Selected Publications in Life Science

For those in computer science, let me remark that more than five author are typically listed in a life science paper because it is quite essential to collect and examine raw data from precious samples, and many excellent people have to be involved in these data collection steps. "*" indicates joint corresponding authors, while "#" joint first authors.

1. Koichiro Doi, Taku Monjo, Pham H. Hoang, Jun Yoshimura, Hideaki Yurino, Jun Mitsui, Hiroyuki Ishiura, Yuji Takahashi, Yaeko Ichikawa, Jun Goto, Shoji Tsuji, <u>Shinichi Morishita</u>: Rapid detection of expanded short tandem repeats in personal genomics using hybrid sequencing. *Bioinformatics* 30(6): 815-822 (2014)

2. Saito TL, Hashimoto SI, Gu SG, Morton JJ, Stadler M, Blumenthal T, Fire A*, <u>Morishita S*.</u> The Transcription Start Site Landscape of *C. elegans*. *Genome Research* Aug;23(8):1348-61 (2013)

3. Qu W, Hashimoto S, Shimada A, Nakatani Y, Ichikawa K, Saito TL, Ogoshi K, Matsushima K, Suzuki Y, Sugano S, Takeda H, <u>Morishita S.</u> Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. *Genome Research*, 22(8):1419-25 (2012)

4. Sasaki, S.#, Mello, C. C.#, Shimada, A., Nakatani, Y., Hashimoto, S., Ogawa, M., Matsushima, K., Gu, S. G., Kasahara, M., Ahsan, B., Sasaki, A., Saito, T., Suzuki, Y., Sugano, S., Kohara, Y., Takeda, H., Fire, A*. and <u>Morishita, S*</u>. Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* 323, 401-4. (2009)

5. Qu W, Hashimoto S, <u>Morishita S..</u> Efficient frequency-based de novo short read clustering for error trimming in next-generation sequencing. *Genome Research* 19(7): 1309-1315 (2009)

6. Kasahara M#, Naruse K#, Sasaki S#, Nakatani Y#, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y, Jindo T, Kobayashi D, Shimada A, Toyoda A, Kuroki Y, Fujiyama A, Sasaki T, Shimizu A, Asakawa S, Shimizu N, Hashimoto S, Yang J, Lee Y, Matsushima K, Sugano S, Sakaizumi M, Narita T, Ohishi K, Haga S, Ohta F, Nomoto H, Nogata K, Morishita T, Endo T, Shin-I T, Takeda H*, <u>Morishita S*</u>, Kohara Y*. The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714-719 (2007)

7. Nakatani Y, Takeda H, Kohara Y, <u>Morishita S.</u> Reconstruction of the Vertebrate Ancestral Genome Reveals Dynamic Genome Reorganization in Early Vertebrates. *Genome Research* 17(9): 1254-1265 (2007)

8. Ohya Y*, Sese J, Yukawa M, Sano F, Nakatani Y, Saito TL, Saka A, Fukuda T, Ishihara S, Oka S, Suzuki G, Watanabe M, Hirata A, Ohtani M, Sawai H, Fraysse N, Latgé JP, François JM, Aebi M, Tanaka S, Muramatsu S, Araki H, Sonoike K, Nogami S, <u>Morishita S*</u>. High-dimensional and large-scale phenotyping of yeast mutants *Proc Natl Acad Sci U S A.* 102(52):19015-20 (2005)

9. Yamada T, <u>Morishita S.</u> Accelerated off-target search algorithm for siRNA. *Bioinformatics*, 21(8):1316-1324 (2005)

10. Taro L. Saito, Jun Sese, Yoichiro Nakatani, Fumi Sano, Masashi Yukawa, Yoshikazu Ohya, and <u>Shinichi Morishita</u> Data Mining Tools for the Saccharomyces cerevisiae Morphological Database. *Nucleic Acids Research* 33(Web Server issue):W589-W591 (2005)

11. Yuki Naito, Tomoyuki Yamada, Takahiro Matsumiya, Kumiko Ui-Tei, Kaoru Saigo*, and <u>Shinichi Morishita*.</u> dsCheck: highly sensitive off-target search software for dsRNA-mediated RNA interference. *Nucleic Acids Research* 2005 33(Web Server issue):W753-W757

12. Jun Sese, Yukinori Kurokawa, Morito Monden, Kikuya Kato and <u>Shinichi Morishita</u>. Constrained

Clusters of Gene Expression Profiles with Pathological Features, *Bioinformatics*, 20: 3137 – 3145, 2004.

13. Yuki Naito#, Tomoyuki Yamada#, Kumiko Ui-Tei, <u>Shinichi Morishita*</u>, and Kaoru Saigo*. siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference. *Nucl. Acids Res.* 2004 32: W124-129.

14. Tomoyuki Yamada and <u>Shinichi Morishita</u>. Computing Highly Specific and Noise Tolerant Oligomers Efficiently, *Journal of Bioinformatics and Computational Biology*, Vol. 2, No. 1, pp21-46, 2004.

