

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

化学物質の有害性評価の迅速化・高度化・標準化に関する研究（H29-化学-一般-001）
網羅的な DNA 付加体解析法を用いた化学物質の DNA 損傷性評価

研究分担者 戸塚 ゆ加里 国立がん研究センター研究所発がん・予防研究分野 ユニット長

研究要旨

我々は新規のヒト発がんリスク評価法として、DNA 付加体の網羅的解析手法(DNA アダクトーム法)の構築に取り組んできた。本年度は、肝発がん性検出評価系の再現性確認と施設間バリデーション試験の実施に用いたラット肝臓サンプルを用い、複数の遺伝毒性/非遺伝毒性肝発がん物質の肝臓における DNA 損傷性の評価を、アダクトーム法により検討した。遺伝毒性ラット肝発がん物質として、2-Nitropropane (2-NP；陽性対照) *o*-Aminoazotoluene (AAT), Dimethylnitrosamine (DMN), 4,4'-Thiodianiline (TDA)、N-Nitrosodiethylamine (NDEA), N-Nitrosodiethanolamine (NDELA), N-Nitrosoethylmethylamine (NEMA), Nitrosodibutylamine (NB), N-Nitrosopyrrolidine (NNP), 3'-Methyl-4-dimethylaminoazobenzene (MDA)、遺伝毒性非発がん物質；Cyclophosphamide (CPA), Nitrofurantoin (NFT), Phenacetin (PCT)、非遺伝毒性肝発がん物質；Monocrotaline (MCT), Phenobarbital (PB)、非遺伝毒性非肝発がん物質；Diazepam (DZP), Disulfiram (DSF), Phenytoin (PHE), Rotenone (ROT), Tolbutamide (TLB), Aspirin (ASA), Triamterene (TRI), Indomethacin (IM), Phenylbutazone (PhB), Promethazine (PMZ), Sulindac (SUL), Tetracycline (TC), Ethionamide (ETH), Theophylline (TEO), Caffeine (CAF), Chloramphenicol (CMP)をそれぞれラットに投与し、24 時間後に肝臓に生成される DNA 付加体を網羅的に解析した。得られたデータを線形判別分析(LDA)により分類したところ、非遺伝毒性非肝発がん物質、遺伝毒性非発がん物質、非遺伝毒性肝発がん物質、遺伝毒性肝発がん物質の 4 つのグループに分離されることがわかった。これらデータの一部を教師データとして用い、機械学習による化学物質の遺伝毒性/発がん性の予測モデルの構築を試みた。試作した遺伝毒性/発がん性予測モデルを用いて、テストデータに対して予測を行ったところ、93%以上の正解率が得られた。また、Leave-One-Out 交差検証を適用した毒性予測モデルの予測精度評価を実施したところ、わずかに正答率が減少した。現在、アダクトームデータに化学物質構造データを追加し、予測精度の向上について検討中である。また、各グループの予測に重要な付加体の探索を行うと同時に、更に化学物質の数を増やし、予測精度の向上について検討を行う予定である。

A．研究目的

既存の *in vitro* 遺伝毒性試験としては、Ames 試験（変異原性試験）、コメットアッセイ（DNA 損傷試験）、小核試験（染色体異常試験）などが簡便な試験法として汎用されている。しかしながら、これらの *in vitro* 試験のみでは化学物質の発がん性の予測は難しく、別の視点から遺伝毒性を評価する試験法を更に追加することが必要であると考えられる。これまで我々は、LC-MS/MS により DNA 付加体を網羅的に解析する方法（アダクトーム法）を用い、DNA 損傷のより詳細な評価を行ない、化学物質の *in vitro* 安全性評価法として妥当かどうかについて確かめてきた。これまでに、遺伝毒性を示すマグネタイトナノ粒子を気管内投与したマウス肺の解析を行ない、マグネタイトナノ粒子が誘発する G:C->A:T 及び G:C->T:A 変異の基となる付加体（エテノデオキシシチジン、-dC）を含む複数の付加体を確認することを報告した。この結果は、アダクトーム法による化学物質の DNA 損傷性評価が有用であることを示唆するものである。今年度は、ラットを用いた *in vivo* モデルを

