Int. J. Nonlinear Anal. Appl.
Volume 11, Special Issue, Winter and Spring 2020, 433-444
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2020.4655

Predicting the Behaviour of the Senescence-Accelerated Mouse (SAM) Strains (SAMPs and SAMR) Using Machine Learning Algorithm

Sura Zaki Al Rashid[a*]

[a]*Software Department, College of Information Technology, University of Babylon,*

Abstract

A primary aspect of human aging is progressive neurological dysfunction. Due to the fundamental variations in aging in mice and humans, it is difficult to obtain and research effective mouse models. There are two types of tissue phenotypes that are distinct; one is the tissue for retina and one for the hippocampus. Each form has three strains. A variational formulation for sparse approximations is introduced in this work, inferring both the kernel hyper-parameters and inducing inputs by maximising a lower bound of probability of true log marginal. In order to account for more complexity with the time series, a model is built on this series with a correlated human model performance. The molecular senescence of the hippocampus and retina, both with accelerated neurological senescence (SAMP10 and SAMP8) models were presented. The purpose of the study is to specify the relationship between these genes or pathways that would provide insight into the mechanism for this phenotype which will be superior to the current incomplete state-of-the-art approximations. Furthermore, the combined study of the essential features of inbred strains and profiling of gene expression can help determine which genes are essential for complex phenotypes. However, the identification, sequencing and gene expression of full-genome polymorphism of inbred mouse strains with intermediate.

Keywords: Sparse Gaussian Process (Classification and Regression), Puma Package, Coregionalisation Model, Senescence-Accelerated Mice strains.

1. Introduction

Aging is outlined by an increase in the probability of death overtime related to characteristic changes in composition. Aging [1] [2], a time-related deterioration of physiological functions, occurs in the

*Corresponding Author: Sura Zaki Al Rashid

Received: February 2020 *Revised:* October 2020

organism's cell, tissue and organs, leading [3][4] to changes such as pigmentation of skin, organ senescence and skin wrinkling. Aging [5] itself cannot be considered an illness, but poses risk for the practical decay of many organs. Particularly, aging is at the centre of age-related diseases such as insanity. The S8 strain is a model of super brain aging of learning and defects of memory. The hippocampus is a crucial part of the visceral brain, precisely governing feeling and noesis [3].

To obtain insights into the means of improving psychological features for reversal of aging in the S8 strain, we tend to additionally examine the structure of the vegetative cells in the of hippocampal region [3]. The gene expression between the senescence-resistant and senescence-prone strains might well be initial indicators for the aging or might indicate accelerated senescence within SAMP8 and SAMP10 strains. There are many dropped genes in the metaphysics of the fascinating cistron (GO) class as concluded by the enrichment analysis of cistron victimisation of the GO[4]. For each of the experimental procedures, the performance of real-world and analytical methodologies was compared to determine organic phenomenon changes; 10 genes were selected from the analysis of the hippocampus (from humans not employed in the analysis of microarray). Of the 10 genes examined, the modifications of zfgh7 expression for genes were fixed with a change bigger than three-fold [4]. The authors faced challenges in computational biology to build a speed model with the highest accuracy for predicting the behaviour gene in both retina and hippocampus [4][5].

The structure of this manuscript is as follows: Section 2 assesses the general aspects of aging, computations aspects and pathway analysis; Section 3 illustrates the methodology adopted for the study. Section 4 presents the results and their discussion. Finally, Section 5 describes some conclusions for this work.

2. METHODS

A brief description of aging and the techniques of machine learning adopted in this work is presented in this section. Subsequently, the use of the adapted analysis of the 16S sequence datasets is presented.

2.1. General aspects of aging

2.1.1. The differences strains in genes expression

The aging of the hippocampus and the retina were analysed; it is noteworthy that specific transcriptional events happened within the retina and hippocampus for each strain in the aging process [6]. In the retina, groups of genes are involved in the perception of brightness and other elements. In the hippocampus, this included genes involved in learning, behaviour and heat-shock response.

In SAMP strains, the S10 strain is used as a model for the study of brain aging and age-related neurodegenerative conditions. Aging brains of SAMP10 mice are characterised by many features: retraction of the dendritic of impaired learning, frontal atrophy, cortical neurons, etc. [7]. Thus, S8 and S10 contribute to a model of senescence acceleration as well as to a promising clinicopathological model of therapeutic intervention for Alzheimer's disease (AD) and other such diseases [8][9][10].

2.2. Computations aspects

2.2.1. Gaussian process (GP)

One of the Bayesian methods for the supervised learning task is a Gaussian Process (GP) which is used for non-linear regression or classification for the assumption that any finite linear combination of training samples will be normally distributed. it is a building block for additional sophisticated machine learning applications[11][12][13]. The GP associates a random variable for a jointly Gaussian

variable that is a set of inputs for the associated random variables. It includes characterised random functions using functions of mean with kernel (covariance function) [14].

