The Human Immunopeptidome Project, a Suggestion for yet another Postgenome Next Big Thing

Arie Admon‡§ and Michal Bassani-Sternberg‡

The time is ripe for staging the Human Immunopeptidome Project, whose goal is to analyze the full repertoires of peptides bound to the HLA molecules, in both health and disease. Mass spectrometry technologies have matured to enable comprehensive analyses of both the membrane-bound and the plasma soluble immunopeptidomes associated with each of the HLA allomorphs and the different diseases. The expected outcomes of such project will include basic understanding of the molecular mechanisms involved with formation of immunopeptidomes, correlating them with their source cellular proteomes, definition of both the consensus motifs and the scope of each allomorphs-specific immunopeptidomes, and most importantly, identification of disease-related HLA peptides, which may eventually serve as biomarkers or immunotherapeutics. Ideally, the Human Immunopeptidome Project will become public and the gathered data will be shared, as soon as possible. Other immunopeptidome projects, of other animals, will follow suit. ***Molecular & Cellular Proteomics 10: 10.1074/mcp.O111.011833, 1–4, 2011.***

Following the completion of the Human Genome Project (1, 2) and the recent advances in molecular immunology and mass spectrometry (3), we believe that the time is ripe for launching yet another postgenome Grand Challenge, namely the Human Immunopeptidome Project. The immunopeptidome is the assortment of thousands of peptides displayed by major histocompatibility complex (MHC)[1] molecules (called in humans: the human leukocyte antigen, HLA). Cells present the MHC molecules with their bound immunopeptidomes at their surface to enable scrutiny of the health-state of the cells by the immune system's circulating T lymphocytes, reviewed in (4). A portion of the HLA molecules transported to the cell surface are released to the plasma with their bound peptides, resulting in the formation of the parallel plasma-soluble HLA peptidome (sHLA). The plasma sHLA peptidomes resemble to a large extent the membranal peptidomes (mHLA) presented at the cells' surface (5, 6). Both of these HLA immunopeptidomes are composed of peptide degradation products of the cellular proteomes, and thus mirror the schemes of protein degradation within their cells of origin (7, 8). The HLA gene cluster is the most polymorphic in the human genome, affecting mostly the peptide binding pockets of the HLA molecules. Therefore, different HLA allomorphs differ in their presented peptide repertoires (4) and each person has a unique haplotype-specific immunopeptidome (9). However, it is thought that the different HLA-I allomorphs of the human population (more than 3000 currently known) can be grouped into a small number of supertypes, each binding subsets of peptides with a related consensus sequence motifs (10, 11). The definition of the HLA peptidomes, presented by specific cells or by each of the HLA allomorphs, was much simplified by the introduction of mass spectrometry to this field (12, 13). The sequence motifs of each HLA can be studied relatively easy, using the pools of peptides recovered from transfected soluble HLA molecules collected from the growth media of cultured cells expressing the sHLA allomorphs without their anchor transmembrane domains (14) and even as tagged sHLA molecule (15), (reviewed in (16)).

Disease Associated HLA Peptides—When cells become sick they produce and degrade disease proteins, resulting in the eventual display of some of degradation products as HLA peptides. Thus, HLA peptidomes were studied extensively as a source for vaccine candidates, while looking for peptides capable of providing protective immunity toward pathogens (17) and cancer, reviewed in (18). Self HLA peptides, conferring undesired anti-self-immune reaction in autoimmune and inflammatory diseases, were also searched for as potential agents for suppression of the pathogenic auto-reactive T cells (19). For example, carriers of HLA B*27:05 have the propensity to develop spondylarthritides (20), (reviewed in (19)). Of special resent interest are the findings, indicating that HLA B*57:01, B*27:05, B*14/Cw*08:02, B*52, and A*25 alleles confer the ability to control HIV infection (21), which implicated specific peptides ligands of these alleles with the ability to elicit effective immune response to the virus.

