A New Animal Model in the Study of UCB Metabolism and Neurotoxicity

Dottorando: Giulia Bortolussi

Coordinatore del Collegio Docenti: Prof. Franco Vittur
Università degli Studi di Trieste

Relatore: Prof. Claudio Tiribelli
Università degli Studi di Trieste

Correlatore: Dott. Andrés Fernando Muro
ICGEB, Trieste

ANNO ACCADEMICO 2007-2008
Ai miei genitori e a Gabriele,
che sono la mia forza.
### Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>5</td>
</tr>
<tr>
<td><strong>RIASSUNTO</strong></td>
<td>7</td>
</tr>
<tr>
<td><strong>LIST OF ABBREVIATIONS</strong></td>
<td>9</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>11</td>
</tr>
<tr>
<td><strong>BILIRUBIN AND DISEASE: AN OVERVIEW</strong></td>
<td>12</td>
</tr>
<tr>
<td><strong>FORMATION OF UnCONJUGATED BILIRUBIN</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>UCB TOXICITY: NEONATAL JAUNDICE, BIND AND KERNICTERUS</strong></td>
<td>14</td>
</tr>
<tr>
<td>NEONATAL JAUNDICE</td>
<td>15</td>
</tr>
<tr>
<td>BIND (BILIRUBIN INDUCED NEUROLOGICAL DYSFUNCTION)</td>
<td>16</td>
</tr>
<tr>
<td>KERNICTERUS</td>
<td>17</td>
</tr>
<tr>
<td><strong>UCB GLUCURONIDATION</strong></td>
<td>17</td>
</tr>
<tr>
<td>The UGT family</td>
<td>18</td>
</tr>
<tr>
<td><strong>UGT1A1 DEFICIENCIES: CRIGLER-NAJJAR SYNDROME AND GILBERT’S SYNDROME</strong></td>
<td>20</td>
</tr>
<tr>
<td><strong>ANIMAL MODEL FOR UGT DEFICIENCY: THE GUNN RAT</strong></td>
<td>21</td>
</tr>
<tr>
<td><strong>GENE THERAPY FOR CRIGLER-NAJJAR DISEASE</strong></td>
<td>24</td>
</tr>
<tr>
<td><strong>GENETIC MANIPULATION: AN OVERVIEW</strong></td>
<td>26</td>
</tr>
<tr>
<td><strong>MICE AS A GOOD ANIMAL MODEL</strong></td>
<td>26</td>
</tr>
<tr>
<td><strong>GENE TARGETING</strong></td>
<td>27</td>
</tr>
<tr>
<td>Inactivating Genes: Knockout</td>
<td>28</td>
</tr>
<tr>
<td>Basic strategies to generate mouse models of knockout genes</td>
<td>28</td>
</tr>
<tr>
<td>Inserting Genes: Knockin</td>
<td>35</td>
</tr>
<tr>
<td>Basic strategy to generate mouse models with knockin thechnique</td>
<td>36</td>
</tr>
<tr>
<td><strong>AIM OF THE STUDY</strong></td>
<td>41</td>
</tr>
<tr>
<td><strong>MATERIALS &amp; METHODS</strong></td>
<td>42</td>
</tr>
<tr>
<td><strong>CHEMICAL REAGENTS</strong></td>
<td>43</td>
</tr>
<tr>
<td><strong>STANDARD SOLUTIONS</strong></td>
<td>43</td>
</tr>
<tr>
<td><strong>ENZYMES</strong></td>
<td>43</td>
</tr>
<tr>
<td><strong>SYNTHETIC OLIGONUCLEOTIDES</strong></td>
<td>44</td>
</tr>
<tr>
<td><strong>RADIOACTIVE ISOTOPES</strong></td>
<td>44</td>
</tr>
<tr>
<td><strong>BACTERIAL CULTURE MEDIA AND STRAIN</strong></td>
<td>44</td>
</tr>
<tr>
<td><strong>MICE</strong></td>
<td>44</td>
</tr>
<tr>
<td><strong>PREPARATION OF DNA</strong></td>
<td>45</td>
</tr>
<tr>
<td>Preparation of genomic DNA from tissue samples</td>
<td>45</td>
</tr>
<tr>
<td>Preparation of genomic DNA from mouse tail biopsies</td>
<td>45</td>
</tr>
<tr>
<td>Preparation of genomic DNA from p96 multiwells</td>
<td>46</td>
</tr>
<tr>
<td>Small-scale preparation of plasmid DNA from bacterial cultures</td>
<td>46</td>
</tr>
<tr>
<td>Large-scale preparations of plasmid DNA from bacterial cultures</td>
<td>47</td>
</tr>
<tr>
<td><strong>ESTIMATION OFNUCLEIC ACID CONCENTRATION AND QUALITY</strong></td>
<td>47</td>
</tr>
<tr>
<td>Spectrophotometric analysis</td>
<td>47</td>
</tr>
<tr>
<td>UV fluorescence of inter-calated ethidium bromide</td>
<td>48</td>
</tr>
<tr>
<td><strong>ELECTROPHORETIC SEPARATION OF NUCLEIC ACID</strong></td>
<td>48</td>
</tr>
<tr>
<td>Agarose gel for DNA separation</td>
<td>48</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Elution and purification of DNA fragments from gels</td>
<td>49</td>
</tr>
<tr>
<td>Enzymatic modification of DNA</td>
<td>49</td>
</tr>
<tr>
<td>Restriction enzymes</td>
<td>49</td>
</tr>
<tr>
<td>Large fragment of E. coli polymerase I and T4 polynucleotide kinase</td>
<td>50</td>
</tr>
<tr>
<td>T4 DNA ligase</td>
<td>50</td>
</tr>
<tr>
<td>Amplification of selected DNA fragments</td>
<td>51</td>
</tr>
<tr>
<td>Preparation of bacterial competent cells</td>
<td>53</td>
</tr>
<tr>
<td>Transformation of bacteria</td>
<td>53</td>
</tr>
<tr>
<td>Sequence analysis for cloning purpose</td>
<td>53</td>
</tr>
<tr>
<td>Construction of the UGT1A1 Targeting Construct</td>
<td>54</td>
</tr>
<tr>
<td>in Vitro site directed mutagenesis</td>
<td>55</td>
</tr>
<tr>
<td>Southern blot analysis</td>
<td>56</td>
</tr>
<tr>
<td>Mini-Southern blot analysis</td>
<td>57</td>
</tr>
<tr>
<td>Hybridization probes preparation</td>
<td>57</td>
</tr>
<tr>
<td>Reutilization of Southern blot membranes</td>
<td>58</td>
</tr>
<tr>
<td>Database and bioinformatics analyses</td>
<td>58</td>
</tr>
<tr>
<td>Results</td>
<td>59</td>
</tr>
<tr>
<td>a) Cloning and sequencing of the UGT1A1 gene</td>
<td>60</td>
</tr>
<tr>
<td>b) Generation of one base deletion in the UGT1A1 Exon 4</td>
<td>63</td>
</tr>
<tr>
<td>c) Preparation of the final UGT1A1 construct</td>
<td>63</td>
</tr>
<tr>
<td>d) Setting up of the screening conditions</td>
<td>66</td>
</tr>
<tr>
<td>Southern blot analysis</td>
<td>66</td>
</tr>
<tr>
<td>PCR analysis</td>
<td>74</td>
</tr>
<tr>
<td>e) Electroporation, selection and screening of G418-resistant clones</td>
<td>76</td>
</tr>
<tr>
<td>f) Characterization and of the G418-resistant clones</td>
<td>76</td>
</tr>
<tr>
<td>g) Blastocyst microinjection to generate chimeras</td>
<td>85</td>
</tr>
<tr>
<td>h) Germ line transmission of the mutated UGT1A1 gene</td>
<td>85</td>
</tr>
<tr>
<td>i) Changing the genetic background of UGT1A1 gene Mutated Mouse to an Homogeneous one</td>
<td>86</td>
</tr>
<tr>
<td>Discussion &amp; Conclusions</td>
<td>87</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>96</td>
</tr>
<tr>
<td>Reference List</td>
<td>97</td>
</tr>
</tbody>
</table>
Hyperbilirubinemia is the most common clinical situation during neonatal life and it is observed in 60% of full-term and 80% of pre-term infants. A combination of factors still not well defined such as: prematurity, infections, genetic disorders, breast-feed under-nourishing, may cause hazardous, toxic levels of UnConjugated Bilirubin (UCB) during neonatal period (neonatal jaundice) that pose a direct threat of brain damage (kernicterus). The deposition of UCB in the Central Nervous System (CNS) causes Bilirubin Encephalopathy (BE) with lifelong motor, auditory and mental impairment. The in vivo knowledge on kernicterus derives almost totally from the investigation on Gunn rat that is a natural model for BE. In this animal model the genetic lesion are closely parallel those present in the Crigler-Najjar syndrome type I and the neuropathological lesions are also similar to those found in humans. The Gunn rat is a mutant strain of Wistar rats that lack the uridin di phospho glucoronosyl transferase (UDPGT) activity toward bilirubin. Although the Gunn rat the classical laboratory model for bilirubin encephalopathy its use for the study of molecular mechanisms involved and the determination of other genes modulating the disease is limited by the existence of different strains and by the impossibility to generate targeted mutations in rats, preventing the in vivo study of the role of other genes in BE (i.e. Mrp1). The aim of my PhD project was to generate a mouse model of hyperbilirubinemia due to a one base deletion in the UGT1a1 gene, identical to the one present in the Gunn rat.

To reach this goal, we took advantage to the “Gene Targeting” technique. This genetic technique uses the homologous recombination to modify an endogenous gene. First we constructed the targeting vector specific for the gene of interest. To target genes in mice, the targeting vector was inserted into mouse embryonic stem cells (ES) in culture. At the same time we set up two screening strategies to verify the presence of the targeted mutation in electroporated ES cells (Southern blot and Multiplex PCR). Two positive clones were identified (A9 and G7). Then the positive ES clones were amplified and injected into the blastocysts. Blastocysts were implanted in to a foster mother to obtain the so-called “chimera.” Chimeric mice have two different populations of genetically distinct type of cells originated from different mouse strains and can be selected by the fur colour. We obtained two chimeras, one deriving from A9 clone and one from G7. These chimeras were mated with wild type mice to check for germ line
transmission. If the modified ES cells made up the reproductive organ, the offspring will inherit the mutated allele (heterozygous).

At the present time we are screening the offspring of the chimeras, to check for germ line transmission. Mating the heterozygous mice, the offspring will have the entire body based on the previously mutated embryonic stem cell (homozygous).

Obtaining this new animal model for bilirubin neurotoxicity (Gunn mouse) will be crucial to understand the mechanisms regulating the disease, together with an improvement of the diagnosis, prediction of the prognosis, and development of new therapeutic strategies.
RIASSUNTO

Più del 60% dei neonati a termine e l’80% dei neonati prematuri sviluppa un ittero fisiologico nella prima settimana di vita a causa dell’immaturità dei processi fisiologici correlati al metabolismo della bilirubina durante il periodo neonatale. Tuttavia, una combinazione di fattori non ancora ben definiti quali: prematurità, disidratazione, sepsi, disordini di tipo genetico e/o malnutrimento durante l’allattamento, possono causare livelli di bilirubina non coniugata (UCB) eccezionalmente alti durante il periodo neonatale (iperbilirubinemia neonatale). Questa condizione fa sì che l’UCB possa attraversare in maniera massiva la barriera ematoncefalica (BBB) e depositarsi in specifiche aree cerebrali, ponendo il neonato a rischio di sviluppare encefalopatia da bilirubina (kermittero). Questa grave patologia (tipica della sindrome di Crigler-Najjar-I) si caratterizza per: sordità e disfunzioni uditive, gravi disordini motori (atetosi, spasticità muscolare e ipotonia), disfunzioni visive e displasia dentale. Il modello animale classico per il kermittero è rappresentato dal ratto Gunn. In questo ceppo di ratti è presente una mutazione nel gene dell’UGT che determina la completa inattività dell’enzima epatico bilirubin-glucuronil transferasi (UGT1A1) responsabile della coniugazione dell’UCB a due acidi glucuronici e la successiva eliminazione della bilirubina coniugata attraverso la bile. Gli effetti fisiologici della mutazione a carico dell’UGT1A1 sono molto simili a quelli riscontrati nei pazienti affetti da sindrome di Crigler-Najjar I. L’esistenza di questo modello animale ha consentito un approccio sperimentale al problema della neurotossicità da bilirubina. Tuttavia il modello del ratto Gunn possiede innumerevoli limitazioni, come: l’esistenza di diversi ceppi e, ancora più importante, l’impossibilità di ottenere ceppi mutanti per altri geni coinvolti nel metabolismo della bilirubina (i.e. Mrp1). Pertanto l’obiettivo del mio progetto di dottorato è stata la generazione di un modello murino di iperbilirubinemia dovuto ad una delezione di una base nel gene UGT1, identica a quella presente nel ratto Gunn.

La tecnica utilizzata per raggiungere tale scopo è denominata “Gene Targeting”; tale tecnica biotecnologica si serve della ricombinazione omologa per modificare uno specifico gene d’interesse. Il primo passo è stata la costruzione di un vettore specifico per il gene di interesse (UGT1a1) contenente la delezione di una base nell’Esone 4. Parallelamente alla costruzione del vettore di targeting abbiamo messo a punto due
tecniche che permettano lo screening dei cloni resistenti alla doppia selezione (positiva e negativa): Soutern blot e Multiplex PCR. Il vettore è stato poi inserito all’intero del genoma murino di cellule staminali embrionali (ES). Dopo lo screening dei cloni resistenti abbiamo identificato due cloni positivi, ovvero che hanno subito ricombinazione omologa del vettore di targeting: A9 e G7. Questi cloni sono stati ulteriormente vagliati per escludere la presenza di eventi di ricombinazione non omologa. Successivamente i positivi cloni sono stati amplificati ed iniettati all’interno di blastocisti. Le blastocisti sono state impantate in una madre adottiva (forster mother), da esse si originano topi chimera. Abbiamo ottenuto due chimere, una dal clone A9 ed una dal clone G7. Queste chimere sono state incrociate a loro volta con topi wilde type per verificare la trasmissione della mutazione per via germinale (in eterozigosi). Al momento stiamo analizzando la progenie delle due chimere ottenute per trovare degli eterozigoti. I topi eterozigoti verranno poi incrociati tra loro per ottenere l’omozigote.

Questo nuovo modello animale permetterà di studiare i meccanismi coinvolti nella neuro-tossicità da bilirubina, analizzando in vivo gli eventi biologici che si sviluppano in caso di mutazione a carico del gene UGT1 ed in futuro potrà anche essere usato per testare nuovi approcci diagnostici e terapeutici per il trattamento di neonati affetti da questa malattia.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE</td>
<td>Bilirubin Encephalopathy</td>
</tr>
<tr>
<td>Bf</td>
<td>Free fraction of unbound Bilirubin</td>
</tr>
<tr>
<td>BIND</td>
<td>Bilirubin Induced Neurological Disfunction</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>CB</td>
<td>Conjugated Bilirubin</td>
</tr>
<tr>
<td>Cre</td>
<td>Causes recombination, recombinase isolated from bacteriophage P1</td>
</tr>
<tr>
<td>CNS I</td>
<td>Crigler-Najjar type I</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic Reticulum</td>
</tr>
<tr>
<td>ES</td>
<td>Embrionic Stem cells</td>
</tr>
<tr>
<td>Flp</td>
<td>Flp, recombinase able to invert DANN isolated from <em>S. cerevisiae</em></td>
</tr>
<tr>
<td>Frt</td>
<td>Flp recombinase recognition site</td>
</tr>
<tr>
<td>GPI</td>
<td>Glucose phospho isomerase</td>
</tr>
<tr>
<td>hs</td>
<td>hours</td>
</tr>
<tr>
<td>HSV-tk</td>
<td>Herpes Simplex Virus-derived thymidine kinase</td>
</tr>
<tr>
<td>jj</td>
<td>Homozygous hyperbilirubinemic Gunn rats</td>
</tr>
<tr>
<td>Jj</td>
<td>Heterozygous normobilirubinemic Gunn rats</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilo bases</td>
</tr>
<tr>
<td>KDa</td>
<td>Kilo Daltons</td>
</tr>
<tr>
<td>LTR</td>
<td>Long Terminal Repeats</td>
</tr>
<tr>
<td>loxP</td>
<td>locus of crossover (x) in P1</td>
</tr>
<tr>
<td>MEF</td>
<td>Mouse Embryonic Fibroblasts</td>
</tr>
<tr>
<td>MWM</td>
<td>Molecular Weight Marker</td>
</tr>
<tr>
<td>Mdr1</td>
<td>Multi Drug Resistance 1</td>
</tr>
<tr>
<td>MLV</td>
<td>Murine Leukemia Virus</td>
</tr>
<tr>
<td>Mrp 1</td>
<td>Multi drug Resistance Protein 1</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple Clonig Sites</td>
</tr>
<tr>
<td>NBD</td>
<td>Nuclotide Binding Domain</td>
</tr>
<tr>
<td>neo&lt;sup&gt;r&lt;/sup&gt;</td>
<td>Neomycin phosphotransferase</td>
</tr>
<tr>
<td>Neo</td>
<td>Neomycin cassette</td>
</tr>
<tr>
<td>NMuLi</td>
<td>N-Mouse Liver-derived cell line</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PCR</td>
<td>Polimerase Chain Reaction</td>
</tr>
<tr>
<td>PGK</td>
<td>Phospho Glicerate Kinase</td>
</tr>
<tr>
<td>PGK-Neo</td>
<td>Phospho Glicerate Kinase promoter + Neomycin resistance gene</td>
</tr>
<tr>
<td>TBS</td>
<td>Total Bilirubin in Serum</td>
</tr>
<tr>
<td>TK</td>
<td>Timidin Kinase</td>
</tr>
<tr>
<td>UDP</td>
<td>Uridine di-phosphate</td>
</tr>
<tr>
<td>UDP-GlcUA</td>
<td>Uridine diphosphogluconate</td>
</tr>
<tr>
<td>UDPGT</td>
<td>Uridin di phosphogloronosyl transferase, or UGT</td>
</tr>
<tr>
<td>UCB</td>
<td>UnConjugated Bilirubin</td>
</tr>
<tr>
<td>UGT</td>
<td>UDP-gucuronosyl transferase</td>
</tr>
<tr>
<td>UGT1a1</td>
<td>UDP-gucuronosyl transferase 1</td>
</tr>
</tbody>
</table>
INTRODUCTION

E’ più facile resistere all’inizio che alla fine.

Leonardo da Vinci
**BILIRUBIN AND DISEASE: AN OVERVIEW**

In the past years, interest in bilirubin encephalopathy has been reawakened by an increase in its prevalence. At birth the newborn is exposed to an increased load of unconjugated bilirubin (UCB) in the presence of immature mechanisms for hepatic uptake, conjugation and biliary secretion (Gourley, 1997). The resulting "physiological" jaundice may be beneficial since UCB is a powerful antioxidant. Some neonates, however, develop much higher plasma UCB concentrations, with deposition of UCB in specific portions of the central nervous system (CNS), causing lifelong motor, auditory and mental impairment (Ostrow et al., 2003). Hyperbilirubinemia is the most common clinical situation during neonatal life and the risk of emergent acute Bilirubin Encephalopathy (BE, formerly kernicterus) accounts for up to 75% of hospital readmissions in the first week of life (Maisels and Newman, 1998; Wennberg et al., 2006). Although permanent clinical sequelae of hyperbilirubinemia should be preventable, particularly if diagnosis is performed at early stages, irreversible BE is still occurring. Around 70% of children with kernicterus die within seven days, while the 30% survivors usually suffer irreversible sequelae, including hearing loss, paralysis of upward gaze, mental retardation, and cerebral palsy with athetosis (Gourley, 1997; Ostrow et al., 2003). There is also evidence that even moderate neonatal jaundice may result in minimal neurological damage, which is unapparent during the neonatal period but leads to developmental or neurological impairment only evident in later life. It is therefore clear that BE is a major health problem, which has been somehow disregarded till now. This prompts for early diagnosis and possibly treatments.