A GP is a stochastic method for the feature space in which the distribution $p(l(m1), l(m2), , l(mn))$ of a function $l(m)$, for the set of points $m1, m2, , mn$ with that space is Gaussian. $l(m) = f(m) + \epsilon$ where ϵ is $\sigma_n^2 I$ noise term. K_l has a full rank with an inverse [8]. GP prior is formulated as (f) latent function as shown in Eq (1)

$$p(fm) = f|m \sim GP(f; mean(m), K_f(m_i, m_j)), \tag{1}$$
$$orp(f|m) = 1/((2\pi)^{(n/2)}[[|K]_f]^{(1/2)})exp(-1/2(f - mean)^T K_l^{(-1)}(f - mean))$$

Where the covariance and the mean functions as Eq(2)

$$mean(m) = f(m), k_f(mimj) \tag{2}$$
$$cov =< (f(mi) - mean(mi))(f(mj) - mean(mj)) > .$$

Recently, a conventional line of research is covariance functions (kernels) because of the extensive Gaussian processes' application, first in machine learning for the geostatistics community, regression and classification [8][15]. The joint amplitude-time of a Gaussian process might be useful in this regards [27].

2.2.2. Covariance functions

The kernel feeds the covariance function where a covariance $K_{x;x^0}$ matrix for each pair of random variables at input points x and x^0 has a GP used for training and manipulation of the $K\_(x[, x]^0)$ using the training set. Computing the covariance matrix(kernel) inversion leads to the complexity of GP as $O(n3)$. Therefore, the GP is prevented from being applied to huge training data points [16] [17] [18]. In this work, the white function and Ornstein-Uhlenbeck (OU) were used with the coregionalisation model[19].

2.2.2.1 Coregionalisation model

The outputs of independent random functions are expressed as linear combinations using a linear model of coregionalisation model[11][20][21]. The resulted model is a GP if the independent random functions are GP or a positive semi-definite kernel [22][23].

Let U outputs ${f_{u(x)}}_{u=1}^{U}$ as well as $x \in R^p$, is shown f_u.

$$f_{u(x)} = \Sigma_{p=1}^{P}(a_{(u,p)} s_{p(x)}) \tag{3}$$

Where the independent functions $s(x)$ and $a_{(u,p)}$ are scalar coefficients with 0 mean, the covariance $k_{p(x;x;')}$ is $cov[s_{p(x)}, s_{(p')(x')}]$, if $p = p^{'}$ else 0 [24] [25]. The Eq(4) explains a cross covariance between $f_{u(x)}$ and $f_{(u')(x)}$ functions.

$$cov[f_{u(x)}, f_{(u')(x')}] = \Sigma_{p=1}^{P}\Sigma_{(i=1)}^{(R_p)}[a_{(u,p)}^{i} a_{(u',p)}^{i} k_p(x, x^{'})] \tag{4}$$
$$= \Sigma_{p=1}^{P}[b_{(u,u')^p} k_p(x, x^{'})]$$

This $s_{p^{i(x)}}$ where $i = 1, \ldots, R_p$ and $p = 1, \ldots, P$ has ((0) mean, $k_p(x, x^{'})$ is $cov[s_p^{i(x)}, s_{(p')}^{(i')}(x)']$ if $i = i^{'}$,and $p = p^{'}$. But $cov[f_{p(x)}, f_{(p')(x')}]$ covariance) is given by $((x, x^{'}))_{(u,u')}$ and the $K(x, x^{'})$ is written as Eq(5).

$$K(x, x^{'}) = \Sigma_{(p=1)}^{P} B_p K_p(x, x^{'}) \tag{5}$$

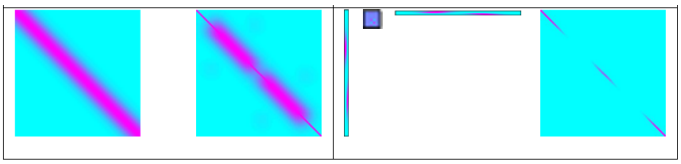
The coregionalisation matrix is $B_p \in R^{(U\times U)}$, and the kernel results from Latent Matric Coregionalisation (LMC), which is a summation of the products of two kernel (covariance) functions; the first model depends on the input vector x and outputs. In the second model, the input dependence with independence of $\{f_{u(x)}\}_{u=1}^{U}$ is called the covariance function $K_p(x, x^{'}))$ [26] [19].

2.2.3. Sparse Gaussian process methods

The Gaussian process (GP) is a famous and sophisticated method for Bayesian non-parametric non-linear classification and regression [8][28]. However, due to its non-parametric nature, it cannot be directly applied to the real world and is intractable for large datasets where it causes computational problems and requires $O(n^3)$ computation time where n is the training data. Therefore, many studies have attempted to obtain sparse approximations of the complete Gaussian process to reduce this complexity down to M^2N [29] [30].

