

5. 参照配列の変更によるマッピングされたリードの変化



ペアエンド処理によるどのように変化したかを確認した。

対象	Liver100%サンプル	ペアエンド処理を実施	
対象配列数(ロード前)		11,795,262	
		mm10	mm9
対象配列数(ペアエンド考慮ロード後)		10,238,689 86.80%	
マイクロサテライト配列除外数		5	
マッピングできた配列		7,329,517 62.14%	6,440,946 54.61%
マッピングできたペア(累計)		22,991,225	19,760,344

Copyright(C)2014-2015 NTT DATA Corporation

6.1. 高精度化のためのマッピングエラー配列の特定① 遺伝子以外の大量マッピング



特定方法

染色体上にくっつけたものの極大値位置をピックアップ

塩基単位マッピング量
(T_match_203)

塩基単位マッピング量移動平均値(T_match_204)

この二つの値の差が大きな該当塩基アドレスを求める。

極大値の上位1000か所に対して、遺伝子と結合し、遺伝子外のアドレスを拾い出す。

エクソン情報(t_refg)と、範囲結合

結合できなかった個所をピックアップ

Copyright(C)2014-2015 NTT DATA Corporation

6.2.高精度化のためのマッピングエラー配列の特定① 遺伝子以外の大量マッピング

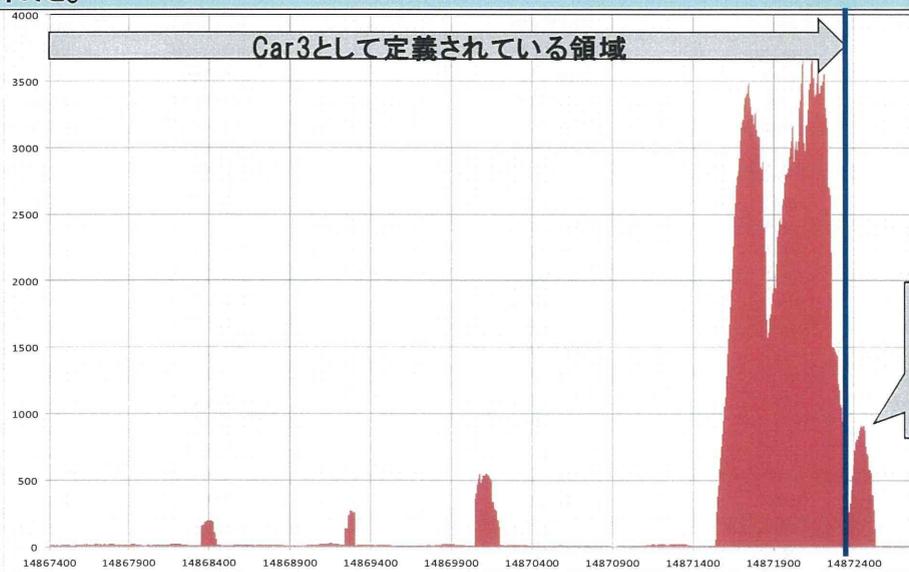
染色体上で、遺伝子以外の個所へのマッピング結果の上位を示します
(積上げ量が整数の上位30位まで)