用い、肝臓をターゲットとした複数の遺伝毒性/非遺伝

毒性非発がん物質の肝臓における DNA 損傷を LC-MS を用いたアダクトーム解析により検討し、DNA 付加体の生成を指標とした有害性評価の検証を行なう。

B．研究方法

雄性 SD ラット（各群それぞれ 5 匹）に遺伝毒性肝発がん物質；2-Nitropropane(2-NP；陽性対照), *o*-Aminoazotoluene(AAT), Dimethylnitrosamine(DMN), 4,4'-Thiodianiline(TDA), N-Nitrosodiethylamine (NDEA), N-Nitrosodiethanolamine(NDELA), N-Nitrosoethylmethylamine(NEMA), Nitrosodibutylamine(NB), N-Nitrosopyrrolidine(NNP), 3'-Methyl-4-dimethylaminoazobenzene(MDA)、遺伝毒性非発がん物質；Cyclophosphamide(CPA), Nitrofurantoin(NFT), Phenacetin(PCT)、非遺伝毒性肝発がん物質；Monocrotaline(MCT), Phenobarbital(PB)、非遺伝毒性非肝発がん物質；Diazepam(DZP), Disulfiram(DSF), Phenytoin(PHE), Rotenone(ROT), Tolbutamide(TLB),

Aspirin(ASA), Triamterene(TRI), Indomethacin(IM), Phenylbutazone(PhB), Promethazine(PMZ), Sulindac(SUL), Tetracycline(TC), Ethionamide(ETH), Theophylline(TEO), Caffeine(CAF), Chloramphenicol(CMP)を各種濃度(表1参照)で投与を行った後24時間後に肝臓を摘出した。DNAを抽出後、DNaseI、ヌクレアーゼP1、アルカリホスファターゼ、ホスホジエステラーゼによりモノデオキシリボヌクレオチドに消化した後、LC-TOFMSに供しDNA付加体の網羅解析を行った。得られたデータはSCIEX社が提供するバイオインフォマティクス解析ソフトウェアを用い、デオキシリボヌクレオチドに特徴的なニュートラルロス(-116.04736)及び各種核酸に特異的なニュートラルロス(-152.0572; dG, -136.0623; dA, -112.0511; dC, -127.0508; dT)を生じたピークを選択的に抽出することで、ノイズなどを抽出しないように系をデザインした。得られたデータを線形判別分析(LDA)により解析した。

次に、得られたデータから、対照群である2-NPおよびMCを除いた、146サンプルのデータを用いた。各曝露群について、全データの3/4を訓練データに、残りの1/4をテストデータに、それぞれ分配し、機械学習の訓練データおよびテストデータの作成を行った。遺伝毒性、肝発がん性、遺伝毒性/肝発がん性を付加体から予測するモデルを、教師あり機械学習手法を用いて試作した。学習アルゴリズムとしては、ランダムフォレストを使用した。さらに、実際に予測モデルを使用する場合には遺伝毒性/肝発がん性の情報がない化学物質に対して実施することを想定して、Leave-One-Out交差検証により予測結果の精度評価を行った。Leave-One-Out交差検証は、ある曝露物質のデータをテスト用とし、他の曝露物質のデータを用いて学習したモデルの精度を評価することを、物質を変えながら繰り返す検証方法であり、毒性が未知の物質に対する予測を模擬した精度評価方法である。

(倫理面への配慮)

