Biomethods, Kitakyushu, Japan), and then cultured for 2 days. From Day 3 to Day 6, 10 nM of RA was added to the EB-medium to promote neuronal differentiation. The EB-medium with RA was exchanged every two days, followed by culturing in EB-medium alone from Day 7 to Day 8. On Day 9, the EBs were seeded onto ornithine-laminin (O/L)-coated 24-well-plates and the medium was changed to neuronal inducing medium (NIM) composed of DMEM/F12 medium, N-2 Supplement, Penicillin-Streptomycin, and 10 ng/ml bFGF on Day 10. NIM was changed every 3 days until Day 22.

Culture and neuronal differentiation of human embryonic stem cells.

Similar to the mESC culture protocol, MEFs were used as feeder cells for the culture and passage of the hESC line KhES3 in the medium composed of DMEM/F12 containing 20% KSR, 100 μM NEAA, 2 mM L-glutamine, 100 μM 2-ME, and 5 ng/ml bFGF. After five times of passages with additional MEFs, the MEFs were eliminated by a brief enzymatic treatment. The hESC colonies left on the dishes were harvested. The hESCs (purity > 99%) were seeded at 1.0×10^5 cells/MSA in the medium containing DMEM/F12, 20% KSR, 100 μM NEAA, 2 mM L-glutamine, 100 μM 2-ME, and 10 μM of ROCK inhibitor Y-27632 (Day 1). The generated EBs were cultured for 7 days in the medium, which was exchanged every two days, followed by growth in the medium without Y-27632 for two days. The growing EBs were cultured for 2 additional days in NIM containing DMEM/F12: Neurobasal® Medium (1:1), N-2 Supplement, B-27® Supplement, GlutaMAX™-I, Penicillin-Streptomycin to promote neuronal differentiation. Then EBs were re-plated onto O/L-coated 24-well-plates at 20 EBs/well. They were cultured for 7 days in neuronal proliferation medium (NPM) containing DMEM/F12: Neurobasal® Medium (1:1), two-fold concentrations of N-2 Supplement, two-fold concentrations of B-27® Supplement, GlutaMAX™-I, Penicillin-Streptomycin, 20 ng/ml bFGF. NPM medium was exchanged every 3 days until Day 18, at which point cells were cultured in neuronal differentiation medium (NDM) containing Neurobasal® Medium, B-27® Supplement, GlutaMAX™-I, Penicillin-Streptomycin, 10 ng/ml BDNF until Day 50. The medium was exchanged every 3 days.

Immunocytochemistry

After fixation for 10 min with 4% paraformaldehyde, the cells were permeabilized in 0.1% Triton X-100 in PBS. The cells were incubated with 1% BSA, which was followed by overnight incubation with primary antibody specific for MAP2 (1:200). After PBS washing, the cells were incubated at room temperature for 1 hr with Alexa 568-labeled secondary antibodies (1:1,000). Nuclei were stained using 2 μg/ml of Hoechst 33342 for 15 min.

Morphological measurement by IN Cell Analyzer.

Image analysis by the IN Cell Analyzer 1000 (ICA, GE Healthcare UK Ltd.,, Buckinghamshire, UK) was performed as following. The microphotographs of 57 fields (0.60 mm^2) per well of 24 well plate, (*i.e.,* 1,368 fields per exposure group) were taken automatically. The fields in a well were created without overlap. The fluorescent signal detected by the 535-nm laser line combined with a HQ620 60 M emission filter was considered to be the MAP2-positive signal of neurons. The fluorescent signal detected using the 360-nm laser line combined with a HQ460 40 M emission filter was considered to indicate the Hoechst33342 positive nuclei. Fluorescence emission was separately recorded in the blue and red channels, and a flat field correction was applied for inhomogeneous illumination of the scanned area for each of the two channels.

A typical merged image of hESC-derivatives is shown in Figure S1A. Hoechst-positive nuclei were recognized using IN Cell Developer Software (GE Healthcare UK Ltd.) and replaced by yellow dots to accurately count the nuclei number (Figure S1B). MAP2-positive signals were also recognized by this software with a threshold appropriate for tracing the MAP2-postive neurites. The replaced pink images are shown in Figure S1C. MAP2-positive signals surrounding nuclei were regarded as the cell bodies of differentiated neural cells. The number of nuclei within the cell body was considered to be the number of MAP2-positive neurons in each field. The areas considered to be cell bodies were subtracted from the MAP2-postive images to generate the image of the neurites as shown in Figure S1D. Then the software automatically replaces the MAP2-positive neurite images with branching morphologies, which indicated neurite-length as green center lines and branching points as blue circles (Figure S1D, S1E, and S1F). As shown in Figure S1E, the cell bodies were successfully distinguished from the neurites. The total length of the MAP2-positive projection was automatically measured on its midline (Figure S1F). The values for neurite length/cell were calculated by dividing the total MAP2-positive neurite length within a field by the number of MAP2-positive neurons. The branching points of MAP2-positive projections were automatically counted as the total number of branching points of the MAP2-positive projections (Figure S1F). The values for branching points/cell were calculated by dividing the total MAP2-positive neurite branching points within a field by the number of MAP2-positive neurons. The average of the neurite length/cell or branching points/cell of all fields in one well of the 24-well plate were indicated as the value of one experiment. Images of mESC-derived neuronal cells were analyzed in the same manner.

To evaluate the accuracy of ICA measurements, the total neurite lengths of 5 randomly selected fields were manually measured using ImageJ (IJ) software (NIH).

The correlations between data from IN Cell Developer Software and from IJ are shown in Figure S2. Although the values obtained by ICA tended to be higher, (approximately 1.73-fold for neurite length and 4.42-fold for branching points), the two values obtained by ICA and IJ were well correlated, which demonstrated the accuracy of the ICA measurement.

Procedure of Bayesian network analysis

The linkages between MeHg and differentiation marker genes were visualized using a network model that was based on the Bayesian algorithm, which was modified from the one defined previous study (Toyoshiba et al. 2004). All the data of mRNA levels were normalized by expressing values as a ratio to the average of DMSO control. The values for MeHg concentration were expressed as following: 1.000, 1 nM of MeHg; 1.100, 10 nM of MeHg; 1.200, 100 nM of MeHg. Calculated values used for Bayesian network analysis are shown in Table S2 and Table S3.

The network was quantified to calculate the posterior probability distribution for the strength of the linkages on the basis of gene expression and chemical exposure datasets. Briefly, a network consists of a collection of P nodes, denoted G1, G2,... GP, with observed values n1, n2,....np. We define βij (i,j=1, 2,..., P) as parameters in the log-linear function form describing the linkage from node i to node j. Mathematically, this is written as

$$E[\log(G_j)] = \sum_{i=1,\neq j}^{P} I_{ij}\beta_{ij}\log(g_i)$$

where $E[\log(G_j)]$ represents the expectation for the natural logarithm of G_j and I_{ij} (i,j = 1,2, ... ,P) is an indicator function that equals 1 if node G_i has a link to node G_j, otherwise it equals 0. If a node has a regulatory effect on node G_i, then that node is referred to as a "Parent of node G_i" and we refer to it as belonging to the set $Pa(G_i)$. The prior distribution for the variance is assumed to be inverse Gamma and assuming that the natural log of G_j follows a normal distribution with mean and standard deviation β_j, posterior distributions for each parameter can be estimated. The posterior distributions for the linkages were derived using Gibbs sampling. Gibbs sampling method has no limitation on the number of possible parents and is easy to cooperate with knowledge information or past experimental result by taking the information into the prior distribution. The network was used to evaluate the ability of the algorithm to produce higher posterior probabilities (P-value) at the correct linkage in the network. In each simulation, Gibbs sampling was performed 66,000 times. The initial Gibbs samplings were considered to be the burn-in period and were removed in estimating and the last 32,000 iterations were used to establish.