No.	染色体番号	アドレス	配列	積上げ量	BlastN結果
1	1	24619050	GTTTAAAGCTCTAGGGAGGGAGCTGCATGAGTGGAGGACGCTTCAGATGAATATATACGG	6774	Mus musculus mitochondrial genome.
2	12	7076300	GCACATGAGCGGGTGGAGGAGGGACACCCGCTGAGCCAGCAGATACCGAATCAACCT	2443	Mus musculus 75K class III scRNA gene, complete sequence
3	3	14872400	AAAACTGGTGTCAAGTGTCTTTTGGTTGGAGGCTGGATGCTGGTTCTGGAAAGGAAGGAGG	915	Mus musculus chromosome 3, clone RP24-493E2, complete sequence
4	10	98077600	TAACTAACATGGTATCCGAGTTGTATACGGTATGCTGGAGAATTTGGATTTCTGTTACTCA	859	Mus musculus castaneus isolate G2 mitochondrion, complete genome
5	12	122152900	ACTTCTACTTGGATGAGTACTATCTTTGGTCTACCAACATTAAGTCATATCTTCTCT	561	mouse beta-2-microglobulin gene(seg.4). Length = 515
6	19	4063450	CAGCCACAGGCTAAATTTTACAAGTATGTCACAGGCTTTTCCAGGGTCCCAATCTCTGAGT	528	Mus musculus chromosome 19, clone RP23-41B18, complete sequence.
7	16	46154750	CATGATCTGAGTAGAAGATTTGCTGGACACCGGACCTGACAGCCAGGGCAGCGAG	511	Mus musculus (clone C7/B9) 5'-adenosyl, homocysteine hydrolase (ahcy) mRNA, complete cds.
8	19	4064300	CTGTACTCTGGTCAGATTTGGGACCTTCTAGTCCACAGAGGGAGACCCCATCAGAGTC	453	Mus musculus chromosome 19, clone RP23-41B18, complete sequence.
9	12	91630650	GCTCTGGCTGGTAAGCTCTCCAGGGCAGGCAGGGGACCAAGTGTCTCACAGCAATTTCC	444	Mouse mRNA for prothrombin
10	19	4065050	CCAGCGACAGCAACAGTCCAGAAAAGGGAGTGGACAGATTTCTCATTTTTAGAAAACA	379	Mus musculus chromosome 19, clone RP23-41B18, complete sequence.
11	4	61959900	AGGAGCTCAATTTGGTAGATGATAAATTTTCTGGCTGCAAGACACGCTCTTCTGGTAGGT	333	Mus musculus phosphoenolpyruvate carboxylase 1, cytosolic, mRNA (cDNA clone MGC46881 IMAGE4973187)
12	12	173158700	GATCTAGGAGAAAGCAGCTGCTCCAGCTTGGATGAGCGGCAATGCTGAGTAGATCAGAAAA	314	Mus musculus phosphoenolpyruvate carboxylase 1, cytosolic, mRNA (cDNA clone MGC46881 IMAGE4973187)
13	15	6498550	AGGAGCTGTTTATGAAATCTCAACCTCTGATGCTGCTGTTTTAATCTACAGTCTCTTC	281	Mus musculus chromosome 15, clone RP23-319G15.
14	11	41499250	CTACAATTTCTGCAATGGAGACAAAGCATTTGGGAATCTATTCTGATGCAATGAAAAAGTTG	271	Mouse DNA sequence from clone IMG-437H1 on chromosome 4.
15	18	20984550	GAATTAATTTCTGATGACAGAGATCAGCACTCTCTAAGAGTATGCTGCAAGAGGATC	270	Mouse DNA sequence from clone RP23-147K3 on chromosome X.
16	12	103478800	TGCCACTCTGTCAGTGAAGAACACATCTGAAGCAGGGAGGGGCTCGGGACCCGGGCTCATGAT	255	Mouse mRNA for catalase (EC 1.11.1.6).
17	12	103467100	CTCGATGTCAGGGGGCATATTTGCTGGGCTAAAGGCATCTGTTCACTTCAGCAAGTAATTAAC	226	Mouse mRNA for catalase (EC 1.11.1.6).
18	19	10063800	GCTCTCTGAGCCGCAAGAAACACTCCCTTGGTTCTGGAGTGGTAGATGATGAATAG	222	Mus musculus BAC clone RP23-370D21 from chromosome 19, complete sequence.
19	18	20985850	AACTACTGGGGGCTGGAGCAATGGCTCAACCAATAGAGCTGGTGTCTTACAGAGGACCA	217	Mus musculus BAC clone RP23-41P19 from chromosome 18, complete sequence.
20	A	1200	GAAAGATCCGAACACAGCGGTGCAATTTGAGAATCAGGCTGTCACATTCGCGACTGATGAAGACAGCATCCT	217	Cloning vector pDG1731 for ectopic integration into the Bacillus subtilis chromosome, complete sequence.
21	3	119709650	AGTGAAGGACGGTGGCTGCTACCAAGATCTTGGAGATCACTACCAAGAAACACTTGTGG	212	Mus musculus ribosomal protein S20, mRNA (cDNA clone MGC102408 IMAGE342252), complete cds.
22	11	100382000	GGAGAAAGGCTGGTATGGGGCTCTGGTCTCAGGAGGGAGGACACAGGCTGATTAG	211	Mus musculus H+ ATP synthase subunit c mRNA, complete cds.
23	6	136489050	TGCTGGCCATTAAGAGATGCAAGCATTTTGAATGGAGGCGACAAAGAGGAAAGGCGCAAGTGA	211	Mus musculus ribosomal protein L38A-like mRNA (cDNA clone MGC102023 IMAGE407930), complete cds.
24	7	83756100	GCTCTGCTACAGACACATGCAAGTGGTTCAGTCCAGACAGCAGTGGCCCGAGCGATACT	168	Mus cookii phosphoprotein mRNA, complete cds.
25	3	81958250	AGTTTGGAGCGAGTGCATATATATAGTATCCCTCAGGCTAAGCTTCAGACAGATTTCTGCT	166	Mus musculus BAC clone RP24-95M16 from chromosome 3, complete sequence.
26	1	95498850	CGAGACATGGTGGCCGACACCGTGCCTGCTCCTCCTCAGATATGATGAGGTGGAGAGATT	165	Mus musculus ribosomal protein L18A, mRNA (cDNA clone MGC49597 IMAGE5011713), complete cds.
27	13	4090450	CCCTACTGGTGGCTGCTGATGATGACAGCAATGTTATAGTACTACAGAGACCAACTAGA	162	Mus musculus BAC clone RP23-193M8 from 13, complete sequence.
28	12	126553400	CAGGTGGACCGGCGAGCAACCAAGTGGCGGGGCTGACAGATCACTGGGGCTACGACAGGG	145	Mus musculus solute carrier family 27 (fatty acid transporter), member 2, mRNA (cDNA clone MGC29997 IMAGE5123650), complete cds.
29	3	105942800	CAATATGATCAGATAACTCCGATACAGAGCAAGGATCAGAAATACCTTACTGGCAATAGGT	145	Mus musculus ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1, mRNA (cDNA clone IMAGE593086), complete cds.
30	3	148632000	CTGTGATAGCAACCCAGTAATTTCTGTGCTTGAATGCTCTTTTAACTACTG	141	Mus musculus chromosome 3, clone RP24-89C15, complete sequence.

単一のリードが、複数個所にマッピングされている場合には、按分されるため、端数ができる。
整数の場合には、単一個所にマッピングされたリードで構成されていると考えられる。