Due to the high UCB plasma levels, babies affected by the genetic pathology of Crigler–Najjar I (CN I) are particularly exposed to BE. CNS I is an autosomal recessive condition caused by absent activity of hepatic UDP-glucuronosyl-transferase (UGT), a protein encoded by the UGT1A1 gene on chromosome 2. The absence of UGT activity results in the progressive accumulation of UCB in the blood and tissues (Wennberg, 2000). Until now phototherapy is the only form of treatment (Cremer et al., 1958).

The understanding of the genetic and molecular mechanisms involved in the cellular damage related to the accumulation of UCB may help for an earlier diagnosis and better treatment of BE, particularly in those subjects more prone to develop the damage but requires the unveiling of several events that remain unknown. In particular: 1) How toxicity is related to UCB levels and what are the molecules involved in
determining its intracellular concentration?; 2) How differential cellular susceptibility is conferred?; 3) What are the molecular mechanisms leading to cell-death?

Although these questions have been rightly formulated for more than 20 years, our comprehension has been limited since: a) most prior studies were done at clinically irrelevant, unrealistically high UCB concentrations (Ostrow et al., 2002) and only recently more meaningful studies have been undertaken (Ostrow et al., 2004); and b) the lack of a reliable genetic animal model of the disease. Although the Gunn rat model has been widely used, this model is limited by the existence of different strains and by the impossibility to generate targeted mutations in rats, not allowing the in vivo study of the role of other genes in BE.

**FORMATION OF UNCONJUGATED BILIRUBIN**

Bilirubin is the oxidative product of the protoporphyrin portion of the heme group of proteins such as hemoglobin, myoglobin and cytochrome P-450. An individual with normal human metabolism generates 250-400 mg/day of bilirubin through the break down of these hemoproteins (Iyanagi et al., 1998). The first step in the degradation of the heme group to bilirubin is cleavage of the α-methene bridge of heme to form biliverdin IXα (Saito and Itano, 1982; Tenhunen et al., 1969; Yoshida et al., 1981; Yoshinaga et al., 1990). This reaction is catalyzed by microsomal heme oxygenase isoform I. O2 and NADPH are required for the cleavage reaction. The cleavage reaction proceeds via a multistep mechanism that depends on reducing equivalents, which are provided by microsomal NADPH-cytochrome P-450 reductase, which contains FAD and FMN as prosthetic groups (Iyanagi et al., 1974; Yoshida and Kikuchi, 1978). The catalytic cycle of the heme oxygenase resembles that of the microsomal cytochrome P-450 mixed function oxidase (Wilks and Ortiz de Montellano, 1993). The central methene bridge of biliverdin at C-10 is reduced by NADPH-biliverdin reductase, a cytosolic enzyme, to form bilirubin. Two isomers of biliverdin, IXα (95-97%) and IXβ (3-5%), are present in human adult bile, whereas the IXβ-isomer is predominant in late fetal bile (Yamaguchi et al., 1994). This suggests that two isoforms of biliverdin reductase are present in the liver. Bilirubin is an unsymmetrically substituted tetrapyrrole dicarboxylic acid. The introduction of CH2 at C-10 of the bilirubin molecule induces a conformational change by rotation of the dipyrrinone groups about the central CH2 group. This allows
intramolecular hydrogen bonding of the propionic acid carbonyl to the amino groups of the dipyrrinone lactam and the pyrrole ring (Nogales and Lightner, 1995) is thought to explain the many unusual properties of bilirubin, such as its high lipid/water partition coefficient and its resistance to hepatobiliary excretion (McDonagh and Lightner, 1985). This conformation of bilirubin is hydrophobic, and has a high affinity for tissues of the central nervous system.

**Figure 1.** Metabolic pathway from the heme group to bilirubin. Panel A) Topology of UGT and MRP2 (cMOAT). Formation of bilirubin from heme. FAD-FMN, NADPH-cytochrome P-450 reductase; HO, heme oxygenase; BR, biliverdin reductase. Panel B) Uptake of bilirubin, conjugation of the bilirubin molecule, and its excretion from the hepatocyte. UGT, UDP-glucuronosyltransferase; Bilirubin-Alb, bilirubin-albumin complex; Bilirubin-L, bilirubin-ligandin complex; MRP2, multidrug resistance protein 2; UDP-GlcUA, UDP-glucuronic acid. The orientation and dimer structure of UGT in the membrane of the ER are depicted. T1 and T2 in the membrane of the ER are the proposed transporters of UDP-GlcUA and glucuronides, respectively. The dashed lines show the alternate pathway for glucuronidation in the ER, based on the conformational model.

**UCB Toxicity: Neonatal Jaundice, BIND and Kernicterus**

During the foetal development, the fetus is protected against bilirubin by the maternal placenta. After the birth, the bilirubin clearance is totally accomplished by the infants. In adult, the production of bilirubin is about of 3-4 mg/Kg die, and in full term newborns the value reaches the 6-8 mg (Gourley, 1997). This difference is principally due to:
1) The minor red blood cell viability (Rubaltelli, 1993) (70-90 days vs. 120 days in adults) (Gourley, 1997)

2) The immaturity of the hepatic bilirubin conjugation (Rubaltelli, 1993). The diglucuronide normally accounts for 80-85% of the bilirubin conjugates in the bile of adult humans, whereas the monoconjugate, easily re-converted in UCB, predominates in newborns.

3) The absence of anaerobic intestinal flora in infants up to 2 months of age, that causes a decreased intestinal UCB degradation (Gourley, 1997).

4) The consequent enhancement to the UCB reabsorption and enterohepatic recycling, causing higher serum levels of UCB (Gourley, 1997; Vitek et al., 2000).

5) The less efficacy of neonatal serum in binding unconjugated bilirubin due to the lesser amount of albumin (Roca et al., 2006; Weisiger et al., 2001).

6) The highest level of expression of hepatic heme oxygenase early after birth, leading to an increase of total serum bilirubin (Drummond and Kappas, 1984; Maroti et al., 2007).

**NEONATAL JAUNDICE**

The Immaturity of most steps of bilirubin metabolism causes a mild, temporary retention of UCB in approximately 60% of healthy term neonates characterized by a serum UCB levels of less than 170 µM (10 mg/dL) within the first days of life (Ostrow et al., 1994). This situation usually resolves spontaneously without sequelae in the first week of life (Ostrow et al., 2003).

The existence of hyperbilirubinemia is clinically evidenced by the deposition of bilirubin and subsequent discoloration of tissues, phenomena named icterus (greek: ikteros) or “jaundice” (french: “jaune”=yellow) (Gourley, 1997). Low, nanomolar concentrations of UCB are beneficial by providing protection from oxidative injuries, such as ischemia (Dore and Snyder, 1999), the risk of coronary artery diseases (Mayer, 2000; Rigato et al., 2005) and cancer (Keshavan et al., 2004).

The antioxidant ability of UCB arises from the redox consuming cycling mechanism that acts between the conversions of UCB in biliverdin. During this step the oxidant species are consumed and the bilirubin regenerated via biliverdin reductase. The pigments may play a role similar to the glutathione cycle in cytoplasm, acting
against the lipophilic reactive oxygen species produced from the cellular membranes, while the GSH-GSSG cycle plays against the cytosolic oxidative species (Sedlak and Snyder, 2004; Tomaro and Batlle, 2002). Remarkably, bilirubin concentrations as little as 10 nM are able to protect cell cultures from 10,000 times higher concentrations of \( \text{H}_2\text{O}_2 \) (Baranano et al., 2002).

Similarly, in Gilbert's syndrome, a very common chronic, mild, fluctuating unconjugated hyperbilirubinemia (see \textit{UGT1a1 deficiencies}), the hyperbilirubinemia may protect against oxidant stress. Rigato in a retrospective study reported that low-mild hyperbilirubinemia is negatively related to the risk of different diseases, such as atherosclerotic disease, cancer, demyelinizatind neuropathies (Rigato et al., 2005). Additionally the \textit{in vivo} neointimal hyperplasia induced by balloon-injury seems to be prevented by hyperbilirubinemia in Gunn rats (jj), due to the antiproliferative properties of the bilirubin pigment that is able to arrest the cell cycle and inhibits the p38 MAP Kinases (Ollinger et al., 2005).

**BIND (Bilirubin Induced Neurological Dysfunction)**

Moderate hyperbilirubinemia, (200-300 \( \mu \)M; 11.7 mg/dL) occurs in at least 16% of infants. The increased entry of UCB into the CNS may cause transitory effects such as hypotonia, lethargy, anorexia, poor suckling and abnormal brainstem evoked potentials (BSAEP), symptoms referred to the deposition of unbound UCB diacid in the central nervous system and development of bilirubin encephalopathy or BIND (Bilirubin Induced Neuronal Dysfunction), which is usually reversible.

The preferred treatment of neonatal jaundice is phototherapy, which converts UCB to photo isomers that can be excreted in bile and urine without conjugation or parenteral administration of tin mesoporphyrin IX, a potent competitive inhibitor of heme oxygenase and thus of bilirubin synthesis. When severely jaundiced neonates respond insufficiently to the above therapies, they are treated by exchange transfusion to physically remove UCB from the circulation (Ostrow, 1987).
**Kernicterus**

In about the 2% of infant (Gourley, 1997) at slightly higher or prolonged serum UCB levels (13-26 mg/dL; 220-440 µM) a severe hyperbilirubinemia occurs (Soorani-Lunsing et al., 2001) and results in permanent neurological sequelae ranging from delay in motor development, impaired cognitive functions, auditory dysfunctions to more severe extrapiramidal motor, auditory and cognitive disorders, termed kernicterus (Hansen, 1994), or even death. The term kernicterus literally means "yellow kern," with kern indicating the most commonly afflicted region of the brain (ie, the nuclear region). Historically, the term refers to an anatomic diagnosis made at autopsy based on a characteristic pattern of staining found in babies who had marked hyperbilirubinemia before they died. Hervieux first described the condition in 1847, and Schmorl first used the term kernicterus as early as 1903. Bilirubin-induced neurologic dysfunction (BIND) refers to the clinical signs associated with bilirubin toxicity (ie, hypotonia followed by hypertonia and/or opisthotonus or retrocollis) and is typically divided into acute and chronic phases. The two terms are commonly used interchangeably, but this use is not technically accurate because one refers to clinical manifestations and the other to an anatomic diagnosis.

The neurological signs of bilirubin encephalopathy and the following kernicterus are selective, preferentially affecting only certain areas of the brain and only certain cells within these regions. The classical pattern of discoloration/damage is symmetric, highly selective involves basal ganglia, hippocampus, pallidum, subthalamic nucleus, horn of Ammon, cranial nerve nuclei, such as the oculomotor, vestibular, and cochlear and the cerebellum (Diamond and Schmid, 1966; Hansen, 1994; Rodriguez Garay and Scremin, 1971).

**UCB Glucuronidation**

Bilirubin bound to serum albumin is rapidly cleared by the liver, where there is a free concentration bidirectional flux of the tetrapyrrole across the sinusoidal-hepatocyte interface. Once in the hepatocyte, bilirubin is bound to several cytosolic proteins, of which only one has been well characterized. The latter component, ligandin, is a small basic component making up to 6% of the total cytosolic protein of the liver. The stoichiometry of binding is one bilirubin per ligandin molecule.
Once in the hepatocyte the propionyl side chains of bilirubin are conjugated to form a diglucuronide. The reaction is catalized by UDP-glucuronosyltransferase-1 (UGT1a1) that utilizes uridine diphosphoglucuronate (UDP-GlcUA) as a glucuronate donor to bilirubin.

In normal bile, the diglucuronide is the major form of excreted bilirubin, with only small amounts of the monoglucuronide or other glycosidic adducts present. Bilirubin diglucuronide is much more water-soluble than free bilirubin, and thus the transferase facilitates excretion of the bilirubin in to the bile. Bilirubin diglucuronide is poorly absorbed by the intestinal mucosa. The glucuronide residues are released in the terminal ileum and large intestine by bacterial hydrolases; the released bilirubin is reduced to the colourless linear tetrapyrroles known as urobilinogens. Urobilinogens can be oxidized to coloured products known as urobilins, which are excreted in the feces. A small fraction of urobilinogen can be reabsorbed by the terminal ileum and large intestine to be removed by hepatic cells and resecreted in bile.

**Figure 2. Bilirubin glucuronidation.** In the hepatocyte the propionyl side chains of bilirubin are conjugated to form a diglucuronide. The reaction is catalized by UDP-glucuronosyltransferase-1 (UGT1a1) that utilizes uridine diphosphoglucuronate (UDP-GlcUA) as a glucuronate donor to bilirubin.

**THE UGT FAMILY**

The vertebrate animal removes numerous xenobiotic and endobiotic compounds from the body by converting them to more water-soluble glucuronides. UDP glucuronosyl transferase proteins (UGTs) are a superfamily of enzymes that catalyze the glucuronidation of numerous chemical compounds (Tukey and Strassburg, 2000). These proteins play an important role in the pharmacokinetic clearance of environmental toxins, endogenous metabolites, and therapeutic drugs. Genetic mutations of the human *UGT1A1* gene cause jaundice, including Crigler-Najjar syndrome and Gilbert syndrome (Tukey and Strassburg, 2000). The UGT1 subfamily of UGT proteins catalyzes the
glucuronidation of the donor substrate (UDP-glucuronic acid) to numerous acceptor substrates (aglycone compounds). The N-terminal aglycone- binding domains of UGT1 proteins are distinct, but very similar to each other. In contrast, the C-terminal UDP-binding domains are identical among all the UGT1 proteins.

**Figure 3. Functional domains of UGT1.** UGT is composed by: an amino-terminal signal peptide (blue box, which is cleaved during synthesis of the polypeptide chain); a variable domain (purple box) encoded by a unique first exon and responsible for the substrate-specificity. Within the common domain: a Nucleotide UDP Binding Domain (NBD, white box) that binds UDP; a stretch of 17 hydrophobic amino acids near the carboxy terminus, that anchors the protein to the lipid bilayer.

The diversity of UGT1 proteins is determined by an unusual genomic organization. The rat and human UGT1 clusters have been characterized (Emi et al., 1995; Gong et al., 2001). However, the number of mouse UGT1 genes and their genomic organization are unknown. Zhang et al identified 14 mouse UGT1 genes by a combination of sequence analyses and cloning experiments. The N-terminal domain of each protein is encoded by a single exon. Fourteen mouse UGT1 variable exons are organized in a tandem array, spanning a region of ∼187 Kb of genomic DNA. Therefore, the C-terminal domains of all UGT1 proteins are identical and are encoder by four constant exons. These four exons span a region of ∼4 Kb of genomic DNA sequences. In addition, there is a consensus 5′ splice site immediately downstream from the last nucleotide of each variable exon, and a consensus 3′ splice site immediately upstream of the first nucleotide of the first constant exon. Therefore, alternative RNA splicing of each variable exon to the first constant exon generates diverse UGT1 mRNAs.

The DNA sequence of UGTs contains an amino-terminal signal peptide, which is cleaved during synthesis of the polypeptide chain (Iyanagi et al., 1986), and a stretch of 17 hydrophobic amino acids near the carboxy terminus, that anchors the protein to the lipid bilayer (Mackenzie, 1986). The stretch of 20-30 amino acids at the carboxy terminus of the UGT contains several basic residues, which are responsible for retaining the UGT in the endoplasmic reticulum (ER) (Jackson et al., 1990; Kinosaki et al., 1993).
Based on identity in amino acid sequence, the UGTs are divided into two families, UGT1 and UGT2 (Burchell et al., 1991). Members of the UGT1 gene family include bilirubin- and phenol-metabolizing isoforms.

They share an identical 245 amino acid carboxy terminus, whereas their N-terminal halves show a striking lack of identity (37-49%) (Burchell et al., 1991; Iyanagi et al., 1986). The N-terminal half of a particular UGT determines substrate specificity for glucuronidation. Members of the UGT2 family include steroid-metabolizing isoforms. Comparison of the amino acid sequences of the proteins encoded by the UGT2 gene family reveals differences in amino acid sequence throughout the length of the protein (Burchell et al., 1991). The C-terminal halves, however, are highly conserved, which may provide the binding site for the common co-substrate, UDP-GlcUA (Iyanagi et al., 1986; Mackenzie, 1986). Efficient excretion of bilirubin across the bile canaliculus requires the conversion of bilirubin to polar conjugates by esterification of the two propionic acid carboxyl groups on the bilirubin molecule. It is believed that glucuronidation occurs on the intramolecular hydrogen-bonded conformation of bilirubin. In this model, the hydrogen-bonded, folded, ridge-tile conformation of bilirubin is stabilized when bound to the active site of the UGT molecule; glucuronic acid is transferred to the hydrogen-bonded carboxyl (or carboxylate) group of bilirubin (Iyanagi et al., 1998).

**UGT1A1 Deficiencies: Crigler-Najjar Syndrome and Gilbert’s Syndrome**

In 1952, Crigler and Najjar described a syndrome characterized by a severe, chronic, non-hemolytic, unconjugated hyperbilirubinemia (Crigler and Najjar, 1952).
Two clinical forms of this syndrome have been described: patients with Crigler-Najjar syndrome (CN) type I have an unconjugated serum bilirubin level of greater than 340 µM, while patients with CN type II have an unconjugated serum bilirubin level in the range of 60-340 µM. Patients with Gilbert's syndrome have an unconjugated serum bilirubin level of 60 µM or below. Gilbert's syndrome is characterized by mild, unconjugated hyperbilirubinemia; this syndrome affects approximately 5% of the population. The molecular bases of CN syndrome and Gilbert's syndrome have been characterized by enzymatic, immunochemical, and molecular genetic analyses. As a result of the elucidation of the UGT1 gene complex in the rat (Emi et al., 1995) and humans (Ritter et al., 1992), several groups have discovered the genetic defects, which cause hyperbilirubinemia. Mutations in the UGT1 gene complex among patients with CN type I have been detected in common Exons 2, 3, 4, and 5; these patients have a defect in all glucuronidation activities catalyzed by UGT1. Mutations in the first Exon (1A1), have also been found among patients with CN type I; these patients have a defect in bilirubin-glucuronidating activity. This finding indicates that the protein encoded by UGT1A1 in humans is the major bilirubin-metabolizing isoform under physiological conditions (Bosma et al., 1994). Among patients with Gilbert's syndrome, two types of genetic abnormality in the UGT1A1 gene have been found. One is a homozygous TA insertion in the TATA box upstream from the 1A1 Exon (Bosma et al., 1995; Kaplan et al., 1997; Monaghan et al., 1997; Monaghan et al., 1996); the other is a heterozygous missense mutation in the coding region (Aono et al., 1995). These findings indicate that there are two inherited patterns of Gilbert's syndrome; one is autosomal recessive, and the other is autosomal dominant. TA insertion in the TATA box also causes very severe hyperbilirubinemia in patients with glucose-6-phosphate dehydrogenase deficiency, an X-linked hereditary disease, which causes hemolytic anemia and which occurs in high incidence (Kaplan et al., 1997). Finally, the position of a particular mutation within the UGT1 gene complex, and the effect of that mutation on glucuronidating activity towards bilirubin, may determine the type of syndrome.

ANIMAL MODEL FOR UGT DEFICIENCY: THE GUNN RAT

The in vivo knowledge on kernicterus derives almost totally from the investigation on Gunn rat that is a natural model for bilirubin encephalopathy. In this
animal model the genetic lesion closely parallel those in Crigler-Najjar syndrome type I (Chowdhury et al., 1993) and the neuropathological lesions are similar to those in humans, with cell loss and gliosis most prominent in the auditory nuclei of the brainstem, the cerebellum, the hippocampus and the basal ganglia (Chowdhury et al., 1993).