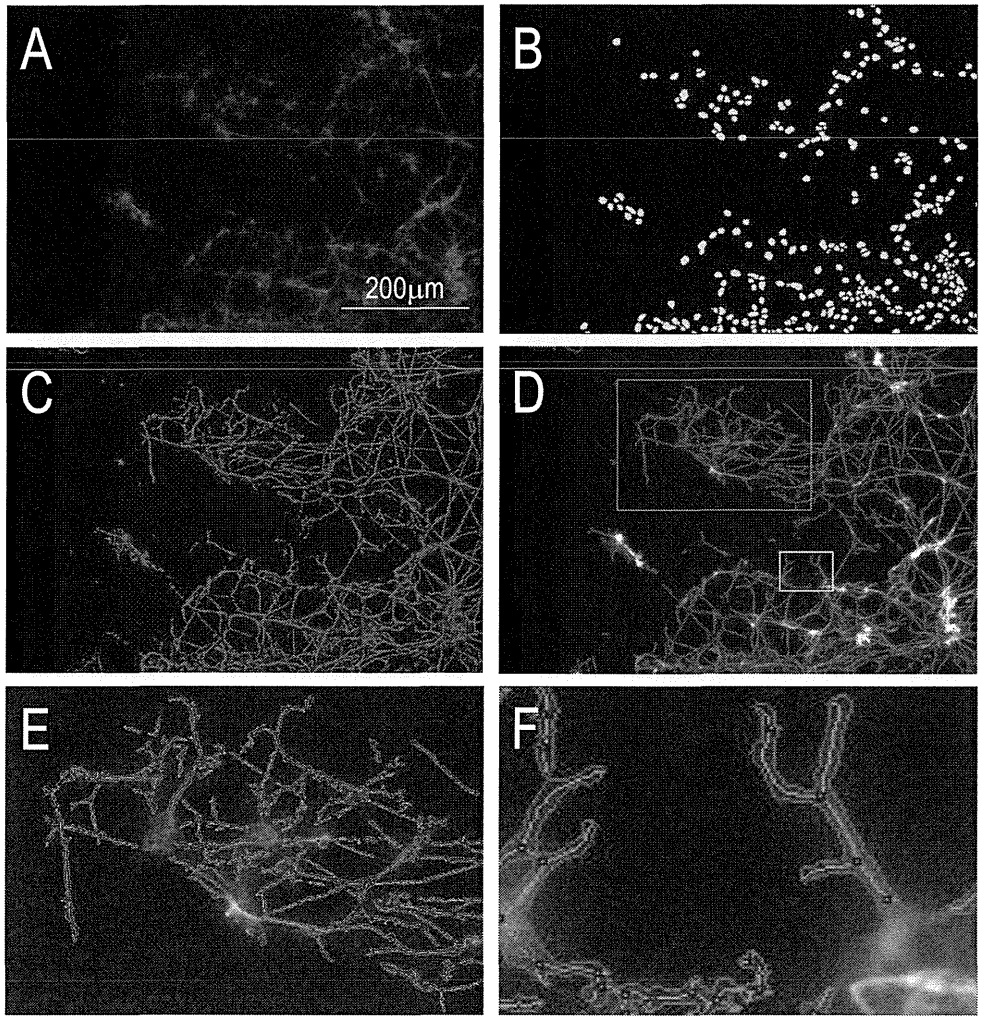


Figure S1. The process of image analysis using the ICA. (A) A typical merged image for MAP2-positive hESC derivatives. (B) Processed images for the recognition of nuclei. (C) MAP2-positive cells. (D) MAP2-positive neurites created using image A. (E) Enlarged image (1.5X) of the inside of the green square in D. (F) Enlarged image (8.5X) of the inside of the yellow square in D. The nuclei were discerned on the basis of a threshold Hoechst signal strength in the blue channel images. Neighboring nuclei were automatically separated (B, yellow spots). MAP2-positive cells were recognized on the basis of the threshold of the Alexa 555 signal strength. The value of the MAP2-positive area was obtained from similar processed images. The values for neurite length and branching points were obtained from processed images (similar to D). Neurites were distinguished from cell bodies, which were determined as MAP2-positive regions surrounding nuclei as shown in E. The median lines of the neurites were traced automatically (F, green lines), and their lengths were considered to be equivalent to neurite lengths. Branching points were determined to be the branching points of the median line (F, blue circle).

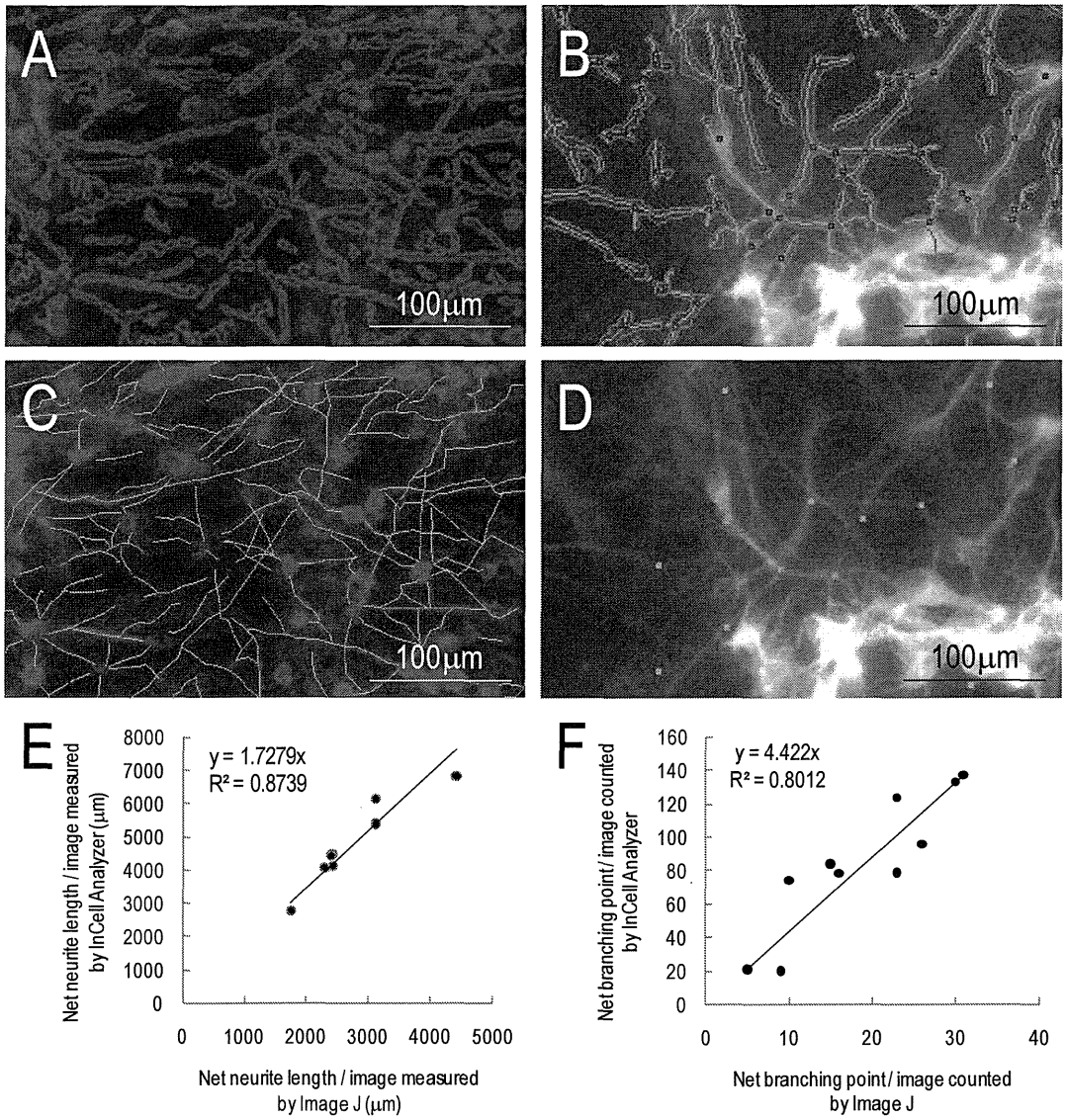


Figure S2. Comparison of ICA and IJ data. The identical areas analyzed by ICA (A) were traced again manually by IJ (C). The total neurites traced are shown as green lines. (E) The correlation between ICA and IJ data for neurite length from 8 photographs. Although the results obtained by ICA were approximately 1.7-times larger than those obtained using IJ, they showed a reliable correlation (R^2=0.8739). In addition, the number of branching points was counted using both ICA (B) and IJ (D). The branching points are indicated by green dots. (F) The correlation between ICA and IJ data of branching points from 10 photographs. The results of ICA were approximately 4.4-times larger than those obtained by IJ, which was probably because the crossing points of neurites were recognized as branching points. In spite of such a large difference, the branching points showed a robust correlation (R^2=0.8012).

Table S1. The primer used for RT-PCR in this study.