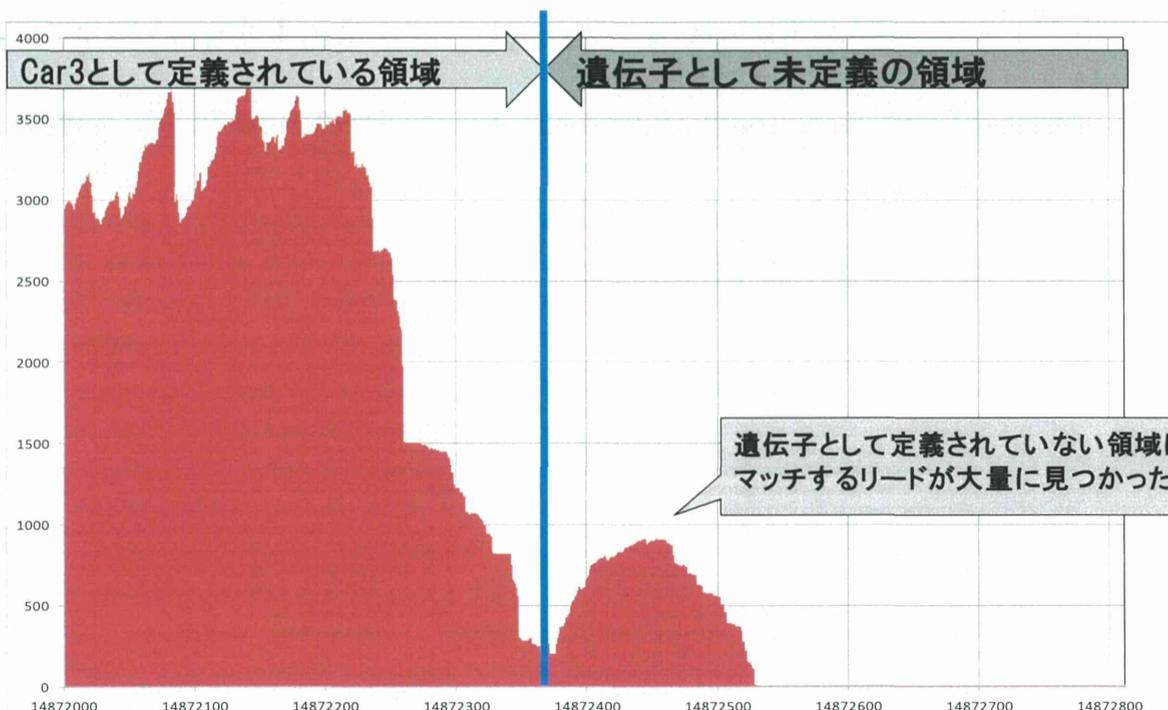
6.2.前ページのテーブル3行目

染色体3のコピーで、遺伝子として定義されていない領域にマッピングされた。



遺伝子として、定義されている領域に隣接して、大量にマッチした。

6.2.前ページの拡大



遺伝子として、定義されている領域に隣接して、大量にマッチした。

Copyright(C)2014-2015 NTT DATA Corporation

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

Alb、Apoa2、ApoEのエクソン位置

Alb	NM_009654	chr5	+	90460888	1	90460888	90461013
Alb	NM_009654	chr5	+	90460888	2	90461720	90461778
Alb	NM_009654	chr5	+	90460888	3	90462699	90462832
Alb	NM_009654	chr5	+	90460888	4	90463983	90464195
Alb	NM_009654	chr5	+	90460888	5	90464852	90464985
Alb	NM_009654	chr5	+	90460888	6	90465873	90465971
Alb	NM_009654	chr5	+	90460888	7	90467343	90467473
Alb	NM_009654	chr5	+	90460888	8	90468480	90468695
Alb	NM_009654	chr5	+	90460888	9	90469501	90469634
Alb	NM_009654	chr5	+	90460888	10	90470740	90470838
Alb	NM_009654	chr5	+	90460888	11	90472004	90472143
Alb	NM_009654	chr5	+	90460888	12	90472460	90472684
Alb	NM_009654	chr5	+	90460888	13	90474566	90474699
Alb	NM_009654	chr5	+	90460888	14	90475268	90475336
Alb	NM_009654	chr5	+	90460888	15	90476459	90476603
Apoa2	NM_013474	chr1	+	171225053	1	171225053	171225121
Apoa2	NM_013474	chr1	+	171225053	2	171225283	171225366
Apoa2	NM_013474	chr1	+	171225053	3	171225665	171225798
Apoa2	NM_013474	chr1	+	171225053	4	171226142	171226379
ApoE	NM_009696	chr7	-	19696243	1	19696243	19697103
ApoE	NM_009696	chr7	-	19696243	2	19697478	19697647
ApoE	NM_009696	chr7	-	19696243	3	19698187	19698253
ApoE	NM_009696	chr7	-	19696243	4	19699010	19699166

これらのエクソンの左端から、30塩基の位置の30塩基の配列を取り出す。

Copyright(C)2014-2015 NTT DATA Corporation

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

一致・不一致の断片数

	Alb	Apoa2	ApoE
30塩基と同じ配列を有するread断片数	33563	7248	14805
上記read配列のうち、マッチした断片数	29346	4585	13234
マッチしなかった断片数	4217 (12.6%)	2663 (36.7%)	1571 (10.6%)

Apoa2は、合致しなかった断片の割合が非常に多かった。

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

Apoa2の断片でマッチしたreadとマッチしなかったreadを比較した。

マッチ成功

```

AGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAGTCC
CAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAGTCCCTC
GCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGC
GCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAGTCCCTT
GCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAGTCCCTC
GCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAG
CCAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAG
CCAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAG
CAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAG
CTTCCAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCTGCGATTCTGTGCTGCAGTCC
CCAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC
CCAGGCTACAGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC
CTGCACTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAG
CCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC
GACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAG
GACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC
GGCCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC
GTGGCCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAG
GGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCCCTCCCG
AGATGGTGACCAGCAGTGGACCATTGGCAGCAGCTTCATGATGGCAGACTATGGCCTAGGAACAGTGCGATTCTGTGCTGCAGTCC

```

マッチ失敗

マッチしなかった断片では、3塩基が減っていた。この位置は、エクソン端なので、スプライシングの際に3塩基ずれる可能性がある。

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

Apoeの断片/Albの断片

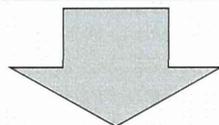
Apoeの断片は、見つけ出したreadの中で、マッチしなかったreadが、10.6%であった。
 特定の配列が生成されていることよりも、読み間違いなどが疑われる。とくに、読み始めや読み終わり近辺での、信頼性の低い領域での読み間違いがあった。

Albの断片は、見つけ出したreadの中で、マッチしなかったreadが、12.6%以下であった。
 特定の配列が生成されていることよりも、読み間違いなどの1塩基の違いが存在した。

Copyright(C)2014-2015 NTT DATA Corporation

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

エクソン跨りで、参照情報と異なるエクソン分割の可能性が示唆された。



エクソン跨りの異なる参照配列の作成方法(パターン発見のアイデア)

大量に存在するはずの遺伝子を対象にする

エクソン端近辺と一致するread断片を拾う

エクソン跨ぎの状況で違いがないかを確認する

対象は、20%以上マッチできていない遺伝子

Copyright(C)2014-2015 NTT DATA Corporation

6.2.高精度化のためのマッピングエラー配列の特定② 大量発現しているはずの遺伝子配列

次の集計をすることで、パターンを見つけることが可能となると考えられる

		一致塩基からの距離								
		1	2	3	4	5	6	7	8	9
マッチング 成功	A	1011	0	978	0	945	941	0	0	0
	G	0	0	0	13	0	0	13	932	0
	C	0	998	0	948	0	0	938	0	928
	T	0	0	0	0	13	13	0	11	11
マッチング 失敗	A	103	1	98	1	1	1	0	86	84
	G	3	3	3	95	3	3	1	3	0
	C	0	99	0	1	0	91	87	0	0
	T	0	1	1	1	92	1	1	1	1

最も存在率が高い塩基の配列を作成し、マッチングを行うことで、微小な抜けなどを解明できる可能性がある。

ACA---CAACGC

ACAGTCCAA

Copyright(C)2014-2015 NTT DATA Corporation

6.3.高精度化のためのマッピングエラー配列の特定③ マッピングできなかった特定配列

以下の方法でマッピングできなかった配列の中で大量に存在する並びを特定し、検討する。

マッピングできなかったリード配列



多い部分配列を見つける。
(マイクロサテライト以外)

