

厚生労働科学研究費補助金(化学物質リスク研究事業)
(H30-化学-一般-004)令和元年度分担研究報告書

生体影響予測を基盤としたナノマテリアルの統合的健康影響評価方法の提案
分担研究課題名：(1)ナノマテリアルによる細胞内網羅的遺伝子発現データベースの構築、
(2)機械学習による生体影響予測の試み

研究分担者 花方信孝 物質・材料研究機構 技術開発・共用部門 副部門長

研究要旨：二年度目は細胞内網羅的遺伝子発現データベースを構築する上で元データとして使用する既存の生命科学系データベースを調査して決定した。生体影響予測の基盤となる機械学習における学習用遺伝子発現サンプル情報のラベリング方法および入力特徴量の調整と検討を行なった。また、マイクロアレイ解析の一色法と二色法のデータを双方利用するため相互のデータ変換法の開発を進めた。他の分担研究と比較するため標準ナノマテリアルとして二酸化チタン曝露時の毒性試験の実施および遺伝子発現マイクロアレイ解析の準備を行なった。

A.研究目的

現在一般社会で開発・使用されているナノマテリアルが社会的に受容されるためには、そのリスクについて十分な安全評価手法が必要である。しかしながら、その手法として代表的な動物実験は費用的にも時間的にも高コストであることに加えて、近年の動物愛護原則から敬遠され、代替手法が求められている。そこで *in vitro* 評価法の一環として、初年度はナノマテリアルによる細胞内網羅的遺伝子発現データベースの構築方法および機械学習による生体影響予測モデルを検討することを研究目的とした。

本年度は、細胞内網羅的遺伝子発現データベースの構築法および機械学習の予測モデルを実際に開発することを目的とした。具体的には、遺伝子発現データベースを構築する上で元データとして使用する既存の生命科学系データベースを決定し、機械学習における学習用サンプル情報のラベリング方法および入力特徴量の調整と検討、マイクロアレイ解析の一色法と二色法のデータ変換法の開発、標準ナノマテリアルとして酸化チタン曝露時の毒性試験の実施および遺伝子発現マイクロアレイ解析を目指した。

B.研究方法

B-1. ZnO 曝露細胞の一色法マイクロアレイ解析

- 細胞：THP-1 細胞および A549 細胞
- 曝露ナノマテリアル：酸化亜鉛 (ZnO)
- 曝露濃度：300 μ g/mL (THP-1 細胞) または

60 μ g/mL (A549 細胞)

- 曝露時間：6 時間または 24 時間
- マイクロアレイ：Agilent SurePrint G3 Human GE Microarray 8x60K Ver. 3.0 (G4851C; GEO platform:GPL21185) 1 枚
- ハイブリダイゼーション：一色法
- マイクロアレイの割り当て：表 1
- マイクロアレイスキャナー：Agilent Surescan G2600D
- 画像数値化処理ソフトウェア：Agilent Feature Extraction v11.5

B-2. マイクロアレイの一色法と二色法の比較

マイクロアレイ解析には一色法と二色法が存在するが、この 2 種類では遺伝子発現を測定する根本原理が異なり遺伝子発現データの取り扱い方に影響するた

表 1 ZnO 曝露実験におけるマイクロアレイ(一色法)

Slide ID	Slide No.	Pos.	B bck	Cy3
AH73	257236319389	1_1	B1	THP-1_Zn0=0 μ g/mL_6hr
AH73	257236319389	1_2	B2	THP-1_Zn0=300 μ g/mL_6hr
AH73	257236319389	1_3	B3	THP-1_Zn0=0 μ g/mL_24hr
AH73	257236319389	1_4	B4	THP-1_Zn0=300 μ g/mL_24hr
AH73	257236319389	2_1	B5	A549_Zn0=0 μ g/mL_6hr
AH73	257236319389	2_2	B6	A549_Zn0=60 μ g/mL_6hr
AH73	257236319389	2_3	B7	A549_Zn0=0 μ g/mL_24hr
AH73	257236319389	2_4	B8	A549_Zn0=60 μ g/mL_24hr

め、同一サンプルに対して実施した一色法と二色法のデータを比較し、そのデータ変換の妥当性を検証することとした。一色法のデータはプローブとハイブリダイズしたサンプル量に比例し、二色法のデータは 2 種