The Gunn rat (Gunn, 1938) is a mutant strain of Wistar rats that lack the uridin di phospho glucoronosyl transferase (UDPGT) activity toward bilirubin. In the homozygous jj animal, a life long hyperbilirubinemia in absence of heamolysis is present (Chowdhury et al., 1991; Iyanagi et al., 1989). In the non jaundiced heterozygous Jj rats the activity of UDPGT is reduced and did not result in retention of bilirubin in the plasma (Schmid et al., 1958). Thus, the Gunn rat is unable to conjugate UCB and the Total Serum Bilirubin (TSB) levels change from 3mg/dL (50 µM) to 20 mg/dL (340 µM) depending on the strain and the diet (normal TSB leves in rat are 0.01-0.05 mg/dL).

**Figure 5. The Gunn rat.** Gunn rats at 17-day post-natal age. The jj hyper-bilirubinemic animals show a clear yellow discoloration of the skin (see ears and tail) and coat, while the heterozygous (Jj) animals show normal skin and coat colour.

In rats, two families of UDPGT are present; the UGT1 and UGT2. The capacity of a tissue to glucoronydate and excrete a substrate depends on the UGT isoforms presence and level of expression (Shelby et al., 2003). The UGT2 family members are encoded from individual genes, each of them containing 6 exons and are responsible for the glucoronydation of steroids. Based on the sequence similarity, the UGT2 family is divided in two subfamilies (Emi et al., 1995): the UGT2A, specific for the olfactory bulb
(2A1), and the UGT 2B, formed by 6 members (2B1-2B6). The presence of a second family of UGT enzymes, explains why in Gunn rats the activity toward bilirubin is undetectable but activity toward several other substrates is normal (Chowdhury et al., 1991).

![Figure 6. The UGT complex in the Gunn rat.](image)

In the Gunn rat a single base deletion in the Exon 4 (common exon, orange box) generates an in-frame stop codon immediately after the deletion, resulting in the synthesis of all truncated proteins encoded by this gene complex. The enzyme most affected by the mutation is UGT1a1 (N-terminal encoded by Exon 1A1, purple box), this leads to hyperbilirubinemia in the rat.

The UGT1 locus spans plus than 120 Kb and forms a gene complex. cDNA cloning of the UGT isoforms of the Gunn rat has shown that the chimeric mRNAs contain variable 5'-portions and a conserved 3'-portion. This suggests that they comprise a family of proteins, the UGT1 family, and that these mRNAs are derived from a single gene by the use of different promoters and spliceosomes (Iyanagi, 1991). The four human UGT protein isoforms that belong to the UGT1 family have identical C termini (Harding et al., 1988; Ritter et al., 1991; Wooster et al., 1991). Emi et al. and Ritter et al. have isolated a large gene complex, from the rat genome and the human genome, respectively, that encodes several isoforms of UGT family 1. The UGT1 gene complex contains several unique promoter regions, which make up the first exon; each of these first exons is capable of splicing with common Exons 2 through 5. This leads to different N-terminal halves, but identical C-terminal halves, of the gene products of the UGT1 gene complex (Emi et al., 1995). Although the carboxyl-terminal halves of each of these
INTRODUCTION – BILIRUBIN AND DISEASE

Genes are identical, the first exon of each UGT1 isoform can be regarded as a distinct gene (UGT1a1, UGT1a2, UGT1a3, ...) (Mackenzie et al., 1997). Sequence analysis of the first Exons has revealed that the proteins encoded by the UGT1 gene complex can be divided into two clusters: the bilirubin-like cluster (A1-A5), and the phenol-like cluster (a6-a9) (Emi et al., 1995). The deletion of a single base pairs in the UGT1 commun Exon 4, a GCT coding for the leucine 413, produce a stop codon that remove the 115C- terminal amino acids of the protein (approximately 13 KDa), responsible for the binding on the endoplasmic reticulum (ER), and generate a truncated form of the enzyme. The protein is unstable and rapidly degraded (Chowdhury et al., 1993; Chowdhury et al., 1991; elAwady et al., 1990; Iyanagi et al., 1989). In Gunn rat the mutation in the Exon 4 (common), causes the simultaneous deficiencies of all UGT1 isoforms (Emi et al., 1995).

Using two-dimensional SDS-PAGE and immunoblotting using isozyme-specific antibodies, each UGT1 isoform in the Gunn rat was associated with the coincident absence of hepatic microsomes (Ikushiro et al., 1995). The bile of homozygous Gunn rats contains a trace amount of bilirubin mono-conjugates (Blanckaert et al., 1977; Wishart, 1978). These results indicate that the UGT2 isoforms, encoded by different genes, do not have significant bilirubin-glucuronidating activity. Among the hepatic UGT1 isozymes in the rat, the enzyme encoded by UGT1A1 is responsible for constitutive bilirubin metabolism. It is known that glucuronidation activity towards bilirubin increases after birth (Wishart, 1978). This suggests that the expression of UGT1A1 is regulated by ontogenetic factors such as glucocorticoids (Wishart et al., 1977) or thyroid hormones (Labrune et al., 1992).

GENE THERAPY FOR CRIGLER-NAJJAR DISEASE

During the past decade, a wide array of gene transfer vectors has been tested for their capacity to correct hyperbilirubinemia in Gunn rats (Bellodi-Privato et al., 2005; Kren et al., 1999; Nguyen et al., 2005; Takahashi et al., 1996; Toietta et al., 2005). These studies revealed that complete and permanent correction was not easily achieved by gene transfer strategies, and most attempts resulted in only partial or transient correction of hyperbilirubinemia (Nguyen et al., 2007). Such failures were consecutive to either immune elimination of virus-infected hepatocytes (Askari et al., 1996; Aubert et al., 2002) or to the weak capacity of some viral vectors to drive high expression of bilirubin UGT1 or to transduce a sufficient proportion of hepatocytes (Sauter et al,
2000; Seppen et al., 2006; Tada et al., 1998). Recently, Toietta et al. (Toietta et al., 2005) reported the first complete cure of Gunn rat after delivery of a gutted adenoviral vector to adult animals. The correction was prolonged for the life of the animals (2 years) without significant toxicity. Nguyen et al. also reported that delivery of murine leukemia virus (MLV)-based retroviral or lentiviral vectors to newborn Gunn rats could stabilize bilirubinemia for at least 40 weeks (Bellodi-Privato et al., 2005; Nguyen et al., 2005). These seminal studies demonstrated the therapeutic validity of an in vivo approach for correcting CN I, a model for gene therapy of metabolic diseases, and paved the way for future clinical application.
GENETIC MANIPULATION: AN OVERVIEW

Animal models have been used in biomedical research for decades. In many cases, aspects of physiology and biochemistry have been investigated. Artificial manipulations have often been restricted to the assessment of the effect of altering the animal’s environment or some characteristic of its phenotype. Some animals, especially Drosophila and mice, have been particularly suitable for genetic analyses and traditional genetic manipulation by carefully selected breeding experiments or exposure of the animals to powerful chemical or radio-isotopic mutagens. A new era in animal research started during the early 1980s, when Drs Mario Capecchi, Martin Evans and Oliver Smithies first reported successful experiments achieving genetic modification in mice using embryonic stem cells (ES cells) (Strachan, 1999). In 2007 they won the Nobel Prize in physiology or medicine for developing technologies that allow any gene in the nucleus to be either partly deleted (knockout) or inserted (knockin) in the germline, and have gone on to generate mice that carry and express that altered gene. These technological progresses have significant consequences for clinical medicine, as they allow detailed examination of gene function in a whole organism where 99% of the genes have a human counterpart (Hacking, 2008).

MICE AS A GOOD ANIMAL MODEL

Despite the fact that 75 million years of evolution divide mice and humans, these mammals maintain a significant degree of genetic and physiological correspondence, especially when compared with the traditional model systems such as flies, worms and yeast. In contrast to many other mammals, mice are fecund, easy to feed, require little space for housing and are suitable to a range of assisted reproductive techniques including cryopreservation which allows the sperm from genetically manipulated lines to be preserved (Kile and Hilton, 2005). It is possible to use a variety of powerful techniques to manipulate murine genomes although the well-characterized genetically identical background of inbred mice means that experiments between laboratories are comparable. Furthermore, the sequencing of the mouse genome and the characterization of a range of strain specific genetic markers allow mutations and genetic alterations to be readily mapped (Waterston et al., 2002). Altogether, their
careful characterization, convenience of captivity and relatively close evolutionary relation to man has guaranteed that mice are the best resources of modelling human disease (Hacking, 2008)(Figure 1).

**THE GENETIC MODEL**

**STRAIN**
- Inbred
- Outbred
- Recombinant inbred
- Consomic

**GENOME**
- Mammalian
- 19 chromosomes
- 28,000 genes
- Human: Mice divergence 75 million years
- 99% genes have human counterpart

**LIFE CYCLE**
- 4 days oestrus
- 20 days gestation
- 4-8 pups per litter
- 7 weeks to sexual maturity
- 2 years life span

**RESOURCES**
- Whole genome sequence
- Well characterized genetic marker for each strain
- Gene trap libraries
- BAC libraries
- ENU point mutation mouse repository

**GENE MANIPULATION**
- Knockout
- Conditional knockout
- Knockin
- ENU mutagenesis
- RNA interference

**ASSISTED REPRODUCTION**
- Germ line cryopreservation
- In vitro fertilization
- Intracytoplasmic injection

*Figure 1. The mouse as a good animal model. Genetic manipulation in the mouse as a model for human physiology and disease (Hacking, 2008).*

**GENE TARGETING**

Gene targeting typically consists in introducing a mutation by homologous recombination. A cloned gene (or gene segment) closely related in sequence to the endogenous target gene is transferred into the appropriate cells. In some of the cells, homologous recombination occurs between the introduced gene and its chromosomal homolog. Gene targeting by homologous recombination has been carried out in some somatic mammalian cells, such as myoblasts (Strachan, 1999). However, the most important application involves mouse ES cells: once a mutation has been engineered into a specific mouse gene within the ES cells, the modified ES cells can then be injected into the blastocyst of a foster mother and eventually a mouse can be produced with the
mutation in the desired gene in all nucleated cells (Capecchi, 1989; Melton, 1994).

Gene targeting in mice is widely used for producing artificial mouse models of human disease. In addition, it provides a powerful general method of studying gene function. The gene in question is selectively inactivated, producing a knockout mouse, and the effect of the mutation on the development of the mouse is monitored carefully. Sometimes there is little or no phenotypic consequence after inactivating a gene that would be expected to be crucially important, such as some genes, which encode a transcription factor know to be expressed in early embryonic development. The lack of a phenotype in such cases is often thought to be due to genetic redundancy (another gene is able to carry out the function of the gene that has been knocked out). As a result, in some cases double or even triple gene knockouts have been generated to analyze gene function (Manley and Capecchi, 1997). To test for the possibility of functional redundancy, Hanks et al. 1995 used a variant of the knockout procedure known as the “knock-in” technique (Hanks et al., 1995). Here follows a description of the two methods.

**INACTIVATING GENES: KNOCKOUT**

Before the development of knockout technology, the ability to study the functional effects of genes was very limited. All the knowledge came from spontaneous mutations isolated in animals and patients, genetic association studies and functional genetic experiments in primary and immortalized cell lines. The problem with these approaches is that isolating spontaneous mutations can be time consuming and expensive. Association studies linking the gene to a disease tell one little about the function of that gene, and the action of the gene within a cell line in vitro may be very different from that within the whole animal. So powerful are the knockout techniques, that national consortia have been created that aim, to map all the functional genetic elements present in mice by systematically generating a knockout or mutated allele for every gene in the genome (Hacking, 2008).

**BASIC STRATEGIES TO GENERATE MOUSE MODELS OF KNOCKOUT GENES**

The theory of the knockout technique is very simple and is based on two fundamental biological properties. First, embryonic stem (ES) have a totipotent nature
that means that they can differentiate into all cell types in the developing embryo (Figure 2). Consequently, if the totipotent ES cells are injected into a developing mouse embryo early enough, at the blastocyst stage, they differentiate in the normal way together with the cells from the blastocyst to create a chimeric mouse derived from two separate germ lines. As a result, these ES cells could be used as a vehicle to introduce manipulated genetic material into the blastocyst to create a transgenic animal (Evans and Kaufman, 1981).

The second biological property, crucial to knockout technology, is the observation that similar sequences of chromosomal DNA can be exchanged in the nucleus in a process called homologous recombination. Usually, homologous recombination generates a unique set of chromosomes for each organism the same as during development when the zygote will recombine segments of paired and similar chromosomes derived from each parent.

Homologous recombination can be used to target specific genes. In spite of this, homologous recombination in mammalian cells is a very rare event (1:1000) and may be difficult to identify against a large background of random integration events (unlike in yeast cells, for example, where it occurs naturally at high frequencies, allowing sophisticated genetic manipulation) (Hacking, 2008).

The frequency of homologous recombination improved, when the degree of sequence homology between the introduced DNA and the target gene is very high. For this reason the DNA introduced is preferably isogenic with the ES cell line that will be used for the gene targeting experiments; that means, both the genomic DNA clones and the ES cells should be derived from the same inbred strain of mice. A complete sequence match, between the targeting vector homology and the locus to be mutated, helps obtaining the maximum frequency of homologous recombination. On the contrary, potential sequence mismatches, resulting from the use of DNA from different mouse strain, can reduce the frequency of homologous recombination or even prevent it at all (Joyner, 1992).

The total amount of sequence homology to be used in the targeting vector should be between 5 Kb and 8 Kb, which represents a balance between increased probability of homologous recombination and ease of vector construction. Smaller segments of homology can reduce the targeting frequency and larger segments can make difficult vector construction as well as the identification of homologous recombinants. Usually, the homologous sequence should be equally divided between the upstream and
downstream arms of homology in the vector. However, if necessary, because of gene structure or cloning limitations, one arm of homology can be short as 1 Kb in length with the other arm balancing the total recommended homology of 5-8 Kb. The arms of homology in the targeting vector should flank any genomic sequences wishing to be deleted (Joyner, 1992).

Figure 2. General strategy for targeting of embryonic stem cells in the mouse. Step 1. Embryonic stem (ES) cells derived from a mouse blastocyst (1) are cultured after which a targeting vector (2) is introduced by electroporation (3). The minority of cells where homologous recombination has occurred (4) will be isolated in culture through both positive and negative selection (5).

Once obtained the electroporated cells, it is necessary to have a powerful method to select the cells undergone to homologous recombination.

To improve identification of the required homologous recombination events, the targeting vector contains a resistance marker gene (typically in the centre of the targeting vector), such as the Neomycin phosphotransferase (neo'), which permits selection of the cells that have incorporated the introduced DNA. This is useful for two reasons, it inactivates the target gene and through the presence of the neo' gene and confers resistance to the aminoglycoside derivative G418. The ES cells where homologous recombination has successfully occurred will grow in culture in the presence of G418, while genetically unmodified ES cells will not grow (Figure 3A).
In addition to this positive selection, there is also the need to negatively select out those cells that have incorporated the plasmid non-specifically at some other part of the genome through random integration. One approach to this problem is to guarantee that a negative selection gene, such as the Herpes Simplex Virus derived thymidine kinase (HSV-tk), is present in all cases of random integration, but in no cases of homologous recombination. This negative selection strategy takes advantage of the difference between viral and mammalian thymidine kinases. Of the two, the viral derived thymidine kinase is less selective in choosing the nucleosides on which it acts. This makes it susceptible to nucleoside analogues such as gancyclovir. This means that ES cells without HSV-tk grow well in the presence of gancyclovir, whereas those expressing HSV-tk die. As a result the ES cells are positively selected for neo\(^r\) expression and G418 resistance where they have undergone homologous recombination. However, ES cells where random integration has occurred will be killed by the negative selection gene HSV-tk in the presence of gancyclovir apart from the presence of the neo\(^r\) gene (Mansour et al., 1988).

**Figure 3. The mechanism for homologous recombination and random integration. Panel A)**

For homologous recombination the positive selection neomycin resistance gene (neo\(^r\)) is inserted into the target gene. The negative selection gene HSV-tk is excluded as it is outside the homologous region of the vector. **Panel B)** In random integration the linearized DNA inserts at its ends which means that the positive (neo\(^r\)) and negative (HSV-tk) genes are still linked together in the ES cell chromosome.
Even after positive-negative selection, many of the ES cells colonies that survive the drug selection will not be homologous recombinants but rather, will contain random integration events where the negative selectable marker was casually lost or damaged. For that reason, all resistant colonies must be screened using Southern blot or polymerase chain reaction (PCR) analysis to identify ES clones with correct gene targeting (Joyner, 1992).

Southern blot analysis requires the use of an “investigative” restriction enzyme digestion and detection with a radioactive probe, called the external probe, that is not contained within the regions of homology included in the targeting vector. For screening purpose, a unique restriction enzyme site is usually introduced in the positive selection marker during the construction of the targeting vector. This restriction enzyme site should be located outside of any “floxed” (see below, BASIC STRATEGY TO GENERATE MOUSE MODELS WITH KNOCKIN THECNIQUE) selectable marker because after Cre expression and deletion of the marker gene, the diagnostic restriction enzyme site will remain. Using this strategy, an investigative digestion and hybridization with an external probe will produce a smaller DNA fragment for the mutant allele compared to the wild type one. This is advantageous for technical reasons because if the mutant band were larger, it would be hard to distinguish it from a partial restriction enzyme cut on Southern blot.

Regarding PCR screening, the basic strategy is to choose PCR primers that will amplify a novel junction fragment created by homologous recombination. The primers are designed in such a way, that one primer binds to the neo' cassette, and the second to a region just past one of the two arms of homology of the targeting vector, within the endogenous locus. It is desirable to include in the screening PCR reaction also a set of primers amplifying the wild type allele (sharing the endogenous primer with the recombinant allele) to have an internal control of the reaction (Joyner, 1992).

Identification of homologous recombinants can require screening of large numbers of neo' ES cells colonies, such large numbers of ES cells lines can be kept frozen on multi-well plates until the use.

Later on, genetically modified ES cells are injected into a blastocyst (Figure 4). Mammalian embryos are extremely “flexible” in the early stages of their development and can not only tolerate abuse or loss of tissue, but can also functionally incorporate cells from other embryos; this ability has been used for several purpose including elucidation of cell lineages, the investigation of cell potential, and the maintenance of mutations produced in ES cells by gene targeting. The extent of contribution of the
foreign cells will depend on their normality, developmental synchronicity with the host embryo, and their capacity to undergo meiosis and gametogenesis (Joyner, 1992).

**Step II**

![Diagram of genetic manipulation process]

**Figure 4. General strategy for targeting of embryonic stem cells in the mouse. Step II.** Genetically modified ES cells are injected into a blastocyst (6) which is then placed in a foster mother (7). The resulting chimeric mice on mating transmit the modified gene to their offspring (8). Identification of the desired progeny is made easier through selection of different coat colours for donors of the ES cells and the recipient blastocysts.

Usually, cells from different mouse strain embryos are combined to produce a composite animal named chimera. C57BL/6 and C129 mice are widely used to produce chimeric mice because they differ in coat colour and other genetic loci, which are useful as markers to estimate chimerism. C57BL/6 embryos proved early on to be compatible hosts for C129 ES cells; chimera formation and most importantly germ line transmission could be efficiently obtained. Successful embryo transfer depends on the quality of embryos and also the suitability of the host maternal environment. Mice are spontaneous ovulators and can be rendered pseudo-pregnant by mating with sterile males during oestrus, that is, they show the hormonal profile of a pregnant female when the stimulus of mating occurs during oestrus.

Introduction of ES cells into the blastocysts can be done using micro-manipulation apparatus composed by: a microscope (under phase or interference
contrast optics), a cooling stage (to prevent stickiness during cell injection and to impart a degree of rigidity to cells and blastocyst), a micro-manipulator to inject the ES cells, micro-instruments for to hold the blastocyst, and a pipette to pick up and inject the ES cells.

After that the injected blastocysts are recovered into a foster mother to produce chimeras.