30塩基程度



部分配列が何に由来しているかをblastn(国立遺伝学研究所)で確認する。

Teradata内に格納された染色体内の全配列との一致箇所を確認する。

どのような状況であれば、一致するのかを検討する。

Copyright(C)2014-2015 NTT DATA Corporation

6.3.高精度化のためのマッピングエラー配列の特定③ マッピングできなかった特定配列



マイクロサテライト以外で、頻出する配列断片(15塩基)を見つけ出す。
15塩基単位で分割済みのデータが蓄積されており、容易に見つけることができる。

染色体上にマッピングできなかった配列断片で、15塩基だけに限った場合にread配列上に、1000個以上存在する配列を見つける。

2931パターン

見つけ出されたread断片配列で、隣接する断片を結合し30塩基の断片配列の配列IDを3個見つける。

8462個

30塩基断片と一致する、染色体上の配列を見つける。
染色体との一致箇所が1か所以内

14185パターン

BLASTNで、当該read断片に対するAnnotationを検索する。
3件チェックしたが、マッチしたものが存在しなかった。
対象として、げっ歯類だけではなく、すべてのものでも同様であった。

数万を超える(1%近く)が長さ30塩基で同じ配列を有しているにも関わらず、BLASTNでは見つからなかった。

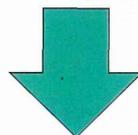
Copyright(C)2014-2015 NTT DATA Corporation

7.処理速度向上の検討



マッピング対象を絞り込むことで、マッピング処理速度の向上を図る。

全染色体配列を対象にマッピング処理を行うことは、未知のマッピング対象の発見には適している。しかし、既知領域での発現量測定したい場合には、無駄な処理を多く含む。



エクソン近傍だけを対象としたマッピングを実施・速度計測を行う。

Copyright(C)2014-2015 NTT DATA Corporation

7.1.実行結果サマリー

マッピング結果のサマリーを示す。

	マッピング時間	マッピングできたペア数	マッピングできたread数
全染色体	33時間39分	22,991,225	7,329,517
Exon近傍のみ	3時間3分	9,316,394	6,287,730

マッピング時間を大幅に更新することができた。

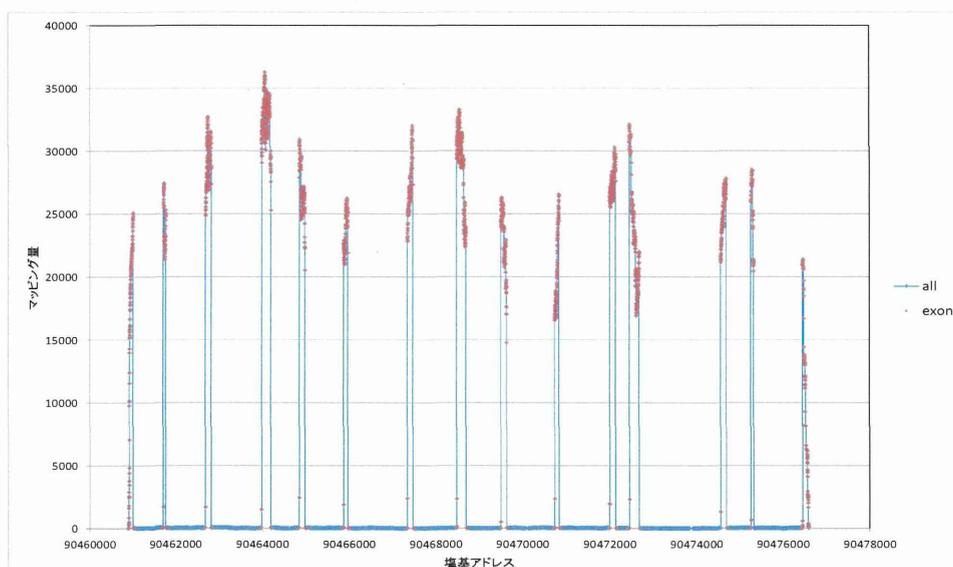
マッピングできなくなったreadが15%ほど存在した。

スプライシング前のRNAは、イントロンを含むので、マッピングできなくなったことによるものと考えられる。

Copyright(C)2014-2015 NTT DATA Corporation

7.2.実行結果 (Alb周辺の状況)

染色体5番のAlb周辺のマッピング量を確認する。



全体でマッピングしたものと、Exon近傍とマッピングしたものでは、Exon間では、もともと小さいので、大きな差異にはなっていない。全体としても、大きな差異は見られなかった。

Copyright(C)2014-2015 NTT DATA Corporation

結果

参照配列をmm9から、mm10に変更したところ、マッチする断片が増加した。

マッチできなかった断片で、大量に存在する配列を詳細にチェックを行ったところ、参照配列と3塩基の違いなど、違いにパターンが見られた。エクソンとイントロンの境目に関して新たな位置の可能性を示すことができた。

マッチングの対象配列をエクソン周辺に絞り込むことによって、高速に、マッチングを実行することができた。また、スプライシング後のエクソンで位置では、ほとんど同じような値を示していた。既知の範囲での定量化を行う際には、高速で実行可能な方法を示すことができた。

Ⅱ. 分担研究報告書

厚生労働科学研究費補助金（化学物質リスク研究事業）
分担研究報告書

「化学物質の有害性評価手法の迅速化、高度化に関する研究
ー網羅的定量的大規模トキシコゲノミクスデータベースの維持・拡充と
毒性予測評価システムの実用化の為のインフォマティクス技術開発ー」
(H24・化学・指定・006)

分担研究課題：「反復暴露実験の分子機序解析による、既存の単回暴露
実験データベースからの反復毒性予測の性能評価」

研究代表者 菅野 純

国立医薬品食品衛生研究所・安全性生物試験研究センター・毒性部・部長

研究要旨

本研究は、先行実施された Percellome(*)トキシコゲノミクス研究を基盤に、化学物質による生体影響の分子メカニズムに依拠した網羅的毒性評価手法を構築し、毒性予測と評価の一層の迅速化、高精度化を進めることを目的とする。これに際し本分担研究では、新型反復暴露実験(**)の分子機序解析による、既存の単回暴露実験データベースからの反復毒性予測の技術開発を行う。

先行研究による単回暴露実験と新型反復暴露実験の比較解析結果から、新型反復暴露が2種類の反応を遺伝子発現に及ぼすことが判明した。それは、過渡反応：暴露の都度に概ね24時間以内に収まる速い変化、及び、基線反応：暴露を重ねるに連れ、日を追って発現値の基線が徐々に移動する緩徐な反応である。さらに平成24年度、25年度に実施した四塩化炭素及びバルプロ酸ナトリウムにおける新型反復暴露実験の解析結果から、新型反復暴露で過渡反応の振幅が増減した遺伝子では、それらの基線反応も同様の増減傾向を示すことを見出した。ただし、四塩化炭素では過渡反応が急速に消失する遺伝子が多いのに対し、バルプロ酸ではそのような傾向は弱いという差異が認められた。

今年度（平成26年度）はクロフィブレートにて同設計の新型反復暴露実験セットを行い、四塩化炭素と類似した過渡反応ー基線反応関係を確認した。またそれらのシグナルネットワーク解析の結果、それらの上流に共通の分子機構が働いている強い可能性が示された。