類のサンプルのプロープへの競合ハイブリダイゼーションの結果である。そのため、一色法で測定した2つのサンプルのシグナル強度を S_A , S_B とすると、二色法のデータへの変換式は S_A/S_B と書ける。一色法のデータから二色法にデータ変換したものと二色法で測定したデータを相関係数などで比較した。

B-3. 二酸化チタンの毒性評価

標準ナノマテリアルとして酸化チタンのナノ粒子を使用することとし、まずは毒性評価を行なうため WST-8 アッセイを実施した。評価を行なった酸化チタンは国立医薬品食品衛生研究所より分与された以下の7サンプル。

- MT-150A (Lot#651105)
- MT-500B (Lot#1880902)
- AMT-100 (Lot#181102)
- TKP-102 (Lot#4190101)
- AMT-600 (Lot#6553)
- TiO₂ (Lot#1001)
- TiO₂ (Lot#1005)

WST-8 アッセイには株式会社同仁化学研究所の Cell Counting Kit-8 (CCK-8) を使用した。細胞は THP-1 細胞と RAW264 細胞を用いて、それぞれ 500,000cells/mL, 70,000cells/mL に調製した。96 ウェルプレートにそれぞれ 50 μ L/well, 100 μ L/well で播種した。THP-1 細胞には酸化チタンを 50 μ L/well 添加し、37 $^{\circ}$ C で 24 時間後に CCK-8 を 10 μ L/well 添加し 4 時間後に波長 450nm で吸光度を測定した。RAW264 細胞は播種から 24 時間後に培地を除去してから酸化チタンを 100 μ L/well 添加し、さらに 24 時間後に CCK-8 を同様に添加して測定した。

なお、酸化チタンのナノ粒子の大きさは動的光散乱光度計 (DLS) により測定した。

C. 研究結果

C-1. ZnO 曝露細胞のマイクロアレイ解析

発現強度 (生データ) : 全 8 アレイの発現強度データの分布を表 2 と図 1 に示す。なお、図はボックスプロットで、縦線の下端が最小値、箱の下端が第 1 四分位値、箱の中線が中央値、箱の上端が第 3 四分位値、縦線の上端が最大値を表す。今回は一色法のデータであるため、アレイ間の発現強度は直接比較することができない。そのためノーマライズ処理を行なう。

発現強度 (ノーマライズ) : ノーマライズ方法は第 3 四分位値を基準とする 75 パーセンタイル法を使用

した。ノーマライズした結果を表 3 と図 2 に示す。発現強度の分布は概ね均等である。

発現強度 (プローブ) : 今回使用したマイクロアレイスライドはスポット数としては 1 アレイあたり 60,901 個あるが、1 つのプローブが複数のスポットにある場合があるので、プローブ単位で発現強度をまとめた。なお、プローブの総数は 58,201 個である。

階層的クラスタリング : 全 8 アレイのデータのうち 8 アレイとも発現比が求まったプローブ 22,003 個の発現強度データで階層的クラスタリングを行なった (図 3)。まず、THP-1 細胞と A549 細胞でクラスタが大きく分かっている。そしてそれぞれの細胞において ZnO に暴露したかしていないかでクラスタが分かれた。6 時間後と 24 時間後の違いは比較的小さいことが分かった。また、A549 細胞よりも THP-1 細胞の方が ZnO の影響が大きいことが明らかになった。

C-2. マイクロアレイの一色法と二色法の比較

今回の一色法で測定したマイクロアレイ解析のデータと前年度に二色法で測定したデータを比較すること

表 2 マイクロアレイ解析発現強度分

サンプル	AH7381	AH7382	AH7383	AH7384	AH7385	AH7386	AH7387	AH7388
スポット数	60901	60901	60901	60901	60901	60901	60901	60901
検出数	33242	43488	39214	42951	38535	35137	35718	38237
検出割合 (%)	54.6%	71.4%	64.4%	70.5%	63.3%	57.7%	58.7%	62.8%
最小値	4	3	3	3	4	4	4	4
第1四分位値	19	13	14	13	16	16	16	15
中央値	101	78	94	65	95	94	102	83
第3四分位値	683	650	617	508	813	753	696	698
最大値	261313	307037	267933	277060	291821	275938	291556	302015
算術平均	1749	2210	2513	1660	2595	2315	2377	2180
標準偏差	8175	11219	11144	8468	11569	10924	10282	10238

