at http://las.perkinelmer.com/content/snps/software.asp.

We chose to focus on the original two-dimensional data instead of angle data, and propose an alternative clustering method based on a mathematical model that includes a latent time variable and logarithmic transformed data. Optimization is based on a pair of one-dimensional normal mixture models having two components. The approach is rapid and accurate, and is therefore considered suitable for a high-throughput system.

2. Mathematical Model

We assume that there are two types of allele for each SNP, which can be denoted by A and B. Four clusters can be identified in the two-dimensional fluorescence data: 0) there is no DNA, 1) the genotype is homozygous AA, 2) the genotype is homozygous BB, and 3) the genotype is heterozygous AB. Let $\boldsymbol{x} = (x_A, x_B)'$ be an observed vector, where x_j is the fluorescence intensity for allele $j \in \{A, B\}$. Consider the following time-dependent response model:

$$\boldsymbol{x} = (\boldsymbol{v} + \boldsymbol{v}_0)t + \boldsymbol{u}, \tag{1}$$

where

$$\boldsymbol{v} = \begin{pmatrix} y_A z_A \\ y_B z_B \end{pmatrix}, \quad \boldsymbol{v}_0 = \begin{pmatrix} \bar{y}_A y_B e^{\lambda_A} \\ \bar{y}_B y_A e^{\lambda_B} \end{pmatrix} \quad \text{and} \quad \boldsymbol{u} = \begin{pmatrix} \bar{y}_A w_A \\ \bar{y}_B w_B \end{pmatrix}.$$

Here, $y_j \sim Bernoulli(p_j)$ is the latent binary response for allele $j \in \{A, B\}$, $y_j = 1$ if allele j is present, otherwise, $y_j = 0$, $\bar{y}_j = 1 - y_j$, $t \sim LN(0, \delta^2)$ is the latent response time, δ^2 can be zero if there is no correlation between heterozygous clusters, $z_j \sim LN(\mu_j, \tau_j^2)$ is the latent response intensity for the case in which allele j is present, $w_j \sim LN(\nu_j, \sigma_j^2)$ is the latent response intensity for the other cases, $\mu > \nu$, $\{y, t, z, w\}$ are mutually independent, and λ_j indicates the background (apparent) response that accounts for the gradient from a homozygous cluster to the heterozygous cluster.

We now turn attention to the parameter λ_j. From Model (1), data belonging to clusters 0 and 2 can be expressed as $(x_A, x_B) = (w_A, w_B)$ and $(w_A + e^{\lambda_A} t, z_B t)$, respectively. The difference between means on the x_A-axis is $e^{\lambda_A + \delta^2/2}$, which is empirically small, so restrict attention to the special case

$$\boldsymbol{v}_0 = 0. \tag{2}$$

Under condition (2), the logarithmic transformed observations, $\tilde{x}_j = \ln(x_j)$ and $\tilde{\boldsymbol{x}} = (\tilde{x}_A, \tilde{x}_B)'$, have the following marginal and joint distribution:

$$f_j(\tilde{x}_j) = p_j \phi^{(1)}(\tilde{x}_j \mid \mu_j, \tau_j^2 + \delta^2) + q_j \phi^{(1)}(\tilde{x}_j \mid \nu_j, \sigma_j^2) \quad \text{for} \quad j \in \{A, B\} \quad \text{and} \quad f(\tilde{\boldsymbol{x}}) = \sum_{l=0}^{3} \xi_l g_l(\tilde{\boldsymbol{x}}) \tag{3}$$

where

$$\begin{aligned}
&\xi_0 = q_A q_B, \quad g_0(\tilde{\boldsymbol{x}}) = \phi^{(1)}(\tilde{x}_A \mid \nu_A, \sigma_A^2)\phi^{(1)}(\tilde{x}_B \mid \nu_B, \sigma_B^2),\\
&\xi_1 = p_A q_B, \quad g_1(\tilde{\boldsymbol{x}}) = \phi^{(1)}(\tilde{x}_A \mid \mu_A, \tau_A^2 + \delta^2)\phi^{(1)}(\tilde{x}_B \mid \nu_B, \sigma_B^2),\\
&\xi_2 = q_A p_B, \quad g_2(\tilde{\boldsymbol{x}}) = \phi^{(1)}(\tilde{x}_A \mid \nu_B, \sigma_A^2)\phi^{(1)}(\tilde{x}_B \mid \mu_B, \tau_B^2 + \delta^2),\\
&\xi_3 = p_A p_B, \quad g_3(\tilde{\boldsymbol{x}}) = \phi^{(2)}(\tilde{\boldsymbol{x}} \mid \boldsymbol{\mu}, \Sigma),
\end{aligned}$$

with $p_j + q_j = 1$, $p_j \geq 0$, $q_j \geq 0$, $\boldsymbol{\mu} = (\mu_A, \mu_B)'$, $\Sigma = diag(\tau_A^2, \tau_B^2) + \delta^2 \mathbf{1}\mathbf{1}'$, and $\phi^{(k)}(\cdot \mid \boldsymbol{\theta}, \Omega)$ is the probability density function of the k-variate normal distribution with mean $\boldsymbol{\theta}$ and covariance matrix Ω. The transformed data can therefore be expressed as a normal mixture model (see McLachlan and Peel, 2000) under condition (2). Note that the homozygous cluster consists of independent variables under Model (3).

3. Estimation and Clustering

Suppose that the parameters of Model (3) are known. The observation can be clustered using Bayes rule (Wolfe, 1970). Let $l = l_A + 2 \cdot l_B$ for $l_j \in \{0, 1\}$ and $j \in \{A, B\}$ and w_l be the posterior probability that $\tilde{\boldsymbol{x}}$ belongs to Cluster l for $l = 0, \cdots, 3$, which is given by the following equation:

$$w_l(\tilde{\boldsymbol{x}}) = Pr(y_A = l_A \quad \text{and} \quad y_B = l_B \mid \tilde{\boldsymbol{x}}) = \frac{\xi_l g_l(\tilde{\boldsymbol{x}})}{f(\tilde{\boldsymbol{x}})}. \tag{4}$$

An individual is classified into cluster l_0, where $l_0 = \arg_{l=0,\cdots,3} \max w_l(\tilde{\boldsymbol{x}})$. The average of the maxima of these posterior probabilities is known as an allocation rate, which assesses the performance of the mixture approach to clustering (e.g. Basford and McLachlan (1985)).

When the parameters of Model (3) are unknown, they can be estimated by the following two steps: 1) estimate each marginal distribution using the EM algorithm (Dempster and Laird *et al.*, 1977; McLachlan and Krishnan, 1997), and 2) estimate the joint distribution and the unknown covariance δ.

Unfortunately, the clustering rule, as based on the posterior probability, can at times provide unreasonable clusters because the variance of Cluster 0 is much smaller than that of Cluster 1 or Cluster 2 in the marginal distribution, and a part of Cluster 0 might be mis-classified as a homozygous cluster. To overcome this difficulty, we consider a shift-logarithmic transformation, so that variances of the marginal distributions can be made equivalent in the following manner: 1) for a given value α, transform the data by $\tilde{x}^{(\alpha)} = \ln(x - \alpha)$, 2) fit the marginal distribution to the transformed data and estimate the variances $\widehat{\tau^2 + \delta^2}^{(\alpha)}$ and $\widehat{\sigma^2}^{(\alpha)}$, 3) repeat operations 1) and 2) until the optimized shift value α^* satisfies $\widehat{\tau_j^2 + \delta^2}^{(\alpha^*)} = \widehat{\sigma^2}^{(\alpha^*)}$, 4) fit the joint distribution to the transformed data $\tilde{\boldsymbol{x}}^{(\alpha^*)}$ where $\boldsymbol{\alpha}^* = (\alpha_A^*, \alpha_B^*)'$ and estimate the covariance, and 5) identify four clusters according to the fitted joint distribution. The existence of the optimized shift value α^* depends on the minimum value of the observations. Thus α^* can be formally written as follows:

$$\alpha^* = \frac{rE(w) - E(zt)}{r - 1}, \quad \text{where} \quad r^2 = \frac{Var(zt)}{Var(w)}.$$

Note that the estimates of the mixing proportions p_j, $j \in \{A, B\}$, might be close to one or zero, which implies that the potential total number of clusters is one or two. In such a case, the validity of the clustering result needs to be assessed not only statistically but also biologically.

4. Examples and Discussion

The Invader assay generates two-dimensional fluorescence intensities. We examined the performance of our method using three examples of SNP data, which were selected from a large number of SNPs available from the Japanese Foundation for Cancer Research (JFCR). All were obtained with informed consent.

Original two-dimensional data are plotted on the left side in Figures 1-3. The fluorescence intensity for allele A (x_A) is plotted on the horizontal axis, that for allele B (x_B) on the vertical axis. Figure 1 shows ideal data, in which the observed results are clearly split, response time (or rate) is thus sufficiently long, and response intensities appear to be saturated. On the other hand, the response time of the data in Figure 2 is shorter and the boundaries among clusters are

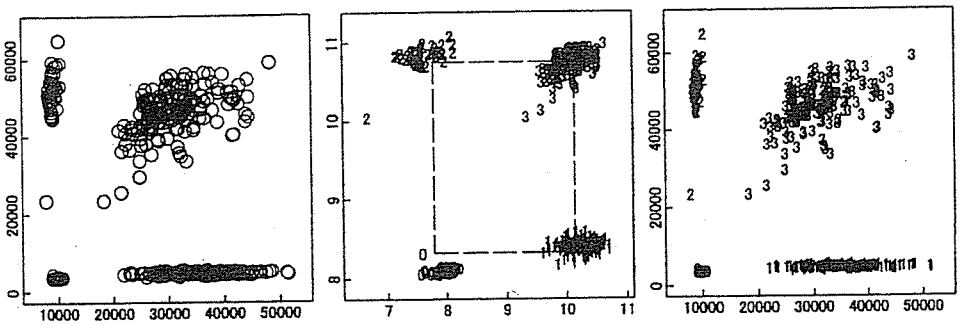

Fig. 1. Example data showing a long response time. Left: scatter plot of original fluorescence intensities. Center: logarithmic transformed observations and clustering results. Right: restored results. The allocation rate of the clustering result is 100.0%. Crosses represent individuals with low posterior probability.