本研究で行う動物実験にあたっては、国立がん研究センターを含む各施設における動物実験に関する指針に則って実施し、可能な限り実験動物の苦痛軽減処置を行う。

C. 研究結果

各種化学物質を投与したラット肝臓DNAのアダクトーム解析を行なった結果を図1に示す。LDA解析を行なったところ、非遺伝毒性非肝発がん物質、遺伝毒性非発がん物質、非遺伝毒性肝発がん物質、遺伝毒性肝発がん物質の4つのグループに綺麗に分離されることがわかった。

1) 遺伝毒性の予測結果

genotoxicity予測モデルを用いて、テストデータの29サンプルに対して予測を行った結果、27サンプルに対して正答となった(93%、表2)。

2) 肝発がん性の予測結果

毒性予測モデルを用いて、テストデータの29サンプルに対して予測を行った結果、28サンプルに対して

正答となった(97%、表3)。

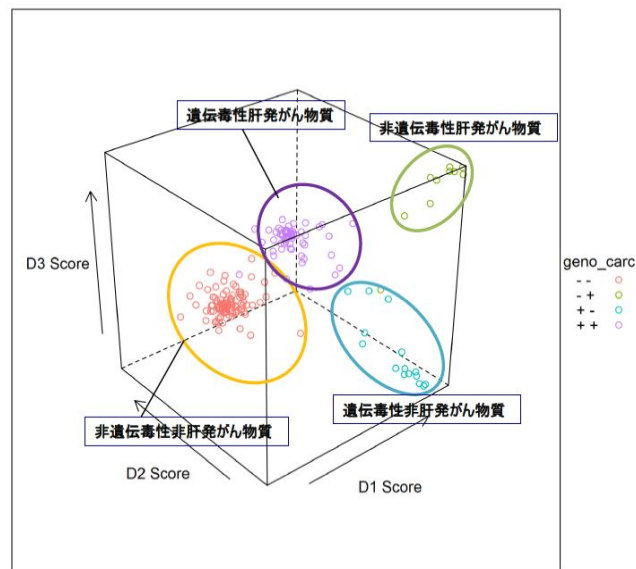
3) 遺伝毒性/肝発がん性の予測結果

毒性予測モデルを用いて、テストデータの29サンプルに対して予測を行った結果、27サンプルに対して正答となった(93%、表4)。

表1 使用した化学物質

	遺伝毒性	発がん性	投与量 (mg/kg bw)
o-Aminobenzoic acid (AA)	+	+	500
Dimethylacetamide (DMA)	+	+	10
4,4'-Thiodianiline (TDA)	+	+	300
2-Nitropropane (2-NP, 降性母質)	+	+	240
N-Nitrosodimethylamine (NDEA)	+	+	70
N-Nitrosodimethylamine (NDEA)	+	+	2000
N-Nitrosodimethylamine (NDEA)	+	+	30
Nitrosodimethylamine (NDMA)	+	+	400
N-Nitrosopyrrolidine (NDP)	+	+	300
β-Methyl-4-dimethylaminoacetamide (M4DA)	+	+	500
Cyclophosphamide (CPA)	+	-	30
Nitrosamine (NT)	+	-	200
Phenacetin (PC)	+	-	550
Mecloretin (MCT)	-	-	20
Phenobarbital (PB)	-	+	50
Diazepam (DZP)	-	-	80
Diazepam (DZP)	-	-	170
Phenacetin (PC)	-	-	550
Racemone (RAC)	-	-	20
Telluramide (TLB)	-	-	830
Aspirin (ASA)	-	-	70
Tetrazolone (TRZ)	-	-	130
Indomethacin (IM)	-	-	1
Phenylbutazone (PhB)	-	-	80
Promethazine (PMZ)	-	-	180
Sulindac (SUL)	-	-	90
Tetracycline (TC)	-	-	270
Ethionamide (ETH)	-	-	440
Theophylline (TEO)	-	-	80
Caffeine (CAF)	-	-	60
Chloramphenicol (CMP)	-	-	830
β-Methyl-4-dimethylaminoacetamide (M4DA, 母質)	-	-	5mg/kg