These methods include choosing a subset of the M-size training data (active) on which a computation is based to speed up GP learning; GP through $M < N$ inducing point f^- be estimated to obtain (SPGP) prior : $p(f) = \int d\bar{f} \prod_n p(f_n|\bar{f})p(\bar{f})$

GP prior	SPGP prior
$N(0, K_N) \approx .p(f)$	$=N(0, K_{NM}K_M^{(-1)}K_{MN} + .\Lambda)$



Where SPGP covariance upended in $O(M^2N) \ll O(N^3)$ is faster than GP. SPGP is equal or similar to GP with covariance that is non-stationary and parameterised by X^-. The data $\{X, y$ summation with noisy σ^2, the predicting variance and mean can be computed in $O(M^2)$and $O(M)$ per test case, respectively [31][32].

The selection of the kernel hyper-parameters and the inducing inputs are required for the Sparse Gaussian process that uses making variables. We applied this technique to regression and found the affected genes in the retina and hippocampus tissues with three strains for young and old humans [33][34][35].

An objective function used in relation to the projected process approximation, Eq(6), is a form

$$F = log[N(y|0, \sigma^2 I + Q_{nn})] \tag{6}$$

of log marginal likelihoods.

In PP, Q_{nn} (an approximation of the true K_{nn}.) is $K_{nm}K_{mm^{(-1)}}K_{mn}$, where Kmn refers to the $m \times n$ cross covariance matrix between inducing and training points.

$$Q_{nn} = diag[K_{nn} - K_{nm} * K_{mm}^{(-1)} * K_{mn}] + K_{nm} * K_{mm}^{(-1)} * K_{mn} \tag{7}$$

$K_{nm} = K_{mm}^T$, and K_{mm} is the $m \times m$ matrix on the inducing inputs.

In SPGP, when modifying the GP prior, the F is obtained when comparing Eq(2) with Eq(3). The inducing inputs X_m show the role of more hyper-parameters of kernel (similar to θ) with parametrised Q_{nn}.

In extra hyper-parameters (similar θ), the inducing input X_m is an important factor where it parametrises the matrix Qnn as continuous optimisation of F heavily parametrised with respect to X_m does not reliably estimate the same Gaussian process model.

2.3. The pathway analysis

Pathway analysis is used to obtain a perception of the motivating biology of the gene expression. It can increase the explanatory power and decrease complexity; gene profiling and high-throughput sequencing are used to examine whether a gene or a list of genes are factors for a phenomenon which can be used to analyse gene ontology (GO) and for further comparisons.

3. Methodology

The methodology focused on a set of important genes and pathway of molecular indices of the behaviour of the retina and hippocampus tissues in SAMs strains. The general stages of this manuscript are described in Fig. 1. Downloading the dataset is the first step followed by computation of the gene expression values from ".cel files". Thereafter, normalisation of the gene expression values training data (Y) for three strains is performed with respect to two factors (old and young age) with two or three replicates with its dimensional $(num\_genes(PorbeID), num\_points)$; in this step, the statistical student's t test was utilised as shown in Eq(8).

$$t = \frac{\bar{x} - u}{s/\sqrt{n}} \tag{8}$$

Where $\bar{x}$ mean of x, s is standard deviation.[36] [4]

In this model, the Y with the parameters were the inputs. In the first step, creating a matrix depended on the conditions (the experiments (time series)), number of tissues, number of strains, number of replicates for young and old human where the matrix dimensionality is (num _ times series, coregionalise _ dim); the X has 27 rows and four columns as in Fig1. The optimisation process was used to infer and optimise the hyper-parameters. Then one of the two models was selected depending on the highest value of likelihood for each gene. For each case there is one model for each gene. After optimisation stage, the Logit and ODDs was computed as Eq(9) and Eq(10) depending the accuracy of one model with the others. Depending on the largest probability value, one model was selected.

$$logit = -log((1/p) - 1) \tag{9}$$

The ratio of these two probabilities are defined as follows:

$$ODDs = p/1 - p \tag{10}$$

- **Prediction making**

The prediction of the variance and mean matrixes were computed using the previous equations that has two steps. The best model is shown and the genes that change their behaviour depending on the retina and hippocampus types were selected with three cases that were plotted separately. After this process, pathways of genes were examined in DAVID.