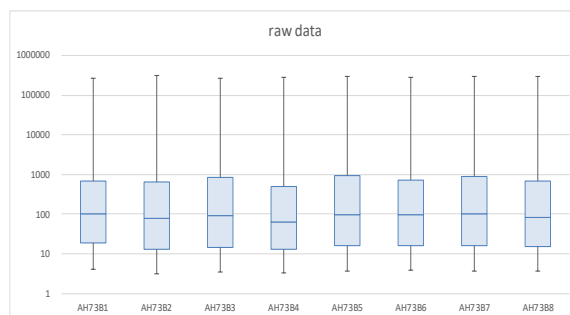


図 1 マイクロアレイ解析発現強度分布 (生データ)。

とした。使用したデータは、プローブ単位でまとめたデータのうち、RefSeq に該当する 25,685 プローブの

データ。二色法で測定した ZnO/ctrl の発現比 (2 を底とする対数) 4 種類 (THP-1 細胞と A549 細胞それぞれについて 6 時間後と 24 時間後) に対応する一色法の発現強度 S_A と S_B から $\ln(S_A/S_B)$ で発現比を求めた。そして、横軸に一色法から求めた発現比を縦軸に二色法による発現比をプロットしたのが図 4 である。

THP-1 細胞の 6 時間後、A549 細胞の 6 時間後と 24 時間後は一色法と二色法のデータが一致する傾向を示している。ピアソンの相関係数を求めるとそれぞれ 0.763, 0.898, 0.884 と高い値であった。しかしながら、THP-1 細胞の 24 時間後では一色法のデータと二色法のデータの相関性が低く、相関係数も 0.255 と低かった。この原因は不明であるが、THP-1 細胞の 24 時間後の ZnO 暴露またはコントロールのサンプルに問題があった可能性がある。

参考のため一色法と二色法のデータで異なるサンプルの発現比の相関係数も求めた (表 4)。THP-1 細胞と A549 細胞の間での相関係数は 0.45 を下回っており発現パターンの差が大きいことが分かる。A549 細胞の 6 時間ごと 24 時間後の間では 0.714 と比較的高く発現パターンが近いことが分かる。一方で THP-1 細胞の 6 時間ごと 24 時間後の間では 0.235 と細胞間の相関係数よりも低く原因は不明である。

