

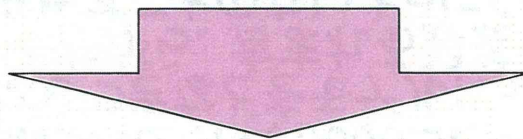
テーマ

平成21年までに独自に開発した、遺伝子発現誘導に基づく化学物質分類技術、及び同期発現遺伝子抽出技術による解析をより一層高精度化させるために、これらのアルゴリズムを、昨年度までに開発・実用化したマイクロアレイデータ高精度化技術(MLANGなど)に適用させる。

1.化合物分類

以前に低発現と考えられていた1コピー以下の遺伝子が、差が小さいが大量に存在するため、化合物分類において、化合物間の距離の主な原因となっていた。

平成22年度の研究において、マイクロアレイ高精度化技術(MLANG)を改良し、上で問題となっていた遺伝子を0発現遺伝子として特定することができた。



低発現の原因が取り除かれたので、高発現の遺伝子が、化合物分類において、化合物間の距離の主要な要因となった。

Apoa2, Apoe, Albなど 発現値が大きすぎて変動が検知できないような遺伝子

想定される誤差の原因

1. 飽和補正における、飽和最大値の推定精度
2. 高濃度の計測誤差

これらの誤差の影響を低減化する手法を含めた化合物分類を作成する

1.1.化合物分類手法の改良項目

- ・ 化合物分類のための想定される誤差
- ・ MLANGの特性と結果の精度
 - MLANGは、双曲線により飽和を補正しているため、高濃度の発現に関して、比較的誤差が大きい。QPCRとの対比より、MAS5を使用した計算での限界は600コピー程度であった。MLANGでは、5000コピーであっても、プローブセット内の各プローブの飽和のばらつきなどをモデルに組み込んであり、状況によっては、推定可能である。しかしながら、含まれる誤差も大きなものになる。この差異が、主要な差異を生み出すことのないようにシグモイドを模した誤差軽減の係数を用いる
- ・ 個体間の変動
 - Percellome実験では、同一の実験条件で、3匹のマウスを用いている。2条件で、3匹同士を比較した際に、絶対コピー数の差異を比較すると、個体差により2条件で差異があるように判定されることがある。個体間で変動が大きな遺伝子が主要な差異を生み出すことがないように、個体間の差異の大きな遺伝子の差異を軽減する係数を用いる。

1.1.化合物分類の改良項目

- ・ 化合物を実験動物に投与する実験では、実験間の差異も無視できない。特に、溶媒が肝臓で代謝されている可能性があり、肝臓と溶媒の間では強い関係があると思われる。
 - ダイオキシン類は脂溶性のため、Corn Oilを溶媒として用いるが、Corn Oilを使用すると、脂肪を輸送、代謝関連の遺伝子が活発に動くことが想定される
- ・ プロジェクト間差異を生み出すと考えられる要因
 - 溶媒
 - 餌種類
 - 季節
 - ・ 餌時間、昼夜の時間は管理されているので大きな変化はないはず
 - マイクロアレイ出荷ロット
 - 機器洗浄
 - 実験者
 - ・ 個人、習熟度

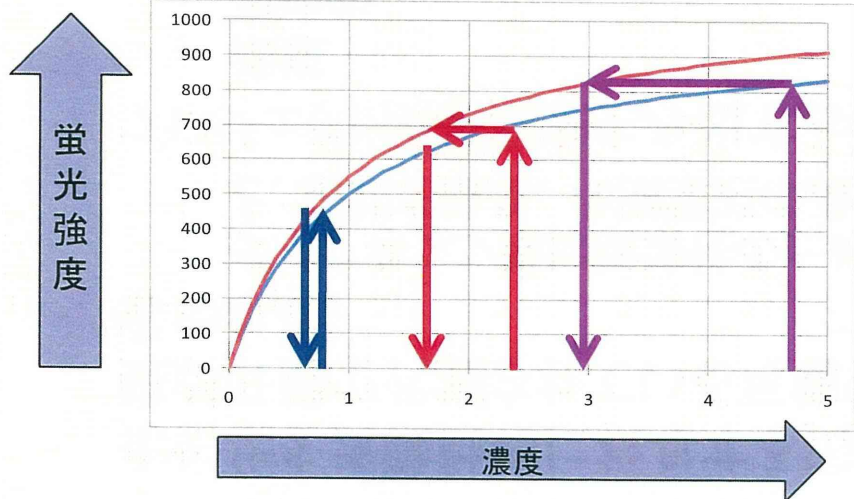
化合物影響とプロジェクト影響を個別に取り出せる工夫を考える

1.2.化合物間距離の定義 誤差による重み付け

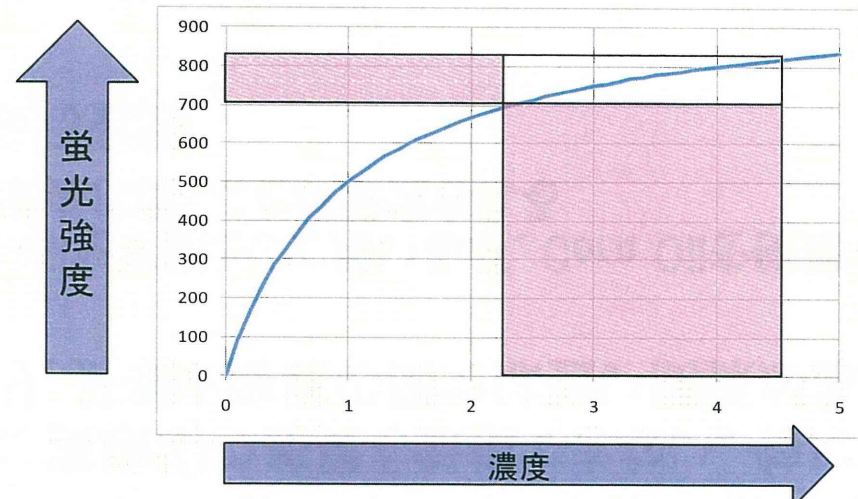
高濃度の誤差の原因

1. 飽和補正における、飽和最大値の推定精度
2. 高濃度の計測誤差

飽和最大値の誤差は、高濃度ほど大きく影響を受ける



蛍光強度の誤差は、高濃度ほど大きく影響を受ける



高濃度のものを双曲線で誤差を戻して、評価する。評価の比重は、化合物分類を実施すると、1遺伝子1個の係数を作成する。

1.2.化合物間距離の定義

各プローブセットの効果係数

プローブセットpの通常発現値 L_p を、溶媒群における全計測値の中央値により求める

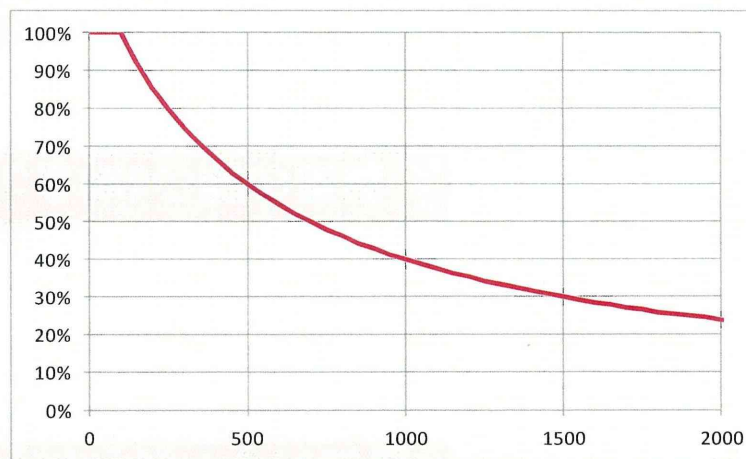
$$L_p = \underset{a,i,j}{\text{median}} (x_{paj})$$

