

**Istituto Metodi Quantitativi - Università L. Bocconi**

Viale Isonzo, 25 - 20135 Milano

Tel. 02-58365632 - Fax 02-58365630

## ***SEMINARIO***

# **"Gene therapy as a new field for statistical modelling"**

**Clelia Di Serio**

**(Centro Universitario di Statistica per le Scienze Biomediche)**

**Università Vita-Salute San Raffaele, Milano**

**Giovedì, 8 Marzo 2007 - ore 16.30**

**Aula 137 - Viale Isonzo 25 - 20135 Milano**

---

### Abstract:

Gene therapy is a form of molecular medicine which treats genetic diseases by replacing a defective gene, responsible for the pathology, with a functional one by means of a vector which is commonly a virus. We here show how gene therapy represents a challenging and modern applied field for developing statistical strategies. In particular we present three different examples of statistical modelling to address three different crucial aspects of gene therapy.

#### 1) STUDY OF THE INTEGRATION DISTANCE DISTRIBUTION.

Tumorigenesis may result from mutations caused by integration of the therapeutic viral vector into the host genome. Depending on whether the virus integrates into or in the vicinity of oncogenes normal transcription can be disrupted, thus accelerating tumors. This process is called insertional mutagenesis.

Statistical modelling is here aimed at understanding the integration process, in terms of "where" and "how" the retrovirus integrates over the whole genome. We provide a new statistical approach to the problem of evaluation of "integration events" position. First we searched for a precise statistical definition of the hypothesis of "random distribution" of integration and for the alternative hypothesis. Then we model the exact distribution of the integration distance, regardless to gene length by using a Beta distribution. We show how this is appropriate and how this allows us to include all the gene orientation.

#### 2) SURVIVAL ANALYSIS TO IDENTIFY LONG AND SHORT SURVIVORS IN ANIMAL MODELS OF IDENTICALLY INBRED MICE.

In the application considered here identically inbred mice are made tumor prone by knocking out the oncosuppressor related gene *Cdkn2a*. These mice develop a variety of tumors with a predictable onset time of 300 days. Bone marrow cells are then extracted from the mice and different vectors are inoculated in them. These cells are transplanted back in the mice, whose survival is then observed.

One goal of this study is to investigate the influence of some covariates related to the integration process on the survival of mice. Survival of the mice is here modelled by means of a Weibull distribution. This distribution actually resulted as the better fit to describe the behaviour of mice stratified in long survivors and short survivors. Unobservable heterogeneity was modelled by including a frailty component in the hazard function.

#### 3) STATISTICAL ANALYSIS OF REAL-TIME PCR DATA

Real-time PCR is one of the most sensitive and reliably quantitative methods for gene expression analysis which is broadly applied in biomedical sciences. However data processing and statistical procedures for the analysis of real-time PCR data are still lacking. We show an example of data from genetic medicine to compare in terms of DNA amplification two different cells which are possible candidates for tumor acceleration.

All the studies are performed in collaboration with Telethon Institute of Genetic Therapy (TIGET).

Reference: [www.cussb.univr.it](http://www.cussb.univr.it)