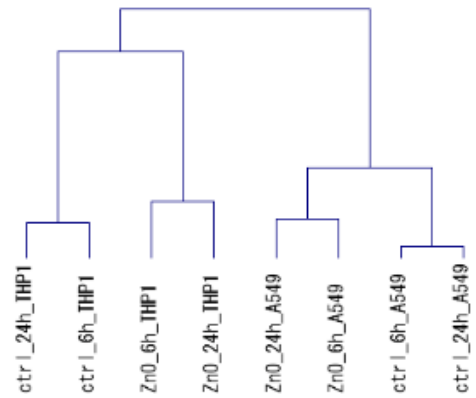


図 3 階層的クラスタリン

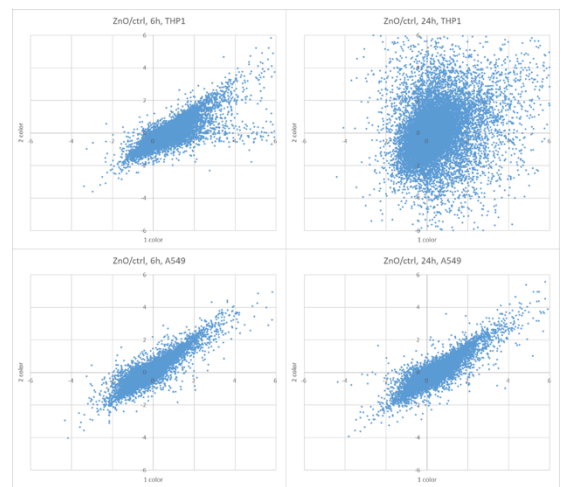


図 4 マイクロアレイ解析の一色法と二色法の比較

表 3 マイクロアレイ解析発現強度分

サンプル	AH7381	AH7382	AH7383	AH7384	AH7385	AH7386	AH7387	AH7388
スポット数	60901	60901	60901	60901	60901	60901	60901	60901
検出数	33242	43488	39214	42951	38535	35137	35778	38237
検出割合 (%)	54.6%	71.4%	64.4%	70.5%	63.3%	57.7%	58.7%	62.8%
最小値	6	5	4	6	4	5	4	5
第1四分位数	28	20	17	26	17	22	18	22
中央値	148	119	111	128	104	129	114	118
第3四分位数	1000	1000	1000	1000	1000	1000	1000	1000
最大値	382452	472686	316311	545421	319670	376275	325505	432948
算術平均	2560	3402	2966	3268	2843	3157	2654	3124
標準偏差	11964	17272	13153	16670	12673	14895	11479	14673

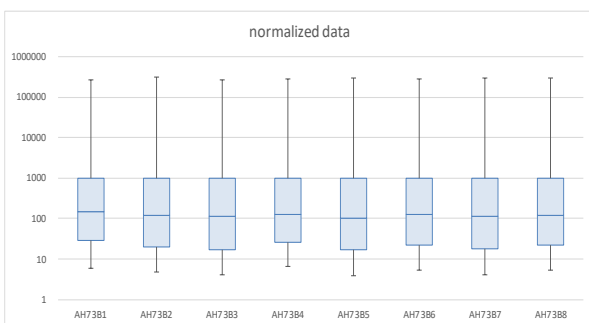


図 2 マイクロアレイ解析発現強度分

表 4 マイクロアレイ解析の一色法と二色法のデータのピアソンの相関係数。

		2color			
		ZnO/ctrl 6h THP1	ZnO/ctrl 24h THP1	ZnO/ctrl 6h A549	ZnO/ctrl 24h A549
1 color	ZnO/ctrl 6h THP1	0.763	0.235	0.344	0.328
	ZnO/ctrl 24h THP1		0.255	0.441	0.419
	ZnO/ctrl 6h A549			0.898	0.714
	ZnO/ctrl 24h A549				0.884

C-3.酸化チタンの毒性評価

標準ナノマテリアルとして酸化チタンのナノ粒子の使用することとした。これにより各分担研究間の比較も可能となる。まずは MMT-8 アッセイによる毒性評価を行なった (図 5、図 6)。浮遊細胞の THP-1 細胞に対しては 7 種類とも顕著な影響は認められなかった。一方で付着細胞の RAW264 細胞においては、MT-

150A, MT-500B, TiO₂-1005 の3種類では影響がみられなかったものの、AMT-100, AMT-600, TKP-102, TiO₂-1001 の4種類では毒性が認められた。THP-1細胞とRAW細胞で実験の日時に差があり、ナノ粒子径はその都度DLSにより測定したが多少のばらつきがある。しかしながら、RAW細胞に毒性を示した酸化チタンとナノ粒子径との間に相関は認められない。