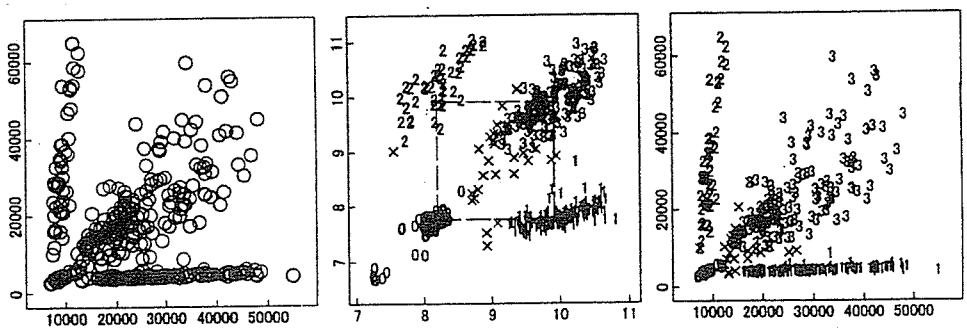

Fig. 2. Example data showing a short response time. Left: scatter plot of original fluorescence intensities. Center: logarithmic transformed observations and clustering results. Right: restored results. The allocation rate of the clustering result is 98.5%. Crosses represent individuals with low posterior probability.

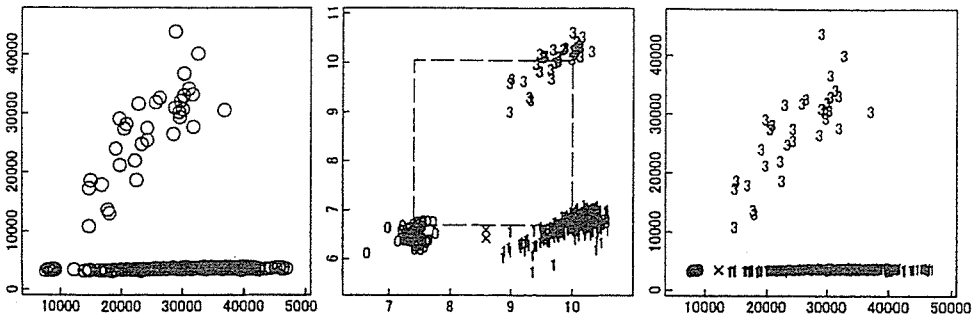

Fig. 3. Example data showing only three clusters. Left: scatter plot of original fluorescence intensities. Center: logarithmic transformed observations and clustering results. Right: restored results. The allocation rate of the clustering result is 99.9%. Crosses represent individuals with low posterior probability.

unclear. Figure 3 illustrates a case in which Cluster 2 is not observed, so there are only three clusters.

For each original data set, the optimized shift value α and its covariance matrix ware estimated. The logarithmic transformed data and original-scale data were then plotted at the center and right of each Figure, respectively. The plotted number expresses the cluster number determined by Bayes rule, and the crosses represent individuals with low posterior probability. The four estimated means are connected by a dotted line. Nine further examples of clustering results are shown in Figure 4. The allocation rate value for each clustering result is superimposed. The individuals with low posterior probability are not distinguished in Figure 4.

Our clustering method is based on Model (3) instead of Model (1). Therefore the estimators of the means, such as the mean of Cluster 0 at the center of Figure 1, are not unbiased. However, we make use of the following four advantages: 1) the joint distribution is expressed in an explicit form; 2) estimators of the marginal distributions can be used for estimation of the joint distribution, such that numerical calculations are easier and faster; 3) regardless of whether or not the real response time is long, four clusters can be properly identified because estimation is primarily conducted on the marginal distributions; and 4) even when there are only three clusters, the joint distribution can be estimated due to the mean structure of the rectangle.

If Condition (2) does not hold, cluster 1 or 2 will be shifted towards the inside of the rectangle, the diagonal of which is defined by clusters 0 and 3 (see the vertical axis of the left top example in Figure 4). In such cases, the marginal distribution might have three components so our proposed method based on two components will not work efficiently. Therefore testing or selecting the number of components on the marginal distribution is useful, as described , for example, by Nakamura and Konishi (1998).

Although our purpose is to classify individuals properly, better estimators of the means and covariance matrices in the normal mixture model might be required for drawing the contour of

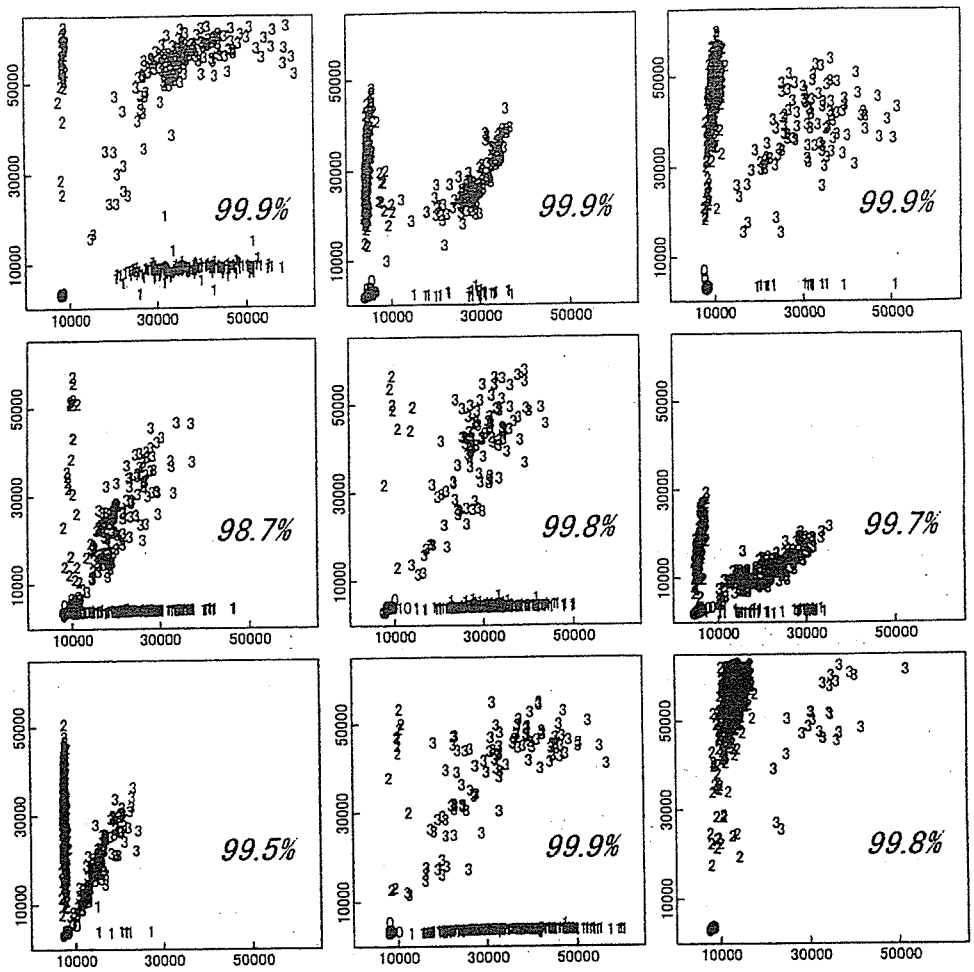

Fig. 4. Nine further examples of clustering results. The allocation rate value for each clustering result is superimposed. Note that individuals with low posterior probability are not distinguished.

the probability density function or finding outliers located in the tail. If the allocation rate is close to one, weighted means and covariance matrices derived from posterior probabilities are useful, which provide almost the same clustering result. Otherwise, the clustering result of model (3) itself might not be reliable and we need to check the goodness of the model fit, the number of potential clusters in the observation and the validity of the experiment.

REFERENCES

Basford, K. E. and McLachlan, G. J. (1985). Estimation of allocation rates in a cluster analysis context. *Journal of the American Statistical Association* **80**, 286-293.

Dempster, A. P., Laird, N, M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, Series B **39**, 1-38.

Fujisawa, H., Eguchi, S., Ushijima, M., Miyata, S., Miki, Y., Muto, T. and Matsuura, M. (2003). Genotyping of single nucleotide polymorphism using model-based clustering. *Bioinformatics*, **20**, 718 - 726.

McLachlan, G. J. and Krishnan, T. (1997). *The EM Algorithm and Extensions*. New York: John Wiley.

McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. Wiley, New York.

Nakamura, N. and Konishi, S. (1998). Estimation of number of components for multivariate normal mixture models based on information criteria. *Japanese Journal of Applied Statistics*, **27**, 165-180, in Japanese.

Oliver, M., Chuang, L.-M., Chang, M.-S., Chen, Y.-T., Pei, D., Ranade, K., de Witte, A., Allen, J., Tran, N., Curb, D., Pratt, R., Neefs, H., de Arruda Indig, M., Law, S., Neri, B., Wang, L. and Cox, D. R. (2002). High-throughput genotyping of single nucleotide polymorphisms using new biplex invader technology. *Nucleic Acids Res.* **30**, e53.

Ranade, K., Chang, M.-S., Ting, C.-T., Pei, D., Hsiao, C.-F., Oliver, M., Pesich, R., Hebert, J., Chen, Y.-D., Dzau, V. J., Curb, D., Olshen, R., Risch, N., Cox, D. R., and Botstein, D. (2001). High-throughput genotyping with single nucleotide polymorphisms. *Genome Res.* **11**, 1262-1268.

Riva, A. and Kohane, I. S. (2002). SNPper: retrieval and analysis of human SNPs. *Bioinformatics* **18**, 1681-1685.

van den Oord, E. J. C. G., Jiang, Y., Riley, B. P., Kendler, K. S. and Chen X. (2003). FP-TDI SNP Scoring by Manual and Statistical Procedures: A Study of Error Rates and Types. *BioTechniques* **34**, 610-624.

Wolfe, J. H. (1970). Pattern clustering by multivariate mixture analysis. *Multivariate Behavioral Res.* **5**, 329-350.