プローブセットpに対して効果係数を次のように定義する

$$\text{effect}_p = 1 + \frac{f(L_p - L_{base})}{I_{max}}$$

ここで関数fは、負のときに0となる関数

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$



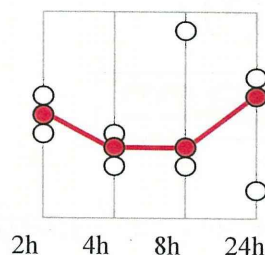
L_{base} 以下の場合には効果係数は100%で、 $I_{max} + L_{base}$ の時に50%となるように設定する。

L_{base} は100コピー、 I_{max} は600コピーと設定した。

効果係数を全組み合わせで共通とすることにより、化合物間の距離を、数学的な距離(三角不等式を満たす)として取り扱えるようにし、化合物間の距離をユークリッド座標系に割り付ける際に不整合を起こさないようにした

1.2.化合物間距離の定義

次のように化合物間距離を定義する



時刻

プローブセットp、化合物a、時刻iにおける代表値

$$x_{pai} = \text{median}_{j=1}^3 (x_{paj})$$

プローブセットp、化合物a、化合物bの間の距離

$$r_{pab}^2 = \sum_{i=1}^4 (x_{pai} - x_{pbi})^2$$

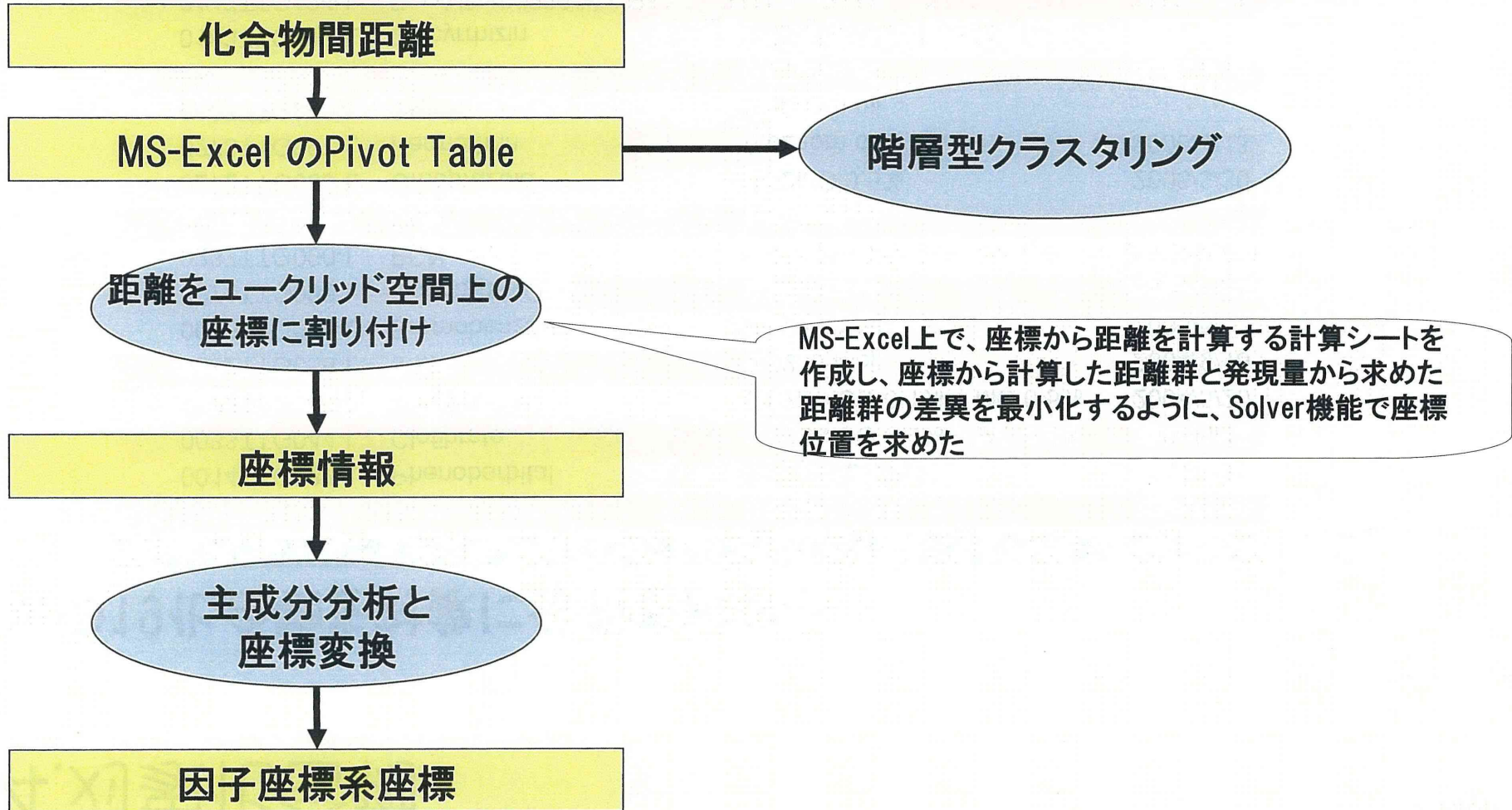
全プローブセットの化合物a、化合物bの間の距離

$$r_{ab}^2 = \sum_{p=1} \frac{r_{pab}^2}{\text{effect}_p^2}$$

化合物とDoseの組み合わせで、一つの反応パターンとして取り扱う。化合物分類では、この組み合わせで、「化合物」と呼ぶこともある

1.3.距離計算後の手順

距離計算後に次の化合物分類の手法を行う



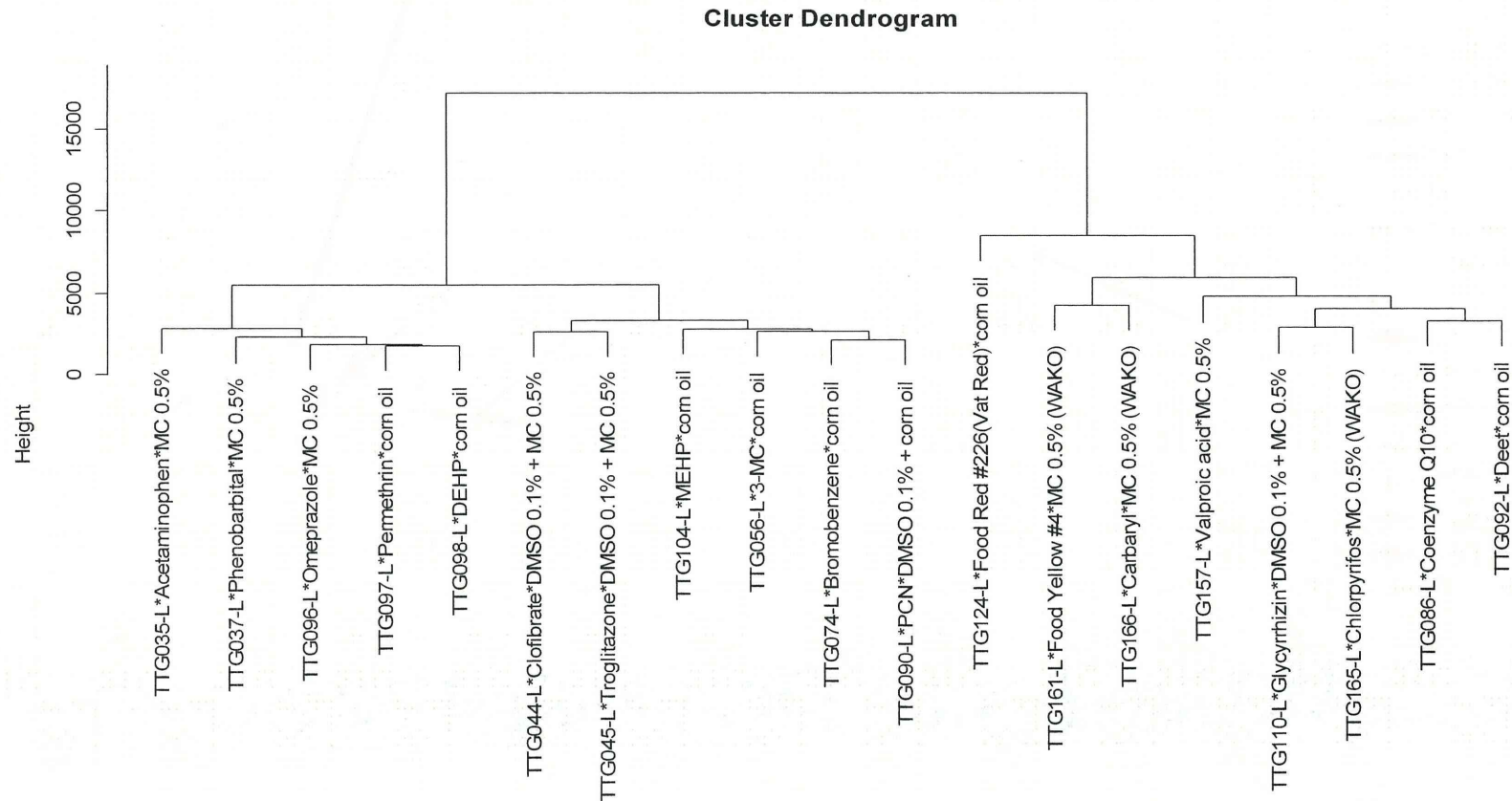
1.4.対象化合物

・ 次の19化合物を対象に分析を試みる

Project No.	Project Name	Chemical	SpF	Vehicle	date
0086TTG035-L		Acetaminophen	0.02	MC 0.5%	2004/2/26
0014TTG037-L		Phenobarbital	0.02	MC 0.5%	2004/2/19
0033TTG044-L		Clofibrate	0.02	DMSO 0.1% + MC 0.5%	2004/5/13
0029TTG045-L		Troglitazone	0.02	DMSO 0.1% + MC 0.5%	2004/5/20
0052TTG056-L		3-MC	0.02	corn oil	2004/9/16
0090TTG074-L		Bromobenzene	0.02	corn oil	2005/5/12
0271TTG086-L		Coenzyme Q10	0.02	corn oil	2005/11/15
0237TTG090-L		PCN	0.02	DMSO 0.1% + corn oil	2005/10/20
0219TTG092-L		Deet	0.02	corn oil	
0272TTG096-L		Omeprazole	0.02	MC 0.5%	2006/3/30
0273TTG097-L		Permethrin	0.02	corn oil	2006/4/13
0107TTG098-L		DEHP	0.02	corn oil	
0108TTG104-L		MEHP	0.02	corn oil	2006/5/16
0183TTG110-L		Glycyrrhizin	0.02	DMSO 0.1% + MC 0.5%	
0147TTG124-L		Food Red #226(Vat Red)	0.005	corn oil	2007/1/25
0245TTG157-L		Valproic acid	0.005	MC 0.5%	2008/12/18
0274TTG161-L		Food Yellow #4	0.005	MC 0.5% (WAKO)	2009/4/23
0275TTG165-L		Chlorpyrifos	0.005	MC 0.5% (WAKO)	2009/7/16
0276TTG166-L		Carbaryl	0.005	MC 0.5% (WAKO)	2009/12/17