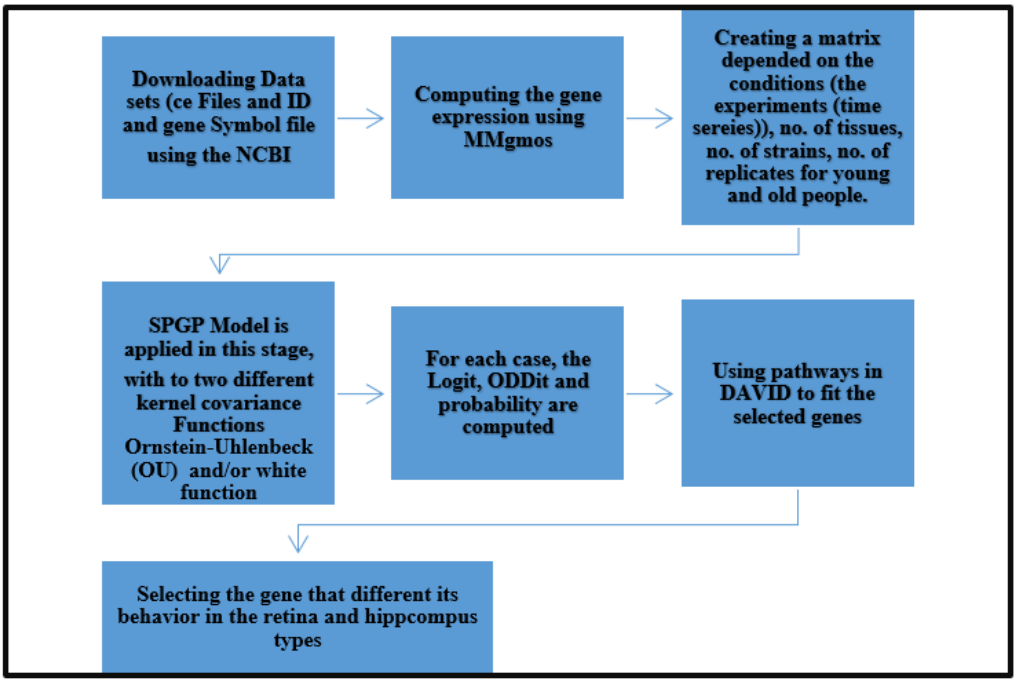


Figure 1: Highlighted boxes represent the sequence of involved steps.

4. Results

4.1. Downloading the dataset

This work demonstrates the human model that has a ID of GEO Series file (GSE6238) which was downloaded for this experiment as shown in Figure 2. The description of dataset in presented in Table I. This explains the details of number of tissues and strains and number of the types for each strain and number of the replicates for each type. The discussed data was deposited in NCBI's Gene Expression Omnibus are available.

Table 1: Description of Dataset

Title of Dataset	GSE6238
Characteristics of Attribute	Real
Characteristics of Data Set	Real
Missing Values	Yes
Number of Attributes	27
Number of Instances	12422
Publication	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6238 [6]

4.2. Results of prepossessing stage

In this stage, two functions were applied (the Mmgmos and RMA functions) for the analysis to obtain the values of gene expression. The Mmgmos method in the Puma package and RMA method in Limma package via R language were used as shown in Figure 3.

4.3. Building the X matrix

The matrix of time series (X) was computed for the correlation among two tissues and three strains (Samp8, Samp10, SAMR) with two or three replicates for young and old participants where X is shown as Table 2.