Adv. & Appl. in Stat. 4(2) (2004), 223-231

CLUSTERING METHOD BY CONNECTED NEIGHBORHOODS AND ITS APPLICATION

KENICHI SATOH[1], HIROKAZU YANAGIHARA[2] and MEGU OHTAKI[1]

[1]Department of Environmetrics and Biometrics
Research Institute of Radiation Biology and Medicine
Hiroshima University, 1-2-3 Kasumi, Minami-ku
Hiroshima 734-8553, Japan

[2]Graduate School of Systems and Information Engineering
University of Tsukuba, 1-1-1 Tennodai, Tsukuba
Ibaraki 305-8573, Japan

Abstract

A clustering method by connected neighborhoods is proposed. Some theoretical properties are derived and the concept is applied to a weighted estimation in order to highlight the major structure of the distribution. Numerical study showed good performance of the estimation for the case when there exist some outliers under the normal distribution and the simple regression model.

1. Introduction

Clustering method is powerful when samples are not homogeneous. It might find some classes among samples and those would be helpful for further progressive analysis. In this paper, a neighborhood around each sample is considered and a relation between two samples is defined according to those neighborhoods. In Section 2, we propose a new

2000 Mathematics Subject Classification: 62H30, 62F35, 62J05.

Key words and phrases: clustering, neighborhood, weighted estimation, outliers.

Received June 22, 2004

clustering method by the relation and discuss some theoretical properties, which shows it to be an extension of the nearest neighbor method in the linkage way (see, Florek et al. [1], Sneath [3] and Johnson [2]). In Section 3, we apply the suggested clustering method to the estimation of major structure of the distribution by the weighted estimation, that is, based on the class sizes. Some illustrative numerical studies demonstrate the good performance of our methodology for the case when samples include some outliers.

2. Theories on Connected Neighborhoods

Let S be a set of n samples $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ and $U(\mathbf{x}_i)$ be a neighborhood around the sample $\mathbf{x}_i$ and assume that $\mathbf{x}_i \in U(\mathbf{x}_i)$ for $i = 1, ..., n$. We define a relation between two samples in the following way.

Definition 1. Consider any two samples $\mathbf{x}_i, \mathbf{x}_j \in S$. If there exist samples with indexes $l_1, ..., l_m$ such that $U(\mathbf{x}_i) \cap U(\mathbf{x}_{l_1}) \neq \varnothing$, $U(\mathbf{x}_{l_1}) \cap U(\mathbf{x}_{l_2}) \neq \varnothing$, ..., $U(\mathbf{x}_{l_m}) \cap U(\mathbf{x}_j) \neq \varnothing$, then the relation between two samples is expressed by $\mathbf{x}_i \sim \mathbf{x}_j$.

Definition 1 implies that there exists the pass which is filled in the gap between $\mathbf{x}_i$ and $\mathbf{x}_j$ by connected neighborhoods. Note that it also holds that $\mathbf{x}_{l_i} \sim \mathbf{x}_{l_j}$ for any $l_i, l_j \in \{i, j, l_1, ..., l_m\}$. Then we give the following Theorem 1 and Algorithm 1 without proofs.

Theorem 1. *The relation defined in Definition* 1 *is an equivalence relation, i.e., it satisfies the following three conditions*: (1) $\mathbf{x}_i \sim \mathbf{x}_i$, (2) *if* $\mathbf{x}_i \sim \mathbf{x}_j$, *then* $\mathbf{x}_j \sim \mathbf{x}_i$, (3) *if* $\mathbf{x}_i \sim \mathbf{x}_j$ *and* $\mathbf{x}_j \sim \mathbf{x}_k$, *then* $\mathbf{x}_i \sim \mathbf{x}_k$. *Hence, it creates the equivalence classes* $C(\mathbf{x}_i) = \{\mathbf{x}_j \mid \mathbf{x}_i \sim \mathbf{x}_j, \mathbf{x}_j \in S\}$ *for* $i = 1, \cdots, n$.

For given $\mathbf{x}_i \in S$, the following recursive algorithm can find all samples belonging to $C(\mathbf{x}_i)$.

Algorithm 1. (0) all samples have the "OUT" status except $\mathbf{x}_i$ that has the "IN" status and let $k = i$, (1) make an index set $I_k\{j \mid$ status of

$\mathbf{x}_j$ is OUT, $\mathbf{x}_j \in C(\mathbf{x}_k)$, $\mathbf{x}_j \in S\}$, (2) unless the index set is null, change the status of $\mathbf{x}_j$ to IN and execute step 1 with $k = j$ for all $j \in I_k$, (3) stop. Finally, the sample set whose samples have IN status is equal to $C(\mathbf{x}_i)$.

From Theorem 1, S can be clustered by some exclusive classes.

Corollary 1. *For given neighborhoods* $U_1, \ldots, U_n$, *there exist* k *different samples with indexes* $l_1, \ldots, l_k$ *such that* (1) $S = C(\mathbf{x}_{l_1}) \cup \cdots \cup C(\mathbf{x}_{l_k})$, (2) $C(\mathbf{x}_{l_i}) \cap C(\mathbf{x}_{l_j}) = \varnothing$ *for* $(\mathbf{x}_{l_i} \neq \mathbf{x}_{l_j})$. *So, it holds that* $n = \sum_{i=1}^{k} n_{l_i}$, *where* n_{l_i} *is the size of the class* $C(\mathbf{x}_{l_i})$, *i.e.,* $n_{l_i} = |C(\mathbf{x}_{l_i})|$.

The generated exclusive classes strongly depend on the size and shape of neighborhoods. For example, if we consider a sphere neighborhood defined by $U_\varepsilon(\mathbf{x}_i) = \{\mathbf{x} \in \Re \mid d(\mathbf{x}_i, \mathbf{x}) \leq \varepsilon^2\}$, where $d(\cdot, \cdot)$ is a metric on $\Re^p \times \Re^p$, then our method and nearest neighbor method are equivalent in the linkage way (see, Florek et al. [1], Sneath [3] and Johnson [2]). But if we consider a cube neighborhood, those methods are completely different. In the sense, our method is an extension of the nearest neighbor method.

3. Application

In this section, we consider a kind of outlier and discuss the weighted estimation method according to the clustering method based on connected neighborhoods. Let $S = S^{(M)} \cup S^{(O)}$, where $S^{(M)} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ is a set of samples from the major part of probability structure and $S^{(O)} = \{\mathbf{x}_{m+1}, \ldots, \mathbf{x}_n\}$ from the outlier, A be a $p \times p$ positive definite symmetric matrix and $d_A(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})' A^{-1} (\mathbf{x} - \mathbf{y})$ be a distance function on $\Re^p \times \Re^p$. We assume that the number of outliers: $n - m$ is unknown but nuisance parameter.

Consider the ε-neighborhood around $\mathbf{x}_i$ given by $U_\varepsilon(\mathbf{x}_i) = \{\mathbf{x} \mid d_A(\mathbf{x}_i, \mathbf{x}) \leq \varepsilon^2, \mathbf{x} \in \Re^p\}$. Then we can easily evaluate the relative

distances of samples among $S^{(M)}$ and/or $S^{(O)}$ in the sense of an expectation.

Theorem 2. *Let* $\mathbf{x}_i$*'s be independent random variables,* $\mathbf{x}_i \in S^{(M)}$ *have mean* μ*, covariance matrix* $\Sigma^{(M)}$ *and* $\mathbf{x}_j \in S^{(O)}$ *have the same mean and different covariance matrix* $\Sigma^{(O)}$ *and assume that* $\Sigma^{(O)} - \Sigma^{(M)}$ *is a positive definite matrix. Then it holds that* $E\{d_A(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})\} \leq E\{d_A(\mathbf{x}_{i_1}, \mathbf{x}_{j_1})\} \leq E\{d_A(\mathbf{x}_{j_1}, \mathbf{x}_{j_2})\}$ *for any* $\mathbf{x}_{i_1}, \mathbf{x}_{i_2} \in S^{(M)}$ *and* $\mathbf{x}_{j_1}, \mathbf{x}_{j_2} \in S^{(O)}$.

From above theorems, it can be concluded that the samples in $S^{(O)}$ are dispersed more widely than those in $S^{(M)}$. Hence, the size of class $C(\mathbf{x}_i)$ for $\mathbf{x}_i \in S^{(O)}$ should be generally less than that of $C(\mathbf{x}_j)$ for $\mathbf{x}_j \in S^{(M)}$ and the information is useful for the estimation on the major structure. In this section, we consider applying the size of classes to a weighted estimation on the variance matrix of the multivariate normal distribution and the coefficients of the simple linear regression model, and evaluate the performance of the methodology numerically.

First, we considered the bi-variate normal distribution with a zero mean vector and the diagonal covariance matrix $\Sigma = diag(\sigma^2, \sigma^2)$, where $\sigma = 0.10$ for $S^{(M)}$ and $\sigma = 0.10, 0.12, ..., 0.50$ for $S^{(O)}$, two sample sizes $n = 20$ and 200 were examined, which corresponded to a small sample case and a large one, respectively, and the sample sizes from outlier part were 40% of n. Figure 1 is an example of a set of observations from bi-variate normal distribution with mean zero and the diagonal covariance matrix $\Sigma = diag(\sigma^2, \sigma^2)$, where $\sigma = 0.10$ for $S^{(M)}$, $\sigma = 0.30$ for $S^{(O)}$ that is the almost middle value in the simulation and the sample size $n = 200$. The dense center of the observations was mainly generated from $S^{(M)}$.