(*) mRNA 発現値を細胞 1 個当たりのコピー数として絶対定量する方法。特許取得済み。

(**) 全動物に同量の検体を反復投与し、遺伝子発現測定直前の投与時に、溶媒群、低用量群、中用量群、高用量群に分けて最終投与を一回行う。

A. 研究目的

本研究は、化学物質による生体影響の分子メカニズムに依拠した毒性評価手法の迅速化、高度化、及びその実用の為のインフォマティクス開発を目的とする。

即ち、先行研究にて構築済みの延べ 5.8 億遺伝子情報からなる高精度トキシコゲノミクスデータベース (TGDB) と単回暴露時の毒性ネットワーク解析技術を基盤に、これらを維持・拡充しつつ、反復暴露のネットワーク解析、及び、その予測評価技術を開発する。ここにインフォマティクス専門家によるシステムトキシコロジーの概念を導入し、網羅的毒性予測評価システムの機能強化と精度向上を図る。

B. 研究方法

(1) 反復暴露影響の分子機序解析による、既存の単回暴露実験データベースからの反復毒性予測の性能評価【菅野】

B1-1: 試薬及び動物 :

本年度 (平成 26 年度) はクロフィブレート (Clofibrate; 分子量 : 242.7、Cas No. : 637-07-0、純度 98.0%、Enzo) について既存データの解析を進めた。また新たに 12 週齢の雄性 C57BL/6J マウス (日本チャールスリバー) に 4 用量 (溶媒 : 0.1% DMSO 含有 0.5% メチルセルロース水溶液) のクロフ

ィブレードを、金属ゾンデを用いて強制経口投与を行い、経時的 (投与 2、4、8 及び 24 時間後) に肝を採取した。

単回暴露 ([0+1]) 時のクロフィブレードの投与量は 0、10、30、100 mg/kg とした。新型反復暴露である 14 日間反復暴露 ([14+1]) 時のクロフィブレードの投与量は全例に対し 70 mg/kg、単回暴露は 0、10、30、100 mg/kg とした。

B1-2: Total RNA の分離精製 :

マウス組織を採取後すみやかに RNA later (Ambion 社) に 4°C で一晩浸漬し、RNase を不活化する。肝は 5mm 径の生検トレパンにより 3 ヶ所を各々別チューブに採取した。その後、RNA 抽出操作までは -80°C にて保存した。抽出に当たっては、RNA later を除いた後、RN easy キット (キアゲン社) に添付される RLT buffer を添加し、ジルコニアビーズを用いて破碎液を調製した。得られた破碎液の 10 µL を取り、DNA 定量蛍光試薬 Picogreen を用いて DNA 含量を測定した。DNA 含量に応じ、臓器毎にあらかじめ設定した割合で Spike cocktail (Bacillus 由来 RNA 5 種類の濃度を変えて混合した溶液) を添加し、TRIZOL (ライフテクノロジーズ・ジャパン社) により水層を得、RN easy キットを用いて全 RNA を抽出した。100ng を電気

泳動し RNA の純度及び分解の有無を検討した。

B1-3:GeneChip 解析：

全 RNA 5 µg を取り、アフィメトリクス社のプロトコールに従い、T7 プロモーターが付加したオリゴ dT プライマーを用いて逆転写し cDNA を合成し、得た cDNA をもとに第二鎖を合成し、二本鎖 DNA とした。次に T7 RNA ポリメラーゼ (ENZO 社キット) を用い、ビオチン化 UTP, CTP を共存させつつ cRNA を合成した。cRNA はアフィメトリクス社キットにて精製後、300-500bp となるよう断片化し、GeneChip ターゲット液とした。GeneChip には Mouse Genome 430 2.0 を用いた。ハイブリダイゼーションは 45°C にて 18 時間行い、バッファーによる洗浄後、phycoerythrin (PE) ラベルストレプトアビジンにて染色し、専用スキャナーでスキャンしてデータを得た。得られた肝サンプルについて、我々が開発した Percellome 手法 (遺伝子発現値の絶対化手法) を適用した網羅的遺伝子発現解析を行った。遺伝子発現データを、我々が開発した「RSort」を用いて、網羅的に解析した。このソフトは、各遺伝子 (probe set: ps) につき、用量、経時変化及び遺伝子の発現コピー数を各軸とした 3 次元グラフにおいて、発現を表す平面の凹凸を評価し、全ての ps を生物学的に有意な順に並び替えるソフトである。また、シグナルネットワークの探索は、Ingenuity Pathways Analysis (IPA) (Ingenuity Systems Inc.) を用いて検討した。

倫理面への配慮

動物実験の計画及び実施に際しては、科学的及び動物愛護的配慮を十分行い、所属の研究機関が定める動物実験に関する指針のある場合は、その指針を遵守している。

(国立医薬品食品衛生研究所は国立医薬品食品衛生研究所・動物実験委員会の制定になる国立医薬品食品衛生研究所・動物実験の適正な実施に関する規程 (平成 19 年 4 月版) 及び国立医薬品食品衛生研究所 遺伝子組換え実験安全管理規則の承認を受けて行う。)

C. 研究結果

平成 26 年度はクロフィブレートについて同様の新型反復暴露実験を行い、[0+1]、[1+1]、[2+1]、[4+1]、及び[14+1]の 5 通りの反復投与期間の長短による過渡反応と基線反応のそれぞれの推移と両反応の関連性の解析を肝について進めた。

過渡反応の有無 (最終日の投与の影響の有無) に関わらず、約 4,000 の遺伝子に、有意な基線反応が認められ、うち 1,700 遺伝子に 1 回から 14 回反復に亘り基線の低下が認められた。この遺伝子群は既知遺伝子情報 (IPA) により、四塩化炭素と共通する特定の経路を構成する遺伝子を高率に含んでいることが判明した。また、過渡反応が基線反応の影響を受けるものは比較的少数であり、受けないものが多いことが判明した。

H24 年度に実施した四塩化炭素、及び平成 25 年度に実施したバルプロ酸ナトリウムの解析結果を加え、これら 3 化学物質の解析結果を総合し、基線反応と過渡反応の組み合わせによる遺伝子リストを完成させ、その経時的な変動を調節する上下流の遺伝

子発現ネットワークを明らかにし、慢性毒性の分子背景の解明を進めた。その結果、①基線反応の強い抑制を示す遺伝子群の発現を支配する上流遺伝子群が3物質に共通していることが明らかになった事、②その更なる上流の候補遺伝子が単回投与から数個示唆された事、③新型反復暴露実験による過渡反応遺伝子群と基線反応遺伝子群の重なり具合が、化学物質固有であることから、この重なりの内容が反復毒性の予測に重要である可能性が示唆された事、が成果としてあげられた。

