Brief Communication Arising

Reverse engineering of the cancer genome

Matthias Wjst

The sequence of a cancer genome may reveal one of the most fascinating experiments in biology where the hostile take over of a cell clone is achieved by properties acquired by new mutations. Two recent Nature reports [1, 2] succeed to provide an inventory of mutations but largely miss the timeline of events, leading to a questionable public health message.

First, tumourigenesis does not start with the fertilised egg [1] but with oogenesis and spermatogenesis. Although this distinction may involve only a few hundred mutations [3] it will have large scale evolutionary and medical consequences [4] as these mutations are being present in all body cells.

The authors further assume rather linear mutation rates by stating that more heterozygous than homozygous mutations in a given chromosomal area are indicating an early loss and duplication event [1]. Even if we assume that such an increased frequency of homozygous mutations is not just a chance finding, local mutation rates may show accelerating or decelerating rates depending on driver mutations, ultimately making this speculation from a single sample premature. By putting the emphasis on ultraviolet light or smoking induced DNA damage, the most frequent cause of mutation - replication error by proliferation - is being downplayed. Somatic substitution with excess of C>T followed by C>A (melanoma) or C>A followed by C>T (lung cancer) may be part of an overall distribution that could approximate the overall SNP distribution in the human genome [5].

Lastly, the extrapolation of 365,000 cigarettes smoked in 50 years and induction of 22,910 mutations ("15 cigarettes leading to one mutation") [2] is an odd biological scenario that gives also a distorted public health message. The sequenced tumour clone may have acquired decisive mutations long before being exposed to smoke during normal lung development [3] and it is far from being clear to what extent the direct tumour progenitor cells have been exposed

to carcinogenic substances [6]. In line with that comes the epidemiological observation that the smoking risk of young lung cancer patients is nearly identical to old lung cancer patients when calculated for the exposure period prior to the diagnosis of cancer [7]. In other words, some may get lung cancer from being exposed to a few hundred cigarettes, while others will never get lung cancer despite a million of cigarettes - clearly a situation where averaging risks does not make any sense.

The absolute number of mutations in a fully dedifferentiated cell line may even be unimportant while the "minimal cancer genome" of driver mutations [8] will be interesting. Highly parallel sequencing will allow to examine repeated samples from the same biopsy sites [3, 9] ultimately allowing a reverse engineering of cancer cells.

References

1 Pleasance, E.D. *et al.*, A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* (2009).

2 Pleasance, E.D. *et al.*, A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* (2009).

3 Greaves, M., *Cancer The evolutionary legacy*. (Oxford University Press, Oxford, 2000).

4 Crow, J.F., Spontaneous mutation in man. *Mutat Res* 437 (1), 5-9 (1999).

5 Sachidanandam, R. *et al.*, A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409 (6822), 928-933 (2001).

6 Maley, C.C. *et al.*, Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 38 (4), 468-473 (2006).

7 Kreuzer, M. *et al.*, Risk factors for lung cancer in young adults. *American Journal of Epidemiology* 147 (11), 1028-1037 (1998).

8 Stratton, M.R., Campbell, P.J., & Futreal, P.A., The cancer genome. *Nature* 458 (7239), 719-724 (2009).

9 Thiberville, L. *et al.*, Evidence of cumulative gene losses with progression of premalignant epithelial lesions to carcinoma of the bronchus. *Cancer Res* 55 (22), 5133-5139 (1995).