Genes	Primer sequences		Product size (bp)
	Forward (5' to 3')	Reverse (5' to 3')	
Mouse			
Nanog	TTTGCCTAGTTCTGAGGAAGC	AGGCAGGTCTTCAGAGGAAG	154
Pou5f1	CACGAGTGGAAAGCAACTCA	CCAAGGTGATCCTCTTCTGC	174
En1	CCGGTGGTCAAGACTGACTC	CGCTTGTCTTCCTTCTCGTT	137
Nodal	CTCTGGCGTACATGTTGAGC	CGTGAAAGTCCAGTTCTGTCC	110
Nes	AGATCGCTCAGATCCTGGAA	CCAAGAGAAGCCTGGGAACT	145
Otx1	AGGCGAGAGGTAGATGGTGA	GAGGCCAGGAGCAGTTCAG	109
Pax6	CGGTGAGCAGATGTGTGAGA	CATGCTGGAGCTGGTTGG	113
Emx2	GACCCGTCCACCTTCTACC	AGCGTTGTGCAAAAGGAAAC	106
Hoxb1	GAACCCAGCACTCTCACTCC	GGTGAAGTTTGTGCGGAGAC	123
Hoxb4	CTGGATGCGCAAAGTTCAC	CTCCTTCTCCAACTCCAGGAC	112
Olig1	CTGCGCGAAGTTATCCTACC	CAGCAGCAGGATGTAGTTGC	111
Olig2	TTACAGACCGAGCCAACACC	GATGGGCGACTAGACACCAG	166
Map2	GAGAAGGAAGCCCAACACAA	CTTCCTCCACTGTGGCTGTT	124
Gapdh	GGTGCTGAGTATGTCGTGGA	GTGGTTCACACCCATCACAA	141
Human			
NANOG	GATTTGTGGGCCTGAAGAAA	AAGTGGGTTGTTTGCCTTTG	155
POU5F1	CTCACCCTGGGGGTTCTATT	TCTCCAGGTTGCCTCTCACT	235
EN1	GTCAAAACTGACTCGCAGCA	GCTTGTCCTCCTTCTCGTTC	130
NODAL	CCAAGCAGTACAACGCCTATC	TGGTGATCTAGGAGCACTCTG	205
NES	AGCGTTGGAACAGAGGTTG	AGGCTGAGGGACATCTTGA	171
OTX1	TCTTCGCCAAGACTCGCTAC	GCACTGGAGAGGACTTCTTCT	180
PAX6	GTCCATCTTTGCTTGGGAAA	CTAGCCAGGTTGCGAAGAAC	111
EMX2	CGATATCTGGGTCATCGCTTC	TGCCAGCTGCTTCCTTTC	183
HOXB1	CTCCGAGGACAAGGAAACAC	CAGTTCTGTCAGCTGCCTTG	175
HOXB4	CTGGATGCGCAAAGTTCACG	CGTGTCAGGTAGCGGTTGTA	138
OLIG1	GTCATCCTGCCCTACTCAGC	CTGCCCAGCAGTAGGATGTAG	107
OLIG2	TATAGATCGACGCGACACCA	GAAAAAGGTCATCGGGCTCT	208
MAP2	GGAGTAACCAAGAGCCCAGA	CTCTGCGAATTGGCTCTGAC	166
GAPDH	ATGGGGAAGGTGAAGGTCG	TGGAATTTGCCATGGGTGGA	170

Table S2. The normalized expression levels of genes in the mESC-derivatives exposed to MeHg.*

Actual MeHg (nM)	1	1	1	1	1	1	10	10	10	10	10	10	100	100	100	100	100	100
MeHg	1.000	1.000	1.000	1.000	1.000	1.000	1.100	1.100	1.100	1.100	1.100	1.100	1.200	1.200	1.200	1.200	1.200	1.200
En1	0.715	1.367	1.254	1.048	0.454	1.192	1.100	0.837	0.633	1.850	1.737	2.724	1.582	1.429	1.172	0.935	2.914	1.252
Nodal	0.653	2.140	0.957	1.112	0.757	0.762	0.757	0.520	0.387	0.791	0.349	0.779	1.797	7.764	2.720	4.422	69.679	2.002
Nes	0.899	0.739	1.090	1.234	1.085	1.175	0.899	0.739	1.090	1.085	1.175	0.920	1.521	1.254	1.558	1.128	1.246	1.455
Otx1	1.288	1.603	1.330	0.629	0.208	0.915	0.463	0.442	0.477	3.662	0.839	4.595	2.387	10.783	1.136	0.394	5.127	2.987
Emx2	0.560	1.276	0.970	0.996	0.526	0.351	1.108	0.677	0.724	0.616	0.491	0.621	0.222	0.274	0.845	0.823	0.068	0.408
Pax6	1.286	1.203	1.290	1.110	1.102	1.102	1.110	0.947	0.863	0.813	0.649	0.704	0.498	0.522	0.536	0.687	0.165	0.362
Hoxb1	2.288	0.825	1.525	20.682	0.206	1.975	0.261	0.166	0.343	5.198	6.380	19.444	16.609	11.751	1.677	0.508	7.128	7.189
Hoxb4	0.996	0.890	0.898	0.958	0.987	0.823	0.964	0.858	0.562	0.765	0.794	0.528	0.511	0.551	0.545	1.004	0.284	0.614
Map2	1.018	0.837	1.080	0.837	0.759	0.808	0.855	0.902	0.787	0.797	0.839	0.674	0.378	0.359	0.401	0.627	0.106	0.426
Olig1	1.687	1.143	1.232	1.478	1.406	1.338	0.928	1.300	1.393	1.084	1.347	1.164	0.889	1.162	1.054	1.135	0.669	1.094
Olig2	1.107	0.602	0.844	0.721	0.734	0.687	0.670	0.717	0.679	0.717	0.751	0.517	0.378	0.486	0.393	0.598	0.160	0.473

*: Normalization procedure was described in the page-4 of this supplement.

Table S3. The normalized expression levels of genes in the hESC-derivatives exposed to MeHg.*

Actual MeHg (nM)	1	1	1	1	1	1	10	10	10	10	10	10	100	100	100	100	100	100
MeHg	1.000	1.000	1.000	1.000	1.000	1.000	1.100	1.100	1.100	1.100	1.100	1.100	1.200	1.200	1.200	1.200	1.200	1.200
EN1	1.159	0.825	0.958	1.048	0.666	1.389	1.040	0.542	1.166	1.293	1.560	0.700	1.322	0.854	1.077	1.448	0.795	1.129
NODAL	0.932	0.915	1.522	0.714	0.932	0.891	1.215	1.292	1.557	0.938	3.444	2.460	4.464	3.368	2.017	0.761	2.112	0.873
NES	1.274	1.253	0.964	0.966	1.054	1.016	1.045	0.973	1.142	1.161	0.928	0.819	1.125	1.004	0.781	1.208	0.635	0.936
OTX1	1.357	0.859	0.688	0.928	1.129	1.225	1.073	0.817	1.246	1.385	0.516	1.059	1.544	1.045	0.879	1.122	0.413	1.322
EMX2	0.891	0.891	0.382	0.501	0.639	0.885	0.639	0.976	0.611	0.680	0.178	0.274	0.573	0.373	0.243	0.617	0.147	0.672
PAX6	0.907	0.848	0.412	0.623	1.049	0.775	0.760	1.044	0.598	0.681	0.240	0.466	0.647	0.564	0.469	0.740	0.243	0.520
HOXB1	0.406	0.834	5.161	1.421	1.928	0.930	1.308	1.274	1.209	1.975	2.663	1.764	1.520	2.430	1.509	1.509	2.300	1.184
HOXB4	0.207	1.039	9.919	6.170	2.876	1.321	3.159	1.897	3.331	1.887	9.283	9.758	5.806	7.686	7.948	1.543	10.607	2.876
MAP2	1.373	1.337	0.553	0.724	1.056	1.337	0.828	0.988	0.913	0.863	0.425	0.521	0.913	0.589	0.464	1.002	0.386	0.714
OLIG1	1.514	1.193	0.969	1.215	1.339	1.183	1.244	1.798	1.095	1.283	1.808	0.988	1.278	0.991	0.745	0.346	0.762	1.125
OLIG2	1.189	1.845	1.242	1.315	1.218	2.121	2.579	2.783	1.052	2.843	2.843	1.977	1.796	1.676	1.423	1.387	0.712	1.580

*: Normalization procedure was described in the page-4 of this supplement.

Int. J. Mol. Sci. **2012**, *13*, 187-207; doi:10.3390/ijms13010187

OPEN ACCESS

International Journal of
Molecular Sciences
ISSN 1422-0067
www.mdpi.com/journal/ijms

Article

Multi-Parametric Profiling Network Based on Gene Expression and Phenotype Data: A Novel Approach to Developmental Neurotoxicity Testing

Reiko Nagano [1,†], Hiromi Akanuma [1,†], Xian-Yang Qin [1,2,†], Satoshi Imanishi [3], Hiroyoshi Toyoshiba [1], Jun Yoshinaga [2], Seiichiroh Ohsako [3] and Hideko Sone [1,*]

[1] Health Risk Research Section, Research Center for Environmental Risk, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba 305-8506, Japan; E-Mails: nagano.reiko@tasc-nt.or.jp (R.N.); akanuma.hiromi@nies.go.jp (H.A.); y_qin@envhlth.k.u-tokyo.ac.jp (X.-Y.Q.); Toyoshiba_Hiroyoshi@takeda.co.jp (H.T.)