1.5.階層クラスタリング

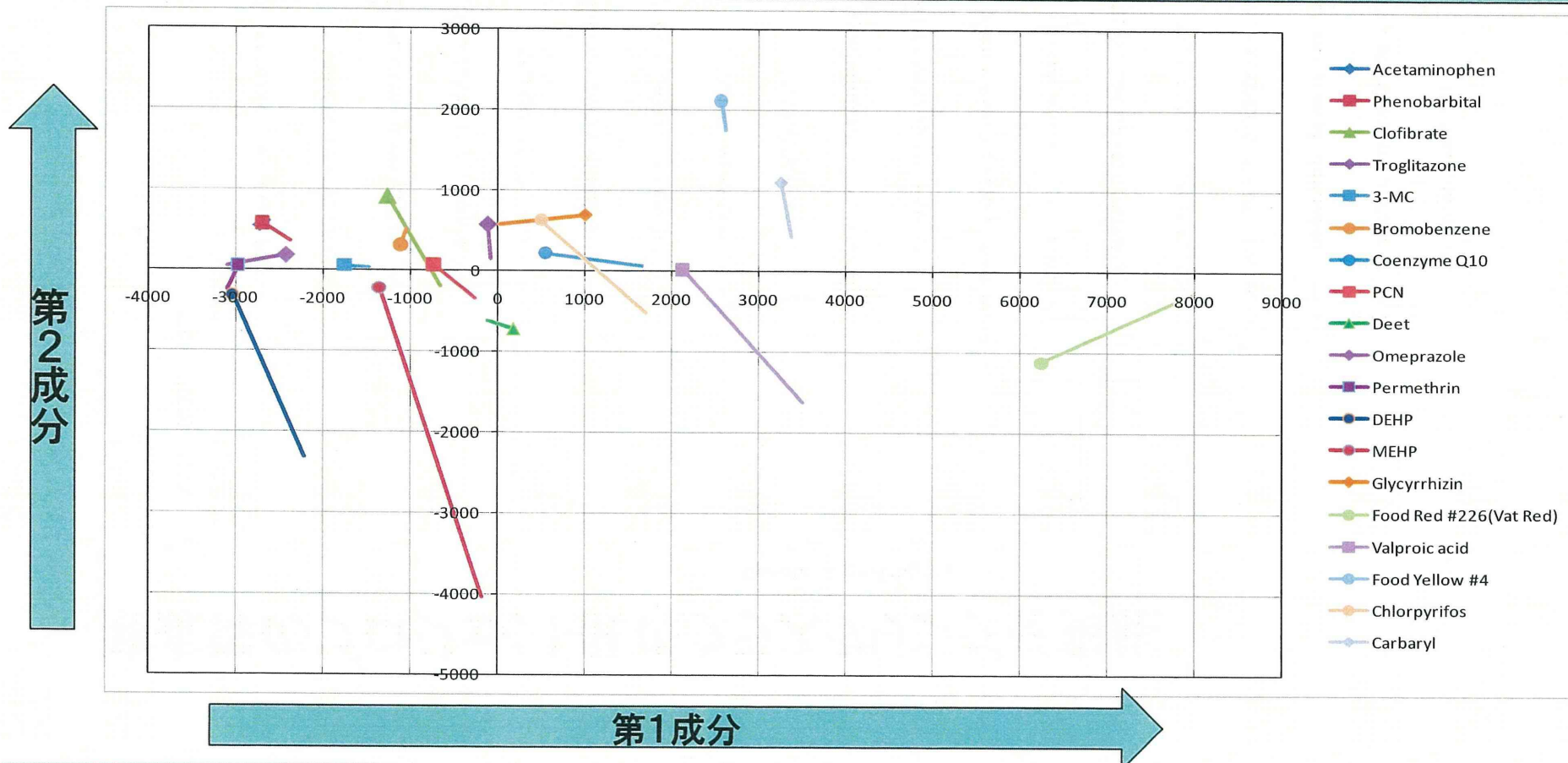
- 溶媒群のプロジェクト間のクラスタリングを実施



実験時期(TTG番号)による影響が強いように見える

1.6.主成分分析による座標変換

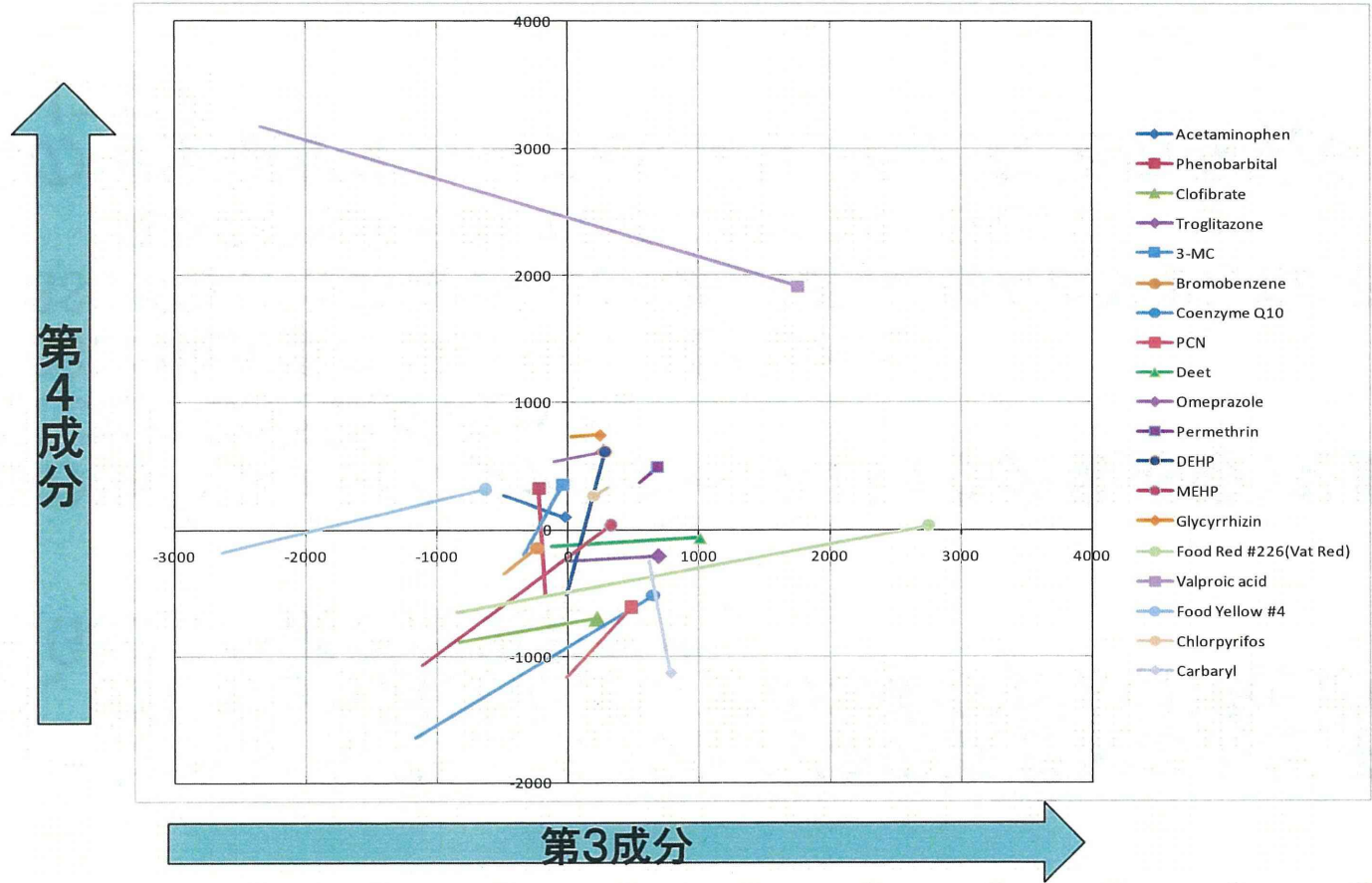
19化合物の溶媒と最大Doseに対する主成分分析結果 第1,2成分



第1成分として、溶媒の差異、特にGSC添加濃度変化による影響が強く表れている

1.6.主成分分析による座標変換

19化合物の溶媒と最大Doseに対する主成分分析結果 第3, 4成分



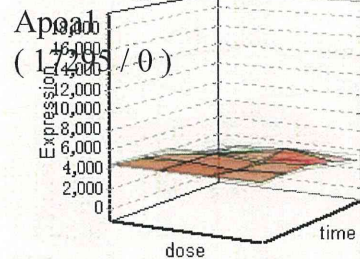
第3成分として、赤色226号、バルプロ酸の影響が現れていると考えられる

-129-

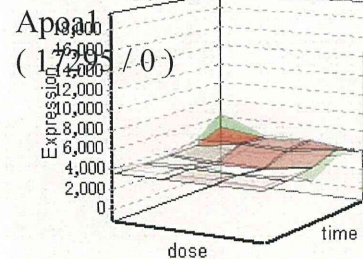
1.7.化合物分類結果の考察

- ・ 次ページから、対象とした19化合物の高濃度の遺伝子(Apoa1、Alb)、及び、GSCのうち濃度の高い方の3種のRNA(THR(GSC)、LYS(GSC)、PHE(GSC))の補正結果を示す。
- ・ GSCのSpF(スパイクファクター:スパイク濃度係数)を変更したタイミングで、大きな変動があった。
 - GSCの低用量のRNAが「0」と推定された
 - GSCが薄くした時点でもっとも激しく現れ、徐々に減少している
 - ・ 実験の経験が反映しているのでは、ないだろうか
- ・ GSCの推定値がずれることにより、チップごとの発現量が全体として動くことになる。
- ・ MLANG補正の詳細を確認し、MLANG補正の精度を向上させることにより、化合物分類の精度を向上させる。