Immunopeptidomics is even becoming useful for diagnostics. In many diseases, such as cancer, viral, bacterial and

From the ‡Faculty of Biology, Technion, Haifa, Israel

Received June 9, 2011
Published, MCP Papers in Press, August 2, 2011, DOI 10.1074/mcp.O111.011833

[1] The abbreviations used are: MHC, major histocompatibility complex; HLA, human leukocyte antigen.

parasite infections, autoimmunity, inflammation and allergy, the affected cells release larger than normal amounts of sHLA molecules to the plasma, reviewed in (22, 23). Therefore, in correlation with the severity of the disease, larger portions of the plasma sHLA immunopeptidomes originate from the disease cells, implying that the sHLA immunopeptidomes can serve as a source for disease biomarkers from this readily available body fluid (5) (reviewed in (6)).

Immunopeptidome Analysis—Analysis of immunopeptidomes makes use of the same tools used for proteomics analyses: capillary reversed-phase chromatography and tandem mass spectrometry (μLC-MS/MS)(13). It begins with extraction of the bound peptides from immunoaffinity purified mHLA or sHLA molecules. Recombinant soluble HLA molecules, lacking their trans-membrane domain, are recovered from the cells' growth medium (14, 24–26). The recombinant soluble HLA molecules can also be tagged with a specific peptide epitopes, which facilitate their selective purification with antitag monoclonal antibodies. This way, immunopeptidomes associated with any HLA allomorph and any desired cultured cell line, can be studied in details (15), reviewed in (16). Furthermore, complex immunopeptidomes can also be defined directly from human plasma, without having to extract the HLA molecules from the tissues. Starting from just a few milliliters of blood, the plasma-soluble HLA molecules can be immunoaffinity purified in sufficient amounts for large-scale analysis (5).

Because the sHLA peptidomes of people are relatively stable, searching for disease associated peptides can be simplified by comparing the patients' sHLA peptidomes before and after treatment, which significantly reduces the tumor load, and again in the unfortunate cases when the disease recurs. HLA peptides correlating in their amounts with the severity of the disease are likely associated with the disease (5, 6).

The Immunopeptidome is a Mirror of the Proteome—The immunopeptidomes reflect the schemes of protein degradation in the cells. A significant portion of MHC peptides are derived from rapidly degrading and short lived proteins (27, 28). It was even suggested that some HLA peptides are encoded by alternative reading frames (reviewed in (29, 30)) and even result from protein splicing (31, 32). Furthermore, proteins that are rapidly degraded within the cells, and are difficult to detect by standard proteomics approaches, can be traced through their resulting MHC peptides (8). Short lived proteins or products of defective ribosome products are the source a significant portion of the MHC peptidome (28), (reviewed in (27)). Thus, analysis of the immunopeptidome may become significant as a rich source of information about the cells' metabolism and regulatory processes. The effects of outside stimuli, such as cytokines and hormones signaling, heat and cold shock stresses, metal poisoning, and oncogenic transformation, were extensively studied in relation to their modulation of the transcriptome and proteome of the cells. Such factor influence also the schemes of protein degradation within the cells (degradome) which can be studied through its products, the MHC peptidomes (reviewed in (7)).

The Human Immunopeptidome Project—Thus, the Human Immunopeptidome Project will attempt to analyze, on a grand scale, the repertoires of peptides associated with the different human diseases, growth conditions, and cell types, and the unique peptidomes bound by each of the HLA allomorphs. The soluble plasma sHLA immunopeptidomes can be extracted and prepared for analysis at the blood collection sites and because the extracted pools of peptides are rather stable, they can be collected anywhere in the world and sent for analysis to specialized mass spectrometry labs. To associate between the immunopeptidomes and the HLA allomorphs presenting them, each allomorph should be expressed in model cultured cells, and the extracted sHLA bound peptides can be similarly identified. It is clear that the expected large data sets will require new bioinformatics tools and large repositories to allow deposition and retrieval of the data, including sequences and relative quantities.