[2] Department of Environmental Studies, Graduate School of Frontier Science, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 270-8563, Japan; E-Mail: junyosh@k.u-tokyo.ac.jp

[3] Center for Disease Biology and Integrative Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8654, Japan; E-Mails: imanishi@m.u-tokyo.ac.jp (S.I.); ohsako@m.u-tokyo.ac.jp (S.O.)

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: hsone@nies.go.jp; Tel.: +81-29-850-2464; Fax: +81-29-850-2546.

Received: 3 October 2011; in revised form: 14 November 2011 / Accepted: 30 November 2011 / Published: 23 December 2011

Abstract: The establishment of more efficient approaches for developmental neurotoxicity testing (DNT) has been an emerging issue for children's environmental health. Here we describe a systematic approach for DNT using the neuronal differentiation of mouse embryonic stem cells (mESCs) as a model of fetal programming. During embryoid body (EB) formation, mESCs were exposed to 12 chemicals for 24 h and then global gene expression profiling was performed using whole genome microarray analysis. Gene expression signatures for seven kinds of gene sets related to neuronal development and neuronal diseases were selected for further analysis. At the later stages of neuronal cell

differentiation from EBs, neuronal phenotypic parameters were determined using a high-content image analyzer. Bayesian network analysis was then performed based on global gene expression and neuronal phenotypic data to generate comprehensive networks with a linkage between early events and later effects. Furthermore, the probability distribution values for the strength of the linkage between parameters in each network was calculated and then used in principal component analysis. The characterization of chemicals according to their neurotoxic potential reveals that the multi-parametric analysis based on phenotype and gene expression profiling during neuronal differentiation of mESCs can provide a useful tool to monitor fetal programming and to predict developmentally neurotoxic compounds.

Keywords: developmental neurotoxicity; embryonic stem cells; high-content screening; Bayesian network modeling; gene expression; multi-parametric analysis

1. Introduction

One of the emerging issues in developmental neurotoxicology is to detect effects of chemicals on fetal programming, which is defined as variations in metabolism, gene expression and genome modification during fetal life that induce or repress the somatic structure and physiological systems after development [1–4]. A significant issue in the prevention of neurodevelopmental deficits of chemical origin is the paucity of testing of chemicals for developmental neurotoxicity [5]. New, precautionary approaches that recognize the unique vulnerability of the developing brain are needed for testing and to control the use of chemicals.

Toxicity testing using embryonic stem cells (ESCs) is an efficient approach for developmental neurotoxicity testing (DNT) [6,7]. Compared with the DNT in animal studies, which are costly, time-consuming, and require considerable numbers of laboratory animals, the ESCs test is unique in that, in a relatively simple cell-line-based assay, it incorporates the entire differentiation route from pluripotent ESCs into differentiated cells [8]. Furthermore, as the neuronal differentiation of ESCs provides insight into the early neurogenesis during embryonic development, several protocols have been developed based on the disturbances of this process to model developmental neurotoxicity [9,10]. A 13-day neural differentiation protocol of mouse embryonic stem cells (mESCs), which is combined with morphological observation, immunocytochemistry, gene expression and flow cytometry, has been applied to assess the developmental neurotoxicity of methyl mercury chloride [9]. More recently, a broad gene expression profile during a 20-day differentiation process of mESCs has been successfully designed, in which transcription-based end points have been used to identify the disturbed neuronal differentiation of mESCs [10]. Developing neurons display plasticity in the type of neurotransmitter phenotype that they can assume, and alterations of synaptic activity and expression of neurotrophic factors can influence the "wiring" of developing neuronal circuits [11]. Consequently, exposure to environmental chemicals that promote or interfere with synaptic activity or expression/function of neurotrophins can result in miswiring, leading to neurobehavioral anomalies. However, a sensitive

method for quantitatively measuring alterations of fetal programming during neuronal differentiation, particularly in the connection between the early disturbances and the later outcomes, has not yet been devised.

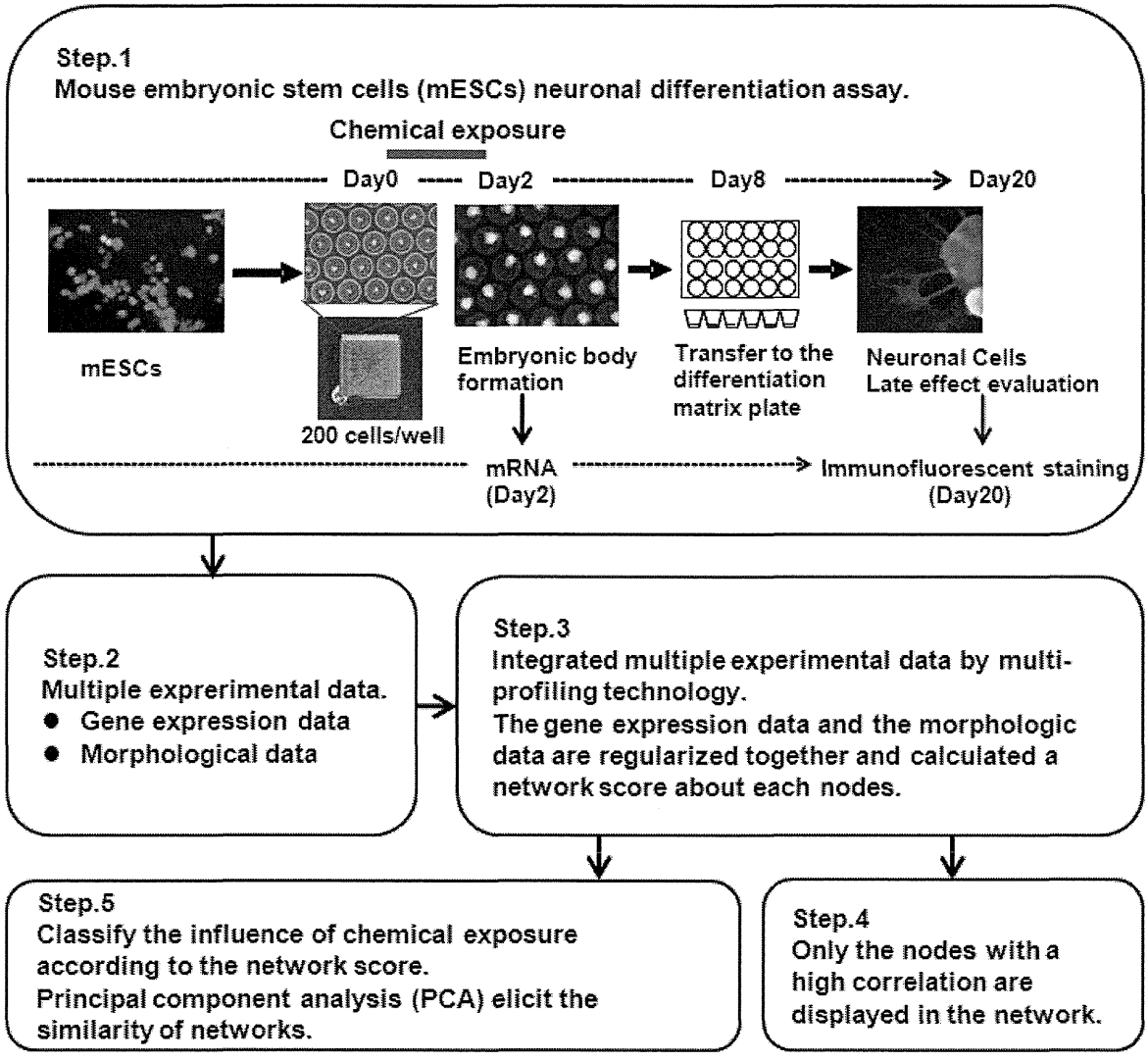
Here, we produced a high-content and sensitive method for quantitatively measuring the developmental neuronal toxicity of 12 environmental chemicals (see Table 1) using mESCs test combined with DNA microarray analysis, morphological analysis and Bayesian approaches. This confers a new predictive insight for chemical screening in a complex cell culture system that mimics early mammalian embryonic development. We performed multi-parametric profiling of gene expression data sampled at the early stage of mESC differentiation and neuronal phenotype data sampled at a later stage of neuronal cell differentiation after embryoid body (EB) formation. Then, these sampled data were analyzed by a Bayesian network analysis (BNA). This analysis can be depicted graphically to represent the probability structure of the causal complex [12–14].