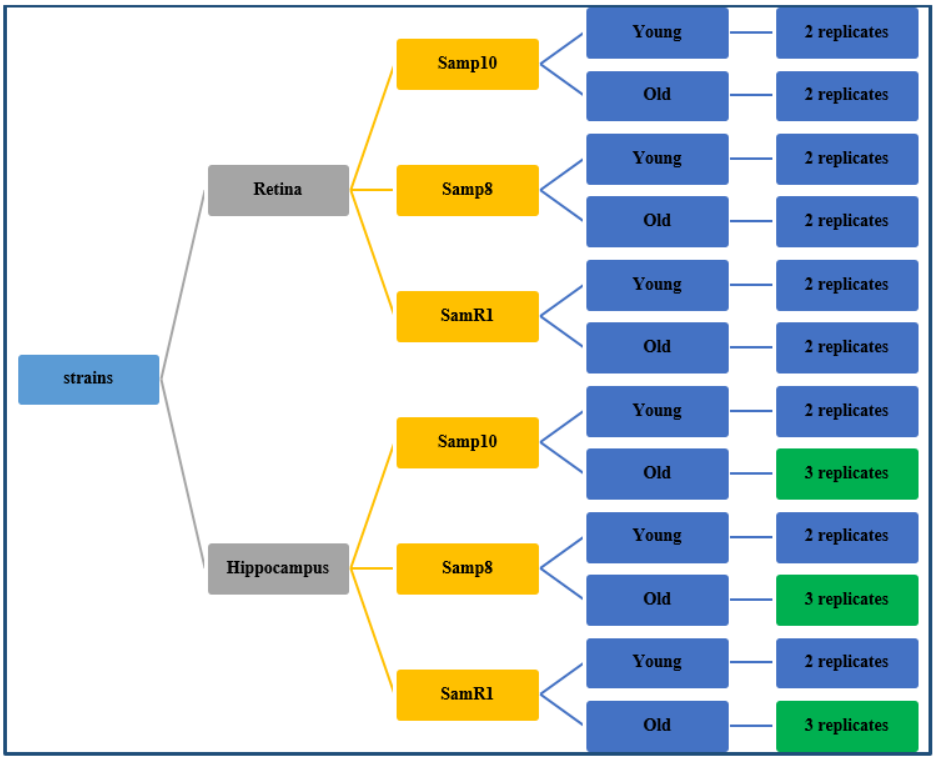


Figure 2: A form of Data Set

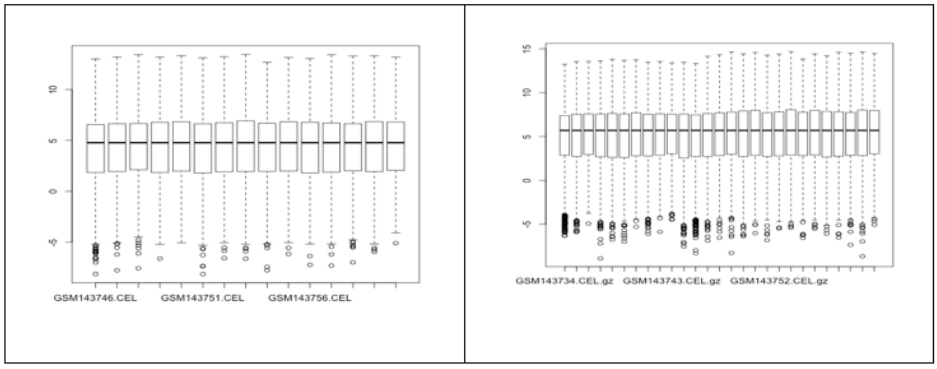


Figure 3: Boxplot of computing the Gene Expression using a) RMA method, b) mmgmos method.

4.4. The Model's results

We applied the model for each gene; in this work, all genes (12422) were trained in two models. After the training process, the maximum of probability was the criteria to select the best model. These models are sorted and one is subsequently selected. In this paper, two models were obtained; the first used the OU with white functions, and the second was applied with only white covariance function. For example, in the first model, the parameters for the gene (ocm), probe ID ("101174_at")*wasused*.

The comparison step used the likelihood value; One model was selected for the same gene (ocm), where the best model is the first; its likelihood is (-23.7141012439). Where there were two models for each gene (12422 genes), the models were sorted the one with the highest probability value for each gene (12422 genes) was selected.

Table 2: A Matrix of time series

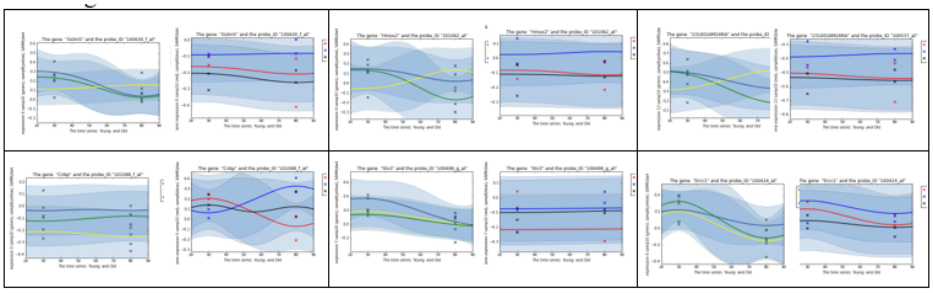
Time series	Samp10& Samp8& SamR	Samp10& Samp8& SamR	(Young &Old with replicates)
30 (young)	0	0	0
30 (young)	0	0	0
30 (old)	0	0	1
30 (old)	0	0	1
30 (young)	0	1	0
30 (young)	0	1	0
30 (old)	1	0	0
30 (old)	1	0	0
30 (young)	1	0	1
30 (young)	1	0	1
30 (old)	1	1	0
30 (old)	1	1	0
80 (young)	0	0	0
80 (young)	0	0	0
80 (old)	0	0	1
80 (old)	0	0	1
80 (old)	0	0	1
80 (young)	0	1	0
80 (young)	0	1	0
80 (old)	1	0	0
80 (old)	1	0	0
80 (old)	1	0	0
80 (young)	1	0	1
80 (young)	1	0	1
80 (old)	1	1	0
80 (old)	1	1	0
80 (old)	1	1	0