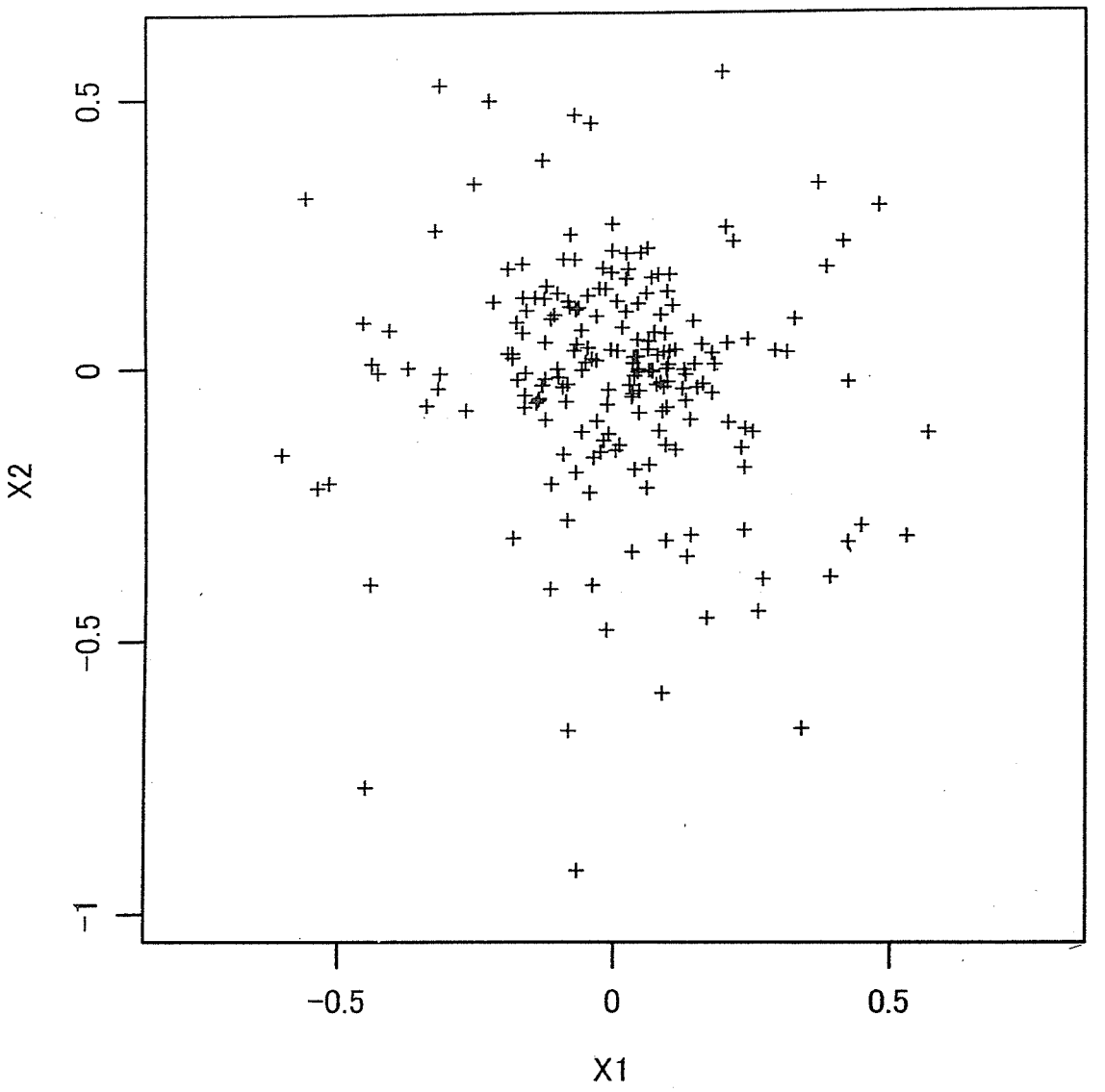

Figure 1. The scatter plot of a set of observations from bi-variate normal distribution with mean zero and the diagonal covariance matrix $\Sigma = diag(\sigma^2, \sigma^2)$, where $\sigma = 0.10$ for $S^{(M)}$, $\sigma = 0.30$ for $S^{(O)}$ and the sample size $n = 200$.

We examined 10,000 repetitions for the following steps and compared the mean square error of the weighted estimators with that of maximum likelihood estimator, (1) two dimensional n observations are generated

from the true model, (2) from the observations the minimum value a_j and the maximum value b_j for jth dimension are found, (3) let A be a diagonal matrix with (c_1^2, c_2^2), where $c_j = (b_j - a_j)$, $j = 1$ and 2, (4) count the ith sample class size n_i for $i = 1, \cdots, n$, (5) if the maximum class size is larger than a half of n, then go to next step, otherwise, consider a larger ε and repeat steps 4-5, (6) three estimators of the covariance matrix are given by $\hat{\Sigma}_j = \underset{i=1}{P_n} w_j(n_i)\mathbf{x}_i\mathbf{x}_i' - \tilde{\mathbf{x}}\tilde{\mathbf{x}}'$, $j = 1, 2$ and 3, where $\tilde{\mathbf{x}} = \underset{i=1}{P_n} w_j(n_i)\mathbf{x}_i$, $w_j(n_i) = v_j(n_i)/\tilde{v}_j$, $\tilde{v}_j = \underset{i=1}{P_n} v_j(n_i)$; here $v_1(n_i) = n_i^{-1}$, $v_2(n_i) = n_i^{-0.5}$ and $v_3(n_i) = \log(1 + n_i)^{-1}$. Note that the maximum likelihood estimator can be described by the uniform weight function $v_0(n_i) = 1/n$ for $i = 1, \cdots, n$.

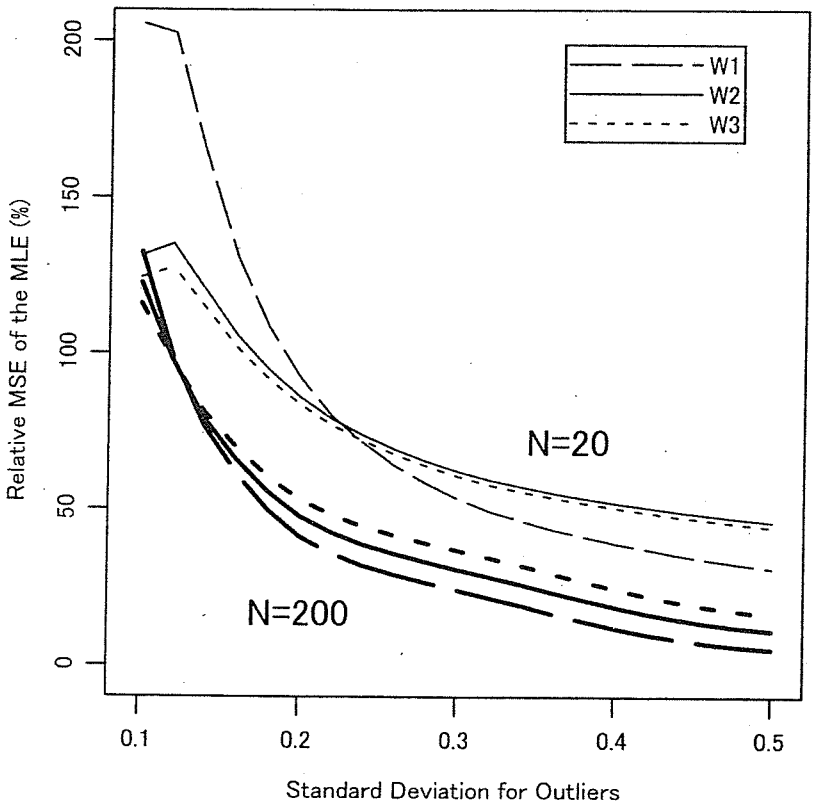

Figure 2. The relative mean square error of the weighted estimators by the connected neighborhood method with that of maximum likelihood estimator on the covariance matrix under bi-variate normal distribution. The fine lines are for the small sample: $n = 20$ and the heavy lines are for the large sample case: $n = 200$.

The weighted estimators were evaluated by the RMSE: relative mean square error to that of maximum likelihood estimator. Figure 2 showed that the RMSE of the weighted estimators decreases as the standard deviation for outliers are larger and it nearly attains zero for the case when the sample size is large, $n = 200$. On the other hand, when the discrepancy between the major structure and that of outliers is small, i.e., the deviation for outliers is small, the estimator given by w_1 took the largest RMSE, suggesting that heavy weight function does not work well without outliers.

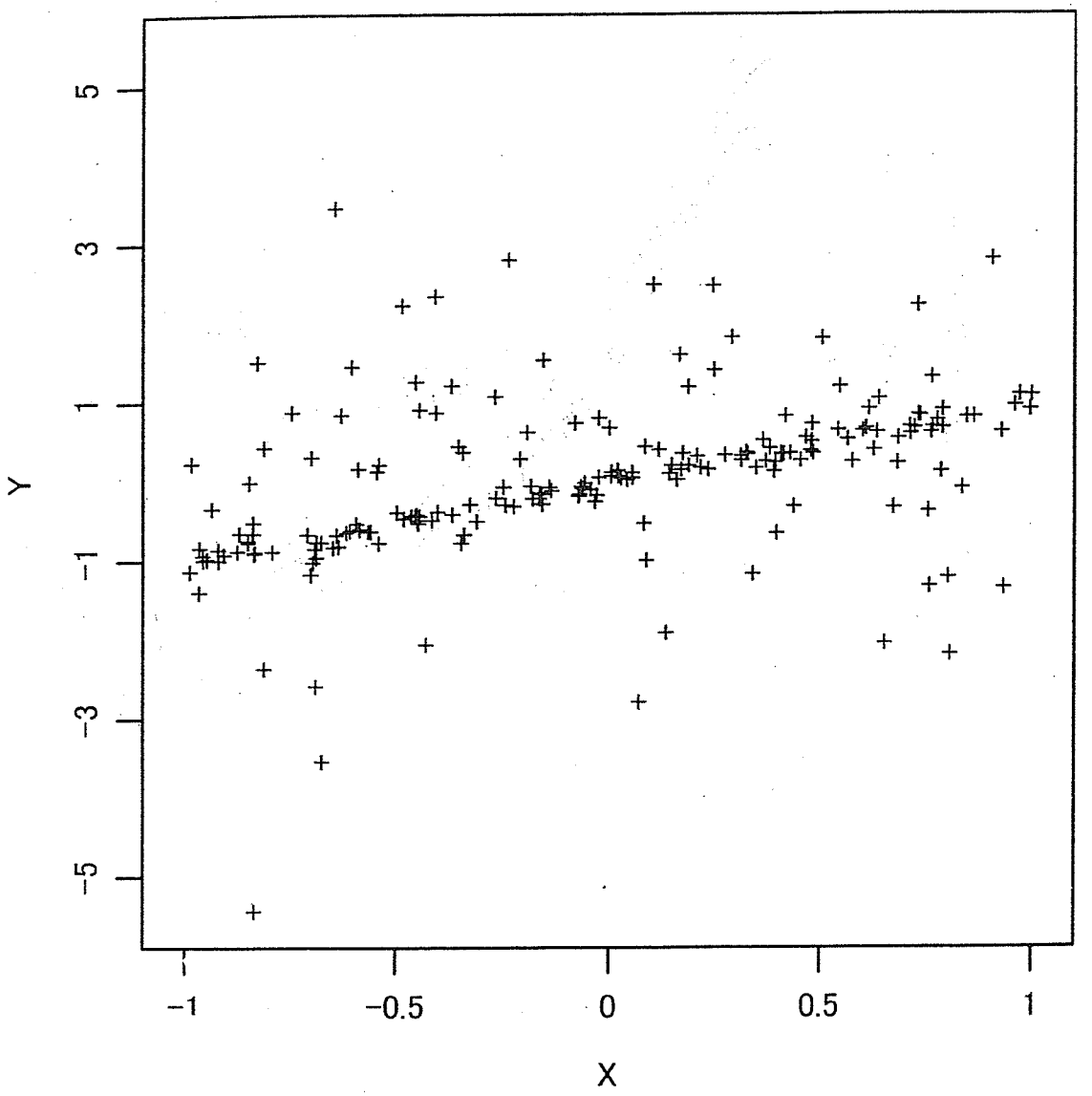

Figure 3. The scatter plot of a set of observations from the simple linear regression model with error standard deviation $\sigma = 0.10$ for $S^{(M)}$, $\sigma = 1.50$ for $S^{(O)}$ and the sample size $n = 200$.