The most suitable and readily visible genetic marker of chimerism is the coat colour. Chimerism combinations of strains, which differ at only one coat colour locus, allow a simple visual appreciation of the degree of tissue contribution of each component in terms of the proportion of the coat that express the ES cell allele (targeted ES cells). This evaluation of chimeric animal is subjective but in general, the degree of coat colour chimerism of a particular animal correlates with the degree of germline contribution. Another method to evaluate chimerism takes advantage of isozyme differences between strains, most commonly the dimeric enzyme glucose phosphate isomerase (GPI), which has three variants distinguishable by electrophoretic mobilities.

All of the ES cells lines in general use were derived from male embryos and thus containing a Y chromosome. There are several reasons for preferring to use male ES cell lines: first, male ES cells produce a higher proportion of phenotypic male chimeras and male chimeras can be mated more rapidly to test for germline transmission; second, female derived XX cell lines are considered to be unstable (Joyner, 1992).

Once chimeras have been produced, they are test-bred to establish the contribution of the targeted ES cells to germline. Chimeras should be mated to mice with genetic markers that will allow a distinction to be made between ES cells-derived and host-blastocyst derived gametes, keeping in mind that only half of the ES cells-derived gametes will carry the targeted allele. For example, mating a chimera made from C129Sv/J-derived ES cells (agouti) and C57BL/6 blastocs (black) (resulting in mixed agouti and black coat phenotype, see picture in Figure 4) with a C57BL/6 wild type mouse: at birth animal can be distinguished on the basis of their coat phenotype. Agouti animals are derived from the C129Sv/J ES cells, half of them will carry the targeted allele (heterozygous for the mutation). Black animals will be pure C57BL/6 (wilde type).

Once the progeny of the chimera, carrying the mutation, have been genotyped by PCR or Southern blot analysis of tail DNA, they can be mated together to produce litters.

Transfer to an inbred background can be then done at any time independently of the founder chimera. A uniform genetic background allows the most precise comparison
to be made between mutant and wild type phenotypes. Transferring the mutant allele to an inbred background is a matter of back-crossings to mice of a desired inbred strain; the more times this is repeated, the more uniform the genetic background becomes. After seven or eight generations, 99% of the loci not linked to the mutant allele will be homozygous, and after 10 generations the mice are considered congenic (Joyner, 1992).

**INSERTING GENES: KNOCKIN**

Although the knockout and mutagenesis techniques have provided great insight into human physiology and disease, they have two significant limitations that have been resolved with the knockin technology (Branda and Dymecki, 2004; Hacking, 2008). The first and major limitation is that many genes have an essential function, and so the elimination of gene activity all over the entire animal can result either in early embryonic lethality, precluding the analysis of gene function at later stages, or, alternatively, a masking of the full phenotype due to genetic compensatory mechanisms. In these cases, knockin strategy may be used in alternative to inactivate a gene in distinct cells and/or at distinct times during development within the context of a normal mouse (a method named conditional gene modification or conditional targeting).

The second limitation, in traditional knockout strategies, is the maintenance of the positive selection cassette within the targeted gene. Its presence can cause a number of problems, such as the disruption of neighbouring gene expression due to strong transcriptional regulatory elements frequently present in selection cassettes (Lerner et al., 1993; Ohno et al., 1994). Whereas the elimination of the positive selection cassette, to avoid the associated deleterious effects, can be done with a second gene targeting step leaving a "clean" mutation in the genome. A relatively simpler strategy is to use Cre or Flp recombinase enzyme, either *in vitro* or *in vivo*, to remove the selection cassette that has been either *loxP*-flanked (“floxed”) or *Frt*-flanked (“flrted”), respectively (see below, BASIC STRATEGY TO GENERATE MOUSE MODELS WITH KNOCKIN THECNIQUE, for a description). In mice, the targeted gene segment will be deleted specifically in the cells expressing the recombinase, while the target gene will remain intact in the other cells of the mouse, where the recombinase is not expressed. This binary design offers great versatility to gene function studies: the role of one gene may be assessed in a variety of lineages, by crossing different recombinase lines (with differing recombinase expression
patterns) to one strain bearing a conditional allele of a single gene; alternatively, the function of different genes may be assessed in a single lineage, by crossing one recombinase line to multiple strains each carrying a conditional allele of a different gene (Branda and Dymecki, 2004).

**Basic Strategy to Generate Mouse Models with Knockin Technique**

The purpose of the knockin technique is to place a functioning gene into the ES cell. In order to do this the positive selection gene, for example coding for neomycin resistance, must be removed from the target gene. The knockin method achieves this through the use of two recombinase enzymes Cre (causes recombination, which was isolated from a bacteriophage P1) and Flp (named for its ability to invert, or "flip", a DNA segment in *S. cerevisiae*). The Cre and Flp enzyme are able to recombine specific sequences of DNA with high fidelity without the need for cofactors (Branda, 2004). For this reason, they have been used successfully to create gene deletions, insertions, inversions, and exchanges in exogenous systems such as flies (Dang and Perrimon, 1992; Golic and Lindquist, 1989; Xu and Rubin, 1993), mammalian cell cultures (O'Gorman et al., 1991; Sauer and Henderson, 1988), and mice (Dymecki, 1996; Lakso et al., 1992; Orban et al., 1992). Cre and Flp recombine DNA at specific target sites, termed loxP (locus of crossover (x) in P1) (Hoess et al., 1982) and Frt (Flp recombinase recognition target) (McLeod et al., 1986), respectively, in both actively dividing and postmitotic cells, as well as in most tissue types. Cre and Flp enzyme, both members of the integrase superfamily, share a common mechanism of DNA recombination that involves DNA strand cleavage, exchange, and ligation (Sadowski, 1995). Although distinct at the nucleotide level, loxP (Figure 5A) and Frt (Figure 5B) sites share an overall structure composed by two 13 bp palindromic sequences, or inverted repeats, divided by an 8 bp asymmetric core, or spacer, sequence. In the presence of two target sites, recombinase monomers bound to the inverted repeats, help DNA synaptic complex formation and recombination between the two sites (Hoess et al., 1985). Strand cleavage, exchange, and ligation occur within the spacers (Amin et al., 1991). Because of the spacer asymmetry, strand exchange is possible only when target sites are connected by synopsis in one orientation (Hoess et al., 1986). Thus, the relative orientation of target sites, with respect to one another, determines the effect of recombination: Cre and Flp can remove a circular molecule from between two directly repeated target sites or
integrate a circular molecule into a linear molecule each possessing a target site (Figure 6A); invert the DNA between two inverted sites (Figure 6B); or exchange sequences distal to target sites present on two linear molecules, such as a pair of non-homologous chromosomes (Figure 6C).

**Figure 5. Cre and Flp recombinase target sites.** The 34 bp Cre (orange arrow) and Flp (green arrow) target sites each contain two 13 bp inverted repeats flanking an 8 bp asymmetrically core sequence, or spacer (nucleotide of the spacer are indicated in bold). The core sequence confers directionality to these sites (bold arrow). By convention, nucleotides are numbered from the center of the site. **Panel A.** the Cre recombinase target site is named loxP. **Panel B.** The Flp recombinase target site is named Frt.

The knockin technology is the starting point for both conditional gene inactivation and molecular fate mapping, tools that have allowed gene function and cell differentiation to be studied in the late steps of embryogenesis and in the adult.

Figure 7 shows an example of the combined use of both the recombinase target sites in the same targeting vector, a method also named Dual-Recombinase Strategy. The targeting vector contains both loxP sites flanking the mutated Exon, and Frt sites flanking the positive selection gene (PGK-Neo). Homologous recombination, occurring in the long and short arms of homology, replaces the endogenous Exon with the mutated form present in the targeting vector (Figure 7B). After the positive-negative selection, the resistant-G418 ES cell clones are injected into blastocyst to generate chimeras. Heterozygous mice carrying the mutated gene are mated to transgenic mice expressing the Frt recombinase (Flp “delete” strain), to remove the positive selection gene (PGK-Neo) leaving the “clean mutation” (Figure 7D).
**Figure 6. Recombination reactions mediated by Cre or Frt.** The relative orientation of target sites with respect to one another determines the outcome of the recombination reaction. **Panel A. Excision/Insertion.** Cre or Frt can excise a circular molecule between two direct target sites, or integrate a circular molecule into a linear one each possessing a target site. **Panel B. Inversion.** Cre or Frt can a DNA fragment between two inverted sites. **Panel C. Reciprocal translocation.** Cre or Frt can exchange sequences distal to target sites present onto linear molecules, such between target sites on nonhomologous chromosomes (I and j) to produce a balanced translocation. Black and green lines represent DNA fragments or chromosomes, with orientation indicated by letters (a-d; e-h; i and j) blue arrows indicate the reversible nature of the reactions.

The presence of loxP sites, flanking the mutated Exon, can be used to generate a “floxed” allele for conditional mutagenesis, mating the “flrted” mice with to transgenic
mice expressing the Cre recombinase (Cre “deleter” strain). If the mutated Exon is an essential region of the gene, then the recombination event results in gene inactivation. The inactivation could be either tissue or time specific.
Figure 7. Dual Recombinase Strategy. (A) Gene structure of a gene; (B) the targeting vector; (C, D, E) the targeted locus are represented. Exon are represented as blue or yellow boxes, introns as straight lines, wavy lines indicate the vector sequences. PGK-Neo and PGK-tk selection cassettes are represented as purple box and pink arrow, respectively. The targeting vector is designed with loxP sites (orange arrows) either side of the mutated Exon; while Frt sites (green circles) either side of the positive selection gene (pink square, PGK-Neo). As the consequence of homologous recombination occurring in the long and short arms of homology (indicated by a dotted black cross), the endogenous Exon D (blue square, A) is replaced by the mutated form (yellow square) present in the targeting vector (B). The targeted locus (C) is characterized by the presence of the Neomycin cassette (purle box, positive selectable gene) and the absence of the tk cassette (pink arrow, negative selectable gene). Chimeric mice are produced following positive and negative selection in ES cells. Once mice carrying the mutated gene are crossed with transgenic mice expressing the Frt recombinase (Flp “deleter” strain), the positive selection gene (PGK-Neo) is removed (D). The presence of loxP sites flanking the mutated Exon can be used to generate a floxed allele for conditional mutagenesis (E).
AIM OF THE STUDY

The aim of this PhD project is to create a mouse model of hyperbilirubinemia to study the molecular mechanisms involved in bilirubin-induced neuronal damage. This mouse model will mimic the human disease known as Criggler-Najjar Syndrome, which is caused by mutations in the UGT1a1 gene. The introduced mutation will be analogous to the one present in the Gunn rats. Those rats have a single nucleotide deletion in Exon 4 that creates an in frame stop codon immediately after the deletion. The gene produces an inactive enzyme as a consequence of the deletion. Similar mutations have also been found in human patients suffering from the Crigler-Najjar Syndrome. We introduced mutation in mice by homologous recombination in Embryonic Stem cells (ES cells), taking advantage of the gene targeting technique.
L’esperto è una persona che ha fatto in un campo molto ristretto tutti i possibili errori.

Niels Bohr
Chemical reagents

General chemicals were purchased from Sigma Chemical, Merck, Life Technologies, AnalaR, BDH Laboratory Supplies, Fluka, Riedel-de-Hael, BD Bioscience, Amersham Biosciences, BioRad, Pharmacia, Analyticals and Invitrogen.

STANDARD SOLUTIONS

All solutions are identified in the text where used except for the following:

- **PBS**: 137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, (pH 7.4).
- **TE**: 10 mM Tris-HCl (pH 7.4), 1 mM EDTA (pH 7.4)
- **10X TBE**: 108 g/l Tris, 55 g/l boric acid, 9.5 g/l EDTA
- **5X DNA sample buffer**: 50% w/v sucrose, 25% w/v urea, 0.025% w/v bromophenol blue, 5X TBE.
- **20X SSC**: 3 M NaCl, 0.3 M sodium citrate, (pH 7).
- **50X Denhardts Solution**: 5 g ficoll (Type 400, Pharmacia), 5 g polyvinylpyrrolidone, 5 g bovine serum albumin (BSA) and ddH$_2$O to 500 ml. Filtered and stored at −20°C.

ENZYMES

- **Restriction enzymes** were purchased from Amersham Pharmacia Biotech or New England Biolabs.

- **DNA modifying enzymes**: Taq Polymerase, T4 DNA ligase and were obtained from Roche Diagnostics, Klenow fragment of E.Coli, T4 Polynucleotide kinase and Calf Intestinal Alkaline Phosphatase were obtained from New England Biolabs. *PfuTurbo DNA polymerase* was from Stratagene. *Pfu DNA Polymerase* from Promega. *Expand High FidelityPLUS PCR System* from Roche. *Expand Long template PCR System* from Roche.

- **Exonucleases**: *RNase A* was purchased from Sigma Chemicals. The 10 mg/ml stock solution of RNase A was prepared in ddH$_2$O and boiled for 10-20 min to destroy
trace amounts of DNase activity. *Proteinase K* was purchased from Sigma chemicals and stocks prepared at concentrations of 10 mg/ml.

All enzymes were used following the manufacturers’ instructions.

**SYNTHETIC OLIGONUCLEOTIDES**

Synthetic DNA oligonucleotides were purchased from Sigma-Genosys.

**RADIOACTIVE ISOTOPES**

$[^\alpha-^{32}P]dCTP$ used in Southen blot analysis was supplied by Amersham Bioscience.

**BACTERIAL CULTURE MEDIA AND STRAIN**

The *E.Coli* K12 strain DH5a was transformed with the plasmids described in this study (pUC19, pBK and pFlrt1) and used for their amplification. Plasmids were maintained in the short term as single colonies on agar plates at 4 °C but for long term storage they were kept on glycerol stocks made by adding sterile glycerol to a final 30% v/v concentration to liquid bacterial cultures. Glycerol stocks were stored at −80 °C. When necessary, from the glycerol stocks an overnight culture of bacteria was grown in Luria-Bertani medium [LB medium: per litre: 10 g Difco Bactotryptone, 5 g Oxoid yeast extract, 10 g NaCl, (pH 7.5)]. Bacterial growth media were sterilised before use by autoclaving. When appropriate, ampicillin or kanamycin were added to the media at a final concentration of 10 mg/ml.

**MICE**

C57BL/6 mice were used to obtain genomic DNA used in Southern blot analysis as a control. Mice were provided by the ICGEB facility (Trieste, Italy).
**Preparation of DNA**

**Preparation of genomic DNA from tissue samples**

Because tissues generally contain large amounts of fibrous material, it is difficult to extract genomic DNA from them in high yield. The efficacy of extraction is greatly improved if the tissue is reduced to powder before homogenization in lysis buffer. Mouse liver was used in this study to obtain genomic DNA for Southern blot analysis. Briefly, mouse liver was removed from the animal and pulverized in liquid nitrogen. After the evaporation of the nitrogen, the powder was mixed with ~10 volumes (w/v) of lysis buffer [10 mM Tris-HCl (pH 8), 100 mM EDTA (pH 8), 0.5% w/v SDS, and 200 mg/ml proteinase K] and incubated 1 hour at 37°C. After that, proteinase K (20 mg/ml) was added at final concentration of 100 µg/ml and the lysate was incubated in a water bath for 3 hours at 50°C. Then, the solution was cooled at room temperature and an equal volume of phenol equilibrated with 0.1 M Tris-Cl (pH 8.0) was added. The tube containing the DNA solution-phenol was mixed by gently inversion to allow the two phases to form, and then centrifuged for 15 min at 6500 rpm. The aqueous solution was transferred in a new tube and the extraction with phenol was repeated. Finally, the genomic DNA in solution was transferred in a new tube and 0.2 volume of 10 M ammonium acetate and 2 volumes of ethanol were added. Mixing by inversion DNA immediately formed a precipitate (genomic DNA tends to wind into a ball). The genomic DNA was removed in one piece with a Pasteur pipette (whose end have been sealed and shaped into a U) and washed twice in 70% ethanol (v/v); after the evaporation of the ethanol the genomic DNA was resuspended in a suitable TE volume, placed at 65°C for one hour and overnight at 37 °C to facilitate the dissolution. The DNA was then checked by electrophoresis on a 0.6% w/v agarose gel, and stored at 4 °C until use.

**Preparation of genomic DNA from mouse tail biopsies**

Mouse tail biopsies (0.5-0.8 cm) were treated with 600 ml of lysis solution [50 mM Tris-HCl (pH 8), 100 mM EDTA (pH 8), 0.5% w/v SDS, and 200 mg/ml proteinase K] and incubated overnight at 58 °C. The mix was centrifuged 15 min at 14000 rpm in a 5415-D rotor (Eppendorf). 500 ml of supernatant were recovered and added to 500 ml
of isopropanol and mixed by inversion. The mix was centrifuged in identical conditions and the pellet was washed with 80% v/v ethanol. The final pellet (genomic DNA) was resuspended in 200 ml of TE pre-warmed at 65 °C, leaving overnight at 37 °C to facilitate the dissolution. The DNA was then checked by electrophoresis on a 0.6% w/v agarose gel, and stored at 4 °C until use.

**PREPARATION OF GENOMIC DNA FROM P96 MULTIWELLS**

This method allows to extracts genomic DNA from eukaryotic cells grown in the individual wells of microtiter plates. Each well yields sufficient genomic DNA for several PCR reactions or for analysis in a single lane of Southern blot hybridization.

Briefly, confluent cells were washed twice with PBS (100 µl) and 50 µl of lysis buffer [10 mM Tris-Cl (pH 7.5); 10 mM NaCl; 10 mM EDTA (pH 8.0) 0.5% (w/v) Sarcosyl; proteinase K 1 mg/ml] were added for each well. The plates were placed in a Tupperware box and incubated 12-16 hours at 60°C in an oven. Then, the plates were allowed to cool at room temperature and 100 µl of NaCl/ethanol solution were added. Plates were stored 1 hour at room temperature to allow the DNA to precipitate. After the ethanolic solution was discarded by gently plate inversion on a bed of paper towels. Genomic DNA remained in the wells was washed twice with 150 µl of cool 70% ethanol (v/v). The plated were allowed to dry at room temperature until the last traces of ethanol have evaporated. The genomic DNA was then resuspended in 30 µlof TE (pH 8.0). By turns, genomic DNA for Southern blot analysis was resuspended in 30 µl of the restriction enzyme mixture desired and incubated overnight at the suitable temperature for the specific restriction enzyme.

**SMALL-SCALE PREPARATION OF PLASMID DNA FROM BACTERIAL CULTURES**

Rapid purification of small amounts of recombinant plasmid DNA was basically performed with the method described in Sambrook et al. 1989 (Sambrook, 1989). Briefly, alkaline lysis of recombinant bacteria was performed by resuspending the bacterial pellet in 100 µl of solution I [50 mM glucose, 25 mM Tris-HCl (pH 8), 10 mM EDTA (pH 8)], 200 µl of solution II (0.2 M NaOH, 1 % w/v SDS) were then added and the
Contents mixed by inversion. 150 µl of solution III [3 M sodium acetate (pH 5.2)] were then added and the contents mixed by inversion. The bacterial lysate was then centrifuged in a 5415-D rotor (Eppendorf) at 14000 rpm and the supernatant transferred to a new tube. An equal volume of 1:1 v/v phenol:chloroform solution was added to the supernatant. The tube was then vortexed and centrifuged as above. An equal volume of chloroform was added to the supernatant. The tube was again vortexed and centrifuged as above. The aqueous phase containing the DNA was then recovered and the DNA pelleted by ethanol precipitation. The final pellet was resuspended in 50 µl of ddH2O and the RNA was eliminated adding 3 µl of RNAse A (10 mg/ml) and incubating the final DNA sample at 37 °C 30 min. An aliquot of such preparation were routinely taken for analysis by restriction enzyme digests and spectrophotometric quantification of DNA.

**LARGE-SCALE PREPARATIONS OF PLASMID DNA FROM BACTERIAL CULTURES**

Large scale preparations of plasmid DNA were carried out by JetStar columns (Genomed GmbH, Germany) according to the manufacturers’ instructions. In order to get a good amount of plasmid, we left a 50 ml of overnight bacterial culture using LB medium. This DNA purified procedure was used to obtain the plasmid DNA utilized for electroporation of ES cells.