Table 1. Summary of 12 test chemicals.

Chemical Name	Ellipsis	Intended Use	Physiological Effect and Toxicity	Mode of Action	Target Protein
Triiodotyronine	T3	Endogenenous hormne	Pseudo thyroid hormone	transcriptional regulation	Thyroid hormone receptor (TR)α, TRβ
Dexamethazone	DEX	Medicinal drug	Pseudo corticosteroid hormone	transcriptional regulation	Glucocorticoid receptor (GR)
17b-Estradiol	E2	Endogenenous hormne		transcriptional regulation	Estrogen receptor (ER)α, ERβ
5a-Dihydrotestosterone	DHT	Endogenenous hormne		transcriptional regulation	Androgen receptor (AR)
2,3,7,8-tetrachlorodibenzo-*p*-dioxin	TCDD	Unintentional chemical	Multi-toxicity	transcriptional regulation	Aryl hydrocarbon receptor (AhR)
Methoprene acid	MPA	Pesticides	Teretogenecity	transcriptional regulation	Retinoid X receptor (RXR)α, RXRβ, RXRγ
Cyclopamine	CPM	Medicinal drug	Teretogenecity	Signal inhibition	Hadgehog signaling pathway
Thalidmide	TMD	Medicinal drug	Teretogenecity and Autism	Unknown	Oxidative stress
4(OH)-2′,3,3′,4′,5′-pentachlorobephenyl 107	PCB	Metabolite of PBC	Multi-toxicity	Unknown	Unknown (ERα, oxidativestress)
Permethrin	PMT	Pesticides	Neuro-toxicity	Unknown	Oxidative stress
Bisphenol A	BPA	Plastic materials	Reproductive and Neuro-toxicity?	Unknown	Unknown (ERα, ERRγ)
Bis(2-ethylhexyl) phthalate	DEHP	Plastic materials	Reproductive and Neuro-toxicity?	Unknown	Unknown [Peroxisome proliferator-activated receptor (PPAR)α, antiTR]

The Bayesian algorithm used in this study was proposed by Toyoshiba *et al.* as a prediction tool for the effect of exposure to chemicals [15]. The TAO-Gen algorithm is based on the assumption of a linear relationship between changes in the expression levels of two genes following chemical exposure [16], which employs the Gibbs sampling method on the search algorithm to estimate posterior probability distribution [17,18]. The advantage of Gibbs sampling is that it samples from a full conditional distribution and it is an efficient and easy sampling procedure. Gibbs sampling is a Markov chain Monte Carlo method, which involves generating a sample from one or several variables with an acceptance probability of one. This process is repeated until the sampled probability distribution is close to the actual distribution. This algorithm can be used to search for key transcription factors of signal transduction during ES cell differentiate process [19].

Figure 1. Experimental steps in this study for the assessment of developmental neurotoxicity.



Therefore, the overall aim of this paper is to make a conceptual and methodological proposal to establish a more efficient approach for DNT (Figure 1). More specifically, two objectives are addressed. The first is to describe the DNT design and to identify multi-parametric profiling networks (MPNs) multiple-index networks for 12 environmental chemicals as examples. These are based on the

gene expression signatures of mESCs and phenotype profiling of neurons differentiated from EBs. The second objective is to suggest an information-predictive approach to detect alterations of fetal programming that can be made operational using BNA. We propose BNA as an operational tool for empirically applying the DNT approach.

2. Results and Discussion

2.1. Phenotype Profiling Based on the Morphology of Differentiated Neuronal Cells by High-Content Image Analysis and Generation of Phenotypic Networks

EBs neurally differentiated into neural cells after transfer to OP/L-coated plates. Effects of the 12 environmental chemicals on neural cell growth and NS morphology are shown in Figure 2. Dexamethazone (Dex), Permethrin (PMT) and 17β-estradiol (E2) significantly increased neurite length, while 4-OH-2′,3,3′,4′,5′-pentachlorobephenyl 107 (PCB), triiodotyronine (T3), Thalidmide (TMD), cyclopamine (CPM) and methoprene acid (MPA) significantly decreased neurite length compared with DMSO control (Figure 2A). In glial fibrillary acidic protein (GFAP) positive glial cells, Dex, 5α-dihydrotestosterone (DHT), bisphenol A (BPA) and PCB significantly increased neurite length, while TMD significantly decreased neurite length (Figure 2B). Chemicals were then classified based on morphological features by MPN analysis to extract and predict their toxicities. 12 phenotypic networks (PNs) were generated from the MPN analysis based on the phenotypes of neuronal cells and NSs. We manually classified three categories out of the12 PNs depending on network structures (Figure 3).

Figure 2. Morphological data of MAP2-positive neurons and glial cells. (**A**) Total length of MAP2-positive neurons per well; (**B**) Total length of glial processes per well. * $P < 0.05$, ** $P < 0.001$ *vs.* the vehicle control (DMSO).