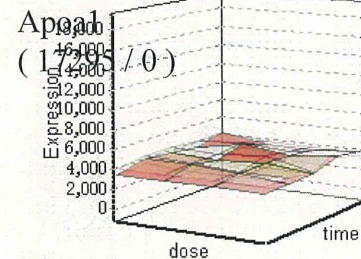
MLG801_TTG035-L



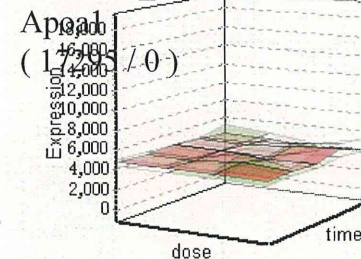
MLG801_TTG037-L



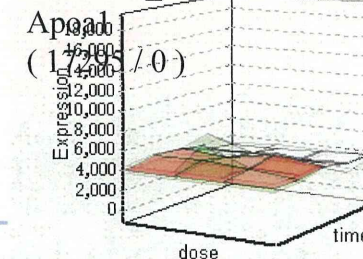
MLG801_TTG044-L



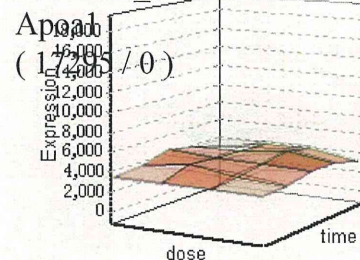
MLG801_TTG045-L



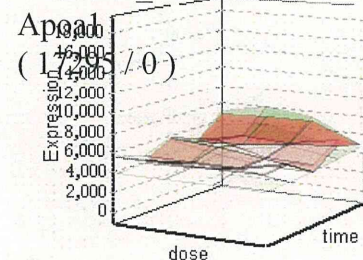
MLG801_TTG056-L



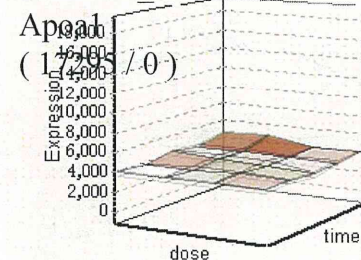
MLG801_TTG074-L



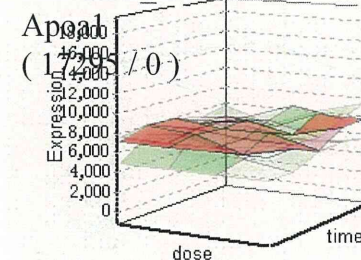
MLG801_TTG086-L



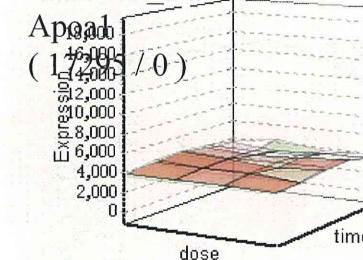
MLG801_TTG090-L



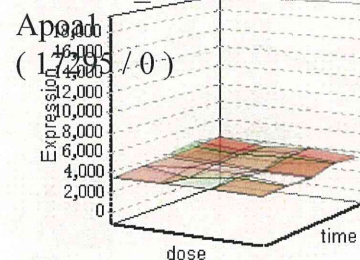
MLG801_TTG092-L



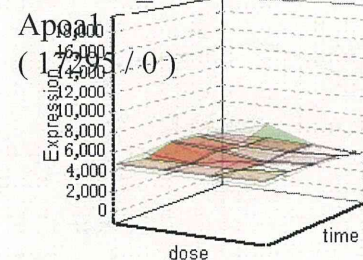
MLG801_TTG096-L



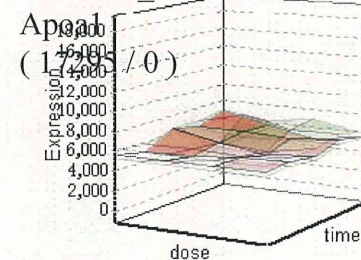
MLG801_TTG097-L



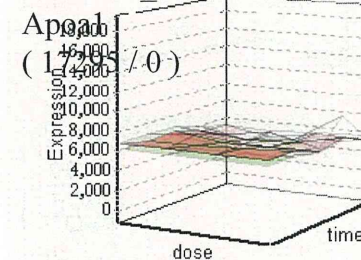
MLG801_TTG098-L



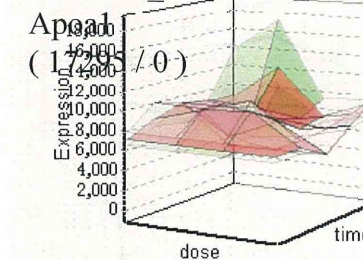
MLG801_TTG104-L



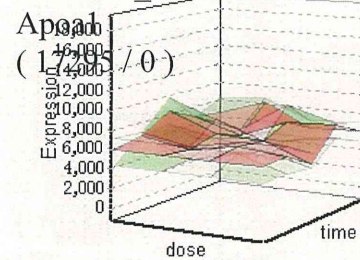
MLG801_TTG110-L



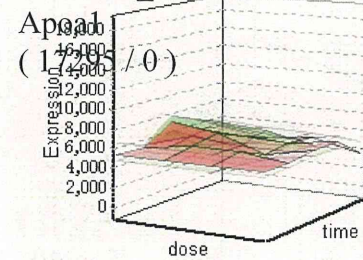
MLG801_TTG124-L