Ideally, the Gathered Data will be Shared and Open—Similarly to the public Human Genome Project (1), the Human Immunopeptidome Project may be open for data retrieval by all interested and for data deposition by the collaborating laboratories. This requires that the acquired information should become public as soon as possible, possibly as was done by the Human Genome Project (33, 34). Possibly less ideally, yet more morally correct, the data will become public only after securing intellectual property rights, to ensure the potential to develop immunotherapeutics by pharmaceutical companies. All efforts should be made to acknowledge and respect the contribution of the depositors (35, 36). Access can be provided both to the raw data and to the analyzed results. Provisions should be made for both secure deposition and for downloading of data, while maintaining patients' anonymity. Furthermore, data will be accepted for deposition only from laboratories adhering to strict ethics requirements, as detailed and reviewed in (37). Attempts should be made to analyze the immunopeptidomes of people of both sexes, and people belonging to diverse ethnic groups, to expand the analyses to rare diseases and to encourage collaboration from different regions of the world. Furthermore, such Immunopeptidome Project may both benefit and contribute to the other post genome Grand Challenges, such as the Human Proteome Project (http://www.hupo.org/research/hpp).

§ To whom correspondence should be addressed: Faculty of Biology, Technion, Haifa, Israel. Tel.: 972-48293407; Fax: 972-48225153; E-mail: admon@tx.technion.ac.il.

REFERENCES

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A.,

Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I., Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., and Chen, Y. J. (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921

2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V. D., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R.-R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z. Y., Wang, A., Wang, X., Wang, J., Wei, M.-H., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S. C., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y.-H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigó, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.-H., Coyne, M., Dahlke, C., Mays, A. D., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) The Sequence of the Human Genome. *Science* **291,** 1304–1351
3. Nilsson, T., Mann, M., Aebersold, R., Yates, J. R., Bairoch, A., and Bergeron, J. J. M. (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Meth.* **7,** 681–685
4. Shastri, N., Schwab, S., and Serwold, T. (2002) Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol.* **20,** 463–493
5. Bassani-Sternberg, M., Barnea, E., Beer, I., Avivi, I., Katz, T., and Admon, A. (2010) Soluble plasma HLA peptidome as a potential source for cancer biomarkers. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 18769–18776
6. Hickman, H. D., and Yewdell, J. W. (2010) Mining the plasma immunopeptidome for cancer peptides as biomarkers and beyond. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 18747–18748
7. Admon, A., Barnea, E., and Ziv, T. (2003) Tumor antigens and proteomics from the point of view of the major histocompatibility complex peptides. *Mol. Cell. Proteomics* **2,** 388–398
8. Milner, E., Barnea, E., Beer, I., and Admon, A. (2006) The turnover kinetics of major histocompatibility complex peptides of human cancer cells. *Mol. Cell. Proteomics* **5,** 357–365
9. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., and Rammensee, H. G. (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351,** 290–296
10. Sidney, J., Peters, B., Frahm, N., Brander, C., and Sette, A. (2008) HLA class I supertypes: a revised and updated classification. *BMC Immunol.* **9,** 1
11. Sette, A., and Sidney, J. (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* **50,** 201–212
12. Rotzschke, O., Falk, K., Deres, K., Schild, H., Norda, M., Metzger, J., Jung, G., and Rammensee, H.-G. (1990) Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells. *Nature* **348,** 252–254
13. Hunt, D. F., Henderson, R. A., Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A. L., Appella, E., and Engelhard, V. H. (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255,** 1261–1263
14. Prilliman, K., Lindsey, M., Zuo, Y., Jackson, K. W., Zhang, Y., and Hildebrand, W. (1997) Large-scale production of class I bound peptides: assigning a signature to HLA-B*1501. *Immunogenetics* **45,** 379–385
15. Hickman, H. D., Batson, C. L., Prilliman, K. R., Crawford, D. L., Jackson, K. L., and Hildebrand, W. H. (2000) C-terminal epitope tagging facilitates comparative ligand mapping from MHC class I positive cells. *Hum. Immunol.* **61,** 1339–1346
16. Hoppes, R., Ekkebus, R., Schumacher, T. N., and Ovaa, H. (2010) Technologies for MHC class I immunoproteomics. *J. Proteomics* **73,** 1945–1953
17. Townsend, A. R. M., Rothbard, J., Gotch, F. M., Bahadur, G., Wraith, D., and McMichael, A. J. (1986) The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell* **44,** 959–968
18. Purcell, A. W., McCluskey, J., and Rossjohn, J. (2007) More than one reason to rethink the use of peptides in vaccine design. *Nat. Rev. Drug*