図1 遺伝毒性肝発がん物質/遺伝毒性非肝発がん物質/非遺伝毒性肝発がん物質/非遺伝毒性非肝発がん物質の肝臓におけるDNA損傷性の評価(LDA解析による)



4) Leave-One-Out 交差検証による予測結果の精度評価

Leave-One-Out 交差検証を適用した毒性予測モデルの予測精度評価を実施した。なお、訓練データとテストデータは対照群を含まない。結果を表5に示すが、Leave-One-Out 交差検証を用いるとわずかに正答率が減少した。具体的には、genotoxicity/carcinogenicityの予測結果は3%、genotoxicityの予測結果は7%、carcinogenicity

の予測結果は7%減少した。

表2 遺伝毒性の予測結果

exposure	sample_no	truth	pred	correct
AAT	1	+	+	1
DMN	5	+	+	1
TDA	2	+	+	1
NDEA	2	+	+	1
NDELA	5	+	+	1
NEMA	5	+	+	1
NB	3	+	+	1
NNP	2	+	+	1
CPA	2	+	+	1
NFT	3	+	+	1
PCT	4	+	+	1
MCT	2	-	+	0
PB	2	-	+	0
DZP	4	-	-	1
DSF	3	-	-	1
PHE	1	-	-	1
ROT	5	-	-	1
TLB	5	-	-	1
ASA	2	-	-	1
TRI	5	-	-	1
IM	3	-	-	1
PhB	5	-	-	1
PMZ	3	-	-	1
SUL	5	-	-	1
TC	5	-	-	1
ETH	2	-	-	1
TEO	4	-	-	1
CAF	2	-	-	1
CMP	4	-	-	1

(灰色：誤答)

表3 肝発がん性の予測結果

exposure	sample_no	truth	pred	correct
AAT	1	+	+	1
DMN	5	+	+	1
TDA	2	+	+	1
NDEA	2	+	+	1
NDELA	5	+	+	1
NEMA	5	+	+	1
NB	3	+	+	1
NNP	2	+	+	1
CPA	2	-	-	1
NFT	3	-	-	1
PCT	4	-	-	1
MCT	2	+	-	0
PB	2	+	+	1
DZP	4	-	-	1
DSF	3	-	-	1
PHE	1	-	-	1
ROT	5	-	-	1
TLB	5	-	-	1
ASA	2	-	-	1
TRI	5	-	-	1
IM	3	-	-	1
PhB	5	-	-	1
PMZ	3	-	-	1
SUL	5	-	-	1
TC	5	-	-	1
ETH	2	-	-	1
TEO	4	-	-	1
CAF	2	-	-	1

CMP 4 - - 1

(灰色：誤答)

表4 遺伝毒性/肝発がん性の予測結果

exposure	sample_no	truth	pred	correct
AAT	1	++	++	1
DMN	5	++	++	1
TDA	2	++	++	1
NDEA	2	++	++	1
NDELA	5	++	++	1
NEMA	5	++	++	1
NB	3	++	++	1
NNP	2	++	++	1
CPA	2	+-	+-	1
NFT	3	+-	+-	1
PCT	4	+-	+-	1
MCT	2	-+	-+	0
PB	2	-+	-+	0
DZP	4	--	--	1
DSF	3	--	--	1
PHE	1	--	--	1
ROT	5	--	--	1
TLB	5	--	--	1
ASA	2	--	--	1
TRI	5	--	--	1
IM	3	--	--	1
PhB	5	--	--	1
PMZ	3	--	--	1
SUL	5	--	--	1
TC	5	--	--	1
ETH	2	--	--	1
TEO	4	--	--	1
CAF	2	--	--	1
CMP	4	--	--	1

(灰色：誤答)