D. 考察

本年度は高脂血症治療薬として使用されているクロフィブレートについても新型反復暴露実験を行った結果、反復暴露により基線の低下が認められた遺伝子群が属する遺伝子発現経路群は四塩化炭素のそれ（単回暴露時に発現変動した遺伝子のほぼ全てについて、基線反応成分（暴露回数を重ねるに連れて発現値のベースライン（基線）が徐々に変動する反応成分）は、過渡反応成分（単回暴露時の2, 4, 8, 24時間のうちに発現が変動する速い変化の成分）が増加する場合は増加、減弱する場合は減少すること）と酷似していることが確認された。他方、過渡反応については、四塩化炭素が基線反応の影響を強く受けたのに対し、クロフィブレートの過渡反応は基線反応の影響を受けないものの比率が多かった。

他方、クロフィブレートは治療薬として利用されているものの、四塩化炭素とバルプロ酸ナトリウムの関係と同様、クロフィブレートとバルプロ酸ナトリウムの比較解析に於いて共通要素や共通特性はあまり見ら

れなかった。

今後、慢性毒性の理解のための過渡反応と基線反応の分子機序解明に、特徴的な基線反応と過渡反応の両極端のパターンを示す四塩化炭素とバルプロ酸ナトリウム、その中間的な反応を示すクロフィブレートの三者の所見の差異を利用し、基線反応を引き起こす上流機構、転写制御領域の解析、基線反応と過渡反応の連鎖・非連鎖の分子基盤、等を分担研究者の協力のもとに推進する。また、単回暴露実験の結果の中から、反復暴露時の基線反応と過渡反応と連関する、即ち、予測に使用可能な、所見の存否を明らかにする。

E. 結論

本分担研究は、ほぼ計画通りに進捗した。新型反復暴露実験を実施した四塩化炭素（平成24年度）、バルプロ酸ナトリウム（平成25年度）及びクロフィブレート（平成26年度）の3化学物質の解析結果から、大筋に於いて、単回暴露時の過渡反応成分と反復暴露時の基線反応成分の基本的な関連性が共通していることを見いだしつつ、各化学物質特有の所見を得た。この過渡反応と基線反応に関する知見は生物学的・毒性学的に新規性が高くエピジェネティクス等の機序の関与が示唆されることから、これを明らかにすることは、反復毒性の分子毒性学的理解の促進、及び、単回暴露実験データベースからの反復毒性予測法を開発するにあたり重要と考える。これらの研究成果は、単回暴露（急性）の毒性ネットワーク解析結果から、反復暴露による生体影響の予測評価技術の開発を推進するとともに、その予測精度の把握と、1乃至4日間の先行反

復暴露を行うプロトコール（上記において [1+1] 或いは [4+1] と表記したもの）による高精度の慢性毒性予測補填の技術開発の可能性をも明らかにするものである。

F. 研究発表

1. 論文発表

Janesick A, Nguyen TT, Aisaki K, Igarashi K, Kitajima S, Chandraratna RA, Kanno J, Blumberg B. Active repression by RAR γ signaling is required for vertebrate axial elongation. Development. (2014);141(11):2260-70.

Tanaka M, Aisaki K, Kitajima S, Igarashi K, Kanno J and Nakamura T, Gene expression response to EWS-FLI1 in mouse embryonic cartilage. Genomics Data 2: 296-298, 2014.

Tanaka M, Yamazaki Y, Kanno Y, Igarashi K, Aisaki K, Kanno J, Nakamura T. Ewing's sarcoma precursors are highly enriched in embryonic osteochondrogenic progenitors. J Clin Invest. (2014);124(7):3061-74.

2. 学会発表

Jun Kanno, “ Taquann” Dispersion Method with Direct Injection Whole-Body Inhalation System for Engineered Nano Materials Toxicity Studies, the 54th Annual Meeting of the Society of Toxicology (2015.3.25) San Diego, USA, poster

Jun Kanno, Ken-ichi Aisaki, Satoshi Kitajima, Percellome Toxicogenomics, 50th Congress of the European Societies of Toxicology (EUROTOX2014)(2014.9.9) Edinburgh, UK, poster

Jun Kanno, Ken-ichi Aisaki, Satoshi Kitajima, Percellome toxicogenomics project as the 3R-toxicology and the foundation of in vitro- and in silico-toxicology, the 9th World Congress on Alternatives and Animal Use in the Life Sciences (WC9) (2014.8.27), Prague, Czech Republic, Oral

相崎健一、北嶋 聡、菅野 純、遺伝子発現から見た毒性学—Percellome トキシコゲノミクスの進捗—、第 36 回日本中毒学会総会・学術集会(2014.7.25) 東京、シンポジウム

北嶋 聡、小川幸男、大西 誠、相磯成敏、相崎健一、五十嵐勝秀、高橋祐次、菅野 純、シックハウス症候群レベルの極低濃度吸入暴露時の海馬 Percellome トキシコゲノミクス—化学構造が異なる 3 物質の比較—、第 41 回日本毒性学会学術年会 (2014.7.3) 神戸、口演

菅野 純、相崎健一、北嶋 聡、Percellome Project の進捗—新型反復暴露による慢性毒性の予測に向けての分子背景の解析—、第 41 回日本毒性学会学術年会 (2014.7.2) 神戸、シンポジウム

北嶋 聡、種村健太郎、菅野 純、毒性の網

羅的把握のための遺伝子発現ネットワーク
描出と動的バイオマーカー抽出、第 41 回日
本毒性学会学術年会(2014.7.2)神戸、シンポ
ジウム

G. 知的所有権の取得状況

1. 特許取得

特許第 5177712 号、2013 年 1 月 18 日登
録、特許権者：国立医薬品食品衛生研究
所、NTT データ、発明者：菅野純、相崎
健一ら、「競合的ハイブリダイゼーション
における遺伝子データの補正方法及び補
正装置」