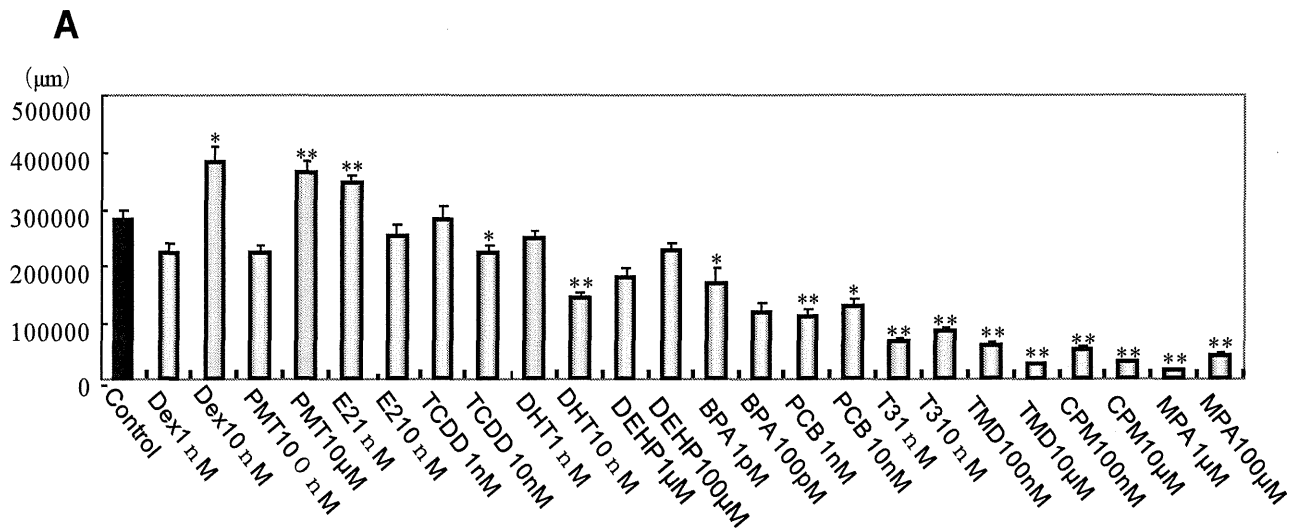


Figure 2. *Cont.*

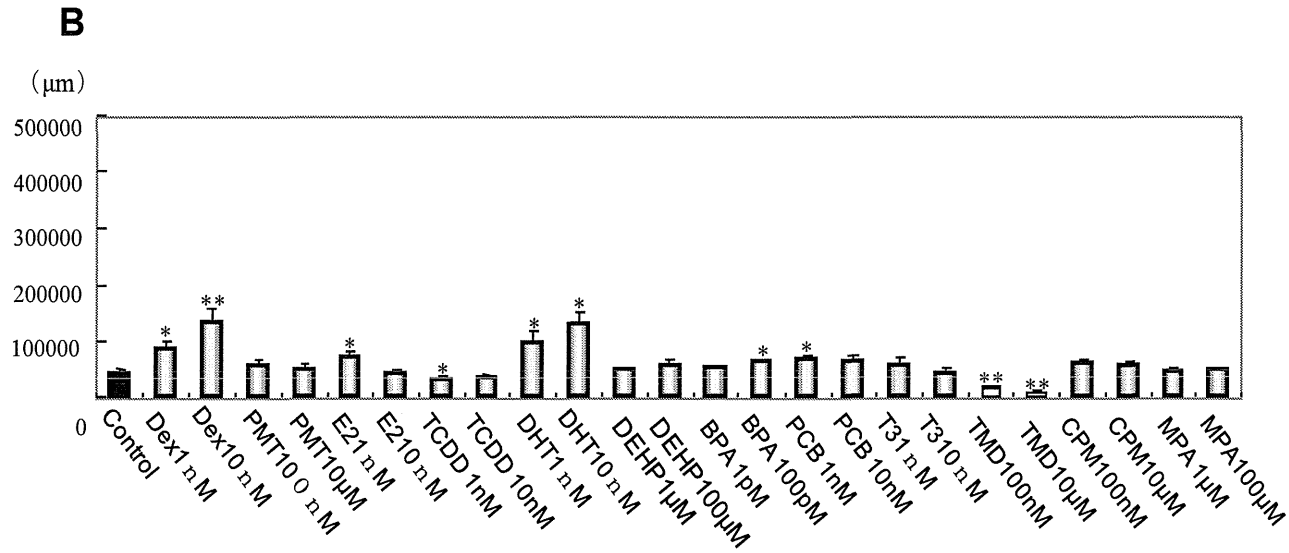
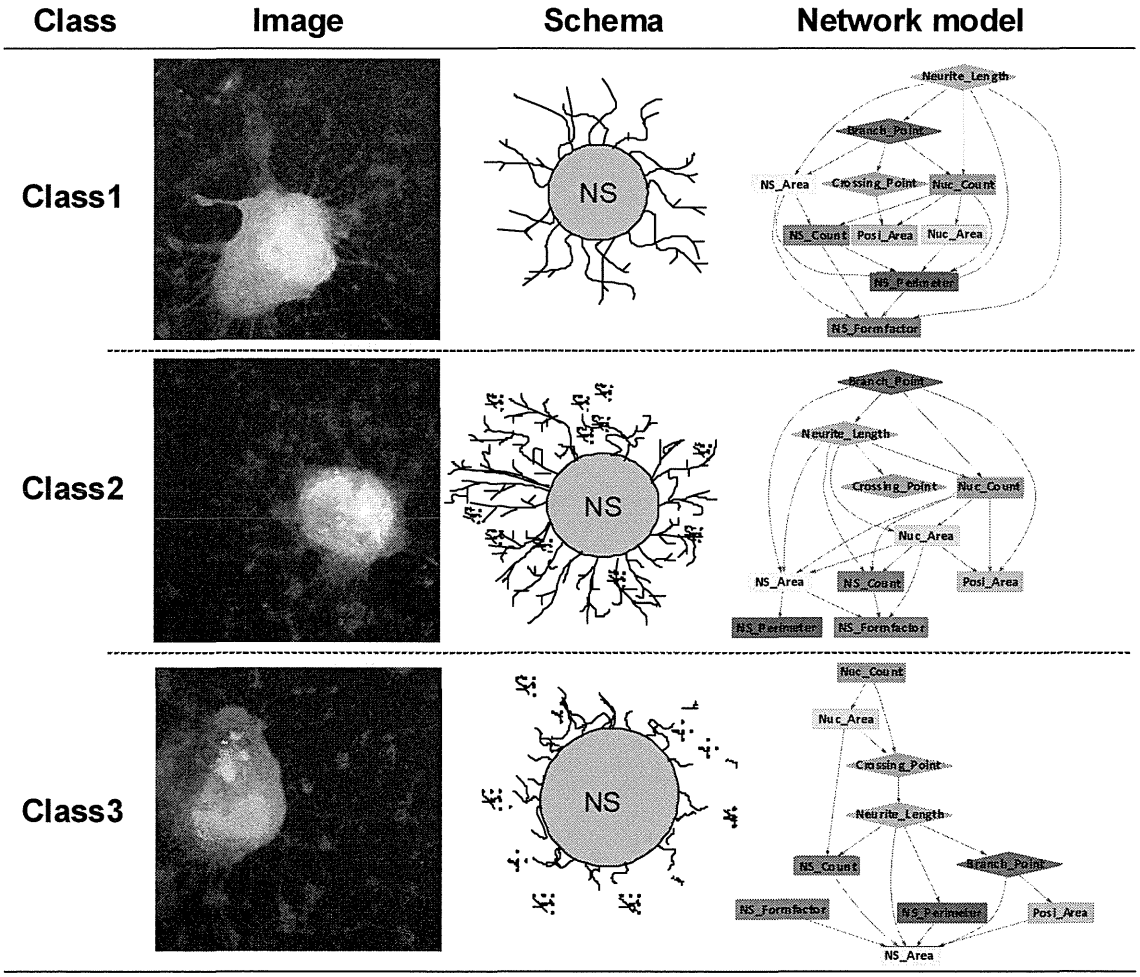


Figure 3. Classification based on morphological imaging and phenotypic feature networks. Class 1: Extension from the turning point is short while the neurite is long; Class 2: Neurite is long and the branch point is complex; Class 3: Neurite is short and there are many nucleus count.



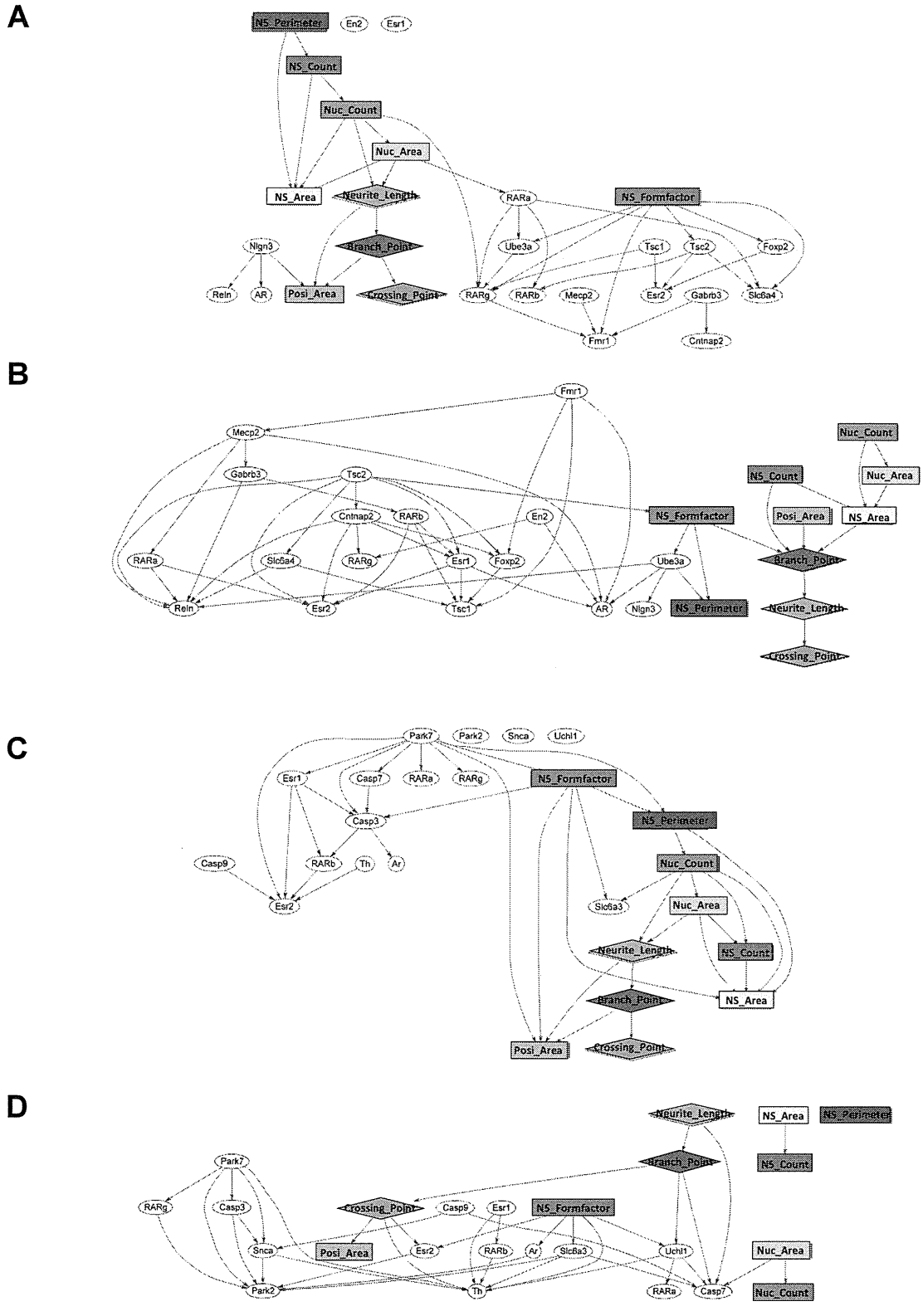
2.2. Generation of a Comprehensive Network Based on Gene Expression and Phenotype Profiling by a Bayesian Network Model

A significant advantage of our unique MPN analysis is that it can predict the correlation coefficient for each pair of nodes, regardless of the data types. Our initial efforts were to derive the interactions between variations of gene expression data after chemical exposure at the early stage of mESC differentiation and effects on the neuronal phenotype data sampled at a later stage of neuronal cell differentiation after EB formation. That is to perform a comprehensive analysis combining data from two different properties. We extracted a discriminative gene group as a gene expression signature from exhaustive genetic profiling, each group was defined by their characteristic category (Table 2) and these gene sets were used in a gene and phenotype interaction network (GPIN) with cell morphological data (Figure 4). To verify whether the MPN analysis can draw out the developmental neurotoxicity, typical examples of DPINs for autism and Parkinson's disease related gene sets exposed to TMD and PMT, respectively, were discussed.