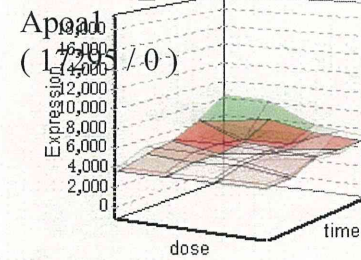
MLG801_TTG157-L



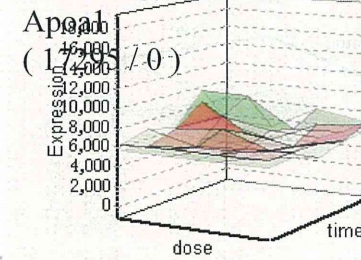
MLG801_TTG161-L



MLG801_TTG165-L

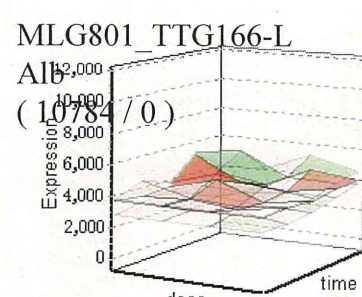
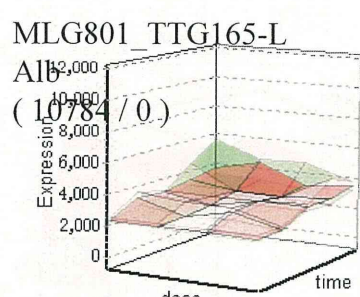
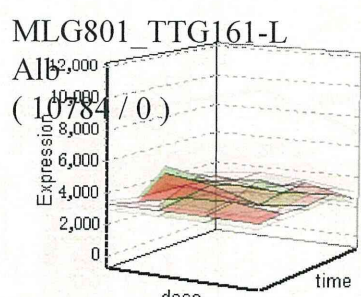
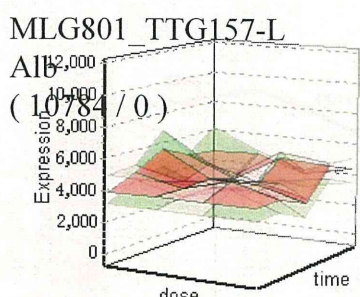
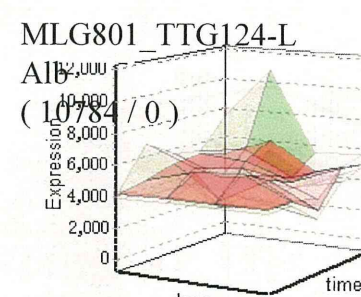
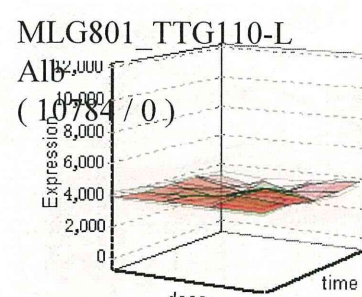
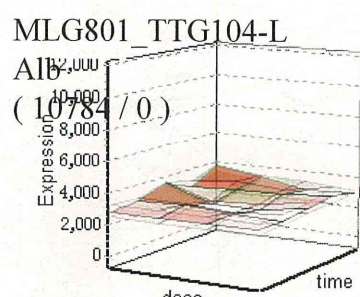
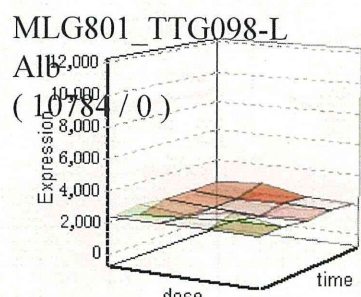
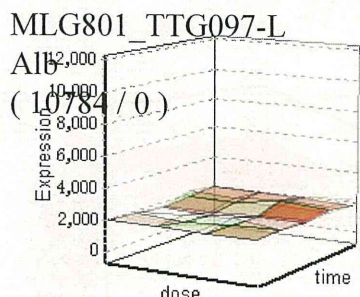
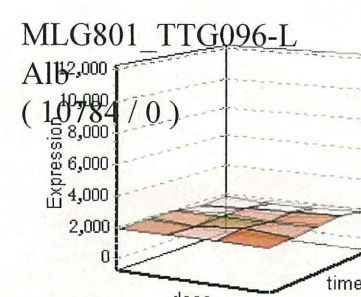
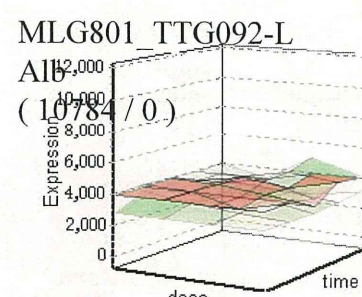
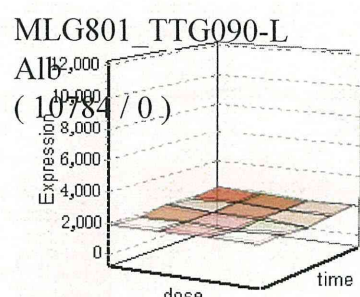
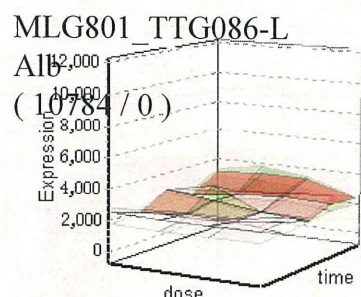
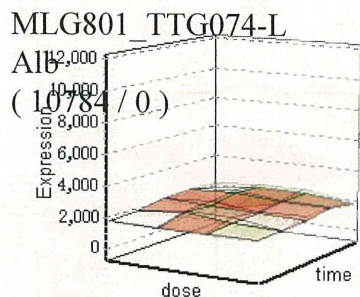
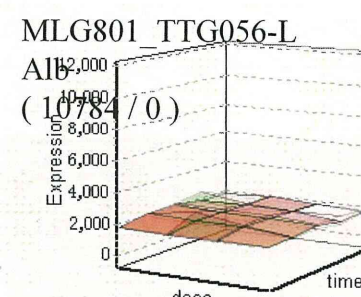
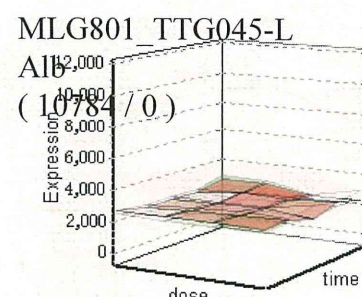
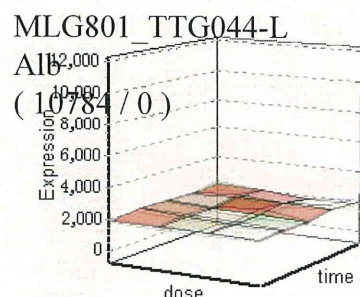
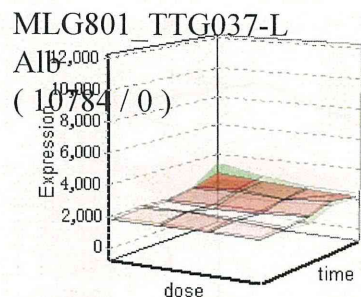
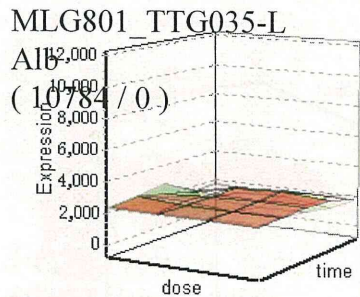