**ESTIMATION OF NUCLEIC ACID CONCENTRATION AND QUALITY**

**Spectrophotometric analysis**

Nucleic acid concentration was determined by measuring the absorbance at 260 nm and 280 nm with a spectrophotometer (Jenway). An optical density of 1.0 at 260 nm is usually taken to be equivalent to a concentration of 50 µg/ml for double stranded DNA, 40 µg/ml for single stranded DNA and RNA, and approximately 20 µg/ml for single-stranded oligonucleotides samples. The ratio of values for optical densities measured at 260 nm and 280 nm is considered as 1.8 for pure sample of DNA and this is
reduced by protein contaminants (Sambrook et al. 1989). Therefore, this value was used to determinate not only the concentration but also the purity of the samples.

**UV fluorescence of intercalated ethidium bromide**

Double-stranded nucleic acids size-fractionated on agarose gels can be visualised after staining with ethidium bromide (0.5 g/ml) since fluorescence of this compound is enhanced by intercalation between bases in a double stranded nucleic acid. Therefore, it has been possible to estimate or semiquantitate DNA on a gel by comparing the intensity of UV-induced fluorescence of the sample with that of a known standard sample. This procedure was mostly performed by eye.

**Electrophoretic separation of nucleic acid**

**Agarose gel for DNA separation**

Double-stranded DNAs (restriction fragments of plasmid or genomic DNA and PCR products) were size fractionated by electrophoresis in agarose gels ranging in concentrations from 0.6% w/v (large fragments) to 2% w/v (small fragments). The gels were prepared melting the right weight of agarose (EuroClone) in 1X TBE containing ethidium bromide (0.5 mg/ml) in the microwave oven. When the agarose gel was solidified, DNA samples, diluted in 1X DNA loading buffer, were loaded into submerged wells of the gels and separated by electrophoresis in 1X TBE buffer at 80-100V for a time depending on the fragment length expected and gel concentration. DNA was visualised by UV transillumination and the result recorded by digital photography. The DNA bands were sized by confront to the molecular weight standards (1 kb DNA ladder, Invitrogen). Horizontal gels were routinely used for visualization of PCR product, fast analysis of DNA restriction enzyme digests, estimation of DNA concentration, or DNA fragment separation prior to elution from the gel.
**ELUTION AND PURIFICATION OF DNA FRAGMENTS FROM GELS**

This protocol was used to purify small amounts (less than 1 µg) of DNA for cloning purpose. The DNA samples were separated by electrophoresis onto an agarose gel as described previously. The DNA was visualized with UV light and the required DNA fragment band was excised from the gel. This slab was cut into pieces, and the QIAquick Gel Extraction Kit (Qiagen) was used according to the manufacturer’s instructions. Briefly, 300 µl of gel solubilisation Buffer QG were added for each 100 mg of the gel slice pieces and incubated at 55 °C for 15 min vortexing every 5 min. The mixture was loaded into a prepared QIAquick column and it was centrifuged at maximum speed for 1 min. The flowthrough was discarded. 750 µl of Buffer PE were added into the spin column and after 2 min, the column was centrifuged in the same conditions twice. The flowthrough was again discarded both times. To elute the bound DNA, 30-50 µl of pre-warmed ddH2O were added onto the centre of the silica matrix of the spin column and the system was centrifuged for 2 min. The amount of DNA recovered was approximately calculated by UV fluorescence of intercalated ethidium bromide in an agarose gel electrophoresis.

**ENZYMATIC MODIFICATION OF DNA**

**RESTRICTION ENZYMES**

Restriction endonucleases were used in the construction and analysis of recombinant plasmids. Each restriction enzyme cuts optimally in a buffer of specific ionic strength. All buffers were supplied by the same company that supplied the enzymes and were used according to the manufacturer’s instructions.

For analytical digests 100-500 ng of DNA were digested in a volume of 20 µl containing 1 U of the appropriate restriction enzyme. The reaction was incubated for 2-3 hours generally at 37 °C. Preparative digestion was made of 2-10 µg DNA using the above conditions and 5 U of restriction enzyme for µg of DNA in 200 µl reaction volume.
**LARGE FRAGMENT OF E.COLI POLYMERASE I AND T4 POLYNUCLEOTIDE KINASE**

These enzymes were used to treat PCR products for blunt-end ligation during construction of recombinant plasmids. The large fragment of DNA Polymerase I (Klenow) is a proteolytic product of E.Coli DNA Polymerase I. It retains polymerisation and 3’-5’ exonuclease activity, but has lost 5’-3’ exonuclease activity. This was useful for digesting specific residues added by Taq DNA polymerase at the 3’ terminus to create compatible ends for ligation. T4 Polynucleotide Kinase catalyses the transfer of phosphate from ATP to the 5’-hydroxyl terminus of DNA. It is useful for the addition of 5’-phosphate to PCR products to allow subsequent ligation.

Briefly, PCR product was incubated at 90°C to inactivate Taq Polymerase. Klenow fragment (5 U) and MgCl₂ to a final concentration to 5 mM were added to 40 µl of PCR product. The mixture was incubated at R.T. for 15 minutes and then the enzyme inactivated at 80 °C for 20 min. EDTA to a final concentration of 0.2 mM, ATP to a final concentration of 1 mM, (10 U) of T4 Polynucleotide Kinase and the proper quantity of 10X Kinase buffer were added to the above mixture and incubated at 37 °C for 30 min. The enzymes were inactivated by incubation at 80 °C for 20 min.

**T4 DNA LIGASE**

T4 DNA ligase catalyses the formation of a phosphodiester bond between adjacent 3’- hydroxyl and 5’-phosphoryl termini in DNA, requiring ATP as a cofactor in this reaction. This enzyme was used to join double stranded DNA fragments with compatible sticky or blunt ends, during generation of recombinant plasmid DNAs.

Generally, 20 ng of linearized vector were ligated with a 5-10 fold molar excess of insert in a total volume of 20 µl containing 1X ligase buffer and 1U of T4 DNA ligase. Reaction was carried out for 2-4 hours at R.T. for sticky end ligations and O.N. at 16 °C for blunt end ligations. In some reactions synthetic oligonucleotides were included in the reaction. In these cases, the amounts added to each reaction to obtain inclusion of oligonucleotides in the resulting plasmid were about 100 fold molar excess over the DNA vector.
AMPLIFICATION OF SELECTED DNA FRAGMENTS

The polymerase chain reaction was performed on genomic DNA or plasmid DNA following the basic protocol accompanying the Taq DNA Polymerase enzyme. The volume of the reaction was usually 30 µl with 1X Taq buffer, dNTP mix (200 µM final concentration), appropriate sense and antisense oligonucleotide primers (15 mM final concentration), and 1-2 U of Taq DNA Polymerase.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ Probe UGT dir</td>
<td>ctgcaagcataagggaccgaga</td>
</tr>
<tr>
<td>5’ Probe UGT rev</td>
<td>ataattgcttcacagcagtc</td>
</tr>
<tr>
<td>Sall UGTDir</td>
<td>cgtgcctgagcttgcatcgttgagctatatt</td>
</tr>
<tr>
<td>RV UGT dir B</td>
<td>tggcatccttttctgacattc</td>
</tr>
<tr>
<td>RV UGT rev A</td>
<td>ccacagtcgatagcagattc</td>
</tr>
<tr>
<td>RV UGT dir C</td>
<td>agggagggaaaaacagagcagtcc</td>
</tr>
<tr>
<td>RV UGT rev B</td>
<td>tggaggtgcagaggatttc</td>
</tr>
<tr>
<td>Sall UGT rev</td>
<td>cgtgcacgcaagttgcttgatcagggacc</td>
</tr>
<tr>
<td>Bam UGT EX4 dir</td>
<td>cgcggatccggcttgctcataaagacattc</td>
</tr>
<tr>
<td>Bam UGT EX4 rev</td>
<td>cgcggatccggctcactaccccccacatc</td>
</tr>
<tr>
<td>BstBI UGT dir</td>
<td>cgtgccttgcaataaaagatgtggtggtggtggtggtggt</td>
</tr>
<tr>
<td>BstBI UGT rev</td>
<td>cgtgccttgcaactagcggcaacaagcagataacct</td>
</tr>
<tr>
<td>3’ Probe UGT dir</td>
<td>tggctcagtctgtgaaggtaaca</td>
</tr>
<tr>
<td>3’ Probe UGT rev</td>
<td>tctctgcatccacagcagctcta</td>
</tr>
</tbody>
</table>

Table 1: List of oligonucleotides used in PCR amplification of the fragments used in the targeting vector.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC19 universal</td>
<td>gtaaaacgacggccagt</td>
</tr>
<tr>
<td>pUC19 rev</td>
<td>aacagctagcagcatg</td>
</tr>
<tr>
<td>UGT1A1 FrgA seq dir</td>
<td>tgcctcagaggtgattgc</td>
</tr>
<tr>
<td>UGT1A1 FrgA-1 seq rev</td>
<td>gcacattacctctggagca</td>
</tr>
<tr>
<td>UGT1A1 FrgA 4977seq dir</td>
<td>gtttctctcgtggtctcat</td>
</tr>
<tr>
<td>UGT1A1 FrgA 5482seq dir</td>
<td>ttgtgaggtggcatgacag</td>
</tr>
<tr>
<td>UGT1A1 FrgB seq dir</td>
<td>gtttctgattccgcagctagt</td>
</tr>
<tr>
<td>UGT1A1 FrgA seq rev</td>
<td>ctatgtagctgttgctctctg</td>
</tr>
<tr>
<td>UGT1A1 FrgB seq rev</td>
<td>atgtatggacactgcaggct</td>
</tr>
<tr>
<td>mUGT1A1 pr7777 dir</td>
<td>tatggaccttgctatggttcc</td>
</tr>
<tr>
<td>mUGT1A1-10387 dir</td>
<td>gcctccatggcgctgcac</td>
</tr>
<tr>
<td>mUGT1A1-11100 dir</td>
<td>taactggagtcaagcctagt</td>
</tr>
<tr>
<td>mUGT1A1-11800 dir</td>
<td>gattctcaagctgccctgctg</td>
</tr>
<tr>
<td>mUGT1A1-13076 rev</td>
<td>ctatgctaggtctgctgcattaatc</td>
</tr>
</tbody>
</table>

Table 2: List of oligonucleotides used in PCR sequencing of the fragments used in the targeting vector.
### Materials & Methods

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGT1A1 Gunn mut dir</td>
<td>gtgaccctgaagtgccttaaatgactgctgatgatttg</td>
</tr>
<tr>
<td>UGT1A1 Gunn mut mut</td>
<td>caaatcatcagcagtcatttaaggacattcaggtggtcac</td>
</tr>
</tbody>
</table>

Table 3: List of oligonucleotides used in point mutation in Exon 4.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ Probe UGT dir</td>
<td>ctgcaagcataaggacccgaga</td>
</tr>
<tr>
<td>5’ Probe UGT rev</td>
<td>ataatagcctcaagcgatgac</td>
</tr>
<tr>
<td>5’ Probe UGT 3453 rev</td>
<td>gtgggtgtgctgcgcgtgcgtgcgtgcgtgcg</td>
</tr>
<tr>
<td>3’ Probe UGT dir</td>
<td>tggctcagttctgtgaggttaaca</td>
</tr>
<tr>
<td>3’ Probe UGT rev</td>
<td>tctctgcatccagccagctcta</td>
</tr>
<tr>
<td>mUGT1A1-13076 rev</td>
<td>ctgattaaggtctgcattaatac</td>
</tr>
<tr>
<td>Neo dir</td>
<td>tgtcaagaccgacctgtccg</td>
</tr>
<tr>
<td>Neo rev</td>
<td>tattcggcaagccagcatcg</td>
</tr>
</tbody>
</table>

Table 3: List of oligonucleotides used in PCR amplification of the Probes used in the Southern blot analysis.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>mUGT1a1-ScreenWT dir</td>
<td>catcttagcttgattgtgtgc</td>
</tr>
<tr>
<td>mUGT1a1-ScreenNEO dir</td>
<td>ggtcttcagtgtgctattgt</td>
</tr>
<tr>
<td>UGT Screen REV 2</td>
<td>ctacttccatttgctacgtcc</td>
</tr>
</tbody>
</table>

Table 5: List of oligonucleotides used in Multiplex PCR.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ Probe UGT dir</td>
<td>ctgcaagcataaggacccgaga</td>
</tr>
<tr>
<td>Neo rev</td>
<td>tattcggcaagccagcatcg</td>
</tr>
</tbody>
</table>

Table 6: List of oligonucleotides used in PCR amplification at the 5’ region of the targeted locus.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence 5’- 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC19 universal</td>
<td>gttaaacgacgacgccagt</td>
</tr>
<tr>
<td>pUC19 rev</td>
<td>aacagctagacgccatg</td>
</tr>
</tbody>
</table>

Table 7: List of oligonucleotides used in PCR amplification of the loxP-Exon 4 region.
**Preparation of bacterial competent cells**

Bacterial competent cells were prepared following the method described by Chung (Chung et al., 1989). *E.Coli* strains were grown overnight in 5 ml of LB at 37°C. The following day, 100 ml of fresh LB were added and the cells were grown at room temperature for 4-5 h until the OD₆₀₀ was 0.3-0.4. The cells were then put in ice and centrifuged at 4 °C and 1000g (Eppendorf, 5810-R rotor) for 15 min. The pellet was resuspended in 10 ml of cold TSS solution [10% w/v PEG molecular weight 4000, 5% v/v DMSO, 35mM MgCl₂, (pH 6.5) in LB medium]. The cells were aliquoted, rapidly freeze in liquid nitrogen and stored at -80°C. Competence was determined by transformation with 0.1 ng of pUC19 and was deemed satisfactory if this procedure resulted in more than 100 colonies (transformation efficiency ≥10⁵ cfu/ug).

**Transformation of bacteria**

Transformation of ligation reactions were performed using 1/2 of the reaction volume. Transformation of clones was carried out using 1ng of the plasmid DNA. The DNA was incubated with 60 ml of competent cells for 30 min on ice. The cells were shock 2 min at 42 °C. After the step of heat shock, 100 ml of LB were added and the bacteria allowed recover for 10-15 min at 37 °C. The cells were then spread on agarose plates containing the appropriate antibiotic. The plates were then incubated for 12-15 hours at 37 °C.

When DNA inserts were cloned into b-galactosidase-based virgin plasmids, 70 ml of IPTG 100 mM and 36 ml of X-Gal (4 % w/v in dimethylformamide) were spread onto the surface of the agarose prior to the plating to facilitate screening of positive clones (white colonies) through identification of galactosidase activity (blue colonies), which indicates the negative clones.

**Sequence analysis for cloning purpose**

Sequence analysis (Sanger method) of the cloning fragments was performed by BMR genomics (PD).
**CONSTRUCTION OF THE UGT1A1 TARGETING CONSTRUCT**

The complete UGT1a1 gene, including the promoter and the 3’ gene-flanking regions, was amplified by PCR (see Table 1) and cloned into the pUC19 vector. The PCR was performed using genomic DNA as template and we used three different Pfu DNA polymerase systems, which has proofreading activity and has a lower rate error than the normal Taq polymerase [Pfu DNA Polymerase (Promega); Expand High FidelityPLUS PCR System (Roche) and the Expand Long template PCR System (Roche)].

The total length of the homology region was about 8 Kb: the long arm (5.1 kb) included Exons 1 to 3 and intervening introns; the fragment containing Exon 4 with its flanking introns (about 0.7 Kb); and the sort arm (2.8 kb) containing those sequences downstream to Exon 4: Exon 5 and downstream sequences. The long arm of homology was divided in three smaller and overlapping fragments (A = 3.1 Kb, B = 2.2 Kb, C = 1.8 Kb) to facilitate the cloning and sequencing procedures. Two additional fragments (5’ Probe = 0.7 Kb and 3’ Probe = 1.0 Kb), which map adjacent to the UGT1a1 sequences used in the construct, were cloned and tested as Southern blot probes. The 3’ Probe has been subsequently utilized to screen the G418-resistant ES clones for homologous recombination in the UGT1a1 locus. All fragments mentioned above were separately cloned. The PCR products were separated in agarose gels, cut from the gel and the DNA, purified as explained previously and cloned in pUC19 (New England Biolabs). The intermediate pUC19 constructs were then sequenced (BMR Genomics, PD) (in table 2 are shown the primers used for sequence analysis).

The absence of mutations was confirmed by aligning the available genomic sequence (C57BL/6) with that obtained from sequencing two or more independent clones, originating from different PCR reactions.

To prepare the final targeting vector, the intermediate pUC19 constructs were digested with the specific enzymes to free the inserts. For cloning purpose we used also pBK clonig vector (kanamycin resistant); this strategy allowed us to skip the purification steps: changing from ampicillin to kanamicin resistance and vice-versa there no need to purify the insert, only the clones having the correct resistance will grow in presence of the specific antibiotic.

Finally, to build the pFlrt1 targeting vector we selected those clones containing the sequence closest to the annotated genomic one and we inserted them one by one checking each new insertion by restriction enzyme analysis. First the short arm oh
homology in the BstBI restriction enzyme site; second the Exon 4 mutated, and finally the long arm of homology (composed of A, B and C fragments previously join together in pBK cloning plasmid). We confirmed that the final vector contained no mutations in the exonic regions inserted.

IN VITRO SITE DIRECTED MUTAGENESIS

In vitro site directed mutagenesis is a rapid and high efficient procedure that allows site-specific mutagenesis in double-stranded plasmids. The method is performed using the high fidelity PfuTurbo DNA polymerase and a temperature cycler.

The double-stranded DNA vector of interest (50 ng) and the two synthetic oligonucleotides primers, each complementary to opposite strands of the vector, containing the desiderate mutations (125 ng of each) were diluted into the PuF DNA Pol reaction mix (2.5 U/ml PfuTurbo DNA polymerase, 1X reaction buffer, 200 µM dNTP mix). The complete mix was introduced in the temperature cycler [Gene Amp PCR System (Applied Biosystems)] and the primers were extended during temperature cycling by the PfuTurbo DNA polymerase [cycling parameters: 1 cycle (initial denaturation step at 95°C for 30 sec) and 18 cycles (denaturation step at 95°C for 30 sec; annealing step at 55°C for 1 min; extension step at 68°C for 1 min/kb of plasmid length)]. At the end of the reactions the incorporation of oligo primers generates a mutated plasmid containing nicks. After the temperature cycling, 10 U/µl of Dpn I restriction enzyme were added to the amplification reaction and an incubation of 1h at 37 °C followed. After the temperature cycling, the amplification product was treated with Dpn I restriction enzyme (10 U/µl of the enzyme, incubation of 1h at 37 °C). Considering that the DpnI endonuclease is specific for methylated DNA and that DNA isolated from almost all E.Coli strains is dam methylated (therefore susceptible to DpnI digestion), this specific restriction enzyme was used to digest the parental DNA template and to select for mutation-containing synthesised DNA. The mutated and nicked vector DNA containing the mutation is then transformed into E.Coli K12 strain DH5α cells having a higher grade of competence (transformation efficiency ≥10^6 cfu/µg). A more efficient transformation procedure was used. The DpnI-treated DNA was incubated with 50 ml of competent cells for 30 min on ice. The cells were shock 45 sec at 42°C and then placed for 2 min in ice. After these steps, 500 ml of NZY+broth [1% NZ amine (casein
MATERIALS & METHODS

hydrolysate), 0.5% yeast extract, 0.5% NaCl, 12.5 mM MgCl₂, 12.5 mM MgSO₄, 20 mM glucose, (pH 7.5)] preheated to 42 °C were added and the bacteria allowed recover for 1 h min at 37 °C with shaking. The cells were then spread on agarose plates containing the appropriate antibiotic. The plates were then incubated for 12-15 hours at 37 °C.

This method was used to create the deletion construct of UGT1a1 Exon 4. Intermediate pUC19 construct containing the wild type Exon 4 was used as the double-stranded DNA template. The two synthetic oligonucleotides primers containing the one base deletion are shown in the Table 3.