Table 2. Lists of 7 gene sets selected for network analysis.

Alzheimer	Autism	Parkinson	Axon Guidance	Pluripotent	Neural Development	Oxidative-Stress
AR	AR	AR	1500003O03Rik	Arid3b	Atbf1	Aass
ApoE	Cntnap2	Casp3	Abl1	Esrrb	Cdyl	Als2
App	En2	Casp7	Ablim1	Fkbp3	Fos	Apoe
Bace	Esr1	Casp9	Cfl1	Hdac2	Gbx2	Ctsb
Casp3	Esr2	Esr1	Cxcl12	Klf4	Gfap	Dnm2
Casp7	Fmr1	Esr2	Efna4	Mybbp1a	Hras1	Fancc
Esr1	Foxp2	Park2	Epha2	Nacc1	Map2	Gpx7
Esr2	Gabrb3	Park7	Ephb1	Nanog	Mapk1	Gpx8
Ide	Mecp2	RARa	Nfatc2	Nfkbib	Mapk3	Gusb
Il1r1	Nlgn3	RARb	Nfatc3	Nr0b1	Nestin	Hprt1
Mme	RARa	RARg	Ntng1	Nr5a2	Pla2g6	Kif9
Psen	RARb	Slc6a3	Sema3a	Pou5f1	Raf1	Noxo1
RARa	RARg	Snca	Sema3b	Rex1	Rhog	Nxn
RARb	Reln	Th	Sema3d	Sall4	Rif1	Park7
RARg	Slc6a4	Uchl1	Sema3f	Smarcad1	Rps6ka1	Ppp1r15b
Tnfrsf1a	Tsc1		Sema3g	Smarcc1	Sall1	Prdx2
	Tsc2		Sema6a	Sox2	Shc1	Prdx6-rs1
	Ube3a		Sema6b	Sp1	Smarcad1	Psmb5
			Sema6d	Spag1	Sox2	Recql4
			Srgap3	Trim28	Tuj1	Scd1
			Unc5d	Zfp281	Map2k1	Slc41a3
				c-Myc		Sod1
						Sod3
						Txnip
						Txnrd1
						Xpa

Figure 4. Typical example of GPINs for autism and Parkinson's disease gene sets. Gene expression and morphological parameters were connected by the strength of the correlation. GPINs of autism related genes and morphological parameters: (**A**) the vehicle control (DMSO) and (**B**) TMD exposure. GPINs of Parkinson's disease related genes and morphological parameters; (**C**) the vehicle control (DMSO) and (**D**) PMT exposure.



In DMSO control GPIN, RARα positively regulates Fmr1 expression via positive regulation of RARγ expression, suggesting that RA induced neural differentiation could maintain Fmr1 expression. On the other hand, Mecp2, the responsible gene of Rett syndrome, negatively related with Fmr1 expression. It is reasonable because Mecp2 plays a role in the transcriptional repression of methylated genes including Fmr1 [20]. However, in TMD-exposed GPIN, Fmr1 was not regulated by RARs, indicating the neural induction by RA was counteracted by TMD. TMD repressed expression of Fmr1 and Mecp2 and MPN analysis also revealed that Fmr1 positively related with Mecp2 in TMD treated EB derivatives. The results mean that TMD repressed expression of Mecp2 via repression of Fmr1 expression. It seemed to contradict the epigenetic silencing of Fmr1 gene by Mecp2. However, Zhang *et al.* reported that Mecp2 mRNA expression level was drastically decreased in the brains of Fmr1 knockout mice, an animal model of fragile X syndrome of autism spectrum [21]. This means the relationship between Fmr1 and Mecp2 is different between normal and pathological neurons. Additionally, Gabrb3, a subunit of GABA A receptor, was positively affected by Mecp2. In Mecp2 deficient mice, subtle dysfunction of GABAergic neurons contributes to numerous neuropsychiatric phenotypes [22]. The relationship of morphological parameters and gene expression parameters was also changed by TMD. RARs became a hub connecting the genes and morphological parameters and NS_formfactor related to expression of some genes independently from other morphological parameters in DMSO control GPIN. This result suggests that RA induced neural differentiation via RARs, thereby, inducing morphological changes. In TMD-exposed GPIN, the morphological parameters were independent from RARs and the expression of Tsc2 related to them via positive connection with NS_formfactor. These results also indicated a counteraction by TMD against the neural induction by RA. Tcs2 is well known to affect cell proliferation and to control cell size and neural development [23]. Therefore, Tsc2 had a high correlativity to morphological parameters.

Parkinson's disease is the result of degeneration of dopaminergic neuron expressing Th. Recently, some research showed that exposure to pyrethroids including PMT could change the dopaminergic system [24,25]. The genes including in the Parkinson set can be divided into three groups, the ubiquitin pathway (Park2, Snca and Uchl1) and the mitochondrial pathway (Park7, Casp3, Casp7 and Casp9) [26] and genes needed for normal dopaminergic activity (Slc6a3 and Th). In DMSO control GPIN, the ubiquitin pathway genes were not connected into the network. The mitochondrial pathway genes were connected positively but no connection was detected affecting the expression of Th. These results mean that the differentiation of Th positive neuron was not affected by both pathways in normal neuronal differentiation. However, in PMT-exposed GPIN, all genes were connected into the network. Th expression was positively related by Park7, RARβ, Slc6a3 and Uchl1 and negatively related by Snca, Esr1, Crossing_point and NS_formfactor. These results suggest the differentiation of Th positive neuron was affected in a complex manner in PMT exposed EB derivatives. Interestingly, Park7, Casp3, Snca, Park2 and Casp9 were connected indicating the ubiquitin pathway and the mitochondrial pathway affected each other as well as they do in dopaminergic neurons of Parkinson's disease. Th positive neuron might die by apoptosis because we detected the increased expression of Casp3 and Casp9 in addition to these results. Although all morphological parameters were connected to GPIN in DMSO control, the NS morphological parameters (NS_area, NS_count and NS _perimeter) were not connected

to other morphological parameters or genes in the PMT-exposed GPIN. The neurite morphological parameters (Neurite_length, Branch_point, Crossing_point and Posi_area) influenced expression of genes in contrast to the NS morphological parameters. Considering the significant increase of total length of Map2-positive neuron (Figure 2A) and no change in the NS morphological parameters by PMT, the PMT-exposed GPIN successfully drew the change of neuronal morphology.