MLG801_TTG166-L



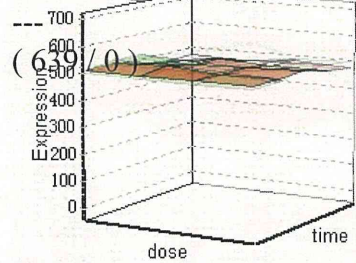
高発現遺伝子の例

TTG124から、GSCを薄めている。

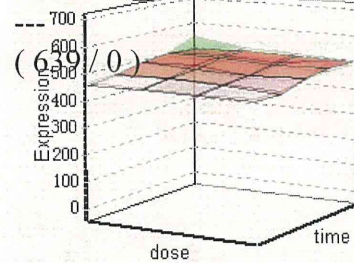


高発現遺伝子の例
TTG124から、GSCを薄
めている。

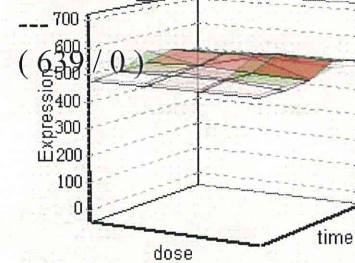
MLG801 TTG035-L



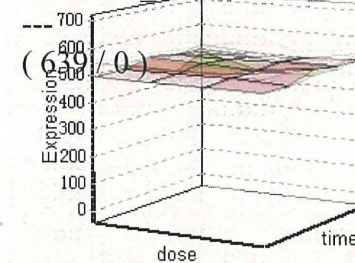
MLG801 TTG037-L



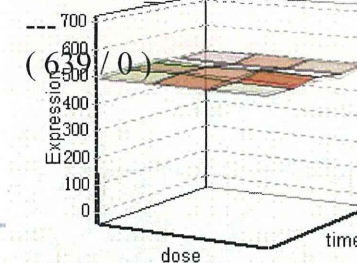
MLG801 TTG044-L



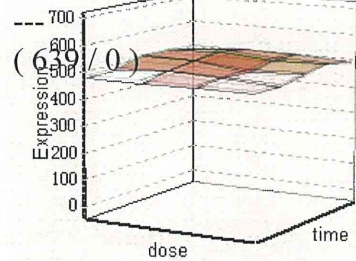
MLG801 TTG045-L



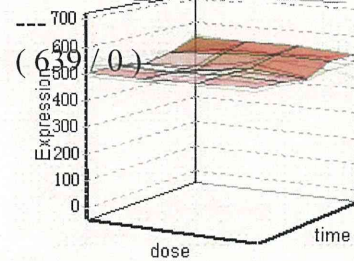
MLG801 TTG056-L



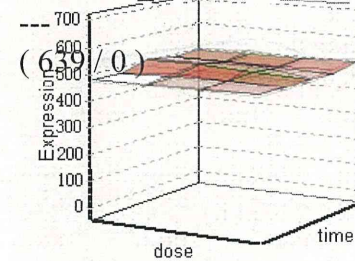
MLG801 TTG074-L



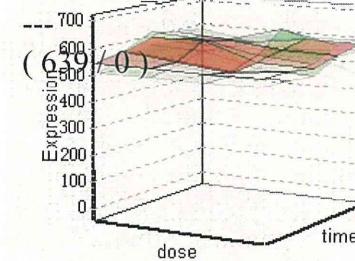
MLG801 TTG086-L



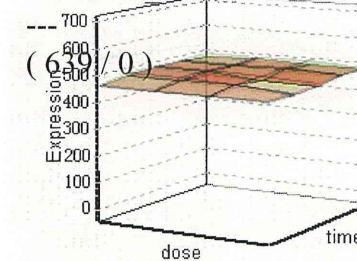
MLG801 TTG090-L



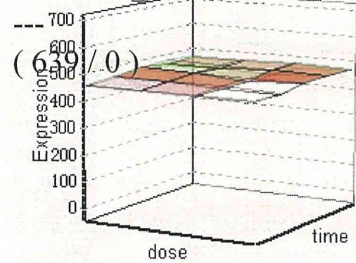
MLG801 TTG092-L



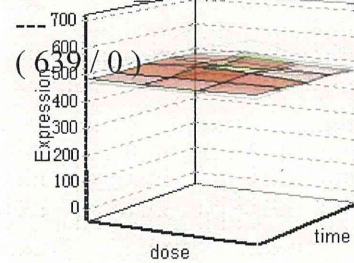
MLG801 TTG096-L



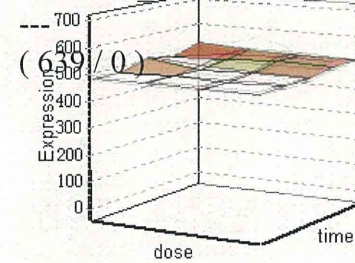
MLG801 TTG097-L



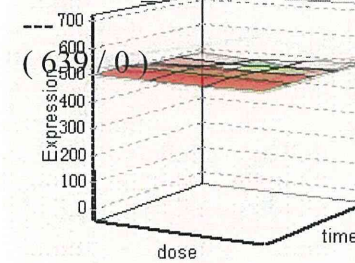
MLG801 TTG098-L



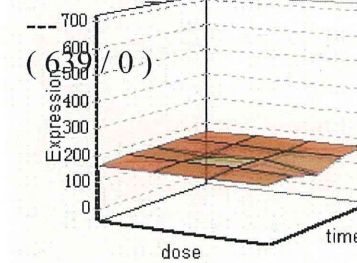
MLG801 TTG104-L



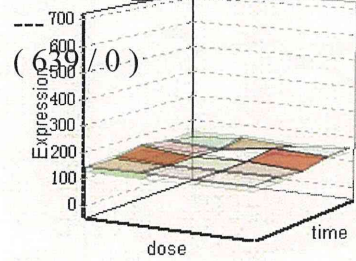
MLG801 TTG110-L



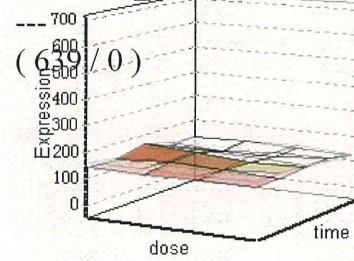
MLG801 TTG124-L



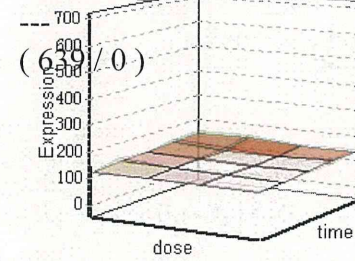
MLG801 TTG157-L



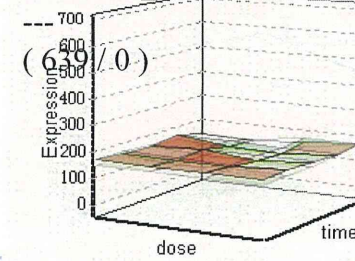
MLG801 TTG161-L



MLG801 TTG165-L

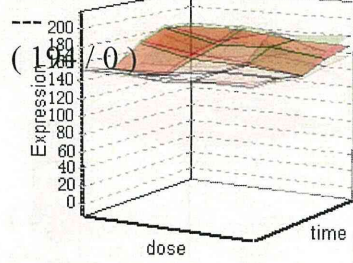


MLG801 TTG166-L

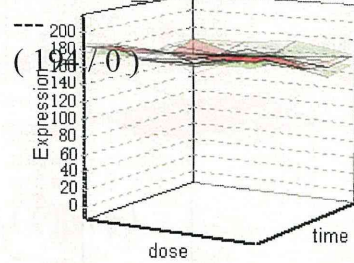


GSC(THR)の推定値
TTG92で変動が大きい