SOUTHERN BLOT ANALYSIS

Southern blot transfer and hybridization was used to analyze G418 resistant clones.

Briefly, genomic DNA was first digested with one restriction enzyme (overnight); after digestion the DNA was checked in an agarose gel for complete digestion and precipitated adding 0.2 volume of 10 M ammonium acetate and 2 volumes of ethanol. After two washes in 70% ethanol (v/v) the DNA was resuspended in TE. The resulting fragments were separated according to size by electrophoresis through a 0.6% (w/v) agarose gel (20V, overnight). The gel was photographed under UV light and washed twice (for 10 minutes each) with 0.2 N HCl for depurination of DNA, then twice (10 minutes each) with Denaturation Solution [1.5 M NaCl; 0.5 M NaOH]. The DNA was transferred from the gel to the nylon membrane (Z-Probe, BioRad) by upward capillary transfer using Denaturation solution (overnight). The day after the blotting was removed. To checked blotting efficiency the membrane was observed under UV light, whereas the agarose gel was stained with a solution containing ethidium bromide. After the membrane was neutralized twice (10 minutes each) with Neutralizing Solution [1 M Tris-Cl; 1.5 M NaCl] and dried at room temperature on a Whatman 3 MM filter paper. The DNA was bound to the membrane by UV cross-linking (2X, 0.25 J). Then, the membrane was prehybridised for at least 1h at 65°C in 20 ml of Pre-hybridization solution [6X SSC, 5X Denhardt’s reagent, 0.5% (w/v) SDS, 100 µg/ml salmon sperm DNA] and hybridized overnight at 65°C to the radiolabelled probe, in 10 ml of a pre-hybridisation solution. The day after the hybridization solution was removed and the membrane was washed with several washing solutions at 65°C: 4X SSC, 0.1% SDS (w/v)
to remove the excess of the hybridisation probe; 2X SSC, 0.1% SDS (w/v) and 0.5X SSC, 0.2% SDS (w/v) with more stringent conditions to eliminate non specific hybridisation. After each wash the membrane was checked for radioactive signal using a geiger (Canberra Packard SRL) to evaluate if continue or suspend the washing. Finally, the membrane was immediately covered with plastic wrap mounted on Whitman 3 MM filter paper and exposed overnight using Cyclone screen (Cyclone, Storage Phosphor System, Canberra Packard SAL).

**MINI-SOUTHERN BLOT ANALYSIS**

This extension of the method was used to analyze DNA restriction fragments originating from the digestion of genomic DNA derived from p96 multiwells containing the G418 resistant clones (see Preparation of genomic DNA from p96 multiwells for a detailed description).

**HYBRIDIZATION PROBES PREPARATION**

The different DNA fragment used as Southern blot hybridisation probes were obtained by PCR from total genomic mouse DNA or pFlrt1 vector (Neo Probe). The specific oligonucleotide primers used for the reactions are listed in Table 3. In all cases, the PCR product were separated in agarose gels and the band corresponding to the desired fragment in each case was cut from the gel and the DNA was purified as explained previously.

Radioactive probes were generated from 50 ng of DNA by the Rediprime II Random Prime Labelling System (Amersham Bioscience) using \( \alpha^{32}P \) dCTP (Amersham Bioscience) and were then purified from unincorporated nucleotides by Nick columns (Amersham bioscience). Both the labelling and the probe purification were carried out according to the manufacturers’ instructions.
REUTILIZATION OF SOUTHERN BLOT MEMBRANES

Removal of radioactive probes from membranes for subsequent re-hybridisation was performed by incubation of the blot in sterile boiling ddH$_2$O containing 0.5% SDS (w/v) solution at 65 °C for approximately 30 min (twice), shaking frequently. The blot was removed from the stripping solution, allowed to cool on a Whatman 3 MM filter paper, and sealed in a plastic bag and stored at 4 °C or R.T. until use. The membranes were analysed by Cyclone screen (Storage Phosphor System, Canberra Packard SRL) to check successful stripping.

DATABASE AND BIOINFORMATICS ANALYSES

UG1A1 gene sequence used to clone the targeting fragments was obtained from GenBank database.
Sequence analysis of the cloning fragments was performed using either ClustaW or CLC programs.
RepeatMasker program was used to highlight repetitive sequence in the Southern blot Probes (5’ and 3’ Probes).
Non è forte colui che non cade mai,
ma colui che cadendo
si rialza.

Johann Wolfgang von Goethe
The objective of the present study is to create a mouse model of hyperbilirubinemia. This mouse model will closely resemble the human pathology named Crigler-Najjar Syndrome I (CNI) and its well-established animal model, the Gunn rat. Those rats have a single nucleotide deletion in Exon 4 that creates an in frame stop codon immediately after the deletion. The gene produces an inactive enzyme as a consequence of the deletion. We introduced mutation in mice by homologous recombination in Embryonic Stem cells (ES cells), taking advantage of the gene targeting technique.

The construction of the mutated mouse strain was divided in various steps:

(a) Cloning and sequencing of the UGT1a1 gene
(b) Generation of one base deletion in the UGT1a1 Exon 4
(c) Preparation of the final UGT1a1 targeting construct
(d) Setting up of the screening conditions
(e) Electroporation, selection and screening of the G418-resistant clones
(f) Characterization of the G418-resistant clones
(g) Blastocyst microinjection to generate chimeras
(h) Germ line transmission of the mutated UGT1a1 gene.
(i) Changing of the genetic background of UGT1a1 gene mutated mouse to an homogeneous one.

a) Cloning and Sequencing of the UGT1a1 Gene

The complete UGT1a1 gene, including the promoter and the 3’ gene-flanking regions, was amplified by PCR and cloned into the pUC19 vector. The PCR was performed using genomic DNA as template and we used the Pfu DNA polymerase system, which has proofreading activity and has a lower rate error than the normal Taq polymerase. The DNA was prepared from the same ES cell line that has been subsequently used for the electroporation (C129Sv/J) in order to avoid a reduction of the recombination efficiency due to the presence of strain-specific polymorphisms.

The total length of the homology region was about 8 Kb (Figure 1): the long arm (5.1 kb) included Exons 1 to 3 and intervening introns; then the fragment containing Exon 4 with its flanking introns (about 0.7 Kb); and the sort arm (2.8 kb) containing those sequences downstream to Exon 4: Exon 5 and downstream sequences. The long
arm of homology was divided in three smaller and overlapping fragments (A = 3.1 Kb, B = 2.2 Kb, C = 1.8 Kb) to facilitate the cloning and sequencing procedures. Two additional fragments (5’ Probe = 0.7 Kb and 3’ Probe = 1.0 Kb), which map adjacent to the UGT1a1 sequences used in the construct, were cloned and tested as Southern blot probes using C57BL/6 and C129Sv/J genomic DNAs. The 3’ Probe has been subsequently utilized to screen the G418‐resistant ES clones for homologous recombination in the UGT1a1 locus. All fragments mentioned above were separately cloned. The absence of mutations was confirmed by aligning the available genomic sequence (C57BL/6) with that obtained from sequencing two or more independent clones, originating from different PCR reactions.

![Figure 1. Scheme of the complete UGT1a1 gene.](image)

The position of the long and short arms of homology used for the targeting construct are indicated. The total length of the homology region is about 8 Kb. The long arm (5.1 kb) was divided in three smaller and overlapping fragments (double‐head arrows, A = 3.1 Kb; B = 2.2 Kb; C = 1.8 Kb) to facilitate the cloning and sequencing procedures. The short homology arm (2.8 Kb) contains those sequences downstream to Exon 4. Two different Southern blot probes (black boxes, 5’ Probe = 0.7 Kb and 3’ Probe = 1.0 Kb), which map adjacent to the UGT1a1 sequences used in the construct, were cloned and tested using C57BL/6 and C129Sv/J genomic DNA.

To reduce the presence of undesired mutations, we initially used the Pfu DNA Polymerase (Promega) to clone the UGT1a1 gene fragments. Subsequently, we checked the DNA sequence of these clones (deriving from the same genomic region) by aligning them with the genomic sequence from the GenBank (C57BL/6 mouse strain). Since we noticed the presence of many mismatches, we tested two other low error DNA polymerase systems: Expand High FidelityPLUS PCR System (Roche) and the Expand Long template PCR System (Roche) to find the one having the lower error rate. However,
there were no significant differences in the error rate among the three PCR systems. To investigate whether these mismatches were errors made by the polymerase strain specific polymorphism, we aligned the sequences derived from the different PCR systems (of the same genomic fragment). The rationale of analyzing fragments derived form independent PCR reactions is the following: if a difference is seen in all of the clones, it means we are dealing with a strain-specific polymorphism but, if the difference is present in some of the clones and not in others, it suggest that this was a PCR error. Surprisingly, we realized that we were in the presence of two different groups of clones, each of them carrying a different set of mismatches (Figure 2). Clones belonging to the same group always showed the same mismatches. These data suggest that the origin of the two different groups could be due to the presence of two different alleles of the UGT1a1 gene, instead of mutations generated by errors in the different DNA polymerase used to amplify each fragment.

Figure 2. Sequence alignments. The genomic sequence of the annotated UGT1a1 gene was retrieved form the GenBank (from the C57BL/6 strain, outlined in light blue) and was aligned with the sequences obtained from the different clones of the same genomic region (each clone deriving from different PCR reactions). We observed two groups of clones (A and B), each of them carrying a different set of mismatches (red squares). Clones belonging to one group always showed the same mismatches.

To build the final pFlrt1 targeting vector we selected those clones containing the sequence closest to the annotated genomic one. We confirmed that all selected clones contained no mutations in the exonic regions.
**b) Generation of One Base Deletion in the UGT1a1 Exon 4**

We cloned the Exon 4 and part of its flanking introns (0.7 Kb) in the pUC19 vector. A single base deletion was inserted at the same position to the one present in the Gunn rat model by using the *in vitro* site direct mutagenesis (as described in Materials and Methods). The presence of the deletion and the absence of undesired mutations were confirmed by sequencing of the entire clone. The one-base deletion produces a shift in the reading frame, creating a stop codon immediately after the deletion. Figure 3 shows the wild type and mutated sequences of Exon 4, and their corresponding reading frame.

![Wild type exon sequence](Val Leu Glu Met Thr Ala)

![Mutated exon sequence](GTC CTG AAA TGA ATG ACT GCC)

**Figure 3. Single nucleotide deletion in UGT1a1 Exon 4 and their corresponding reading frame.** Wild type (in blue) and mutated sequences (in orange) are shown. Immediately after the one base deletion (G, bold letter in wild type sequence) an in frame stop codon is produced. The gene encodes an inactive enzyme as a consequence of the deletion.

**c) Preparation of the Final UGT1a1 Construct**

The UGT1a1 genomic fragments were cloned into the pFlrt1 targeting vector, which is specifically designed for generating constructs to be used in homologous recombination in ES cells (Figure 4). It contains: a) unique cloning sites (MCS-Multiple Cloning Site and BsmI Restriction site), useful for the cloning of the long and short arms of homology; b) a Neomycin cassette [G418 positive selectable gene driven by the Phospho Glycerate Kinase (PGK) promoter, which is active in ES cells] flanked by Frt sites (recognized by the Flp recombinase); c) a unique BamHI cloning site flanked by loxP sites (recognized by the Cre recombinase), for cloning the mutated Exon 4 of UGT1a1; and d) a TK cassette (gancyclovir negative selectable gene driven by the PGK promoter), used to enrich in gancyclovir resistant clones favouring the detection of homologous recombination events. The long UGT1a1 arm of homology was cloned in the
Sall site, the mutated Exon 4 in the BamHI site, and the short arm of homology in the BstBI site.

The scheme in Figure 5 shows the general strategy used for the targeted disruption of the mouse UGT1a1 locus. Homologous recombination occurs between the homology arms present in the construct and the UGT1a1 gene, replacing the wild type Exon 4 by the mutated version present in the targeting construct. Together with the mutated Exon, the Neomycin cassette (NeoR) is also introduced. This cassette confers resistance to the antibiotic G418, allowing only the growth of cells expressing the Neomycin gene product. The targeted locus is also characterized by the absence of the thymidine kinase (TK) cassette (negative selectable gene). The Neo cassette has flanking Frt sites, recognized by the Flp recombinase. It is subsequently removed by crossing heterozygous mice carrying the UGT1a1 mutation with the Flp "deleter" strain.

**Figure 4. Scheme of the pFlrt1 targeting vector.** The pFlrt vector (7083 bp) is specifically designed for generating constructs to be used in homologous recombination in ES cells. It contains a Neomycin cassette (purple box, G418 positive selectable gene driven by the PGK promoter, active in ES cells) flanked by Frt sites (green circles, recognized by the Flp recombinase), a cloning site flanked by loxP sites (orange arrows, recognized by the Cre recombinase), other useful unique cloning sites (yellow box, MCS-Multiple Cloning Sites), and a TK cassette (pink arrow, gancyclovir negative selectable gene driven by the PGK promoter). The long UGT1a1 arm was cloned in the Sall site, the mutated Exon 4 in the BamHI site and the short arm in the BstBI site.
The resulting pFlrt1 targeting vector was checked for correct insertion of the fragments by restriction enzyme cuts and, in some cases, sequencing of the regions close to the cloning sites (data not shown).

Figure 5. Targeted disruption of the mouse UGT1a1 locus. The structure of the UGT1a1 gene (A), the targeting construct (B) and the targeted locus (C and D) are represented. Exons are represented as blue or yellow boxes, introns as straight lines, wavy lines indicate vector sequences. PGK-Neo and PGK-TK selection cassettes are represented as purple and pink boxes, respectively. As the consequence of homologous recombination occurring in the long and short arms (indicated by a dotted black cross), the endogenous Exon 4 (blue square, A) is replaced by the mutated form (yellow square) present in the targeting vector (B). The targeted locus (C) is characterized by the presence of the Neomycin cassette (purple box, positive selectable gene) and the absence of the TK cassette (pink arrow, negative selectable gene). The Neo cassette is subsequently removed by crossing of the heterozygous mice carrying the UGT1a1 mutation with a Flp “deleter” strain (D).
D) SETTING UP OF THE SCREENING CONDITIONS

To screen the G418-resistant clones originated from the ES cells electroporation with the targeting vector, we set up the optimal condition for two techniques: Southern blot and PCR reaction.

SOUTHERN BLOT ANALYSIS

To determine whether highly repetitive sequences were present in the selected probes (5’ and 3’ Probes), we performed a “reverse” Southern blot experiment (Figure 6).

**Figure 6. “Reverse” Southern blot. Panel A) Ethidium bromide gel.** The plasmids (in pUC19 vector) containing the different inserts were cut with EcoRI and HindIII to free the insert and run in an agarose gel. **Panel B) “Reverse” Southern blot analysis.** The gel shown in panel A was blotted onto a nylon membrane and hybridized against radioactively labelled total genomic DNA as a probe. While single copy sequences are low represented in the genome and will not produce a signal after hybridization and exposure of the membrane, the over represented highly repetitive sequences will be predominantly labelled and in the case repetitive sequences were present in the DNA fragments blotted on the membrane, they will be detected. We observed no signal with the 3’ Probe, Fragment B, C and the Short Arm. On the contrary, a clear signal (red circles) was observed for the 5’ Probe and the Fragment A. C57BL/6 genomic DNA digested with HindIII was used as an internal positive control of the technique.
Unlabelled plasmids containing the different inserts (5’ and 3’ Probes, and fragments derived from the targeting construct) were digested with EcoRI and HindIII restriction enzymes to free the inserts, were run in an agarose gel and blotted onto a nylon membrane. Total C57BL/6 genomic DNA was radioactively labelled and used as the hybridization probe against the membrane. While single copy sequences are low represented in the genome and will not produce a detectable signal after hybridization and exposure of the membrane, the over represented high copy regions will be predominantly labelled and, in the case repetitive sequences were present in the DNA fragments blotted on the membrane, they will be detected. We observed no signal with the 3’ Probe, Fragments B, C, Exon 4 and the short arm. On the contrary, a clear signal was observed for the 5’ Probe and the Fragment A suggesting the presence of repetitive sequences in those fragments.

Therefore, to set up the conditions of the Southern blot analysis in the locus we initially digested C57BL/6 genomic DNA with different restriction enzymes and used the 3’ Probe to verify the pattern of the bands in the 3’ region of the UGT1a1 locus. As the 5’ Probe and Fragment A, showed to recognize highly repetitive sequences in the genome we selected a DNA segment of the Fragment B (namely B Probe = 1.1 Kb, internal to the homology region) to check for homologous recombination in the 5’ end of the targeted locus.

To evaluate the efficiency of hybridization of the 3’ Probe we performed a Southern blot using C57BL/6 genomic DNA digested with different restriction enzymes (EcoRV, BamHI and HindIII; see Figure 7). We observed partial digestion of the genomic DNA with EcoRV and HindIII (15 Kb and 7 Kb bands, respectively, indicated by the white triangles), while BamHI showed an unknown band of 2.8 Kb (black triangle). The HindIII restriction enzyme turned out to be the more suitable one for the Southern blot analysis at the 3’ end of the homology region in the UGT1a1 locus.

Since we observed a considerable background level in the Southern blot analysis, we analyzed the sequence of the 3’ Probe with the RepeatMasker program, with the aim of detecting highly repeated sequences not detected in our previous “Reverse Southern” experiment (Figure 6).
Figure 7. Southern blot using the 3’ Probe. Panel A) Scheme of the mouse UGT1a1 locus. The relative position of the restriction enzymes sites (EcoRV, BamHI, HindIII) is shown. The coloured lines indicate the size of the restriction fragments originated from the digestion of C57BL/6 genomic DNA and detected by Southern blot using 3’Probe (black box). Thick lines represent expected bands, while thin lines represent partial digestions. Panel B) Southern blot analysis. C57BL/6 genomic DNA digested with EcoRV, BamHI and HindIII restriction enzymes were run in an agarose gel, then blotted onto a nylon membrane and hybridized against radioactively labelled 3’ Probe. The expected bands are indicated with red triangles. We observed partial digestion of the genomic DNA with EcoRV and HindIII restriction enzymes (15 Kb and 7 Kb bands, respectively, white triangles), while BamHI showed an unexpected band of 2.8 Kb band (black triangle).
We found that the 3’ Probe contains an Alu sequence from position 324 to 447 and an LTR sequence from position 441 to 1065 (Figure 8A). For this reason we decided to shorten the 3’ Probe eliminating all sequences downstream to position 230.

To improve the labelling efficiency of the 3’ shortened Probe (230 bp) we prepared a concatamer composed of three tandem copies of the shortened version of the 3’ probe, (690 bp in total, Figure 8B). This fragment was used as a Probe in the Southern blot analysis.

Figure 8. Scheme of the 3’ Probe. Panel A) 3’ Probe sequence analyzed with the RepeatMasker program. The analysis of the 3’ probe with the RepeatMasker program showed the presence of highly repetitive sequences within the 3’ probe: an Alu sequence (blue box) from position 324 to 447 and an LTR sequence (light blue box) from position 441 to 1065. For this reason we decided to shorten the probe eliminating all sequences downstream to position 230. Panel B) Scheme of the Concatamer of the 3’ Probe. To improve the labelling efficiency of the shortened 3’ Probe (230 bp) we prepared a concatamer composed of three tandem copies of the shortened version (690 bp in total). This fragment was used as a probe in Southern blot analysis.

We tested the 3’ Probe concatamer against genomic DNA from C57BL/6 and C129Sv/J (same ES cell line used for the electroporation) digested with HindIII restriction enzyme (Figure 9). Similarly to what previously seen (Figure 7), we observed an additional 7 Kb band in the C57BL/6 genomic DNA. Analysis of C129Sv/J genomic DNA showed only the expected 6 Kb band (Figure 9A, red triangle in lane 1). As mentioned above, the 7 Kb band could be the result of a partial digestion of the DNA. An alternative possibility could be the absence of one HindIII site in one of the two alleles, producing a 7 Kbp band instead of the expected 6 Kbp one.
RESULTS

To have a rough idea of the lower detection limit of the Southern Blot analysis using the 3’ Probe, we performed a calibration curve using HindIII-digested C129Sv/J genomic DNA (Figure 9B). The determination of the sensitivity is necessary to know whether the amount of DNA obtained from the G418 ES cell clones grown in a p96 multiwell plate, an amount that is highly variable and depends on the number of cells per well, could be detected.