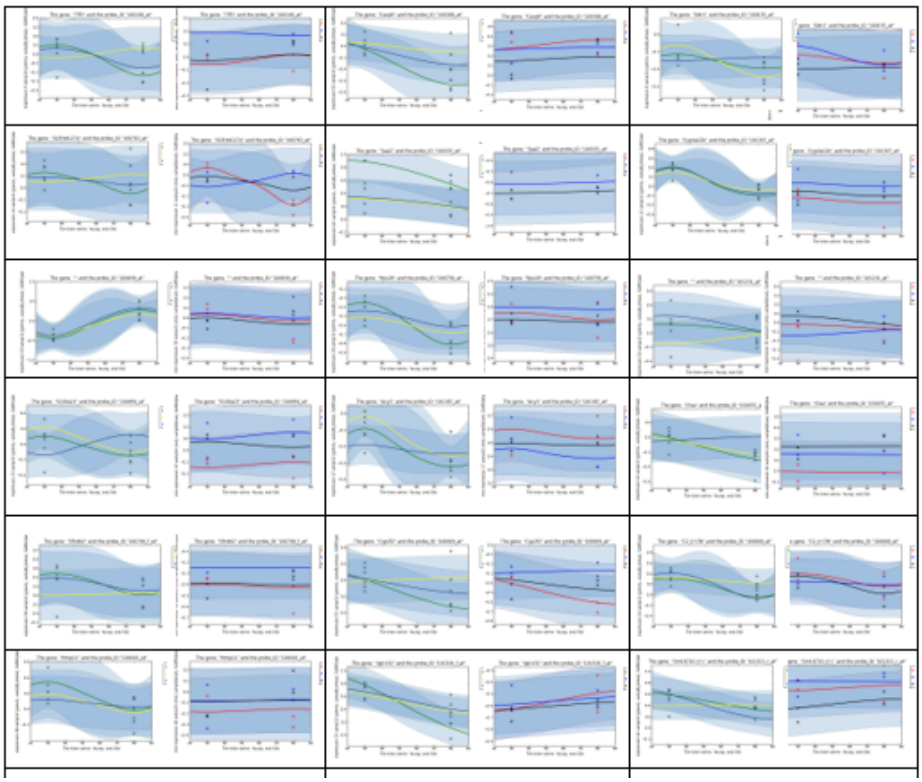
- **Results of posterior step**

The posterior mean and variance were computed. There are parameters necessarily built to compute it; $X_newhastestpoints(testdata)$.

Thereafter, DAVID was used to check the genes. The majority of the changes of the age-linked genes were strain specific with only some usual pathways found for accelerated neurological and normal aging. Screening of polymorphism guided to the identification of mutations that have a direct influence on important disease processes, including a mutation in a fibroblast growth factor gene and a mutation in and ectopic expression of the gene which is involved in inflammatory response. This work examined the expression levels of two differentially genes. The analysis was performed on SAMP8, SAMP10, and SAMR1 mice. There were more expressions of genes in the retina and hippocampus of SAMP8 than SAMR1 and more abundant genes in the retina and hippocampus of

SAMP8 mice than SAMR1. Increased signal in Samp8 and Samp10 retina and hippocampus was relative to SAMR. There were dropped genes in the GO categories.