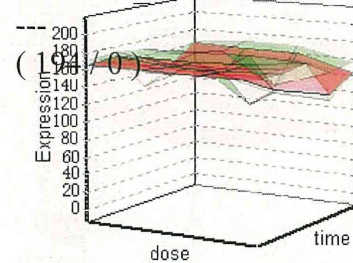
MLG801 TTG035-L



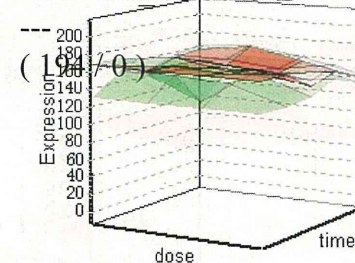
MLG801 TTG037-L



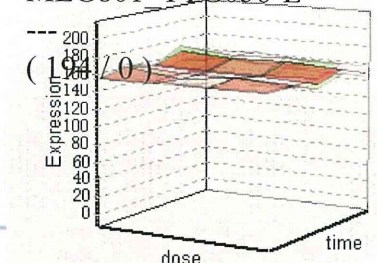
MLG801 TTG044-L



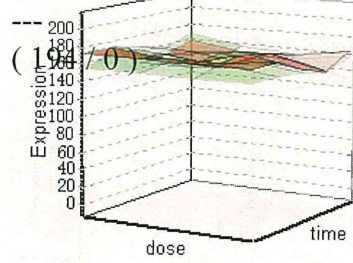
MLG801 TTG045-L



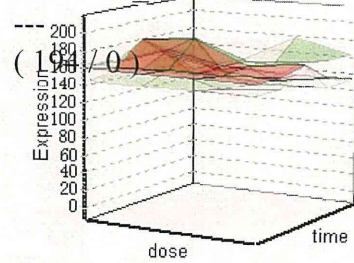
MLG801 TTG056-L



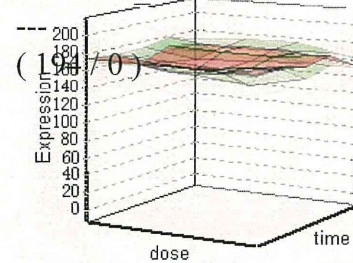
MLG801 TTG074-L



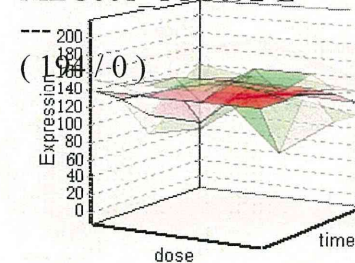
MLG801 TTG086-L



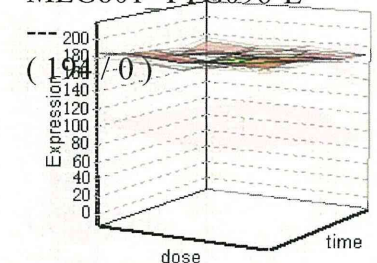
MLG801 TTG090-L



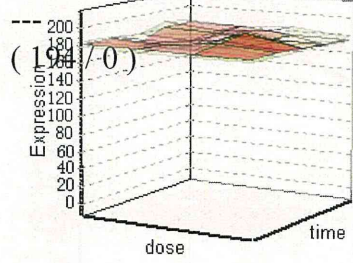
MLG801 TTG092-L



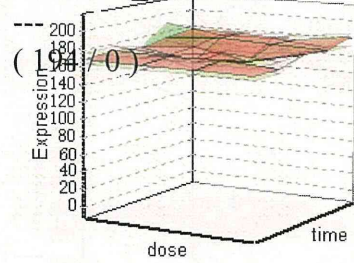
MLG801 TTG096-L



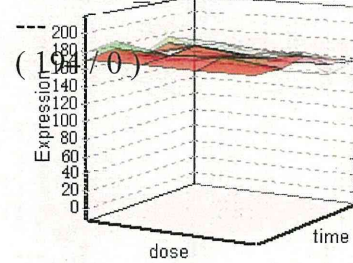
MLG801 TTG097-L



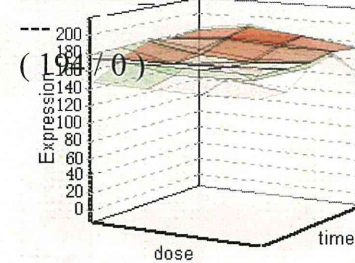
MLG801 TTG098-L



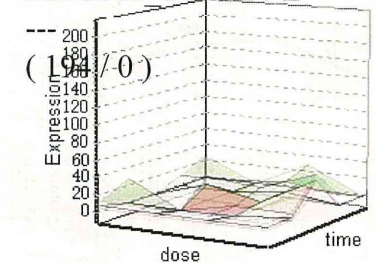
MLG801 TTG104-L



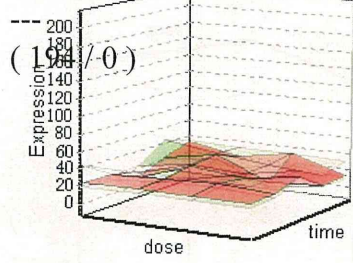
MLG801 TTG110-L



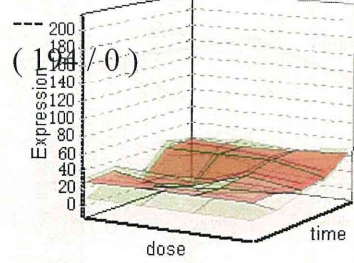
MLG801 TTG124-L



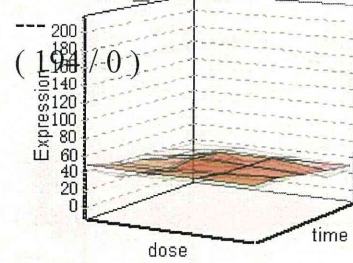
MLG801 TTG157-L



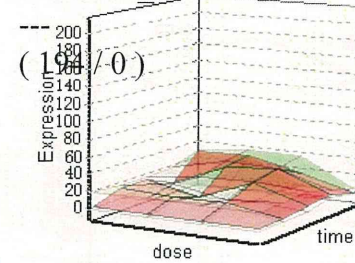
MLG801 TTG161-L



MLG801 TTG165-L

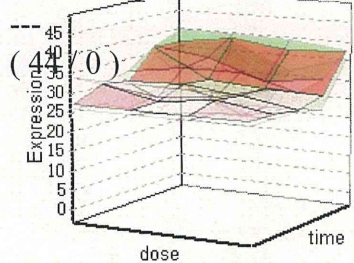


MLG801 TTG166-L

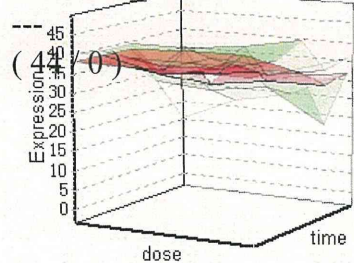


GSC(LYS)の推定値
TTG124以降で変動が
大きい

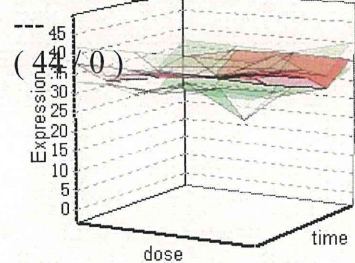
MLG801 TTG035-L



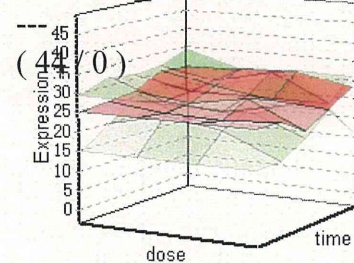
MLG801 TTG037-L



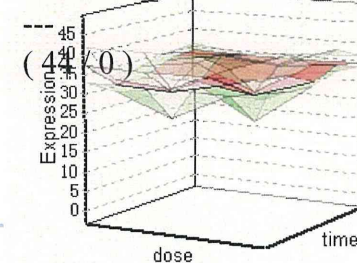
MLG801 TTG044-L