Figure 9. Southern blot analysis using the 3’ Probe concatamer. Panel A) Southern blot analysis. Genomic DNA from C57BL/6 and C129Sv/J mice was digested with HindIII restriction enzyme, run in an agarose gel, then blotted onto a nylon membrane and hybridized against the radioactively labelled concatamer of the 3’ Probe. We observed an additional 7 Kb band in the C57BL/6 genomic DNA (white triangle in lane 2), which was not observed in the analysis using C129Sv/J genomic DNA that showed only the expected 6 Kb band (red triangle in lane 1). Panel B): Calibration curve. To determine the lower DNA amount detectable by Southern blot using the 3’ Probe concatamer we run different amounts of HindIII digested C129Sv/J genomic DNA (from 0.5 µg up to 6 µg) in an agarose mini-gel system. The gel was blotted onto a nylon membrane and hybridized against the radioactively labelled Probe. We were able to detect a specific signal even when 0.5 µg of cut DNA were loaded (lane 2) but only after 48 hs exposure, while the signal in the other lanes, having higher amounts of DNA, was visible after overnight exposure (lanes 3-5).

We run different amounts of HindIII digested C129Sv/J genomic DNA (from 0.5 µg up to 6 µg) in an agarose mini-gel system. After hybridization of the membrane, we were able to detect a specific signal even when 0.5 µg of cut DNA were loaded (Figure
RESULTS

9B, lane 2), but only 48 hs after exposure, while the signal in the other lines, having higher amounts of DNA, was visible after overnight exposure (Figure 9B, lanes 3-5).

To test in advance DNA-preparation conditions similar to those of the G418-resistant ES cell clones (cultured in p96 multiwell, allowing the preparation of variable and low amounts of genomic DNA), we tested the 3’ Probe concatamer in Southern blot using genomic DNA from NMuLi (liver-derived mouse cells) and C129Sv/J ES cells cultured in p96 multiwells (Figure 10). After DNA preparation, the genomic DNA was digested with HindIII restriction enzyme, run in an agarose gel, then blotted onto a nylon membrane and hybridized against radioactively labelled 3’ Probe concatamer. We were able to detect a specific signal after an overnight exposure (Figure 10, red triangle, 6 Kb). We also observed a higher molecular weight band of 7 Kb (Figure 10, white triangle) in the lanes loaded with NMuLi genomic DNA, probably due to partial digestion or the absence of a HindIII site in one of the two UGT1a1 alleles.

Figure 10. Setting up of the conditions to perform mini-Southern blot analysis with the 3’ probe. Genomic DNA from N-muli (liver-derived mouse cells) and C129Sv/J ES cells (cultured in p96 multiwells) was digested with HindIII restriction enzyme and run in an agarose mini-gel system. The gel was blotted onto a nylon membrane and hybridized against radioactively labelled 3’ Probe concatamer. We were able to detect a specific signal after an overnight exposure (red triangles, 6 Kb). We also observed a higher molecular weight band of 7 Kb (white triangle) in the lines loaded with NMuLi genomic DNA, probably due to partial digestion or the absence of a HindIII site in one of the two UGT1a1 alleles.
Subsequently, we tested a 1.1 Kb DNA segment, contained in the B Fragment (See Figure 1) as the probe (B Probe) in Southern blot analysis. To evaluate the hybridization efficiency of the B Probe we performed a Southern blot using C57BL/6 genomic DNA digested with different restriction enzymes (BstEII, BamHI, HindIII) and C129Sv/J genomic DNA digested with HindIII (see Figure 11A for a detailed description).

As already noticed in previous Southern blot analysis using the 3’ Probe, we observed two bands in the digestion of the C57BL/6 genomic DNA with HindIII (an additional 7 Kb, white triangle in Figure 11B), while C129Sv/J showed the expected band of 6 Kb. Also, we performed a calibration curve to determine the lower genomic DNA amount detectable with the B probe (Figure 11C). Two different amounts of HindIII digested C129Sv/J genomic DNA (4.5 µg and 6 µg) were run in an agarose mini-gel system, then blotted on to a nylon membrane and hybridized against radioactively labelled B Probe. We were able to detect a specific signal after 48 hs exposure when 4.5 µg (lane 2) and 6 µg (lane 3) were loaded.
Figure 11. Southern blot using the B Probe. Panel A) A scheme of the mouse UGT1a1 locus. The relative position of the restriction enzymes sites (BstEII, BamHI, HindIII) are shown. The coloured lines indicate the size of the restriction fragments originated from the digestion of genomic DNA (C57BL/6 and C129Sv/J) and detected in Southern blot using B Probe (red box). Thick lines represent the expected bands, while thin lines represent partial digestions. Panel B) Southern blot analysis. C57BL/6 and C129Sv/J genomic DNA were digested as indicated, run in an agarose gel, blotted onto a nylon membrane and hybridized against radio-labelled B Probe. The expected bands are indicated with red triangles. We observed partial digestion of the C57BL/6 genomic DNA with HindIII (7 Kb, white triangle, lane 4), while C129Sv/J DNA showed the expected band of 6 Kb. Panel C) Calibration curve. To determine the lower genomic DNA amount detectable with the B Probe two different amounts of HindIII digested C129Sv/J genomic DNA (4.5 µg and 6 µg) were run in an agarose gel and treated as described for Panel B. We were able to detect a specific signal in the 4.5 µg (lane 2) and 6 µg (lane 3) after 48 hs exposure.
**PCR analysis**

In parallel to the Southern blot analysis, we set up an alternative method for the detection of recombinant clones based on PCR of genomic DNA.

We selected two sets of screening primers: one for the wild type allele and one for the recombinant one, sharing the reverse primer (see Figure 12A for detailed description).

The upstream primer of the wild type allele is located in the junction between the fragment containing the Exon 4 and the fragment containing the short arm of homology. This location allows us to amplify only the wild type allele (in this position the sequence of the recombinant allele is different to that of the wild type one, as it the BamHI site was inserted for cloning purposes into the targeting vector). The upstream primer of the recombinant allele is located at the 3’ end of the Neomycin cassette. The downstream primer, located outside the homology region, in the 5’ end of the fragment containing the 3’ Probe, was used as a reverse primer for both alleles.

Having in mind that the Multiplex PCR reaction favours smaller products, the primers were accurately selected to obtain a product of the recombinant allele smaller in size than that of the wild type one (3.4 Kb and 3.7 Kb, respectively).

Subsequently, we set up the condition to efficiently detect both the recombinant and the wild type alleles in the same reaction tube (Multiplex PCR). We performed a PCR with constant amounts of genomic DNA (C129Sv/J, wild type allele) and increasing amounts of the targeting vector (recombinant allele) (Figure 12B). We observed that using a low amount of targeting vector (ratio 0.2, Figure 11B, lane 1) we obtained a similar intensity of the two bands; while increasing the amount of the targeting vector, the recombinant band was preferentially amplified and competed out the wild type band at increased ratios showing less amount of product (Figure 11B, lanes 2 to 5).
Figure 12. Multiplex PCR. Panel A) A Scheme of the mouse UGT1a1 wild type locus and targeted locus. The position of the primers is indicated by coloured triangles. The upstream primer (red triangle) of the wild type allele is located in the junction between the fragment containing the Exon 4 and the fragment that contains the short arm of homology. This location allows us to amplify only the wild type allele (in this location the sequence of the recombinant allele is different to that of the wild type one.). The downstream primer (red triangle), located outside the homology region in the 5’ end of the fragment containing the 3’ Probe, was used as reverse primer for both alleles. The upstream primer of the recombinant allele (purple triangle) is located at the 3’ end of the Neomycin cassette (purple box). Panel B) Multiplex PCR analysis. The two sets of primers described above were used in a PCR reaction containing constant amounts of C129Sv/J genomic DNA (wild type allele) and increasing amounts of the targeting vector (recombinant allele). This system allows us to detect both recombinant (3.4 Kb band) and wild type allele (3.7 Kb band) in the same PCR test tube. Lines 7 and 8 are control lines using only genomic DNA and targeting vector, respectively.
**RESULTS**

**e) ELECTROPORATION, SELECTION AND SCREENING OF G418-RESISTANT CLONES**

The full pFlrt1 targeting vector was sent to the Telethon ES cell facility (DIBIT San Raffaele, Milano) conducted by Drs. Mario Bianchi and Lorenza Ronfani, where the electroporation of ES cells and selection of G418-resistant clones were performed. After the electroporation, the G418-resistant clones were amplified in the absence of feeder layer cells and aliquots of them were sent to our laboratory in p96 multiwell plates, to prepare genomic DNA and perform the screenings. Two different strategies were used to perform the analysis:

- Southern blot analysis, using the 3’ Probe (external probe), the B Probe (internal probe) and the Neo Probe sequentially;
- Multiplex PCR analysis, using together primers specific for the wild type and recombinant alleles.

**f) CHARACTERIZATION AND OF THE G418-RESISTANT CLONES**

Genomic DNA of the G418-resistant clones was prepared and digested with HindIII restriction enzyme in a final volume of 30μl. 5μl of the digestion were kept for PCR analysis. The remaining 25 μl were loaded in an agarose gel for Southern blot analysis (mini-Southern blot analysis).

Figure 13A shows a scheme of the mouse UGT1a1 wild type locus and targeted locus. We screened ∼200 G418-resistant clones and we found three positive ones: A9 and D12 were positive both in Southern blot analysis (Figure 13B) and PCR analysis (Figure 14, panel A, B, D and E ); G7 was identified only by PCR analysis (Figure 14, panels C and F), because of technical problems with the Southern blot technique.

Surprisingly, the recombinant band was less intense than the wild type one, both in Southern blot and PCR. We quantified the Southern blot bands after background subtraction (red squares in Figure 13B): the A9 clone showed a recombinant band intensity of 17 % and wild type band 83%; the D12 clone showed a recombinant band of 5 % and wild type band 95%. The recombination event is a rare one, and normally occurs in one of the alleles generating heterozygous cell lines. Therefore, Southern blot analysis of those cells normally produces two bands (one per allele), having equal or almost equal intensity. In some cases it is possible to observe some bias towards the
RESULTS

wild type allele because of the presence of some residual MEFs, used as feeder layer, at the moment of amplifying the cells for the preparation of DNA. The quantification of the bands indicated that in the A9 clone 1/3 of the cells are apparently recombinant, whereas in D12 clone 1/10 of the cells are apparently recombinant. A possible explanation could be that these clones were heavily "contaminated" with MEF cells from the feeder layer, used to grow the ES cells, or, alternatively, that the ES cell clones were not pure and contaminated with wild type ES cells (somehow growing in G418-containing medium). For this reason we performed a sub-cloning by limiting dilution of the cells, to these three apparently positive clones in order to remove the suspected contamination. We screened the sub-clones by Southern blot and PCR.

Figure 15A shows the results of the Southern blot using genomic DNA derived from A9, D12 and G7 sub-clones digested with HindIII and hybridized with the 3’ Probe. Only A9 and G7 sub-clones showed the expected 8 Kbp recombinant band (suggesting the presence of the recombinant allele) and the 6 Kbp wild type band. We screened 89 sub-clones for A9 and 72 for G7, of these 11 and 20 respectively were apparently negative, indicating that the sub-cloning procedure purified the recombinant clones from the wild type ones. Most of the negative clones contained less DNA hampering the visualization of the recombinant band.

However, Southern blot quantification of the bands (blue squares in Figure 15A) showed that the bands were not equimolar (A9: 33% recombinant, 67% wild type; G7: 28% recombinant, 72% wild type). On the contrary D12 was totally negative (only the 6 Kb wild type band).

To verify the absence of undesired genomic rearrangements in the locus and to confirm the homologous recombination events, we performed additional Southern blot analysis digesting the genomic DNA of the positive clones with a number of different restriction enzymes and hybridized them with different probes (3’ Probe, B Probe and Neo Probe). Figure 15B shows Southern blot analysis using genomic DNA derived from A9, G7+ (recombinant sub-clones), G7- (non recombinant G7 sub-clone, apparently negative in previous Southern blot analysis) digested with BglII, HindIII, NdeI, PstI, SphI and XbaI restriction enzymes, and hybridized with the 3’ Probe. The size of the expected wild type and recombinant bands using the 3’Probe in the Southern blot are shown in the Table of Figure 13B. All restriction cuts generated the expected band-pattern.
Figure 13. Southern blot using the 3’ Probe. Panel A) A Scheme of the mouse UGT1a1 wild type locus and targeted locus. The relative position of HindIII restriction sites and the size of the fragments originated from the digestion of C129Sv/J genomic DNA and detected in Southern blot using 3’ Probe concatamer (light blue box) are represented. Double-headed lines indicate the expected band for the wild type (6 Kb) and the correctly targeted locus (8 Kb). Panel B) Southern blot analysis of G418-resistant ES cell clones. Genomic DNA of the G418-resistant ES cells was prepared and analyzed as described in Figure 9. All clones showed the 6 Kb band of the wt allele. Only the A9 and D12 clones (indicated with the red rectangles) showed an 8 Kb band (red triangle) suggesting the presence of the recombinant allele. C57BL/6 genomic DNA digested with HindIII was used as a control (light blue arrow). Red squares highlight the regions used for Southern blot bands quantification.
**Figure 14. PCR analysis of G418-resistant ES cells.** DNA of the G418-resistant ES cells was digested with HindIII to reduce the length of DNA fragments and subsequently analyzed by PCR. PCR reactions were performed as described in Figure 11. Positive clones (A9, D12 and G7, are highlight in red) show the 3.7 Kb wild type band (orange triangle) and the 3.4 Kb recombinant one (red triangle) when using the specific primers separately (Panel A, B, C) or together in the same PCR reaction tube (Panel D, E, F), confirming the Southern blot results obtained in Figure 11.

Subsequently, we performed Southern blot analysis digesting A9, G7+ (recombinant sub-clones), G7- (non recombinant G7 sub-clone) DNA with other restriction enzymes (BclI, BglII, BglIII, KpnI, SphI) and hybridizing the blotted membrane with the internal B Probe. We observed the expected recombinant and wild type bands (Figure 16).

To further confirm the presence of the recombinant allele we stripped the membranes showed in Figure 15A and hybridized them with a radioactive labelled probe of the Neomycin cassette (Neo Probe) (see Figure 17A for a detailed description).
RESULTS

Figure 15. Southern blot using the 3’ Probe concatamer. Panel A) Southern blot analysis of G418-resistant clones obtained after A9, D12 and G7 sub-cloning. Genomic DNA of the sub-clones was digested with HindIII and was analyzed by Southern blot. Most of the A9 and G7 sub-clones showed the expected 8 Kb recombinant band (red triangle), while all the D12 sub-clones showed only the 6 Kb wild type band (orange triangle). Blue squares indicate the clones used for the quantification of the bands.

Panel B) Southern blot analysis of the recombinant clones with a set of restriction enzymes. Genomic DNA of the sub-clones (A9, G7+ (recombinant clones), G7- (non recombinant clone) was digested with different restriction enzymes (BglI, HindIII, NdeI, PstI, SphI, XbaI) and analyzed by Southern blot. All restriction cuts generated the expected band pattern listed in the table. Wild type and recombinant bands are indicated with orange and red triangles, respectively. The G7- clone (sub-clone of G7 not showing the recombinant band in previous Southern blot analysis) was used as a negative control. □ XbaI produces similar bands in the wt and recombinant alleles. △ G7+: DNA is missing due to technical problems.
RESULTS

Figure 16. Southern blot using the B Probe. Panel A) A Scheme of the mouse UGT1a1 wild type locus and targeted locus. The relative position of Hind III restriction sites, (HindIII) the size of the Hind III digested fragments and detected in Southern blot using B Probe (red box) are represented. Double-headed lines indicate the expected bands for the wild type (0.4, 2.1 and 6 Kb) and the targeted locus (0.4, 2.1 and 8 Kb). Panel B) Southern blot of A9, G7+ (recombinant clone), G7- (non recombinant clone). The same membrane used in Figure 13B was hybridized with the B probe. All restriction cuts generated the expected band pattern listed in the table. Wild type and recombinant bands are indicated with orange and red triangles, respectively. G7- was used as a control for the wild type allele.

As expected, only the A9 and G7 sub-clones showed the expected 8 Kb band corresponding to the targeted allele (Figure 17B). On the contrary, clones derived from D12 sub-cloning showed two higher molecular weight bands (of 10 and 12 Kb) indicating non-homologous recombination events in the genome.

In parallel we also performed the Multiplex PCR analysis of A9, D12 and G7 sub-clones (Figure 18). All the A9, D12 and G7 sub-clones showed the expected 3.7 Kb wild type band, but only the A9 and G7 sub-clones showed the 3.4 Kb recombinant band. These results confirmed those observed above by Southern blot analysis indicating that all the D12 sub-clones were negative.

Since the external 5’ Probe contained repetitive sequences not allowing its use in a Southern blot of genomic DNA, we checked for the correct homologous recombination at the 5’ end of the targeted UGT1a1 locus by PCR followed by Southern blot analysis (Figure 19A).
**Figure 17. Southern blot using the Neo Probe. Panel A). A Scheme of the mouse UGT1a1 wild type locus and targeted locus.** The relative position of the HindIII restriction sites (HindIII), the size of the restriction fragments the expected bands for the targeted locus (8 Kb). The Neo gene is present only in the targeted allele. **Panel B) Southern blot of clones derived from the A9, D12 and G7 positive clones.** The membranes used in Figure 13A were stripped and hybridized against radioactively labelled Neo Probe. A9 and G7 sub-clones showed the 8 Kb band corresponding to the targeted allele (red triangle), while all the D12 sub-clones showed two higher molecular weight bands (10 and 12 Kb respectively) indicating non homologous recombination events in the genome.
**Figure 18. Multiplex PCR of A9, D12 and G7 sub-clones.** Genomic DNA of the sub-clones was digested with HindIII and used to perform Multilex PCR analysis. Positive clones (A9 and G7) are highlighted in red. All the A9, D12 and G7 sub-clones showed the expected 3.7 Kb wild type band (orange triangle) but only A9 and G7 sub-clones showed the 3.4 Kb recombinant band (red triangle). All the D12 sub-clones screened were negative.

The PCR was designed to exclusively amplify the recombinant allele: the forward primer is complementary to the 5’ flanking region of the targeted locus and the reverse primer complementary to the Neomycin gene. Using this strategy, analogous to the one used for the analysis of recombinant clones using a primer complementary to the 3’ flanking region, only the product of homologous recombination will be amplified. PCR products of A9, G7+ (recombinant clones), G7- and D12 (non recombinant clones) were run in an agarose gel (Figure 19B, ethidium bromide). Only A9 and G7 sub-clones showed the recombinant band of 7.8 Kb. G7- and D12 showed no PCR product of the expected size.