2. 実用新案登録

なし

3. その他

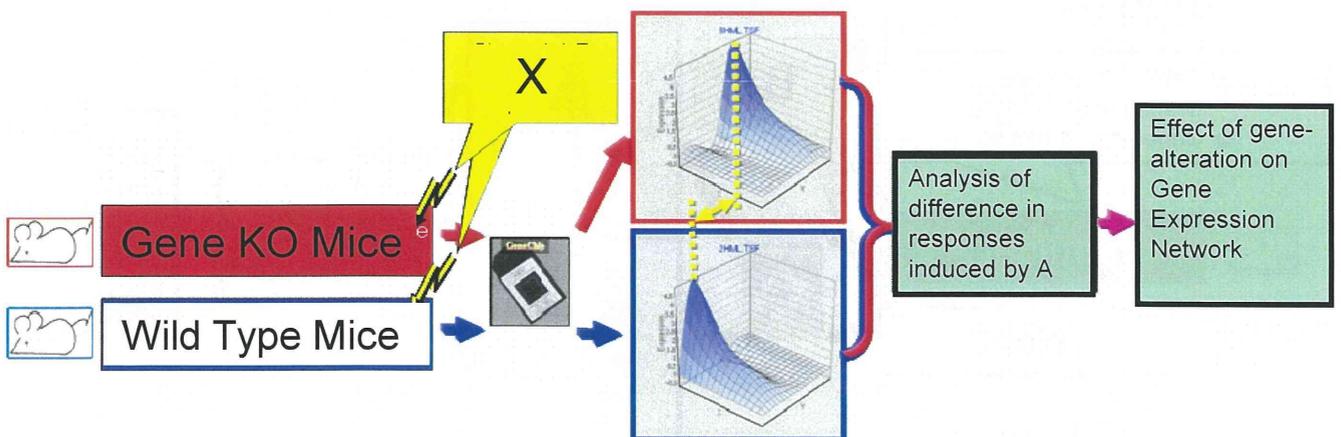
なし

平成24年度厚生労働科学研究補助金 化学物質リスク研究事業
化学物質の有害性評価手法の迅速化、高度化に関する研究
-網羅的定量的大規模トキシゲノミクスデータベースの維持・拡充と
毒性予測評価システムの実用化の為にインフォマティクス技術開発-

反復暴露実験の分子機序解析による、 既存の単回暴露実験データベースからの 反復毒性予測の性能評価

国立医薬品食品衛生研究所・安全性生物試験研究センター・毒性部
菅野 純

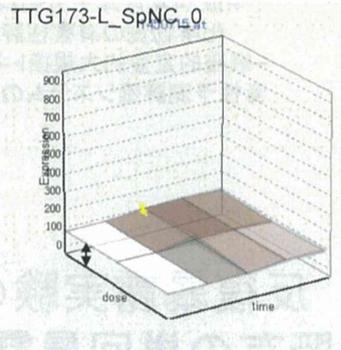
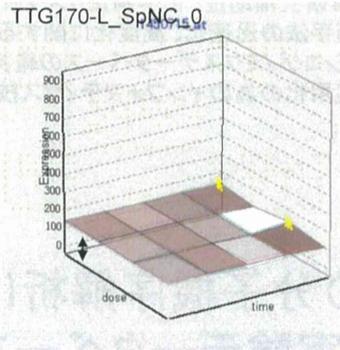
Gene Knockout Mice



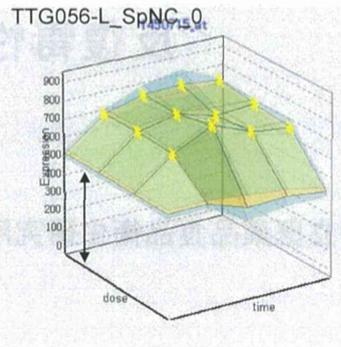
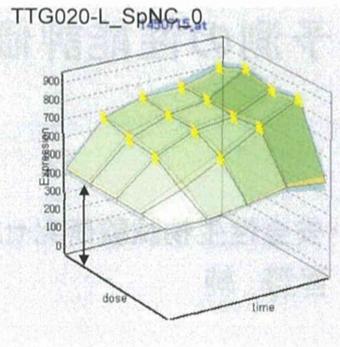
TCDD

3-MC

AhR KO mouse



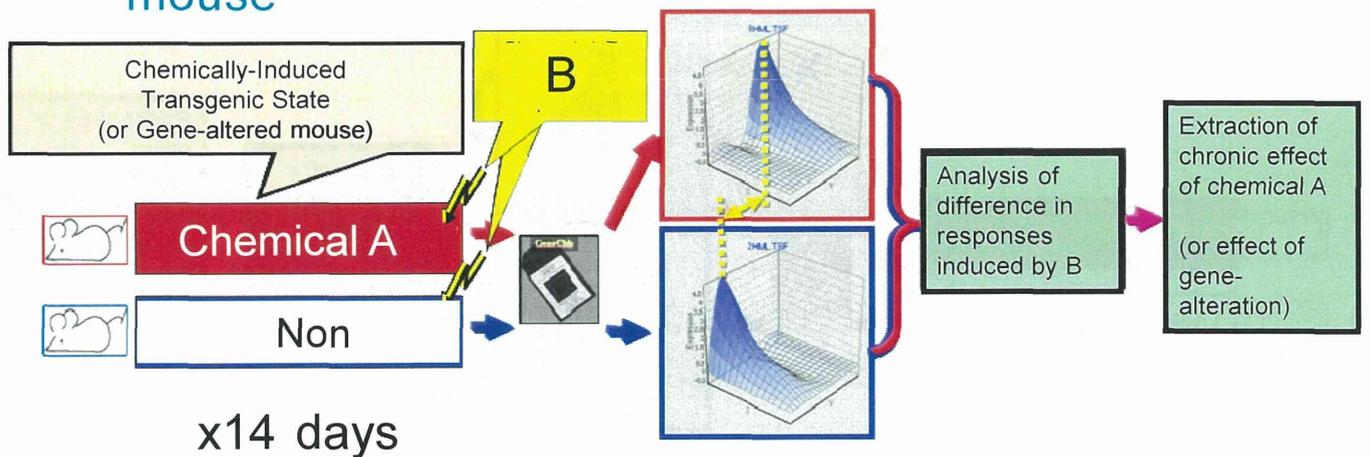
Wild type mouse



Cyp1a2
 NM_009993
 cytochrome P450, family 1, subfamily a, polypeptide 2

New Concept of Repeated Dose

Consider “chronic/ repeated exposure”
 as
 “Chemically-Induced Transgenic State”,
 which is equivalent to Gene Knockout / transgenic
 mouse