表5 DNA アダクトーム解析の毒性予測結果の精度評価の比較

	訓練データとテストデータに当該曝露物質を含む交差検証			Leave-One-Out 交差検証		
	正答数	合計サンプル数	正答率	正答数	合計サンプル数	平均的な正答率
geno_carc	27	29	93%	131	146	90%
geno	27	29	93%	125	146	86%
carc	28	29	97%	131	146	90%

D. 考察

29種の遺伝毒性/非遺伝毒性非肝発がん物質を投与したラットの肝臓からDNAを抽出し、アダクトーム法を用いてDNA付加体の網羅解析を行なった。得られたデータを用いてLDA解析を行った。その結果、遺伝毒性肝発がん物質、非遺伝毒性肝発がん物質、遺伝毒性非肝発がん物質、非遺伝毒性非肝発がん物質の4つのグループに分離できた。アダクトームデータの一部を教師データとして化学物質の遺伝毒性/発がん性を予測するモデルを機械学習手法(アルゴリズムとしてはランダムフォレストを使用)を用いて試作した。試作された遺伝毒性/発がん性予測モデルを用いてテストデータに対して予測を行ったところ、93%以上の正解率が得られた。また、Leave-One-Out交差検証を適用した毒性予測モデルの予測精度評価を実施したところ、わずかに正答率が減少した。予測精度をさらに向上させるためには、説明変数の加工、学習アルゴリズムの変更、チューニングなどの学習方法を改良する必要がある。さらに、訓練データの増加も必要であると思われる。特に、今回、いずれの予測モデル

においても予測に失敗した化学物質物質(MCT, PB)は、非遺伝毒性発がん物質である。この結果は、データが他の化学物質に比べ極端に少ないことが予測度を低くしている可能性を示唆している。今回、この2種しか非遺伝毒性発がん物質の暴露は行なっておらず、同カテゴリーのサンプルを追加することが、予測精度を上げるためにも重要であると考えられる。

E. 結論

遺伝毒性ラット肝発がん物質として、2-NP(陽性対照)、AAT, DMN, TDA, NDEA, NDELA, NEMA, NB, NNP, MDA、遺伝毒性非発がん物質として、CPA, NFT, PCT、非遺伝毒性肝発がん物質として、MCT, PB、非遺伝毒性非肝発がん物質としてDZP, DSF, PHE, ROT, TLB, ASA, TRI, IM, PhB, PZ, S UL, TC, ETH, TEO, CAF, CMPをそれぞれラットに投与し、24時間後に肝臓に生成されるDNA付加体を網羅的に解析した。得られたデータを用いてLDA解析を行ったところ、非遺伝毒性非肝発がん物質、遺伝毒性非発がん物質、非遺伝毒性肝発がん物質、遺伝毒性肝発がん物質の4つのグループに分離されることがわかった。次に、得られたデータの一部を使用して、遺伝毒性、肝発がん性、遺伝毒性/肝発がん性を付加体から予測するモデルを、教師あり機械学習手法を用いて試作した。学習アルゴリズムとしては、ランダムフォレストを使用した。アダクトームデータの一部を教師データとして化学物質の遺伝毒性/発がん性を予測するモデルを機械学習手法(アルゴリズムとしてはランダムフォレストを使用)を用いて試作した。試作された遺伝毒性/発がん性予測モデルを用いてテストデータに対して予測を行ったところ、93%以上の正解率が得られた。また、Leave-One-Out交差検証を適用した毒性予測モデルの予測精度評価を実施したところ、わずかに正答率が減少した。今回、いずれの予測モデルにおいても予測に失敗した化学物質物質(MCT, PB)は、非遺伝毒性発がん物質である。同カテゴリーのサンプルを追加することが、必要であると考えられた。

現在、アダクトームデータに化学物質構造データを追加し、予測精度の向上について検討中である。また、各グループの予測に重要な付加体の探索を行うと同時に、更に化学物質の数を増やし、予測精度の向上について検討を行う予定である。