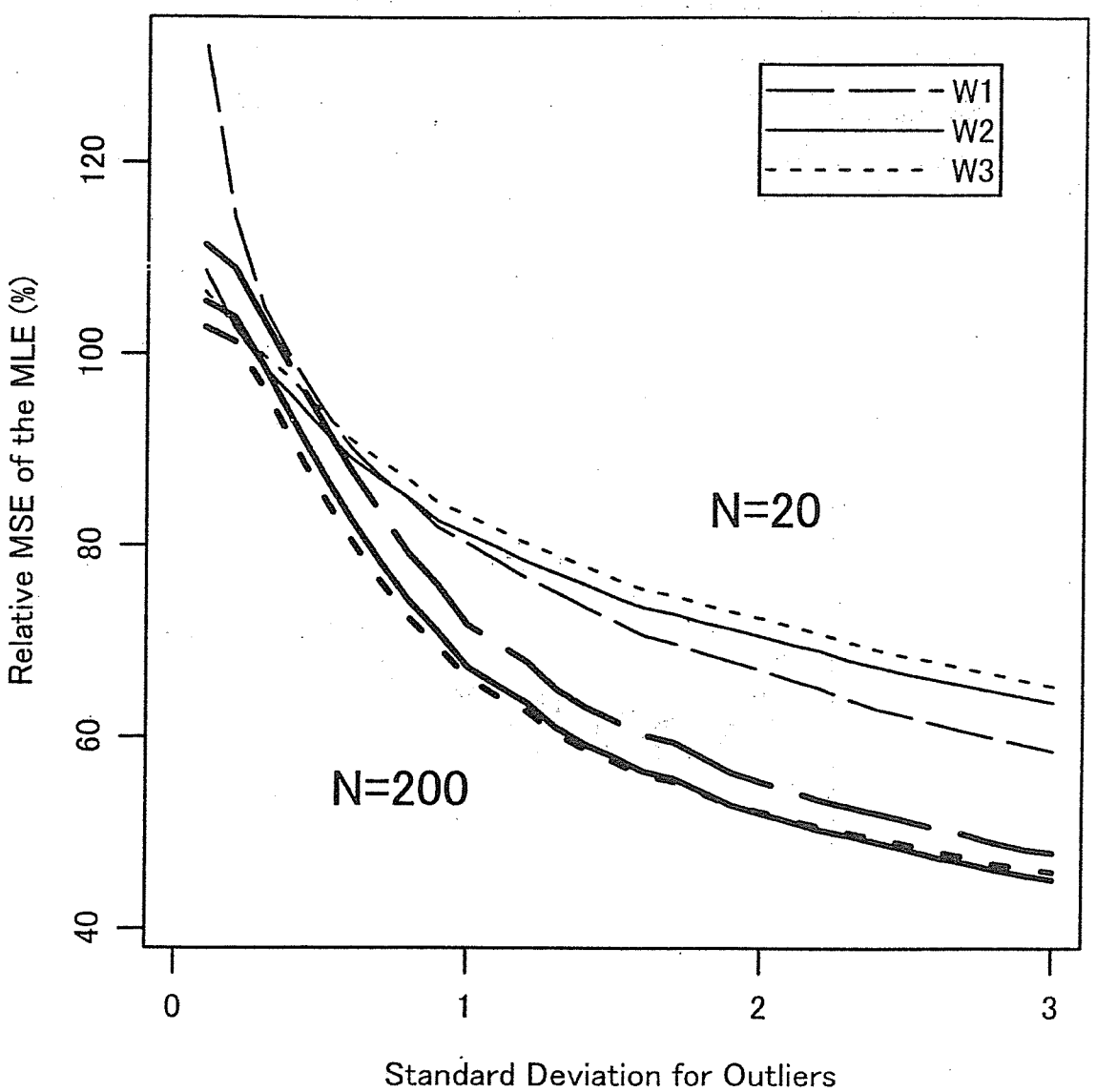

Figure 4. The relative mean square error of the weighted estimators by the connected neighborhood method with that of maximum likelihood estimator on the coefficients vector under the simple linear regression model. The fine lines are for the small sample: $n = 20$ and the heavy lines are for the large sample case: $n = 200$.

Secondly, the weighted estimation of the coefficients on the simple linear regression model, $y_i = \beta' \mathbf{z}_i + e_i$, $e_i \sim N(0, \sigma^2)$, $i = 1, ..., n$ was considered, where y_i is ith response variable, β is an unknown coefficients parameter vector, $\mathbf{z}_i = (1, x_i)'$, x_i is an explanatory variable, the error term has $\sigma = 0.1$ for $S^{(M)}$ and $\sigma = 0.1, 0.2, ..., 3.0$ for $S^{(O)}$. Figure 3 is the scatter plot of a set of observations from the simple linear regression model with error standard deviation $\sigma = 0.1$ for $S^{(M)}$, $\sigma = 1.5$ for $S^{(O)}$ and the sample size $n = 200$. Although $\mathbf{x}_i = (x_i, y_i)'$ does not hold in Theorem 2, it might hold for the those samples whose explanatory variables are close enough. The weighted estimator of the coefficient vector is given by $\hat{\beta}_j = (X'W_jX)^{-1}X'W_j\mathbf{y}$ for $j = 1, 2$ and 3, where $X = (\mathbf{z}_1 \cdots \mathbf{z}_n)'$ is the design matrix, $\mathbf{y} = (y_1, \cdots, y_n)'$ and $W_j = diag\{w_j(n_1), \cdots, w_j(n_n)\}$. We also examined 10,000 repetitions for the above steps 1-6 and compared the mean square error of the weighted estimators with that of maximum likelihood estimator. Although Figure 4 represented the similar trend to the result under the bi-variate normal distribution, the convergence of the RMSE was very slow and it attained at 50% even when the error standard deviation for $S^{(O)}$ was 30 times as large as that of $S^{(M)}$.

In conclusion, our suggested method is helpful to grasp the major structure for the case when there are some outliers and the lighter weight functions w_j, $j = 2$ and 3 are better in the sense that those are stable even when there are no outliers.

References

[1] K. Florek, J. Lukaszewiez, J. Perkal, H. Steinhaus and S. Zubrzchi, Sur la liason et la division des points d'un ensemble fini, Colloq. Math. 2 (1951), 282-285.

[2] S. C. Johnson, Hierarchial clustering schemes, Psychometrika 32 (1967), 241-254.

[3] P. H. A. Sneath, The application of computers to taxonomy, J. Gen. Microbiol. 17 (1957), 201-226.

0021-972X/04/$15.00/0
Printed in U.S.A.

The Journal of Clinical Endocrinology & Metabolism 89(8):4113–4118

doi: 10.1210/jc.2004-0069

Nur-Related Factor 1 and Nerve Growth Factor-Induced Clone B in Human Adrenal Cortex and Its Disorders

LIANGYING LU, TAKASHI SUZUKI, YOSUKE YOSHIKAWA, OSAMU MURAKAMI, YASUHIRO MIKI, TAKUYA MORIYA, MARY H. BASSETT, WILLIAM E. RAINEY, YUTAKA HAYASHI, AND HIRONOBU SASANO

Departments of Pathology (L.L., T.S., Y.Y., Y.M., T.M., H.S.), Pediatric Surgery (L.L., Y.H.), and Second Department of Internal Medicine (O.M.), Tohoku University School of Medicine, 980-8575 Sendai, Japan; and Division of Reproductive Endocrinology (M.H.B., W.E.R.), University of Texas Southwestern Medical Center, Dallas, Texas 75390-9032

Nerve growth factor-induced clone B (NGFI-B; NR4A1) and Nur-related factor 1 (Nurr1; NR4A2) are members of NGFI-B family of orphan receptors. We recently demonstrated induction of CYP11B2 (aldosterone synthase) by Nurr1 and NGFI-B, suggesting possible important roles of these transcriptional factors in the regulation of adrenocortical steroidogenesis. Therefore, we immunolocalized Nurr1 and NGFI-B in various human adrenal specimens to study their biological significance. In nonpathological adrenal glands (n = 25), Nurr1 and NGFI-B immunoreactivities were detected at high levels in the fetal definitive zone or postnatal zona glomerulosa. NGFI-B immunoreactivity was increased according to development in the zona fasciculata, reaching a level similar to that in the zona glomerulosa in adult adrenal cortex. In adrenocortical neoplasms (n = 44), Nurr1 immunoreactivity was higher in aldosteronoma than in Cushing's adenoma or adrenocortical carcinoma. NGFI-B immunoreactivity was also higher in aldosteronoma than in adrenocortical carcinoma, but was not significantly different among the types of adenoma. Both Nurr1 and NGFI-B mRNA expressions were correlated with their immunoreactivities in adrenocortical neoplasms (n = 23), and mRNA expression of Nurr1 was significantly ($P < 0.0001$) associated with that of CYP11B2. These results suggest that the expression of Nurr1 and NGFI-B plays an important role in human adrenal cortex and its neoplasms, including possible regulation of steroidogenesis. (*J Clin Endocrinol Metab* 89: 4113–4118, 2004)

HUMAN ADRENAL CORTEX is composed of three distinct zones, *i.e.* the zonae glomerulosa, fasciculata, and reticularis. These three zones produce distinct steroid hormones, such as aldosterone in the zona glomerulosa, cortisol in the zona fasciculata, and dehydroepiandrosterone and dehydroepiandrosterone sulfate in the zona reticularis (1). This functional zonation results from the zone-specific expression of steroidogenic enzymes (2). It is also well known that adrenocortical neoplasms excessively produce various corticosteroids and are generally associated with an abnormal expression of steroidogenic enzymes (3–5). Therefore, it is very important to examine the possible regulation of adrenocortical steroidogenesis to obtain a better understanding of functions of the human adrenal cortex and its disorders.