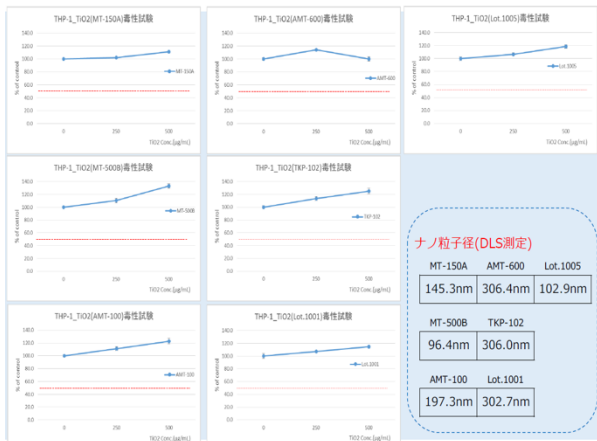


図5 THP-1細胞に対する酸化チタンの毒性評価。

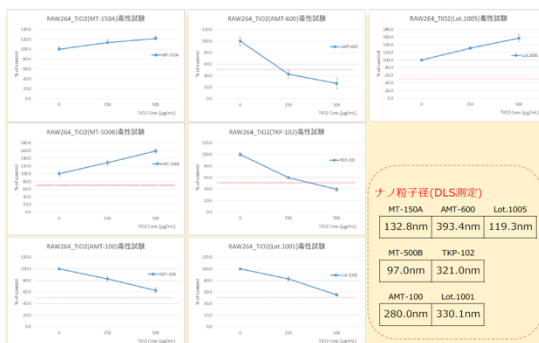


図6 RAW264細胞に対する酸化チタンの毒性評価。

D.考察

網羅的遺伝子発現データとしてマイクロアレイ解析のデータを利用する際には原理的に一色法と二色法の2種類のデータがあり、それらを双方ともに扱うにはデータの変換が必要となる。原理的には一色法はマイクロアレイ上のプローブ量が一定であることを前提としており、製造上の誤差が小さい必要がある。一方で二色法は製造上の誤差があっても問題ない方法であり、2種類のサンプルを異なる蛍光色素で標識し競合ハイブリダイゼーションを行なうことで発現比を測定する。前年度に実施した二色法のデータに対して本年度は同じサンプルについて一色法によるマイクロアレイ解析を行い、二色法と比較した。一色法のデータか

ら二色法のデータに変換した結果、4つのケースのうち3つでは良好な対応関係が見られた。しかしながら、残り1つのケースでは適切な対応関係が得られなかった。原因は実験サンプルの問題と思われるが、データ変換について良好な対応が見られる場合でもピアソンの相関係数は高々0.76~0.90となっており、一色法から二色法にデータ変換した場合にはこれくらいの誤差を含むことに留意する必要がある。一方で二色法から一色法への変換についても考えると、二色法における2つのサンプルのうち1方の発現強度について既知であれば、変換は可能と思われる。

生命科学系データベースとしてマイクロアレイ解析の生データが登録されているGEOデータベースを選択したが、サンプル条件などのラベル付けが課題である。今のところ手動でラベルを付けているが、データのメタ情報からうまくラベルを生成する方法を検討する必要がある。また、前年度に行なったGene Ontology エンリッチメント解析のような結果をラベルとして利用することも検討しているが、有意な結果がないケースがあるので、利用は限定的となる。

機械学習を実行する上で、ヒトの遺伝子の数はRefSeqデータベースに限っても26,000個ほどもあるため、入力特徴量としては扱いにくい。そのため顕著な遺伝子に限定して1/10程度の量に減らしてモデルを作成し機械学習を実行させている。この部分の効率的なデータの選択が本研究において重要である。

実測データとしてZnOを曝露した細胞の系を扱っていたが、他の分担研究と協調するためにも共通となる標準ナノマテリアルを使うことが適切として酸化チタンのナノ粒子を用いることとした。今回は7種類に酸化チタンについて毒性評価を行なったが、THP-1細胞とRAW264細胞で異なる結果を示すナノ粒子もあった。これらについてマイクロアレイ解析も実測し、機械学習のモデルを検証するデータとして利用したい。

以上を踏まえて、次年度は機械学習における学習用サンプル情報のラベリング方法および入力特徴量の調整・検討の結果とマイクロアレイ解析の一色法と二色法のデータ変換法の開発を元に、中途であった最小規模のデータベースの構築を完成させ、拡大・拡張する。機械学習の実行にあたってはアルゴリズムの最適化が解決すべき課題となる。標準ナノマテリアルとして酸化チタン曝露時の実測した遺伝子発現マイクロアレイ解析結果を用いて生体影響予測の精度を検討す

る。最終的には完成したデータベースと生体影響予測機能を提示する。

E.結論

本年度は、細胞内網羅的遺伝子発現データベースの構築法および機械学習の予測モデルを開発することを目的として、網羅的遺伝子発現データの基礎となるマイクロアレイ解析の一色法と二色法について検討し、一定の精度で一色法から二色法へのデータ変換を行なうことができた。機械学習における学習用サンプル情報のラベリング方法および入力特徴量の削減方法について検討した。分担研究間で共通して使用する標準ナノマテリアルとして酸化チタンのナノ粒子を採用し曝露時の毒性試験を実施した。

F.研究発表

本年度はなし

G.知的所有権の出願・登録状況

本年度はなし