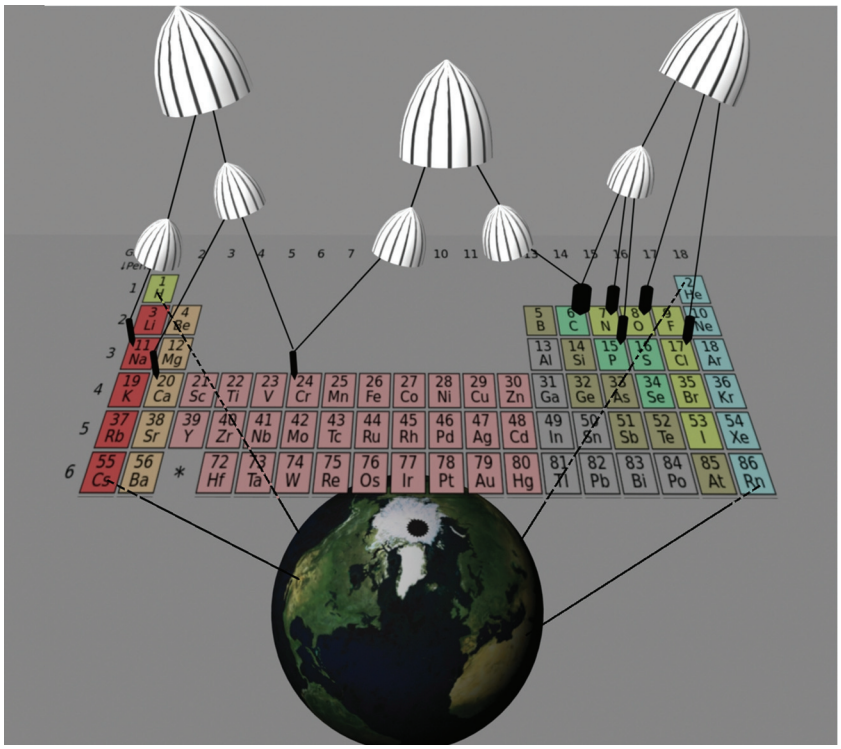

Fig. 3 Illustration of a portion of the environmental genome of industrial products (illustrated by the top genome structures).

Table 1 Complexity measures of example chemical genomes, Fig. 2 (using one half of mirror image of genome)

Chemicals	Synthesis integration stages, y	Complexity of product, x	Structural building unit (SBU) for both sides of environmental genome, n	Ecotoxicity, kg 2,4-D equivalent per kg chemical
Benzoyl chloride	4	2.5	1.25	2.2
Paroxetine	15	119.5	1.38	Not available
Nylon	7	14.5	1.47	5.6

adjusted life years (DALYS) per kg chemical) can also be derived from the same EGIP data.[18]

Below the environmental genomic nodes in Fig. 2 and 3 are distinctive manufacturing sequences that consist of a moderate number of unit processes necessary to produce each chemical or material in a manufacturing plant. From our studies, there are generally 50–200 processes per chemical/material product, each consisting of 10–20 unique unit processes (reactors, distillation columns, pumps, membrane units, *etc*.). Further, each of the manufacturing plant unit processes requires 10–20 pieces of basic information and calculations (such as heat of formation, reflux ratio, process heuristics, concentrations, *etc*.) to estimate the energy and emissions from each chemical or material manufacturing plant. Thus the environmental genome of industrial products for the approximately 100 000 chemicals and materials in commerce contains approximately 0.5 billion to 8 billion elements of information. This is in the same size range as the human genome (3 billion chemical bases). Mapping these EGIP elements (the summation of all chemical/material manufacturing plants) generates the chemical emissions to air, water, and terrestrial environments and hence the global impact on humans and ecosystems. Subsets of these manufacturing plants provide the regional or local influence on the environment and human health, such as manufacturing zones (for example, West Virginia, Singapore, the Ruhr region, *etc*.). Thus a more refined structure for EGIP may be described in which the two dimensional pyramidal shape is overlain on geographic information of these regional manufacturing facilities (longitude, latitude, elevation) to describe a three dimensional structure of the environmental genome. It is noted that a similar pyramidal shape may also exist for all products in which macrostructure is built (semiconductors, bicycles, hairdryers, automobiles, *etc*.). The environmental genome information might then tie manufacturing clusters to genetic or acquired human health. Information in the environmental genome can also be used to examine more globally, the systems chemistry (the study of macro relationships of chemical molecular building) concepts currently governing the entire industrial system used to meet society's needs for products.

Some fraction of the health of our global society is influenced by this environmental genome (in addition to other non-shared environmental factors of life style choices). For the human genome, there is also significant influence on human health, as well as some small influence of society on the human genome led by proteomics and genomics research (Chadwick, 1999[19] Biobank UK, 2006,[20] Collins and Jegalin, 1999[21]). By contrast, the EGIP is characterized by a substantial two-way influence between molecular structure product manufacturing and society. In contrast to the human genome, the environmental genome has two inseparable effects on the quality of life in societies,

(a) the largely positive effect of meeting human needs through products in commerce and

(b) the consequential effect of chemical emissions to the environment, as illustrated in Fig. 1.