Nerve growth factor-induced clone B (NGFI-B; NR4A1) and Nurr1 (Nur-related factor 1; NR4A2) belong to a NGFI-B family of nuclear hormone receptors as well as neuron-derived orphan receptor 1 (NR4A3) (6). These nuclear receptors activate transcription by binding to the NGFI-B-responsive elements (NBREs) located in the promoter region of target genes (7, 8) and regulate various cellular functions, such as the differentiation of neural cells (9, 10), the apoptosis of T lymphocytes in the thymus (11), and the modulation of retinoic acid signal transduction (12). The expression of Nurr1 and NGFI-B has been previously detected in murine adrenal glands (13). Very recently, we demonstrated that the human CYP11B2 (aldosterone synthase) gene, which is a key enzyme of aldosterone production, contains NBRE in the promoter region, and its expression was markedly induced by Nurr1 or NGFI-B (14). These *in vitro* data suggest important roles for Nurr1 and NGFI-B in the human adrenal cortex, including the regulation of steroidogenesis. However, a detailed examination of the expression of these nuclear receptors has not been reported in the human adrenal gland and its disorders. Therefore, after the previous *in vitro* study (14), we immunolocalized Nurr1 and NGFI-B in nonpathological and pathological specimens of human adrenal cortex. In addition, we examined mRNA expression of Nurr1 and NGFI-B in adrenocortical neoplasms using real-time RT-PCR and examined the correlation with CYP11B2 mRNA expression.

Abbreviations: GAPDH, Glyceraldehyde-3-phosphate dehydrogenase; 3βHSD2, 3β-hydroxysteroid dehydrogenase type 2; NBRE, nerve growth factor-induced clone B-responsive element; NGFI-B, nerve growth factor-induced clone B; Nurr1, Nur-related factor 1.

JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.

Materials and Methods

Human adrenal specimens

Sixty-nine human adrenal specimens were examined in this study. Twenty-five specimens of nonpathological adrenal glands were obtained from autopsy files (11–36 wk gestation and 1 d to 68 yr of age) from Tohoku University Hospital (Sendai, Japan). Forty-four cases of adrenocortical tumors (14 aldosteromas, 10 Cushing's adenomas, 10 nonfunctioning adenomas with no clinical hormonal abnormalities, and 10 adrenocortical carcinomas) were retrieved from the surgical pathology files of Tohoku University Hospital. Adrenocortical carcinomas were histologically diagnosed based on the criteria of Weiss (15). These

specimens were fixed in 10% formalin for 24–48 h at room temperature and embedded in paraffin wax.

Twenty-three cases of adrenocortical neoplasms were also available for real-time RT-PCR analysis (eight aldosteronomas, six Cushing's adenomas, six nonfunctioning adenomas, and three adrenocortical carcinomas). Specimens for RNA isolation were snap-frozen and stored at −80 C.

Research protocols for this study were approved by the ethics committee at Tohoku University School of Medicine.

Immunohistochemistry

Rabbit polyclonal antibodies for Nurr1 (sc-991) and NGFI-B (1600045) were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA), and Geneka Biotechnology (Montréal, Canada), respectively. Utilization of these antibodies for immunohistochemistry has been reported previously (14).

Immunohistochemical analysis was performed employing the streptavidin-biotin amplification method using a Histofine Kit (Nichirei, Tokyo, Japan). Antigen retrieval was performed by heating the slides in an autoclave at 120 C for 5 min in citric acid buffer (2 mM citric acid and 9 mM trisodium citrate dehydrate, pH 6.0). The dilutions of the primary antibodies used in this study were: Nurr1, 1:250; and NGFI-B, 1:200. The antigen-antibody complex was visualized with 3,3′-diaminobenzidine solution [1 mM 3,3′-diaminobenzidine, 50 mM Tris-HCl buffer (pH 7.6), and 0.006% H_2O_2] and counterstained with hematoxylin. Immunohistochemical preabsorption tests for Nurr1 and NGFI-B were performed for negative controls of immunohistochemistry. Normal rabbit IgG was also used in place of the primary antibodies as a negative control.

Evaluation of immunoreactivity

After completely reviewing immunohistochemical sections, relative immunoreactivity for Nurr1 and NGFI-B in each zone of adrenocortex was evaluated by an H scoring system, as described by McCarty *et al.* (16) with some modifications (17). Briefly, adrenocortical cells were counted in each zone, and H-scores were subsequently generated by adding together 2 × the percentage of strongly stained nuclei, 1 × the percentage of weakly stained nuclei, and 0 × the percentage of negative nuclei, giving a possible range of 0–200. The H scores were independently and blindly evaluated by three of the authors (T.S., T.M., and H.S.) to obtain immunohistochemical data objectively, and the mean of the three values was used for analysis. The adrenals were classified into the following age groups in this study: 11–36wk gestation (n = 5), 1 d to 5 months of age (n = 4), 11 months to 8 yr of age (n = 5), 10–18 yr of age (n = 5), and 27–68 yr of age (n = 6). The relative immunoreactivity of tumor cells was also evaluated by H-scoring system described above. Statistical significance was evaluated using a Bonferroni test, and $P < 0.05$ was considered significant.

Real-time RT-PCR

Total RNA was carefully extracted from 23 specimens of adrenocortical neoplasms with guanidinium thiocyanate, followed by ultracentrifugation in cesium chloride. An RT kit (SuperScript II Preamplification System, Invitrogen Life Technologies, Inc., Grand Island, NY) was used in the synthesis of cDNA.

The Light Cycler System (Roche, Mannheim, Germany) was used to semiquantify the mRNA levels of Nurr1, NGFI-B, and CYP11B2 in 22 adrenocortical neoplasms by real-time PCR (18). Settings for the PCR thermal profile were: initial denaturation at 95 C for 1 min, followed by 40 amplification cycles of 95 C for 0 sec, annealing at 66 C (Nurr1, NGFI-B, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH)) or 60 C (CYP11B2) for 15 sec, and elongation at 72 C for 15 sec. The primer sequences used in this study are as follows: Nurr1: forward, 5′-AACCCTGACTATCAAATGAGTG-3′; reverse, 5′-CAATGCAGGAGAAGGCAGAAAT-3′ (19); NGFI-B, forward, 5′-TCTGCTCAGGCCTGGTGCTAC-3′; reverse, 5′-GGCACCAAGTCCTCCAGCTTG-3′ (20); CYP11B2: forward, 5′-TCCTGCTCTTCTTGCATCTGG-3′; reverse, 5′-TTTGCCCTGCAAATGGTTG-3′ (21); and GAPDH: forward, 5′-TGAACGG-

TABLE 1A. Relative immunoreactivity of Nurr1 in nonpathological human adrenal cortex

Age group (no. of cases)	Fetal type		Adult type		
	Definitive zone	Fetal zone	Glomerulosa	Fasciculata	Reticularis
Fetus					
11–36 wk gestation (n = 5)	151 ± 19.8	26.0 ± 4.74[a]			
After birth					
1 d–5 months (n = 4)	153 ± 19.0	28.7 ± 9.82[a]			
11 months–8 yr (n = 5)			132 ± 8.95	58.8 ± 16.5[b]	15.3 ± 7.76[b,c]
10–18 yr (n = 5)			150 ± 14.8	69.8 ± 17.9[b]	7.20 ± 3.43[b,c]
27–68 yr (n = 6)			147 ± 16.9	77.3 ± 13.2[b]	13.8 ± 6.65[b,c]

The relative immunoreactivity of each zone in the adrenal cortex was evaluated by the H scoring system (0–200). Data are the mean ± SEM.
[a] $P < 0.001$ *vs.* definitive zone.
[b] $P < 0.001$ *vs.* zona glomerulosa.
[c] $P < 0.001$ *vs.* zona fasciculata.

TABLE 1B. Relative immunoreactivity of NGFI-B in nonpathological human adrenal cortex

Age group (no. of cases)	Fetal type		Adult type		
	Definitive zone	Fetal zone	Glomerulosa	Fasciculata	Reticularis
Fetus					
11–36 wk gestation (n = 5)	122 ± 12.6	24.2 ± 3.61[a]			
After birth					
1 d–5 months (n = 4)	125 ± 24.3	20.8 ± 5.83[a]			
11 months–8 yr (n = 5)			143 ± 12.5	53.0 ± 16.7[b]	16.7 ± 5.54[b]
10–18 yr (n = 5)			147 ± 12.4	113 ± 8.09[c,d]	12.7 ± 2.40[b,e]
27–68 yr (n = 6)			167 ± 8.87	132 ± 7.94[c,f]	24.8 ± 12.0[b,e]

The relative immunoreactivity of each zone in the adrenocortex was evaluated by the H scoring system (0–200). Data are the mean ± SEM.
[a] $P < 0.001$ *vs.* definitive zone.
[b] $P < 0.001$ *vs.* zona glomerulosa.
[c] $P < 0.05$ *vs.* zona glomerulosa.
[d] $P < 0.05$ *vs.* age 11 months–8 yr.
[e] $P < 0.001$ *vs.* zona fasciculata.
[f] $P < 0.01$ *vs.* age 11 months–8 yr.