15. Taro L. Saito, Miwaka Ohtani, Hiroshi Sawai, Fumi Sano, Ayaka Saka, Daisuke Watanabe, Masashi Yukawa, Yoshikazu Ohya, and <u>Shinichi Morishita</u>. SCMD: *Saccharomyces Cerevisiae* Morphological Database. *Nucl. Acids. Res.* 32: D319-D322, 2004.

16. Miwaka Ohtani, Ayaka Saka, Fumi Sano, Yoshikazu Ohya, and <u>Shinichi Morishita.</u> Development of Image Processing Program for Yeast Cell Morphology, *Journal of Bioinformatics and Computational Biology* Vol. 1, No. 4 695-709, 2004.

17. Jun Ogasawara and <u>Shinichi Morishita</u>. Fast and Sensitive Algorithm for Aligning ESTs to Human Genome. *Journal of Bioinformatics and Computational Biology*, Vol.1, No.2, 363-386, 2003.

18. Toshihiko Honkura, Jun Ogasawara, Tomoyuki Yamada and <u>Shinichi Morishita</u>. The Gene Resource Locator: gene locus maps for transcriptome analysis. *Nucleic Acids Research*, 2002, Vol. 30, No. 1 221-225.

Selected Publications in Computer Science

For those in life science, let me note that no more than five authors are involved in a computer science paper presumably because we need only pencils, memo pad, and computers to do research. In computer science, we first submit a paper to a conference to exchange our idea timely and then submit a revised version to a journal. It typically takes more than one year to publish a paper in a journal after its submission.

1. Kazuki Ichikawa, <u>Shinichi Morishita</u>: A Simple but Powerful Heuristic Method for Accelerating k-Means Clustering of Large-Scale Data in Life Science. *IEEE/ACM Trans. Comput. Biology Bioinform.* 11(4): 681-692 (2014)

2. Taro L. Saito and <u>Shinichi Morishita</u>. Relational-Style XML Query. *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data (ACM SIGMOD)*, Vancouver, 303-314 (2008)

3. Yasuhiko Morimoto, Hiromu Ishii, and <u>Shinichi Morishita</u>. Efficient Construction of Regression Trees with Range and Region Splitting. *Machine Learning*, Kluwer Academic, 45, pages, 235-259, 2001. (Conference version: *Proceedings of VLDB '97*, pages 166-175, 1997)

4. Takeshi Fukuda, Yasuhiko Morimoto, <u>Shinichi Morishita</u> and Takeshi Tokuyama. Data Mining with Optimized Two-Dimensional Association Rules. *ACM Transactions on Database Systems (TODS)*, Volume 26 , Issue 2, pp. 179 – 213, June 2001. (Conference version: *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data (SIGMOD'96)*, pp. 13-23, 1996)

5. <u>Shinichi Morishita</u> and Jun Sese. Traversing Itemset Lattices with Statistical Metric Pruning. *Proc. of ACM SIGACT-SIGMOD-SIGART Symp. on Database Systems (PODS)*. pp.226-236, May 2000.

6. <u>Shinichi Morishita</u>. Computing Optimal Hypotheses Efficiently for Boosting. *Progress in Discovery Science*, Springer, 471-481, 2000.

7. Takeshi Fukuda, Yasuhiko Morimoto, <u>Shinichi Morishita</u>, and Takeshi Tokuyama. Mining Optimized

Association Rules for Numeric Attributes. ***Journal of Computer and System Sciences*** – Special issue on the ***15th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems***, Academic Press. 58(1): 1-12, 1999.

(Conference version: ***15th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems***, pp. 182-191, 1996)

8. Takeshi Fukuda, Yasuhiko Morimoto, Shinichi Morishita, and Takeshi Tokuyama. Constructing Efficient Decision Trees by Using Optimized Numeric Association Rules. ***Proceedings of VLDB '96***, pages 146-155, Bombay, India, September 1996.

9. Kazukuni Yoda, Takeshi Fukuda, Yasuhiko Morimoto, Shinichi Morishita, and Takeshi Tokuyama. Computing Optimized Rectilinear Regions for Association Rules. ***Proceedings of the Third Conference on Knowledge Discovery and Data Mining (KDD'97)***, pages 96-103, Los Angels, August 1997.

10. Shinichi Morishita. Avoiding Cartesian Products for Multiple Joins. ***Journal of the ACM***, Volume 44, Number 1, pp. 57-85, January 1997.

(Conference version. ***Proceedings of 11th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems***, pp. 368-379, 1992)

11. Shinichi Morishita. An Extension of Van Gelder's Alternating Fixpoint to Magic Programs. ***Journal of Computer and System Sciences*** – Special Issue on the ***12th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems***, Academic Press, Volume 52, Number 3, pp. 506-521, June 1996

(Conference version: ***12th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems***, pp. 123-134, 1993)

12. Marcia A. Derr, Shinichi Morishita, and Geoffrey Phipps. The Glue-Nail Deductive Database System: Design, Implementation, and Evaluation. ***VLDB Journal***. 3(2):123-160, 1994

(Conference version: Design and Implementation of the Glue-Nail Database System'. ***Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data (SIGMOD'93)***, pp. 147-156, Washington DC, May 1993)