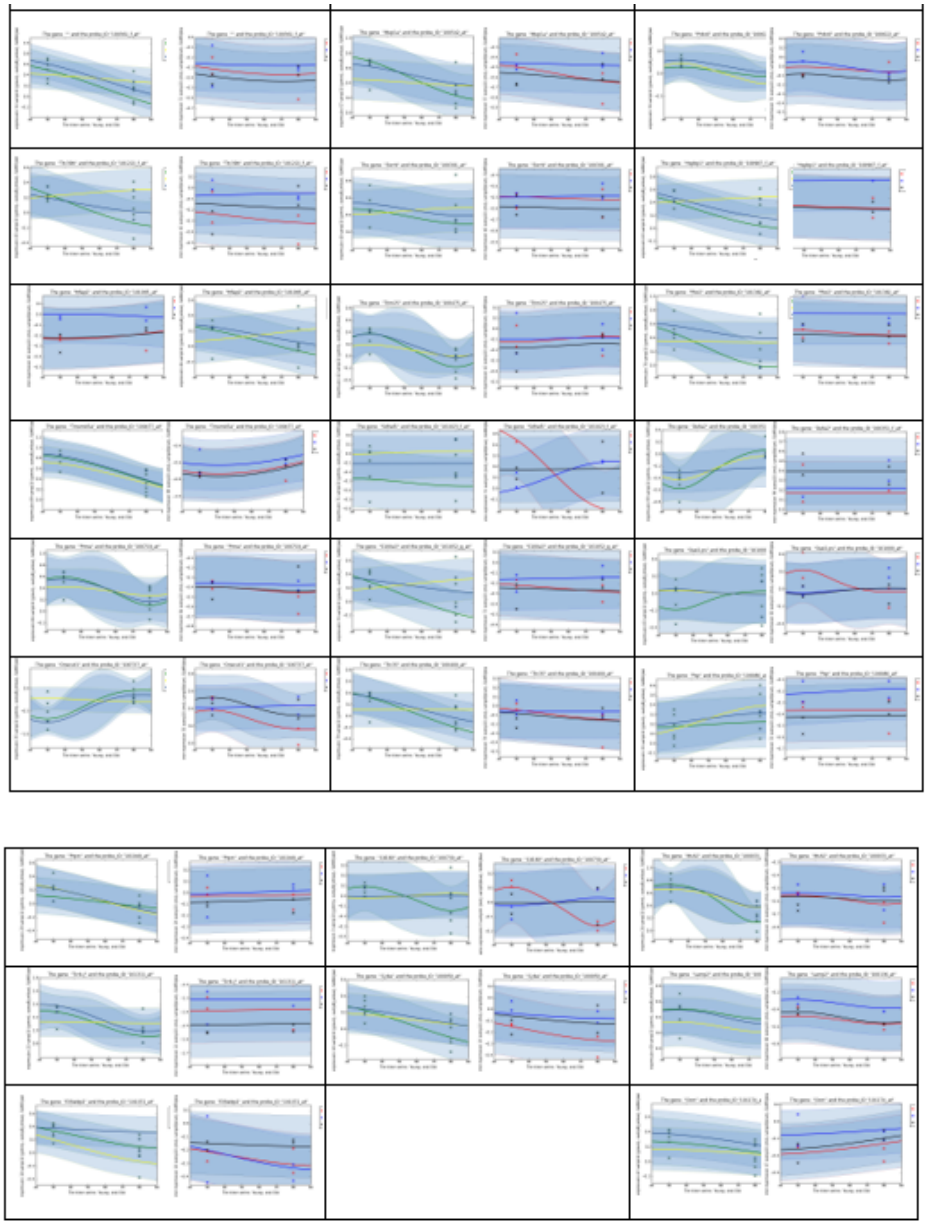


Figure 4: The results of models explains the behaviour of genes expression.

5. Conclusions

There are multiple tissues and processes involved in the aging process, and a set of genes and pathways are also involved in the same. In the pathology of SAMP8, SAMP10 and SAMR mice, gene expressions with some important candidates were associated. The multiple strain tissue analysis with the integration of gene expression derived from many experiments was essential for decoding the molecular biology of the aging process in humans and animals. Furthermore, relating the obtained knowledge from all genome sequencing of many strains with analyses of gene expressions with phenotypic contrast will help enhance the knowledge regarding the process of aging. The used Sparse Gaussian process considered multiple replicates. For validation, a widely acceptable gene ontology was demonstrated as well as a tool for functional annotation was used to indicate highlighted genes and their obtained characteristics using this research. These results led to the achievement that this model developed basing SGP can successfully predict genes. These results can be helpful for further

research and analysis.

References

[1] T. Onodera et al., "Depressive Behavior and Alterations in Receptors for Dopamine and 5-Hydroxytryptamine in the Brain of the senescence Accelerated Mouse (SAM)-P10," Jpn. J. Pharmacol., pp. 312–318, 2000.
[2] I. Akiguchi et al., "SAMP8 mice as a neuropathological model of accelerated brain aging and dementia: Toshio Takeda's legacy and future directions," Neuropathology, vol. 37, no. 4, pp. 293–305, 2017.
[3] F. J. Chen et al., "Icariin delays brain aging in senescence-accelerated mouse prone 8 (SAMP8) model via inhibiting autophagy," J. Pharmacol. Exp. Ther., pp. 121–128, 2019.
[4] T. Sheela and L. Rangarajan, "An Approach to reduce the large feature space of Microarray Gene Expression Data by gene clustering for efficient sample classification," vol. 2, no. 8, pp. 1–12, 2018.
[5] S. A. Fattah, H. A. Lafta, and S. Z.alrashid, "B-pred: An intelligent and adaptable medical diagnosis system based on bagging machine learning," Int. J. Sci. Technol. Res., vol. 9, no. 3, pp. 1325–1331, 2020.
[6] T. a Carter et al., "Mechanisms of aging in senescence-accelerated mice.," Genome Biol., vol. 6, no. 6, p. R48, 2005.
[7] Y. Sanada et al., "Senescence accelerated mice as a new mouse model for spontaneous osteoarthritis," Osteoarthr. Cartil., vol. 26, no. 2018, p. S47, 2018.
[8] C. K. Rasmussen, C.E adn Williams, Gaussian processes for machine learning., vol. 14, no. 2. 2006.
[9] K. Niimi and E. Takahashi, "Differences in saccharin preference and genetic alterations of the Tas1r3 gene among senescence-accelerated mouse strains and their parental AKR/J strain," Physiol. Behav., vol. 130, pp. 108–112, 2014.
[10] A. A. AL-Mshanji and S. Z. AL-Rashid, "Improving clustering algorithm for gene expression data using hybrid algorithm," Compusoft, vol. 8, no. 9, pp. 3422–3430, 2019.
[11] S. Z. Al-Rashid, "Studying the effect of Mouse models for Gene Expression using Coregionalization Models in Gaussian process," 4th Sci. Int. Conf. Najaf, SICN 2019, pp. 210–215, 2019.
[12] M. Alvarez and N. Lawrence, "Computationally efficient convolved multiple output gaussian processes," J. Mach. Learn. Res., vol. 12, pp. 1459–1500, 2011.
[13] A. A. K. Akalaitzis and N. D. Lawrence, "A Simple Approach to Ranking Differentially Expressed Gene Expression Time Courses through Gaussian Process Regression," BMC Bioinformatics, 2011.
[14] M. Marczyk, R. Jaksik, A. Polanski, and J. Polanska, "Adaptive filtering of microarray gene expression data based on Gaussian mixture decomposition.," BMC Bioinformatics, vol. 14, p. 101, 2013.
[15] A. K. Al-Mashanji and S. Z. Al-Rashi, "Computational Methods for Preprocessing and Classifying Gene Expression Data- Survey," 4th Sci. Int. Conf. Najaf, SICN 2019, pp. 121–126, 2019.
[16] A. Kalaitzis, "Learning with structured covariance matrices in linear Gaussian models," no. February, 2013.
[17] H. Topa, Á. Jónás, R. Kofler, C. Kosiol, and A. Honkela, "Gaussian process test for high-throughput sequencing time series: Application to experimental evolution," Bioinformatics, vol. 31, no. 11, pp. 1762–1770, 2015.
[18] S. Al-Rashid, M. Arifur, N. H. Al-aaraji, N. D. Lawrence, and P. R. Heath, "Increasing Power by Sharing Information from Genetic Background and Treatment in Clustering of Gene Expression Time Series," J. Univ. Babylon, Pure Appl. Sci., vol. 26, no. 4, pp. 253–267, 2018.
[19] X. Emery and M. Peláez, "Reducing the number of orthogonal factors in linear coregionalization modeling," Comput. Geosci., vol. 46, pp. 149–156, 2012.
[20] E. J. Cooke, R. S. Savage, P. Kirk, R. Darkins, and D. L. Wild, "Bayesian hierarchical clustering for microarray time series data with replicates and outlier measurements," BMC Bioinformatics, vol. 12, no. 1, p. 399, 2011.
[21] S. D. Oman and B. Vakulenko-Lagun, "A Remark on the Use of a Weight Matrix in the Linear Model of Coregionalization," Math. Geosci., vol. 44, no. 4, pp. 505–512, 2012.
[22] S. D. Oman, B. Vakulenko-Lagun, and M. Zilberbrand, "Methods for descriptive factor analysis of multivariate geostatistical data: a case-study comparison," Stoch. Environ. Res. Risk Assess., vol. 29, no. 4, pp. 1103–1116, 2015.
[23] S. Z. Al-Rashid and N. H. Al-aaraji, "Bayesian Models with Coregionalization to Model Gene Expression Time Series for Mouse Model for Speed Progression of ALS Disease," Eur. J. Sci. Res., vol. 132, no. 1, 2015.
[24] X. Liu, M. Milo, N. D. Lawrence, and M. Rattray, "A tractable probabilistic model for Affymetrix probe-level analysis across multiple chips," Bioinformatics, vol. 21, no. 18, pp. 3637–3644, 2005.
[25] J. Hensman, N. D. Lawrence, and M. Rattray, "Hierarchical Bayesian modelling of gene expression time series across irregularly sampled replicates and clusters.," BMC Bioinformatics, vol. 14, p. 252, 2013.