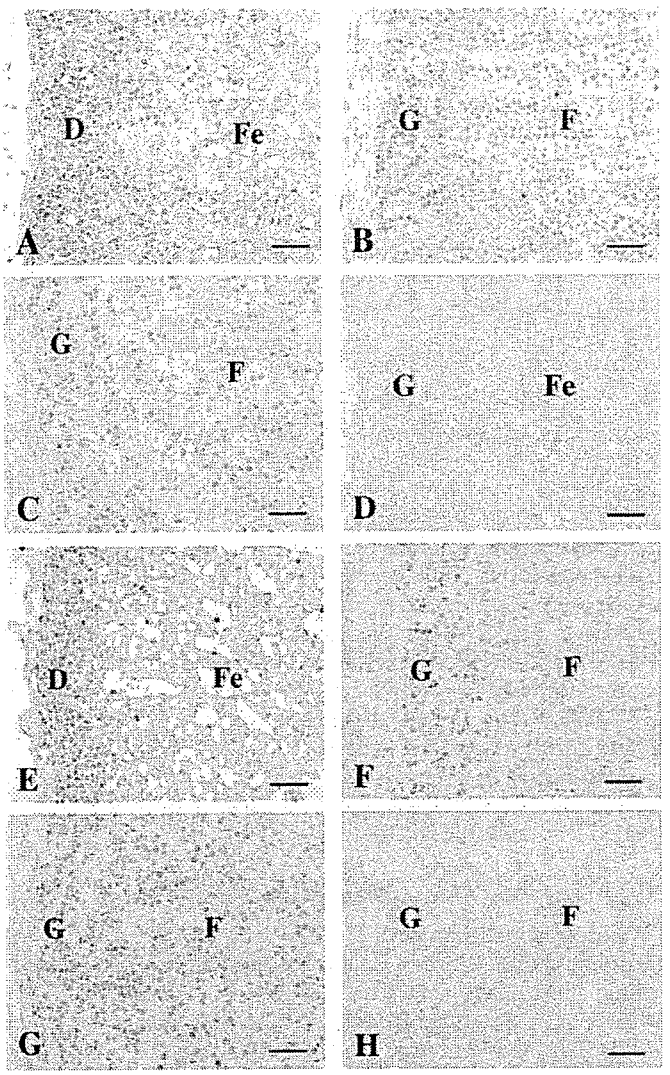

FIG. 1. Immunohistochemistry for Nurr1 (A–D) and NGFI-B (E–H) in the adrenal cortex. A, Immunoreactivity for Nurr1 was detected at a high level in the nuclei of cortical cells in the definitive zone (D; 11 wk gestation). Fe, Fetal zone. B and C, Nurr1 immunoreactivity was detected at a high level in the zona glomerulosa (G; B, 5 yr of age; C, 49 yr of age). F, Zona fasciculata. D, Immunohistochemical preabsorption test for Nurr1 showed no specific immunoreactivity. E, NGFI-B immunoreactivity was detected at a high level in the definitive zone (14 wk gestation). F, At 11 months of age, immunoreactivity for NGFI-B was high in the zona glomerulosa, but low in the zona fasciculata. G, NGFI-B immunoreactivity was high in both zonae glomerulosa and fasciculata in the adrenocortex at 28 yr of age. H, Immunohistochemical preabsorption test for NGFI-B. No specific immunoreactivity was detected. *Bar*, 50 μm.

GAAGCTCACTGG-3′; and reverse, 5′-TCCACCACCCTGTTGCTGTA-3′ (22). To verify amplification of the correct sequences, PCR products were purified and subjected to direct sequencing. Nonpathological adrenal tissues were used as positive controls for Nurr1, NGFI-B, and CYP11B2. Negative control experiments lacked cDNA substrate to check for the possibility of exogenous contaminant DNA, and no amplified products were detected under these conditions. The mRNA level for Nurr1, NGFI-B, and CYP11B2 in each case has been summarized as a ratio of GAPDH, and subsequently evaluated as a ratio (percentage) compared with that in the positive controls (nonpathological adrenal glands = 100%).

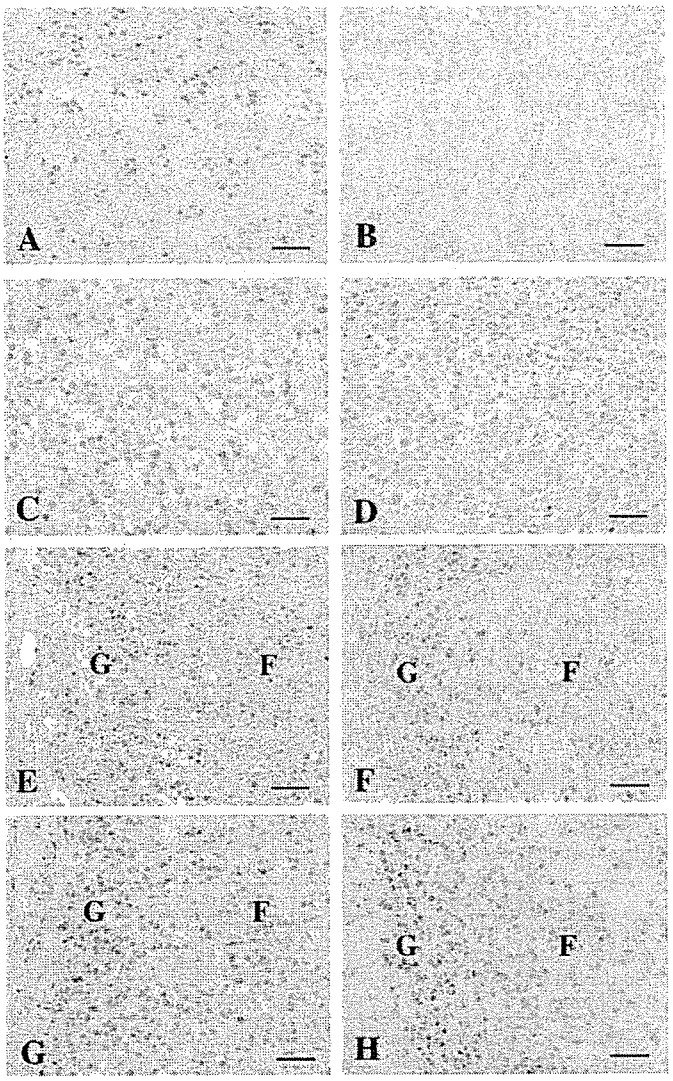

FIG. 2. Immunohistochemistry for Nurr1 and NGFI-B in adrenocortical adenoma (A–D) and its attached nonneoplastic adrenal cortex (E–H). A–D, Nurr1 immunoreactivity was detected at a high level in the nuclei of tumor cells in aldosteronoma (A), whereas a great majority of tumor cells were negative for Nurr1 in Cushing's adenoma (B). NGFI-B immunoreactivity was detected at a high level in aldosteronoma (C; same case as in A) or Cushing's adenoma (D; same case as in B). E–H, Nurr1 (E and F) and NGFI-B (G and H) immunoreactivities were not significantly different in the attached nonneoplastic adrenocortex regardless of the type of adenoma. E–H were taken from the same slide as A–D, respectively. *Bar*, 50 μm.

Results

Nonpathological adrenal cortex

Results for Nurr1 immunoreactivity in nonpathological human adrenocortex are summarized in Table 1A. Nurr1 immunoreactivity was detected in the nuclei of cortical cells, and its relative immunoreactivity was significantly ($P < 0.0001$) higher in the definitive zone (151 ± 19.8) than in the fetal zone (26.0 ± 4.74) of the fetal adrenal (Fig. 1A). At 11 months to 8 yr of age, immunoreactivity for Nurr1 was significantly higher in the zona glomerulosa (132 ± 8.95) than in the zonae fasciculata (58.8 ± 16. 5) and reticularis (15.3 ±

7.76; $P < 0.0001$, respectively). Nurr1 immunoreactivity in each zone was not significantly changed among the age groups examined (11 months to 8 yr, 10–18 yr, and 24–62 yr; Fig. 1, B and C).

Results for NGFI-B immunoreactivity in nonpathological adrenal cortex are summarized in Table 1B. NGFI-B immunoreactivity was detected in the nuclei of cortical cells, and its relative immunoreactivity was significantly ($P < 0.0001$) higher in the definitive zone (122 ± 12.6) than in the fetal zone (24.2 ± 3.61; Fig. 1E). At 11 months to 8 yr of age, NGFI-B immunoreactivity was significantly higher in the zona glomerulosa (143 ± 12.5) than in the zonae fasciculata (53.0 ± 16.7) and reticularis (16.7 ± 5.54; $P < 0.0001$, respectively; Fig. 1F). Immunoreactivity for NGFI-B in the zona fasciculata was significantly increased in adolescent and adult age groups (10–18 yr, 113 ± 8.09; 27–68 yr, 132 ± 7.94) compared with that the 11 months to 8 yr of age group ($P < 0.05$, and $P < 0.01$, respectively; Fig. 1G), whereas NGFI-B immunoreactivity in the zonae glomerulosa and reticularis was not significantly changed.

Adrenocortical tumor

The results for Nurr1 and NGFI-B immunoreactivity in adrenocortical tumors are summarized in Table 2A. Nurr1 relative immunoreactivity was significantly higher in aldosteronoma (121 ± 9.91; Fig. 2A) than in Cushing's adenoma (62.1 ± 12.2; Fig. 2B) and adrenocortical carcinoma (60.0 ± 19.0; $P < 0.001$, respectively). NGFI-B immunoreactivity was also higher ($P < 0.05$) in aldosteronoma (122 ± 14.3; Fig. 2C) than in adrenocortical carcinoma (76.4 ± 13.6), but its difference among the types of adrenocortical adenoma did not reach statistical significance (Fig. 2D).

The results for Nurr1 and NGFI-B immunoreactivity in attached nonneoplastic adrenal cortex of adenoma are summarized in Table 2, B and C. The relative immunoreactivity of Nurr1 and NGFI-B in attached nonneoplastic adrenocortex of adenoma was not significantly changed regardless of the type of adenoma examined (Fig. 2, E–H).

mRNA expression of Nurr1, NGFI-B, and CYP11B2 in adrenocortical adenoma

mRNA expression for Nurr1, NGFI-B, and CYP11B2 was detected as a specific single band (352, 358, and 121 bp, respectively) and was semiquantified by real-time RT-PCR. mRNA expression of Nurr1 and NGFI-B was detected in all adrenocortical adenomas examined, and the range of mRNA levels was 0.160–354% for Nurr1 and 12.0–223% for NGFI-B (nonpathological adrenal glands = 100%, respectively). As shown in Fig, 3, A and B, the mRNA levels of Nurr1 and NGFI-B were significantly correlated with the relative immunoreactivity (for Nurr1: $r = 0.896$; $P < 0.0001$; for NGFI-B: $r = 0.787$; $P < 0.0001$). mRNA expression of Nurr1 was significantly associated with that of CYP11B2 ($r = 0.765$; $P < 0.0001$; Fig. 3C), whereas no significant association was detected between NGFI-B and CYP11B2 mRNA levels ($r = 0.213$; $P = 0.3301$; Fig. 3D).