Since many PCR products of different sizes were observed, we performed a Southern blot using the radioactively labelled 5’ Probe (Figure 19B, Southern 5’ Probe) and that one of Neomycin. In the case of the analysis of a PCR reaction, the complexity of the analyzed fragments is very low (just a few bands amplified, corresponding to a minor portion of the genome), contrasting the high complexity of the genomic DNA containing a high proportion of repetitive sequences. Therefore, we used the 5’ Probe for the PCR analysis, even if it may contain some repetitive sequences, as suggested by the “reverse” Southern blot experiment (Figure 6).
Figure 19. Analysis at the 5’ end of the targeted UGT1a1 locus. Panel A) A Scheme of the mouse UGT1a1 locus. The relative position of the PCR primers (purple triangles) and the size of the PCR product (double-headed purple line), position of the 5’ Probe (black box) and the Neo Probe (purple box) are represented. Panel B) PCR followed by Southern blot analysis. PCR products of A9, G7+ (recombinant), G7- and D12 (non recombinant) clones were run in an agarose gel (left). Only the A9 and G7 sub-clones showed the recombinant band of 7.8 Kb. The G7- and D12 clones showed no PCR product of the expected size. Since many products of different sizes were amplified, we performed a Southern blot using the radioactively labelled 5’ Probe. Only the A9 and G7 sub-clones showed the expected 7.8 Kb recombinant band. Subsequently we stripped the membrane and hybridized it with the radioactively labelled Neo Probe to confirm the presence of the Neomycin cassette in the PCR product. As expected only the recombinant sub-clones A9 and G7 were positive.
RESULTS

Only the A9 and G7 sub-clones showed the expected 7.8 Kb recombinant band. Subsequently we stripped the membrane and hybridized it with the radioactively labelled Neomycin Probe to confirm the presence of the Neomycin cassette in the PCR product (Figure 19B, Southern Neo Probe). As expected, only the recombinant sub-clones A9 and G7 were positive.

Subsequently, we sequenced the region of the positive clones (A9 and G7) corresponding to the Exon 4 fragment (0.7 Kb, see Figure 1), to verify the presence of the introduced mutation (one base deletion) and loxP sites in the recombinant allele.

Previous to the injection of blastocysts, to verify whether the number of chromosomes of the positive ES cell clones was the expected one, the Telethon ES cell Facility performed the chromosome counting of the A9 and G7 positive cell clones.

At the same time and to cover all possibilities, the Telethon ES cell Facility performed a second round of electroporation and we screened the G418-resistant clones (∼150 clones), but we didn’t find positive ones (data not shown).

g) BLASTOCYST MICROINJECTION TO GENERATE CHIMERAS

The microinjection of the positive clones (A9 and G7) into the blastocysts was performed by the Telethon ES cell facility (DIBIT San Raffaele, Milano). Up to now, they obtained one female chimera from the A9 clone and one male chimera from the G7 clone (Figure 20, panel A and B).

h) GERM LINE TRANSMISSION OF THE MUTATED UGT1A1 GENE

Chimeras were sent to our animal house and mated with wild type mice (C57BL/6 and FVB) to check for germ line transmission of the mutated UGT1a1 gene and to initiate the changing of the genetic background (see below).

At the present time we are screening 14 pups obtained from G7 chimera (Figure 20, panel C and D). Meanwhile A9 chimera has not given germ line transmission.
Figure 20. A9 and G7 chimeras. Panel A) Female chimera from the A9 clone. Panel B) Male chimera from the G7 clone. Panel C) Pups derived from G7 chimera mated with female C57BL/6 mouse. Panel D) Pups derived from G7 chimera mated with femalFVB e mouse.

1) Changing the Genetic Background of UGT1a1 Gene Mutated Mouse to an Homogeneous One

Heterozygous mice having the UGT1a1 mutation will be mated to obtain the homozygous mice and to have a first idea of the effects of the mutation. In parallel, heterozygous mice will be mated to the FlipE “delete” strain to remove the Neomycin cassette from the UGT1a1 locus.

The colony will be amplified and the UGT1a1 mutation will be transferred to a homogeneous genetic background by successive backcrosses with FVB wild type mice.
DISCUSSION & CONCLUSIONS

Per aspera sic itur ad astra.

Seneca. Hercules furens, atto II, v. 437
The main aim of this Thesis is the development of a new mouse model of chronic unconjugated hyperbilirubinemia at risk of bilirubin induced neurological dysfunction, to study the molecular mechanisms involved in bilirubin toxicity. This mouse model will closely resemble the human pathology named Crigler-Najjar Syndrome I (CNI) and its well-established animal model, the Gunn rat. Both CNI patients and Gunn rats have mutations in the UGT1a1 gene that encodes for the hepatic enzyme responsible for bilirubin conjugation. This leads to inability to conjugate bilirubin. The mutation introduced in the mouse UGT1a1 gene will be equivalent to that present in the Gunn rat. This strain of Wistar rats have a single nucleotide deletion in the Exon 4 of the UGT1a1 gene, that creates an in frame stop codon immediately after the deletion. As a consequence of the deletion, the gene produces an incomplete and thus inactive enzyme. Due to lack of conjugation, the biliary excretion of bilirubin is abolished, which results in a consistent, permanent elevation of plasma unconjugated bilirubin levels. Unconjugated bilirubin (UCB) values in CNI patients are 20-40 fold higher than those measured in normal individuals, while in Gunn rats are even more elevated (400-1000 fold depending on the strain). UCB at extremely elevated levels accumulate in specific areas of the brain as the basal ganglia, hippocampus, several nuclear clusters of the brainstem and cerebellum. Deposition of the UCB in the brain produces a characteristic yellow staining of the brain nuclei and results in irreversible brain damage also named Bilirubin Encephalopathy (BE). The American Academy of Pediatrics (AAP) describes: 1) Bilirubin Encephalopathy (BE) as "the clinical central nervous system findings caused by bilirubin toxicity to the basal ganglia and various brainstem nuclei"; 2) Bilirubin-Induced Neurological Dysfunction (BIND) as the spectrum of permanent sequelae seen (often in the context of subtle, non-classical sequelae) as well as a non-validated scoring system for evaluating the severity of acute bilirubin encephalopathy in regards to the need for treatment; and 3) Kernicterus (KI) as a definition be reserved for the chronic and permanent neurological sequelae. KI is a severe and life-threatening condition; 70% of children with KI die within seven days, while the 30% survivors usually suffer irreversible neurological damages.

Although the Gunn rat is the classical laboratory model for bilirubin encephalopathy its use, as a model for the study of molecular mechanisms involved and the determination of other genes modulating the development of the disease, is limited for the following reasons: first, the gene targeting technology is limited to mice being
extremely difficult to obtain rat mutant strains of other genes of interest. Second, since
the development of the gene targeting technique mice represent the preferential animal
model for human diseases making this animal the well-characterized in aspects ranging
from pharmacology to physiology and toxicology. Third, on what specifically concerns
the Gunn rat, strain-specific differences in bilirubin-induced neuronal damage obtained
in the various laboratories make the interpretation of the results very complex.
Therefore, it is of crucial importance the development of a mouse model of Crigler-
Najjar disease in order to analyze the functions of other genes involved in the
development and/or protection of neurological damages caused by elevated levels of
bilirubin in the newborn.

By taking advantage of the “Gene Targeting” technique we are creating a mouse
model of the human Crigler–Najjar syndrome, a strain devoid of UGT1a1 activity,
analogous to the Gunn rat model. To increase the efficiency in recombination, it is
necessary to use isogenic DNA in the targeting construct to avoid reduction of
recombination efficiency due to the presence of strain specific polymorphisms. During
the initial times of gene targeting, fragments of genomic DNA from phage libraries
containing C129Sv/J genomic DNA were normally used as the source of isogenic DNA.
However, with the reduction in the error rate of new generation polymerases used in
PCR, many laboratories are performing DNA cloning using these proofreading
polymerases. The lower error rate is due to the use of enzymes having proofreading
activity, or mixes of two polymerases, one having high processivity and the other one
with proofreading activity.

Therefore, to clone the UGT1a1 gene (including the 3’ gene-flanking regions) by
PCR we utilized genomic DNA from the same ES cell line used for the electroporation as a
template, and we used a proofreading polymerase system to reduce the presence of
undesired mutations the PCR.

Those systems are all designed for performing high-fidelity DNA synthesis, and
should produce less mutations than traditional Taq DNA polymerases (without
proofreading activity), for this reason they are proposed for the cloning of genes. For
example, the Pfu polymerase has 3’ to 5’ exonuclease proofreading activity, analogous to
the bacterial DNA polymerase correcting nucleotide miss-incorporation errors.

Since the sequence annotated in the Genbank is from the C57BL/6 mouse strain
and no information is available regarding the UGT1a1 gene sequence, we designed a
strategy based on the sequence alignment of clones obtained from independent PCR
reactions with the C57BL/6 sequence. Mismatches being present in all clones may suggest the presence of a strain-specific polymorphism, while mismatches present in only one or few clones may represent PCR errors.

We first used the Pfu DNA polymerase to clone the UGT1a1 gene, but since sequence alignment of the Pfu-amplified fragments with the annotated genomic sequence evidenced the presence of many mismatches, we decided to test two other Pfu DNA polymerase systems, to find that with the lowest error rate. No significant difference in the error rate of the three DNA polymerase systems tested was observed. Interestingly, sequence alignment of DNA sequences (from the same genomic fragment), deriving from different PCR system reactions, revealed two different groups of clones, each of them carrying a specific set of mismatches: clones belonging to one group always showed the same mismatches. These findings suggest that the origin of the two groups may be due to the presence of different alleles of the gene instead of mutations produced by the polymerase. The mechanisms that maintain heterozygosity at the UGT1a1 locus in an inbred strain are not known, but these observations are supported by the presence of two bands in the Southern blot analysis of Hind III-digested C57Bl/6 liver genomic DNA.

To generate a null mutation in the UGT1a1 gene different approaches were achievable and the different procedures were examined: 1) to completely delete the UGT1a1 Exon 4, which is common to all mRNA forms coded by the UGT gene complex, affecting all encoded enzymes; 2) to mutate the UGT1a1-specific gene promoter and Exon 1, to affect only the UGT1a1 enzyme, responsible for bilirubin glucuronidation; 3) introduce the base deletion in UGT1a1 Exon 4, analogous to the one present in the Gunn rat and in many Crigler-Najjar patients. As described before, we choose the latter option to exactly mimic the molecular defect present in the Gunn rat. We expect to observe a similar phenotype and to take advantage of the available experience with the rat model. However, in the case the inserted mutation will not produce the expected phenotype, we will adopt a backup strategy since we have the possibility to remove the entire Exon 4, by using the loxP sites flanking the Exon 4. When direct repeats of loxP sites are present in the genome, the intervening sequence is removed in the presence of the Cre recombinase (see INTRODUCTION Section). By mating heterozygous mutant mice with a “deleter” Cre transgenic strain, the completed “floxed” sequence will be deleted from the genome. Alternatively, the use of tissue specific and/or inducible Cre transgenic strains may allow the temporal and spatial control of the Exon 4 deletion.
Electroporation of ES cells with the targeting construct and G418 selection generates a high number of clones, with a minimal proportion of homologous recombinant ones. Clearly, the task of detecting the right events needs to be very efficient. Therefore, in parallel to the construction of the targeting vector we set up the optimal conditions to screen the G418-resistant clones originated from the ES cells electroporation with the targeting vector by two techniques: Southern blot and PCR.

Since the majority of G418-resistant clones have the targeting vector randomly inserted, the screening strategy has to differentiate between random and site-specific recombination. The screening strategy is based in the appearance in the genome of new sequences or restriction enzyme sites (“diagnostic” restriction enzyme digestion), present in the targeting construct but not in the wild type allele. These sequences are normally present in the case of random and site-specific recombination. Therefore, the use of probes or primers recognizing sequences of the target locus not present in the targeting construct (“external primers and probes”) is mandatory to differentiate homologous recombination from random integration events. In order to determine the specificity of the selected probes (5’ and 3’ external Probes), we first performed a “reverse” Southern blot experiment. This practice was essential for the subsequent analysis, because the Probes used must recognize only the specific fragment and not random sequences present in the genome. The “reverse” Southern blot revealed that the 5’ Probe recognizes highly repetitive sequences in the genome and cannot be used for Southern blot analysis. Since the 3’ Probe did not, it was selected to check for homologous recombination in the 3’ end of the targeted locus. Together with the probes we also tested the fragments included in the targeting vector (Fragment A, B and C, Exon 4 and the Short Arm). Fragment B did not recognize highly repetitive sequences in the genome, and we therefore selected a DNA segment contained in this fragment (namely B Probe, internal probe) to check for homologous recombination at the 5’ end of the targeted locus.

Both Probes were tested with genomic DNA digested with different restriction enzymes. In both the cases HindIII restriction enzyme turn out to be the more suitable one for Southern blot analysis.

Since we noticed a background noise higher than expected in the Southern blot using the 3’ Probe, we analyzed the sequence with a RepeatMasker program that highlighted the presence of highly repetitive sequences, not evident with the “reverse” Southern analysis. By shortening the probe and preparing a concatamer we strongly improved the labelling-efficiency of the probe and its sensitivity.
An all the cases in which we used genomic DNA derived from C57BL/6 mouse strain we observed an additional 7 Kb band (instead of the predicted 6 Kb band observed in the C129Sv/J genomic DNA) that could be the result of the absence of one HindIII restriction site in one of the two alleles. Alternatively, the second band could be the result of a partial digestion of the DNA but this possibility is less probable since all different experiments performed always produced the two bands having the same proportion. The presence of the two bands was also observed when genomic DNA from N-MuLi cells was tested.

Considering the fundamental importance to obtain a clear and definitive outcome from the screening of the G418-resistant ES cell clones, we set up an alternative method for the detection of the positive clones based on PCR. A powerful multiplex PCR screening was developed allowing the simultaneous detection of both the alleles (wild type and recombinant) in the same PCR reaction test tube.

After the electroporation of the ES cells with the targeting vector, the G418 resistant clones were screened with the two techniques described above. We screened ~200 clones and we initially found three positive ones: A9, D12 and G7. Unexpectedly, the recombinant band was less intense than the wild type one both in Southern blot and PCR analysis (the recombinant band was 5-17% of that of the wild type one, instead of the expected 1:1 proportion). A possible explanation could be that the DNA preparation of these clones was heavily contaminated with DNA from MEF cells (from the feeder layer) or alternatively, with the growth of wild type ES cell in G418 medium. To check for possible contamination we performed a sub-cloning of the three positive clones by limiting dilution and we then screened the sub-clones by Southern blot and PCR analysis. Only A9 and G7 sub-clones turned out to be positive, while D12 sub-clones were totally negative. Quantification of the purified sub-clones showed a notable increase in the proportion of the recombinant band (~ 30% of the wild type). Additionally, the discovery of some non-recombinant sub-clones among A9 and G7 sub-clones demonstrated that, in addition to some plausible contamination with MEF DNA, the sub-cloning procedure effectively purified the recombinant clones from the wild type ones.

To verify the absence of undesired genomic rearrangements in the UGT1a1 locus and to confirm the homologous recombination events, we performed: 1) additional Southern blot analysis digesting the genomic DNA of the positive clones with a number of different restriction enzymes and hybridized them with different probes (3’Probe, B
DISCUSSION AND CONCLUSIONS

probe and Neo Probe); 2) Multiplex PCR analysis; and 3) Long Range PCR analysis followed by Southern blot to check for correctness of the homologous recombination event at the 5’ end of the locus. All these analysis clearly demonstrated the positivity of A9 and G7 sub-clones, while D12 persisted to be negative.

The homologous recombination event at the UGT1a1 locus was further confirmed by sequencing the loxP-Exon 4-loxP region to verify the presence of the one-base deletion of the two positive clones. The absence of gross chromosomal aberrations was verified by chromosome counting. Then, the A9 and G7 clones were micro-injected into balstocysts and we obtained one female chimera from the A9 clone and one male chimera from the G7 clone. These two chimeras were mated with wild type mice to check for germ line transmission of the mutated UGT1a1 gene and to start changing the genetic background. At present the G7 chimera produced two litters, with FVB and C57BL/6 females. In both cases the pups were of agouti coat colour indicating germ line transmission of the ES cells genome. We are screening the progeny of the G7 chimera looking for heterozygous pups; meanwhile the A9 chimera has not given germ line transmission.

These results will pave the way for additional studies to be performed in the near future. They include: 1) mating heterozygous mice to obtain the homozygous mutant animals, to assess the effects of the mutation; 2) mating the heterozygous mice with transgenic mice expressing the Flp recombinase to remove the Neomycin cassette from the UGT1a1 mutated locus; 3) changing the genetic background into an homogeneous one.

Regarding the first point, two scenarios may arise: a) the introduced mutation may not produce a visible phenotype. In this case, we will remove the entire Exon 4 taking advantage of the loxP sites present in the targeted locus; b) the introduced mutation, including the neomycin cassette, produce a drastic phenotype leading to premature death of the mice. As a matter of fact, in 2008 Tukey et al. (Nguyen et al., 2008) created a knockout mouse in which the UGT1a1 locus was disrupted introducing the neomycin cassette into the Exon 4. Although this resulted in extremely high levels of UCB in plasma in these mice (40-60 fold higher than the wild type mice) it also induced a fatal neonatal phenotype within two weeks after birth. This might be related to the presence of the Neomycin cassette, which can cause a number of additional, unexpected problems, such as the disruption of neighbouring gene expression due to strong transcriptional regulatory elements frequently present in selection cassette (Lerner et
al., 1993; Ohno et al., 1994). Alternatively, the early mortality may be associated to the genetic background used in that study (50% C129Sv/J and 50% C57BL/6).

As mentioned before (Point 2 of future steps), been conscious of this unlucky event, we already planned to remove the positive selection cassette with a second in vivo gene-targeting step, mating the heterozygous mice with the Flp deuter strain, leaving a “clean” mutation in the genome. The deletion of the complete Exon 4 should also produce an inactive UGT1a1 product, as it produces the deletion of the C-terminal region of the protein due to a change of the natural reading frame after Exon 3, introducing a stop codon early in the Exon 5.

We expect that the phenotype of the homozygous mice lacking the UGT1a1 enzyme will resemble the rat counterpart, with serum UCB levels reaching the maximal concentration in the neonatal period and staying stable well above the physiological levels throughout the animal life, as already observed in the Gun rat. Moreover, we expect to observe many of the clinical signs and histopathological lesions found in humans suffering kernicterus, and partially already observed in the jaundiced rats.

Furthermore, the extensive knowledge acquired in the bilirubin-induced neurotoxicity using the rat model will be exploited to plan the experimental design used to characterize the mutant mice. However, minor differences in the phenotype might directly indicate species-specific variations between rats and mice in the bilirubin metabolism.

The establishment of the hyperbilirubinemic UGT1a1 mutant mice will provide a unique tool to study the events associated with bilirubin–induced neurological damage. The availability of a mouse where the plasma level of UCB is genetically determined and sufficiently high to induce BE, similarly to what observed in Crigler–Najjar patients, will allow us to have a unique insight on the role of the different players reported to be involved in bilirubin neurotoxicity. It will also be crucially important to validate in vivo the protective effects of other potentially important transporters (i.e. Mrp1) or pathways identified in our laboratory by in vitro experiments on neuronal cells exposed to UCB [X(C)\(\cdot\)]. Furthermore, these animals might be used in the future to screen drugs able to prevent or ameliorate damage [glycoursodeoxycholic acid (GUDCA) and polytheinilglycol (PEG)] as well as situations where the damage will be worsened (hypoxia, hypoglicemia, infection, etc). The re-insertion of a wild type copy of the UGT 1a1 gene in the hepatocytes of mutant mice will be the gene therapy challenge of the Gunn mouse model.
To summarize, obtaining this new animal model for bilirubin neurotoxicity (Gunn mouse) will be crucial to understand the mechanisms regulating the disease, together with an improvement of the diagnosis, prediction of the prognosis, and development of new therapeutic strategies.
Prof. Claudio Tiribelli

Aranda Alejandro
Bellarosa Cristina
Berengeno Andrea
Boscolo Elena
Chavez Norberto
Corsucci Sabrina
Di Maso Vittorio
Gazzin Silvia
Giraudi Pablo
Kazemi Mayam
Leal Sandra
Mazzone Graciela
Pascut Devis
Qaisiya Mohammed
Robert Celeste
Rosso Natalia
Sukovati Caecil
Zabetta Carlos

Andrés Fernando Muro

Costessi Luisa
Iaconcig Alessandra
Moretti Federico
Morgan Marcos
Nedeljkovic Mirjana
Porro Fabiola

Ronfani Lorenza
REFERENCE LIST:


Ostrow, J.D. (1987). Bile pigments and jaundice (Marcel Dekker, Inc.).