F. 研究発表

1. 論文発表

- 1) Mimaki S, Watanabe M, Kinoshita M, Yamashita R, Haeno H, Takemura S, Tanaka S, Marubashi S, Totsuka Y, Shibata T, Nakagama H, Ochiai A, Nakamori S, Kubo S, Tsuchihara K. Multifocal origin of occupational cholangiocarcinoma revealed by comparison of multilesion mutational profiles. *Carcinogenesis*. 2019 in press.
- 2) Gi M, Fujioka M, Totsuka Y, Matsumoto M, Masumura K, Kakehashi A, Yamaguchi T, Fukushima S, Wanibuchi H. Quantitative analysis of mutagenicity and carcinogenicity of 2-amino-3-methylimidazo[4,5-*f*]quinoline

in F344 *gpt* delta transgenic rats. *Mutagenesis*. 2019; 34(3): 279-287.

- 3) Totsuka Y, Lin Y, He Y, Ishino K, Sato H, Kato M, Nagai M, Elzawahry A, Totoki Y, Nakamura H, Hosoda F, Shibata T, Matsuda T, Matsushima Y, Song G, Meng F, Li D, Liu J, Qiao Y, Wei W, Inoue M, Kikuchi S, Nakagama H, Shan B. DNA Adductome Analysis Identifies N-Nitrosopiperidine Involved in the Etiology of Esophageal Cancer in Cixian, China. *Chem Res Toxicol*. 2019; 32(8): 1515-1527.
 - 4) Dertinger SD, Totsuka Y, Bielas JH, Doherty AT, Kleinjans J, Honma M, Marchetti F, Schuler MJ, Thybaud V, White P, Yauk CL. High Information Content Assays for Genetic Toxicology Testing: A Report of the International Workshops on Genotoxicity Testing (IWGT). *Mutation Res*. 2019; 847: 403022.
 - 5) Totsuka Y, Wakabayashi K. Biological significance of aminophenyl-*-*carboline derivatives formed from co-mutagenic action of *-*carbolines and aromatic amines and its effect on tumorigenesis in humans: A review. *Mutation Res*. 2020; 850-851: 503148.
- ### 2. 学会発表
- 1) Totsuka Y, Exploration of Esophageal Cancer Etiology using DNA Adductome Analysis, 6th ACEM-48th JEMS、東京(2019年11月)
 - 2) Iwamura K, Shimada H, Matsuda T, Kato M, Elzawahry A, Nagai M, Endo O, Totsuka Y. Whole genome sequencing analysis elucidates the association between environmental factors and human cancer development. 6th ACEM-48th JEMS、東京(2019年11月)
 - 3) Ono H, Nagai M, Narushima D, Hamamoto R, Totsuka Y, Kato M. Detection of DNA adducts by nanopore sequencing using deep learning. 6th ACEM-48th JEMS、東京(2019年11月)
 - 4) Totsuka Y, Whole genome sequencing analysis elucidates the association between environmental factors and human cancer development、第78回日本癌学会学術総会、京都(2019年9月)
 - 5) Totsuka Y, Exploration of Esophageal Cancer Etiology using Comprehensive DNA Adduct Analysis (DNA Adductome Analysis) 2nd Hebei International Forum on Theory and Practice of Cancer Prevention and Control、石家庄(2019年7月)
 - 6) Totsuka Y, How Adductomics Can Inform Cancer Etiology, Mutgraph meeting、リヨン(2019年7月)
 - 7) 戸塚ゆ加里、ナノマテリアルの遺伝毒性評価の動向 JRC 会議に参加してー MMS 定例会、京都

(2019年6月)

- 8) 戸塚ゆ加里、発がん性評価法としての DNA アダクトーム解析の展望、第46回日本毒性学会学術年会、徳島(2019年6月)

G. 知的財産権の出願・登録状況

1. 特許取得

該当なし。

2. 実用新案登録

該当なし。

3. その他

該当なし。