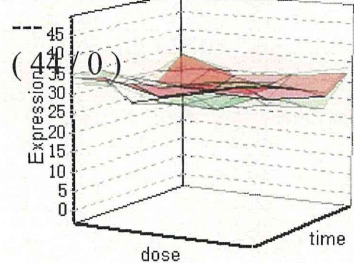
MLG801 TTG045-L



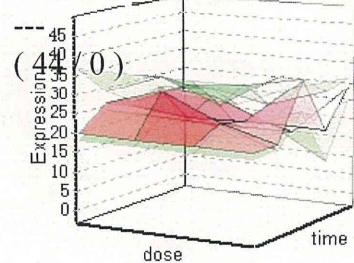
MLG801 TTG056-L



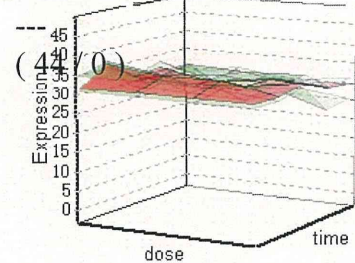
MLG801 TTG074-L



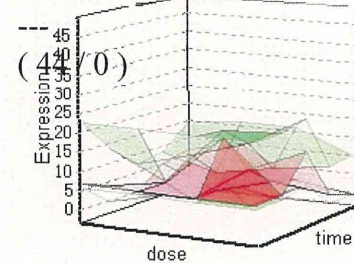
MLG801 TTG086-L



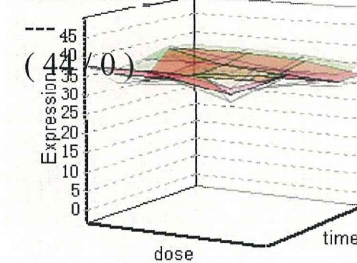
MLG801 TTG090-L



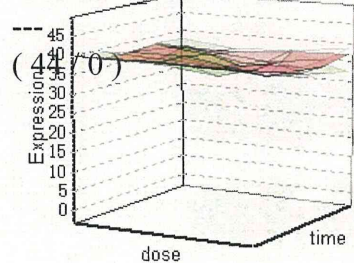
MLG801 TTG092-L



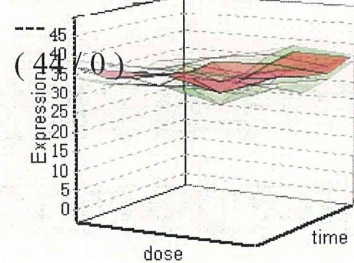
MLG801 TTG096-L



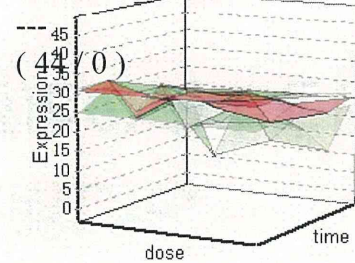
MLG801 TTG097-L



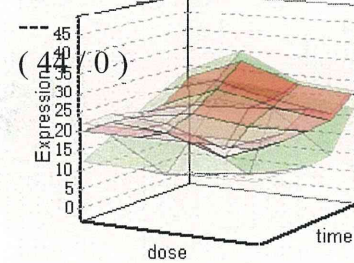
MLG801 TTG098-L



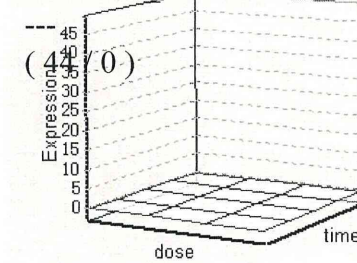
MLG801 TTG104-L



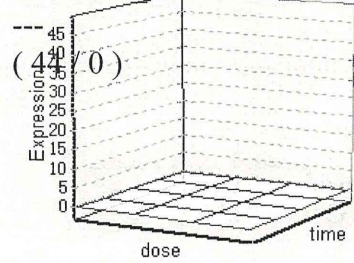
MLG801 TTG110-L



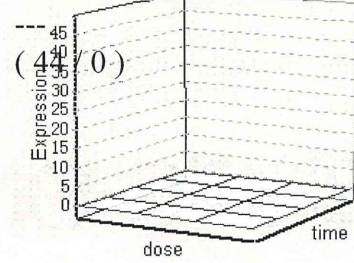
MLG801 TTG124-L



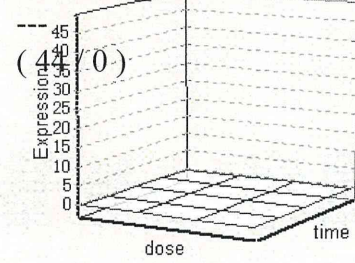
MLG801 TTG157-L



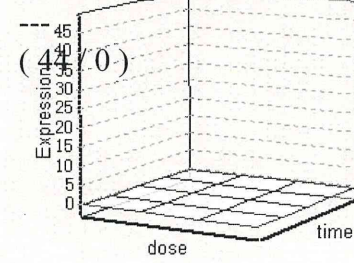
MLG801 TTG161-L



MLG801 TTG165-L



MLG801 TTG166-L



GSC(LYS)の推定値

TTG124以降で0と推定
されている

2.MLANG SpFの精度へ影響

- ・ 化合物分類の実施により、MLANG補正において、SpFの濃度が影響を与えていることが判明した。
- ・ 影響の原因を特定し、精度の向上を図る。

2.1.SpFの精度への影響の調査

- ・ **完全ブランクとGSCのみを計測したチップの間の差異を調査する。**
 - 大きな差異を発生しているプローブは、GSCの影響、又はデブリ/泡などと考えられる。
 - 大きな差異を発生しているプローブを特定し、次のいずれに該当するか確認する。
 - ・ 適切な補正処理が行われている可能性(従来からMLANG上の係数として存在する)
 - ・ デブリの可能性(同一座標近くの値分布を確認する)
 - ・ 泡の可能性(同一座標近辺の値分布を確認する)
 - ・ もっと短い部分一致の可能性
- ・ **もっと短い部分一致の情報の追加で多くのプローブを救える場合には、係数を作成する一致長さの閾値を短くすることを検討する**