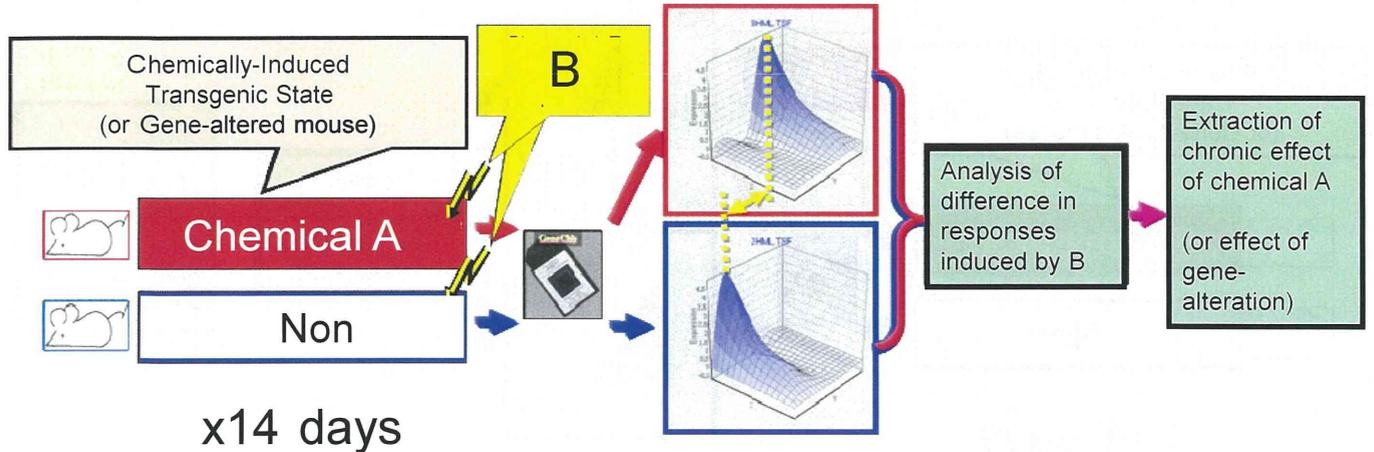


New Concept of Repeated Dose

慢性(反復)暴露を、

【化学物質投与により誘導された遺伝子発現改変状態】と考える。

慢性(反復)暴露反応を、遺伝子改変動物と比較できる。



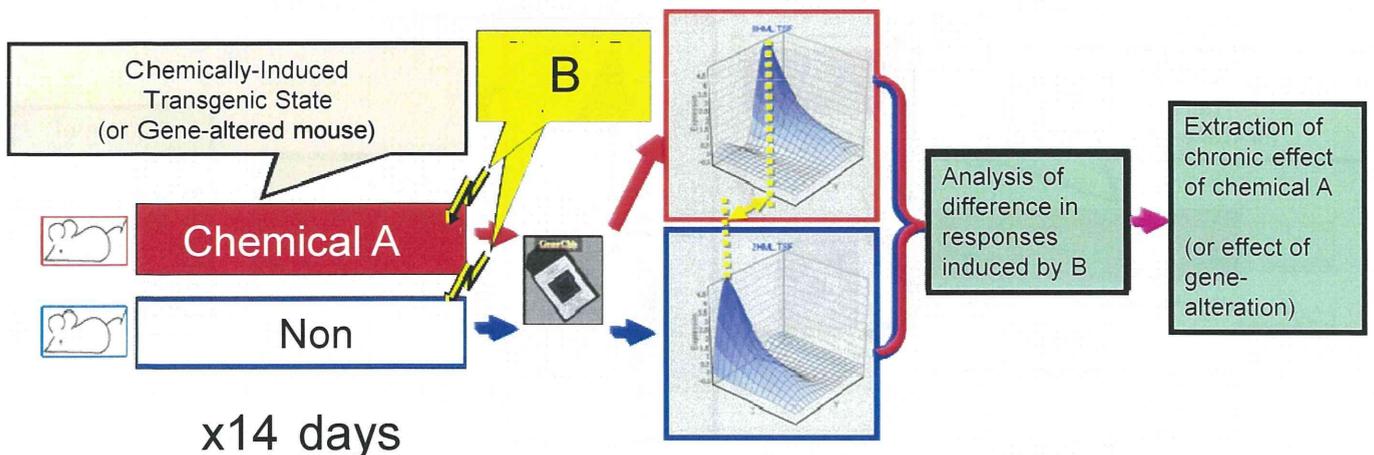
New Concept of Repeated Dose

慢性(反復)暴露を、

【化学物質投与により誘導された遺伝子発現改変状態】と考える。

A+B protocol

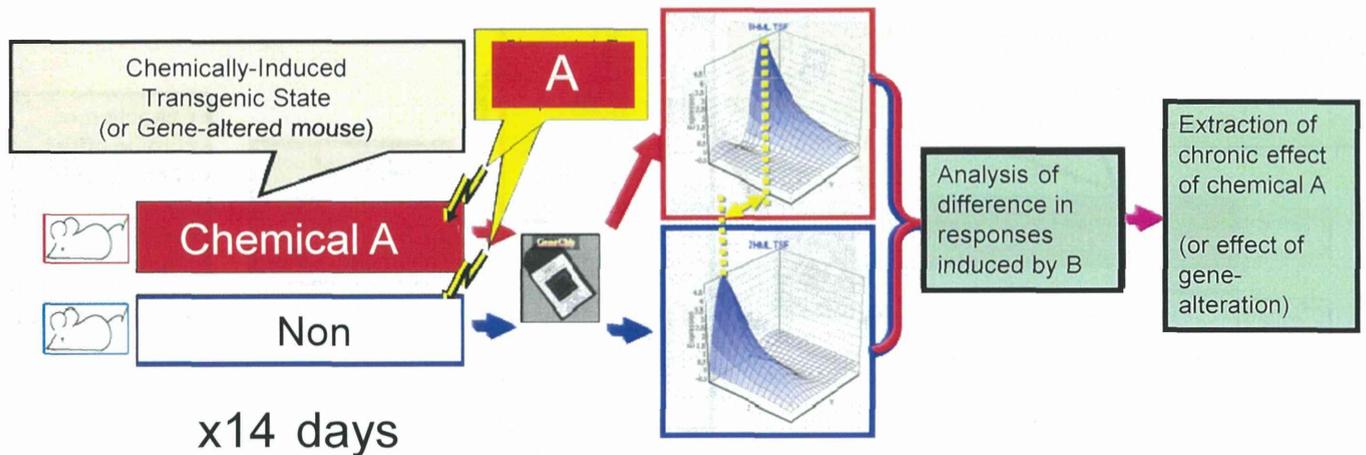
慢性(反復)暴露反応を、遺伝子改変動物と比較できる。



New Concept of Repeated Dose 拡張

慢性(反復)暴露を、化学物質投与により誘導された遺伝子発現改変状態と考える。

↓
化学物質 A の反復影響を A で観察する。



New Concept of Repeated Dose 拡張

慢性(反復)暴露を、化学物質投与により誘導された遺伝子発現改変状態と考える

化学物質 **A+A' protocol**