Both of these impacts on society occur, but are often distributed separately since manufacturing is geographically separated from product use. The consequential effects belong to each product and derive from the thermodynamic laws governing the conversion of natural resources into products (built as molecular structure or macrostructure). These environmental effects can thus never be zero.

Conclusions

The environmental genome of industrial products (EGIP) has a significant impact on global human health, as does the human genome. While the structures of these two genomes are different (double helix *versus* pyramidal) both are comparable in informational size and change with time (evolution). Both vary within the respective populations (members of the global society for the human genome or manufacturing plants within an industrial sector for the EGIP). In both cases, the underlying genomes are a large body of information that must undergo substantial further research in order to translate understanding into change (proteomics *versus* life cycle research). Green chemistry is measured in how it can or does improve the environmental genome of industrial products and the impact on the environment. However, without the complete underlying map of the environmental genomes for industry, the efficient, integrated utilization of green engineering and green chemistry principles (Abraham and Nguyen,[22] Anastas and Warner,[23] Winterton,[24] Overcash and Cunningham[25]) to improve public health is substantially restricted. In conclusion, *the human genome is held in awe as it created each of us. The industrial environmental genome must be respected for it is what we as humans have truly created*.

Acknowledgements

Years of contributions by students, post-docs, and colleagues to mapping individual chemical supply chains and developing the tools to apply engineering principles to life cycle inventory analysis are sincerely thanked.

References

1 *Webster's Third New International Dictionary of the English Language, unabridged*, Merriam-Webster, Inc., Springfield, MA, 1993.
2 NSTC, *Materials Genome Initiative for Global Competitiveness*, National Science and Technology Council, Washington, DC, 18pp, 2011.
3 TSCA, *Chemical Inventory for New Chemical Registration*, U.S. EPA, Washington, DC, 1985.
4 EINECS, *European Inventory of Existing Commercial chemical Substances*, European Chemical Agency, Helsinki, Finland, 1990.
5 A. Luch, Nature and nurture – lessons from chemical carcinogenesis, *Nat. Rev. Cancer*, 2005, **5**, 113–125.
6 F. Collins, The case for a US prospective cohort study of genes and environment, *Nature*, 2004, **429**, 475–477.
7 P. Lichtenstein, N. Holm, P. Verkasalo, A. Iliadou, J. Kaprio, M. Koskenvuo, E. Pukkala, A. Skytthe and K. Hemminki, Environmental and heritable factors in the causation of cancer, *N. Engl. J. Med.*, 2000, **343**(2), 78–85.
8 D. Briggs, Environmental pollution and the global burden of disease, *Br. Med. Bull.*, 2003, **68**, 1–24.
9 V. Weaver, T. Buckley and J. Groopman, Approaches to environmental exposure assessment in children, *Environ. Health Perspect.*, 1998, **106**(suppl. 3), 827–832.
10 A. Hill, The environment and disease: association or causation?, *Proc. R. Soc. Med.*, 1965, **58**, 295–300.
11 M. Peters, K. Timmerhaus and R. West, *Plant Design and economics for Chemical Engineers*, McGraw Hill, New York, NY, 2003.
12 G. Ulrich and P. Vasudevan, *Chemical Engineering Process Design and Economics*, Process Publishing, Durham, NH, 2004.
13 C. Wild, Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology, *Cancer Epidemiol., Biomarkers Prev.*, 2005, **14**(8), 1847–1850.
14 M. Vrijheid, R. Slama, O. Robinson, L. Chatzi, M. Coen, *et al.*, The Human Early-Life Exposome (HELIX): Project Rationale and Design, *Environ. Health Perspect.*, 2014, **122**(6), 535–544.
15 T. Simmons, *EXPOsOMICS Project*, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, 2013.
16 Centers for Disease Control and Prevention, *Exposome and Exposomics*, 2012, Retrieved 5 March 2013, http://www.cdc.gov/niosh/topics/exposome/.
17 M. Overcash and E. Griffing, *Life cycle inventory (lci) database*, Environmental Clarity Inc., http://www.environmentalclarity.com, 1998–2015.
18 TRACI, *Tool for the Reduction and Assessment of Chemical and other environmental Impacts*, U.S. EPA, Cincinnati, OH, 2011.
19 R. Chadwick, The Icelandic database – do modern times need modern sagas, *Br. Med. J.*, 1999, **319**(7207), 441–444.

20 Biobank UK, UK Biobank: Report of the Integrated Pilot Phase, UK Biobank Coordinating Centre, Cheshire, UK, http://enquires@ukbiobank.ac.uk, 2006.
21 F. Collins and K. Jegalin, Deciphering the code of life, *Scientific American*, 1999, Dec: 86-91.
22 M. Abraham and N. Nguyen, Green Engineering: Defining the Principles, *Environ. Prog.*, 2003, **22**(4), 233–236.
23 P. Anastas and J. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, New York, 1998.
24 N. Winterton, Twelve more green chemistry principles? *Green Chem.*, 2001, G73–G75.
25 M. Overcash and V. Cunningham, *Pollution Prevention Technology*, Today Series, Amer. Inst. of Chemical Engineers, New York, N.Y., 1993, 289pp.