Discov. **6,** 404–414
19. Marcilla, M., and López de Castro, J. A. (2008) Peptides: the cornerstone of HLA-B27 biology and pathogenetic role in spondyloarthritis. *Tissue Antigens* **71,** 495–506
20. Benjamin, R., and Parham, P. (1990) Guilt by association: HLA-B27 and ankylosing spondylitis. *Immunol. Today* **11,** 137–142
21. The International HIV Controler Study (2010) The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. *Science* **330,** 1551–1557
22. Campoli, M., and Ferrone, S. (2008) Tumor escape mechanisms: potential role of soluble HLA antigens and NK cells activating ligands. *Tissue Antigens* **72,** 321–334
23. Tabayoyong, W. B., and Zavazava, N. (2007) Soluble HLA revisited. *Leuk. Res.* **31,** 121–125
24. Barnea, E., Beer, I., Patoka, R., Ziv, T., Kessler, O., Tzehoval, E., Eisenbach, L., Zavazava, N., and Admon, A. (2002) Analysis of endogenous peptides bound by soluble MHC class I molecules: a novel approach for identifying tumor-specific antigens. *Eur. J. Immunol.* **32,** 213–222
25. Corr, M., Boyd, L. F., Frankel, S. R., Kozlowski, S., Padlan, E. A., and Margulies, D. H. (1992) Endogenous peptides of a soluble major histocompatibility complex class I molecule, H-2Lds: sequence motif, quantitative binding, and molecular modeling of the complex. *J. Exp. Med.* **176,** 1681–1692
26. Margulies, D. H., Ramsey, A. L., Boyd, L. F., and McCluskey, J. (1986) Genetic engineering of an H-2Dd/Q10b chimeric histocompatibility antigen: purification of soluble protein from transformant cell supernatants. *Proc. Natl. Acad. Sci. U.S.A.* **83,** 5252–5256
27. Yewdell, J. W. (2007) Plumbing the sources of endogenous MHC class I peptide ligands. *Current Opin.Immunol.* **19,** 79–86
28. Yewdell, J. W., Anton, L. C., and Bennink, J. R. (1996) Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J. Immunol.* **157,** 1823–1826
29. Starck, S. R., Ow, Y., Jiang, V., Tokuyama, M., Rivera, M., Qi, X., Roberts, R. W., and Shastri, N. (2008) A distinct translation initiation mechanism generates cryptic peptides for immune surveillance. *PLoS One* **3,** e3460
30. Yewdell, J. W., and Hickman, H. D. (2007) New lane in the information highway: alternative reading frame peptides elicit T cells with potent antiretrovirus activity. *J. Exp. Med.* **204,** 2501–2504
31. Hanada, K., Yewdell, J. W., and Yang, J. C. (2004) Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427,** 252–256
32. Vigneron, N., Stroobant, V., Chapiro, J., Ooms, A., Degiovanni, G., Morel, S., van der Bruggen, P., Boon, T., and Van den Eynde, B. J. (2004) An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304,** 587–590
33. Collins, F. S., Green, E. D., Guttmacher, A. E., and Guyer, M. S. (2003) A vision for the future of genomics research. *Nature* **422,** 835–847
34. Collins, F. S., Morgan, M., and Patrinos, A. (2003) The Human Genome Project: Lessons from Large-Scale Biology. *Science* **300,** 286–290
35. The Wellcome Trust 2003: Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy-communications
36. Birney, E., Hudson, T. J., Green, E. D., Gunter, C., Eddy, S., Rogers, J., *et al.* (2009) Prepublication data sharing. *Nature* **461,** 168–170
37. Cambon-Thomsen, A. (2004) The social and ethical issues of post-genomic human biobanks. *Nat. Rev. Genet.* **5,** 866–873