[26] C. Chan, A. Al-Bashabsheh, Q. Zhou, T. Kaced, and T. Liu, “Info-Clustering: A Mathematical Theory for Data Clustering,” IEEE Trans. Mol. Biol. Multi-Scale Commun., vol. 2, no. 1, pp. 64–91, 2016.
[27] Ghane, Mahdi, Zhen Gao, Mogens Blanke, and Torgeir Moan. ”On the joint distribution of excursion duration and amplitude of a narrow-band Gaussian process.” Ieee Access 6 (2018): 15236-15248.
[28] H. Il Suk, Y. Wang, and S. W. Lee, “Incremental sparse Pseudo-input Gaussian process regression,” Int. J. Pattern Recognit. Artif. Intell., vol. 26, no. 8, 2012.
[29] I. Bal and E. Es, “Sparse Gaussian Processes for Large-Scale Machine Learning,” p. 12, 2010.
[30] N. Lawrence, M. Seeger, and R. Herbrich, “Fast sparse Gaussian process methods: The informative vector machine,” Adv. Neural Inf. Process. Syst., 2003.
[31] M. A. Álvarez and N. D. Lawrence, “Sparse Convolved Multiple Output Gaussian Processes,” pp. 1–33, 2009.
[32] M. Lázaro-Gredilla, J. Com, C. E. Rasmussen, and A. U. Es, “Sparse Spectrum Gaussian Process Regression Joaquin Quiñonero-Candela Aníbal R. Figueiras-Vidal,” J. Mach. Learn. Res., vol. 11, pp. 1865–1881, 2010.
[33] L. Wernisch, “Analysing gene expression data using Gaussian Processes,” vol. 1, no. 1, 2003.
[34] H. Bijl, J. W. van Wingerden, T. B. Schön, and M. Verhaegen, “Online sparse Gaussian process regression using FITC and PITC approximations,” IFAC-PapersOnLine, vol. 48, no. 28, pp. 703–708, 2015.
[35] M. J. Hosseini and S. I. Lee, “Learning sparse Gaussian graphical models with overlApping Blocks,” Adv. Neural Inf. Process. Syst., no. Nips, pp. 3808–3816, 2016.
[36] D. A. Díaz-Uriarte Ramón and S. Alvarez, “Gene selection and classification of microarray data using random forest,” BMC Bioinformatics, vol. 7, no. 1, p. 3, 2006.