TABLE 2A. Relative immunoreactivity of Nurr1 and NGFI-B in adrenocortical tumors

Type of tumor (no. of cases)	Nurr1	NGFI-B
Aldosteronoma (n = 14)	121 ± 9.91	122 ± 14.3
Cushing's adenoma (n = 10)	62.1 ± 12.2[a]	105 ± 13.0
Nonfunctioning adenoma (n = 10)	91.8 ± 12.6	90.0 ± 19.6
Carcinoma (n = 10)	60.0 ± 19.0[a]	76.4 ± 13.6[b]

The relative immunoreactivity was evaluated by the H scoring system in each case (0–200). Data are the mean ± SEM.
[a] $P < 0.01$ *vs.* aldosteronoma.
[b] $P < 0.05$ *vs.* aldosteronoma.

Discussion

Aldosterone is produced in the zona glomerulosa of the adrenal cortex through an interaction of several steroidogenic enzymes, including P450 side-chain cleavage (CYP11A), 3β-hydroxysteroid dehydrogenase type 2 (3βHSD2), CYP21, and CYP11B2. Among these enzymes, CYP11B2 is a specific enzyme for aldosterone biosynthesis, and it is expressed exclusively in the zona glomerulosa (23).

TABLE 2B. Relative immunoreactivity of Nurr1 in the adjacent nonneoplastic adrenal cortex of adenomas

Type of adenoma (no. of cases)	Glomerulosa	Fasciculata	Reticularis
Aldosteronoma (n = 14)	124 ± 10.8	72.4 ± 9.62[a]	17.1 ± 3.94[a,b]
Cushing's adenoma (n = 10)	143 ± 10.8	88.6 ± 12.5[a]	24.2 ± 6.21[a,b]
Nonfunctioning adenoma (n = 10)	124 ± 11.3	60.9 ± 9.61[a]	9.38 ± 4.26[a,b]
Nonpathological adrenal (27–68 yr; n = 6)[c]	147 ± 16.9	77.3 ± 13.2[a]	13.8 ± 6.65[a,b]

The relative immunoreactivity of each zone in the adrenal was evaluated by the H scoring system (0–200). Data are the mean ± SEM.
[a] $P < 0.001$ *vs.* zona glomerulosa.
[b] $P < 0.001$ *vs.* zona fasciculata.
[c] Data were taken from Table 1A.

TABLE 2C. Relative immunoreactivity of NGFI-B in the attached nonneoplastic adrenal cortex of adenomas

Type of adenoma (no. of cases)	Glomerulosa	Fasciculata	Reticularis
Aldosteronoma (n = 14)	149 ± 8.97	111 ± 12.1[a]	22.6 ± 12.0[b]
Cushing's adenoma (n = 10)	151 ± 11.3	119 ± 10.5[a]	22.4 ± 4.30[b]
Nonfunctioning adenoma (n = 10)	143 ± 7.4	109 ± 9.80[a]	25.8 ± 4.46[b]
Nonpathological adrenal (27–68 yr; n = 6)[c]	167 ± 8.87	132 ± 7.94[a]	24.8 ± 12.0[b]

The relative immunoreactivity of each zone in the adrenal was evaluated by the H scoring system (0–200). Data are the mean ± SEM.
[a] $P < 0.05$ *vs.* zona glomerulosa.
[b] $P < 0.001$ *vs.* zona glomerulosa. $P < 0.001$ *vs.* zona fasciculata.
[c] Data were taken from Table 1B.

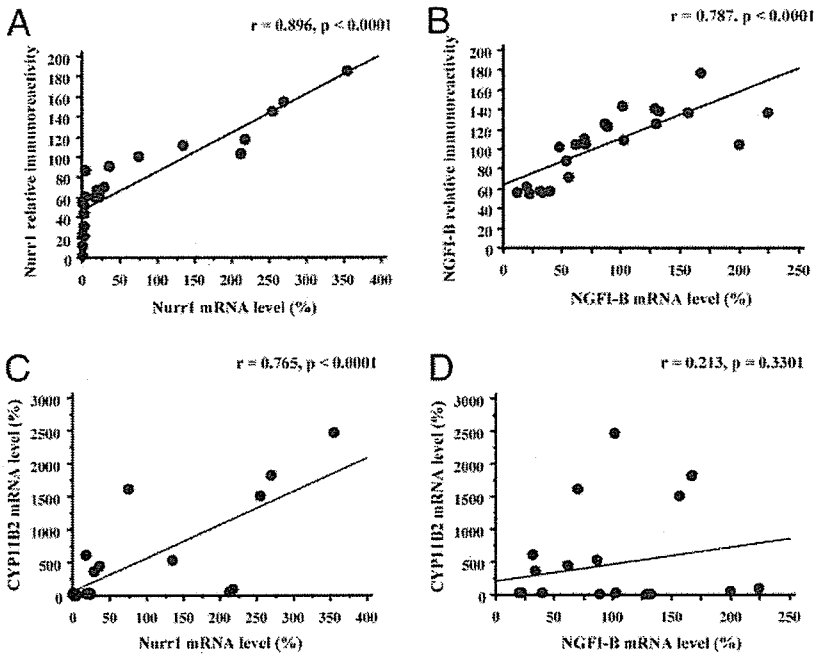

FIG. 3. A and B, Association between the mRNA level and the relative immunoreactivity of Nurr1 (A) or NGFI-B (B) in 23 cases of adrenocortical adenomas. Significant positive associations were detected (for Nurr1: r = 0.896; $P < 0.0001$; for NGFI-B: r = 0.787; $P < 0.0001$). C and D, Association between the mRNA level of Nurr1 (C) or NGFI-B (D) and that of CYP11B2 in 23 adrenocortical adenomas. CYP11B2 mRNA was significantly correlated with Nurr1 mRNA (r = 0.765; $P < 0.0001$), but not with NGFI-B mRNA (r = 0.213; $P = 0.3301$).

Very recently, Bassett *et al.* (14) demonstrated that the CYP11B2 gene has two functional NBREs in the promoter region and was markedly up-regulated by Nurr1 and NGFI-B. In this study Nurr1 immunoreactivity was present at a high level in the zona glomerulosa or aldosteronoma, and mRNA expression of Nurr1 was significantly correlated with that of CYP11B2 in adrenocortical neoplasms. Therefore, it is suggested that Nurr1 expression plays an important role in aldosterone production through the induction of CYP11B2 in the zona glomerulosa of adrenal cortex or aldosteronoma. However, it is also true that Nurr1 immunoreactivity was detected at low levels in the zona fasciculata or adrenocortical neoplasms in addition to aldosteronoma in this study. Activation of Nurr1 depends on two so-called activation functions (AF1 and AF2), located at the N- or C-terminal regions, and it was partly regulated by the phosphorylation (24). Therefore, posttranslational modifications of Nurr1 are also considered to play some role in the zone-specific expression of CYP11B2 in the adrenal gland. In adrenocortical adenomas expressed Nurr1 mRNA at low levels, relative immunoreactivity of Nurr1 was variably detected (Fig. 3A). In these cases, evaluation of Nurr1 mRNA in tumor tissues may be reflected by heterogeneous expression of Nurr1 in neoplastic cells and/or amounts of stroma within the samples.

NGFI-B immunoreactivity was detected at high levels in the zonae glomerulosa and fasciculata in the nonpathological adrenal gland and in various types of adrenocortical adenoma. The zona fasciculata is mainly involved in cortisol production, and CYP11A, 3βHSD2, CYP17, CYP21, and CYP11B1 (11β-hydroxylase) are expressed in this zone. Previous studies demonstrated that human and mouse CYP21 gene promoters contain NBREs, and induction of CYP21 transcription by NGFI-B has been proposed (20, 25). In addition, Bassett *et al.* (26) recently demonstrated that 3βHSD2 contains an NBRE in the promoter region and was significantly up-regulated by NGFI-B. On the other hand, NGFI-B had no effect on the induction of CYP11B1 (14) and CYP17 (26), which are not expressed in the zona glomerulosa (2, 23, 27). Therefore, NGFI-B may be partly involved in aldosterone and/or cortisol production through the regulation of some related enzyme expressions in the adrenal cortex and its neoplasms.

In fetal adrenal glands, immunoreactivity of steroidogenic enzymes is known to become generally discernible after 23 wk gestation in the definitive zone (28), and the definitive zone is considered to become steroidogenically active in the late phase of pregnancy (29). However, Nurr1 and NGFI-B immunoreactivities were detected in the definitive zone in all fetal adrenals examined (from 11–36 wk gestational) in our study. Therefore, the expression of Nurr1 and NGFI-B is postulated to occur before the expression of steroidogenic enzymes in the definitive zone in the fetal adrenals. Rainey *et al.* (30) reported that NGFI-B mRNA expression was very low in the fetal adrenal gland (15–20 wk gestation) compared with that in the adult adrenal by microarray and Northern analyses. These data are not necessarily consistent with our present results, but may be due to the different gestational ages examined or the different examination methods used. In addition, the fact that microarray and/or Northern analysis required a whole adrenal specimen may contribute to this difference from the present immunohistochemical study, because the definitive zone is markedly thin and much smaller in volume than the fetal zone in human adrenal.

Autonomous neoplastic production of cortisol in Cushing's adenoma or at least some nonfunctioning adenoma patients results in adrenocortical atrophy with suppression of steroidogenic enzyme in the zonae fasciculata/reticularis of the adjacent nonneoplastic adrenocortex through inhibition of ACTH secretion. In addition, the expression of steroidogenic enzymes is markedly decreased, except for CYP21, in the zona glomerulosa of the adjacent nonneoplas-