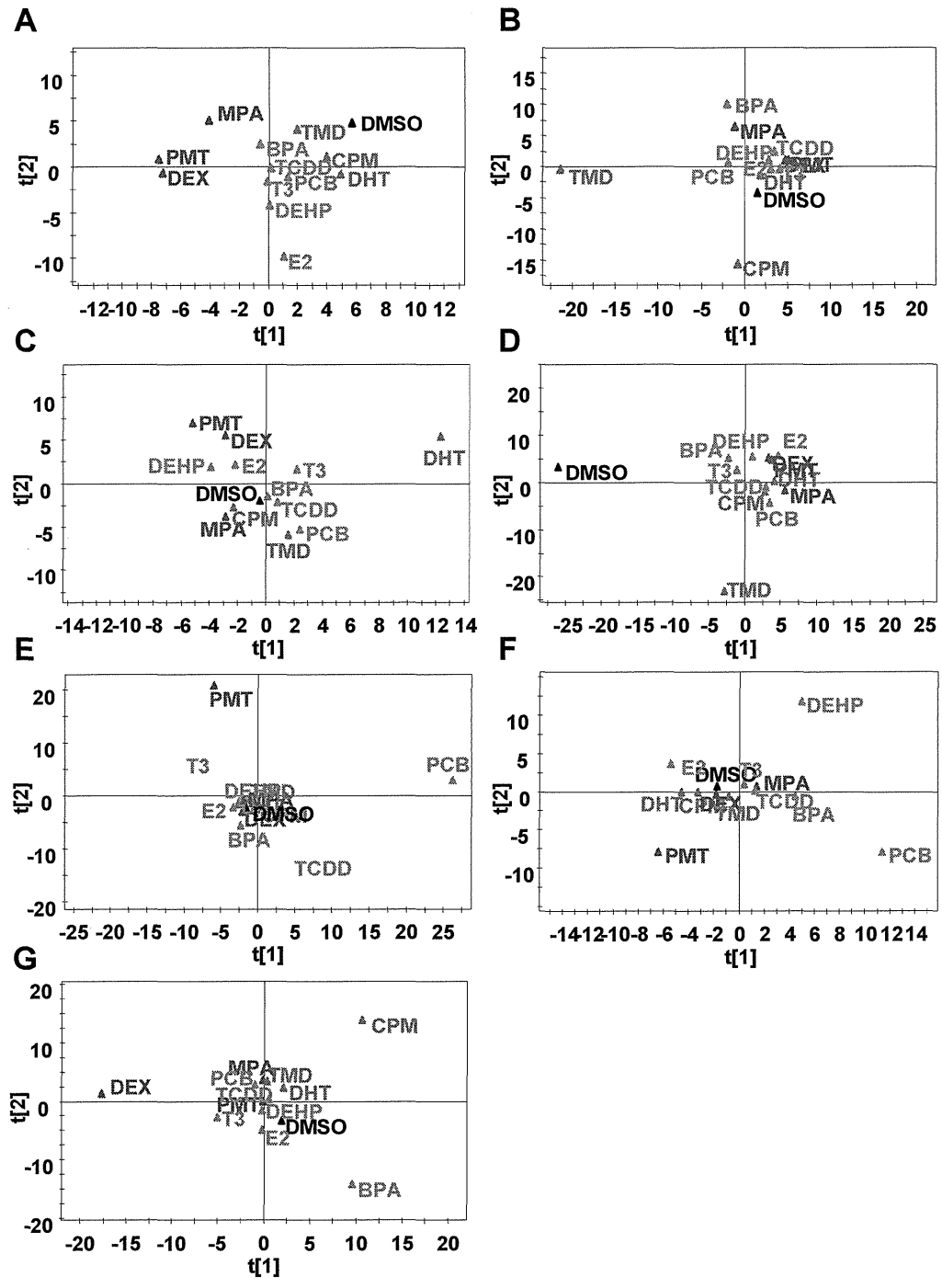
The comparison of TMD-exposed GPIN or PMT-exposed GPIN with DMSO control GPIN for Autism set and Parkinson's disease set could be understood without contradicting known pathological pathways. Therefore, we propose that our MPNs approach could draw out the risk of chemicals. The gene expression profiling data of our study have been published on the Profiles of Chemical Effects on Cells (pCEC) system [27], which is a toxicogenomics database with a toxicoinformatics system for risk evaluation and toxicity prediction of environmental chemicals [28] and produced by the National Institute of Environmental Studies, Japan. The microarray data have also been released on the GEO data base [29].

2.3. Classification of Chemicals Based on the Values of the Parameters of the Comprehensive Networks

The genomic data and cell morphological data were converted to the same matrix vector and were used to analyze GPIN. Principal component analysis (PCA) based on the probabilistic relationship data of the GPIN showed that all variance between the 12 chemicals could be described using the first and second principal components (PCs) (Figure 5). The two dimensional PCA plot showed four different groups: DMSO control (black), TMD group (CPM and DHT, green), BPA group [2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), PCB, T3, bis(2-ethylhexyl) phthalate (DEHP) and E2, blue] and MPA group (PMT and DEX, red) were derived for the Alzheimer's disease related gene set. The same color coding was used for other experiments, which enabled us to visually recognize changes to the grouping of chemicals. When the largest variable variation was placed in the vertical axis (PC1) and the second variation in the horizontal axis (PC2), the two-dimensional plot showed the position of each chemical. PMT and DEX were located near, but separated from, DMSO in Alzheimer set and Parkinson set. The toxic effects of DEX were reported in animal model of Alzheimer' disease [30] and Parkinson' disease [31] although we found no report about PMT in Alzheimer's disease. In Alzheimer set, E2 was located further away from DMSO than DHT and the opposite positioning was detected in Parkinson set. It might reflect the sexual differences of the diseases as the risk of Alzheimer's disease is higher in females [32] and that of Parkinson's disease is higher in males [33]. Because the responsible genes of gender specific Autism spectrum were involved in the Autism set, such gender dependent differences might not be detected in present data. In Autism set, TMD was more separate from DMSO than the others. Indeed other than TMD, the chemicals show no evidence of involvement in autism at present. In the Axon guidance set, all chemicals were almost equally distant from DMSO. As shown in (Figure 2A), all chemicals influenced the total length of Map2-positive neuron at high dose. Therefore, this result is reasonable. In the pluripotent set, PMT and PCB were separated from the others indicating that these chemicals affected the differentiation from ES cells. In fact, PMT and PCB are also located away from DMSO in neural development set. The characterization of chemicals according to their neurotoxic

potential reveals that the method described in this current study—that the MPN analysis based on phenotype and gene expression profiling during neuronal differentiation of mESCs—can provide a useful tool to monitor fetal programming and to predict developmentally neurotoxic compounds.

Figure 5. PCA based on Bayesian network parameters. PCA were applied to the Bayesian network parameters based on phenotypic and global gene expression profiling to evaluate the neurotoxicity of 12 environmental chemicals. Score plots based on (**A**) Alzheimer's disease related gene set; (**B**) Autism related gene set; (**C**) Parkinson's disease related gene set; (**D**) Axon guidance related gene set; (**E**) Pluripotent related gene set; (**F**) Neural development related gene set; and (**G**) Oxidative stress related gene set.



2.4. Discussion for Future Work

ESCs test combined with transcriptomics for the assessment of development toxicity has been well studied in recent years [8,34]. However, studies based on the genotype-phenotype profiling are rare. Cell phenotypes are complex and difficult to quantify in a high throughput fashion. The lack of comprehensive phenotype data can prevent or distort genotype-phenotype profiling. Our study described a unique approach to perform multiple phenotype profiling using gene expression data from the early stage of mESC differentiation and morphological data of neuronal cell differentiation after EB formation. Our method provided numerous advantages: (i) Our method can predict multiple phenotype profiles, which could help researchers to reveal different aspects of complex diseases and facilitate treatment design; (ii) Our method can provide a quantitative phenotype description of the sample characteristics; (iii) Our method can extrapolate the profiling to classes beyond those represented in the training data. This is an advantage over traditional classification methods. In contrast, traditional regression methods cannot be directly applied to microarray datasets from different platforms and cannot predict relationships between early events and late phenomena during the differentiation of ES cells into neuronal cells. However, our method can be applied to other types of genomics data such as proteomics or metabolomics. The present study focuses on linear gene-phenotype associations, but more complex relationships can also be devised depending on the data characteristics. Our multi-parametric profiling method for constructing interfering networks of the gene expression data and cellular phenotypic data is only one of many possible approaches. As mentioned above, our MPN analysis can predict the correlation coefficient for each pair of nodes, regardless of the data types. Therefore, our informatics approach and experimental design is also an efficient tool for data integration, mining and network analysis for the other model systems. However, another important issue for the future will be the validation of a larger set of chemicals at a broad concentration range to identify the specific and mechanistically defined markers for differential environmental chemicals.

ES cell-based assays are a promising platform to assess developmental toxicity, because they are capable of recapitulating many of the differentiation states and rely on signaling pathways present in development. We used a neuronal differentiation assay of mESC to assess the activity of groups of environmental chemicals, most of which have *in vivo* toxicity data. The results of this study demonstrated that a subset of tested chemicals are effective in this assay, and that as a statistical analysis, BNA, identified predictive models of detecting fetal programming in the mESC differentiation for a subset of the tested chemicals. Chandler *et al.* demonstrated evaluation of environmental chemicals using a mESC adherent cell differentiation and cytotoxicity assay, showing that genes involved in reactive oxygen species signaling pathways were strongly associated with decreased ES cell differentiation [35]. However, their approaches are linear regression or categorical approaches and are not identical with our approaches. Our approach is unique in linking early gene expression events to the later cellular phenotype features by BNA, which has become popular among biological scientists [36]. Many studies using BNA focus on basic physiological and developmental phenomenon based on cell proliferation [37]. In contrast, our study targets effects of early exposure on late-onset phenotypes, in accordance with the principles of fetal programming against environmental chemicals. In this regard,