DEBATE continued

Safety of embryo cryopreservation: facts and artefacts

Embryo cryopreservation and development: facts, questions and responsibility

Emmanuel Dulioust¹,³, Marie-Claire Busnel², Michele Carlier², Pierre Roubertoux² and Maurice Auroux¹

¹Biologie de la Reproduction et du Développement, CHU de Bicêtre, Université Paris-Sud, 94275 Le Kremlin-Bicêtre, and ²Genétique, Neurogenétique, Comportement, CNRS UPR 9074, Institut de Transgenose, 3B rue de la Ferollerie, 45071 Orleans Cedex 02, France

³To whom correspondence should be addressed

This Debate was previously published published on Webtrack 53 on February 15, 1999

Four years ago, the results of an experimental study led us to conclude that in the mouse, cryopreservation of preimplantation embryo induced post-natal effects and thus could not be considered as neutral for the long-term development (Dulioust et al., 1995). This raised some concern in the scientific community and in the public about the safety of embryo cryopreservation in human assisted procreation. It has even been said that we had withdrawn our conclusions (Human Fertilisation and Embryology Authority, 1996); this was unfounded. Since then, only few studies have focused on this issue. Recently, a severe attack of our work was published (Testart, 1998). Here, we examine these criticisms and discuss the state of the art in the light of recent clinical and experimental studies.

Methodological criticisms

Principles of the study

According to Testart (1998), our research was limited to some kind of statistical fishing, lacking rational hypotheses, since ‘it appears that the clinical risk of freezing human embryos is only to slightly decrease the pregnancy rate’ (Testart, 1998). In other words, while acknowledging that the embryo’s ability to implant may be impaired by the freeze–thaw process, it is supposed that no other feature is altered. We thought that this common viewpoint was far from being substantiated and that, besides more detailed and long-term clinical investigations, additional experimental studies were needed.

Screening for various possible effects is common in toxicology, especially in reproductive and developmental toxicology. More precisely, our reflection and strategy were based on several facts and hypotheses: (i) freezing, cryoprotectants and finally, cryopreservation as a whole induce in cells a wide spectrum of chemical and metabolic disorders (Dulioust, 1995); these effects could in turn alter, through several pathways, developmental determinants like nuclear or mitochondrial DNA; (ii) although neither the practice of embryo cryopreservation in humans and cattle nor previous experimental studies have evidenced genotoxic or teratogenic effects, other works have reported genetic effects either in bacteria or somatic cells (Dulioust, 1995). This suggested the possibility of a still undetected genotoxic potential; (iii) assessment of post-natal development after embryo cryopreservation has mainly focused on patent anomalies at birth or in early life, and has not been carried out on adult and senescent individuals. However, many mutations can have mild or delayed phenotypic effects, and numerous clinical and experimental observations show that developmental disturbances can respect viability and macro-morphological appearance while altering physiological processes, including behaviour and cognitive functions, through biochemical or microstructural anomalies (Auroux, 1997). Besides classic morphological examination, neural and behavioural screening has been recommended, for instance, by the National Institutes for Mental Health (NIMH), for transgenic and knock-out mice. There was thus a serious gap between the wide range of potential long-term effects of embryo cryopreservation and the means by which they had been investigated until then.

These facts, plus our past experience, prompted us to use, besides basic physical examination, sensori-motor and behavioural tests in an experiment comparing mice derived from cryopreserved versus non-cryopreserved embryos. We used classical tests (Carlier et al., 1983; Roubertoux et al., 1992) most of which are cited in a recently published list of behavioural tests (Crawley et al., 1997). In addition, we assessed mandible morphology, a classic teratological test, as an indicator of micromalformations (Festing, 1972; Bailey, 1985). We also programmed a long-term study, postulating that some anomalies could be undetected in youth but become detectable in adults or in old subjects. Furthermore, the experiment was carried out in parallel on two hybrid strains. This caution had two purposes. Since it is well known that genetic background can modulate the susceptibility to environmental modifications, effects, for instance embryo tolerance to cryopreservation (Schmidt et al., 1985, 1987; Dinnyes et al., 1995), we wished to be able to detect a possible variability of the results in relation with genotype. We also wanted to compare the intensity of eventual effects of cryopreservation with that of genotype-related differences. Similarly, sex-related differences have been documented in various situations (Auroux, 1997). So, in our experiment three main factors, genotype, sex and cryopreservation, were considered...
and various tests or evaluations were performed at different periods up to 67 weeks of age.

**Statistical analysis**

Testart has apparently understood that our analysis only consisted in numerous paired comparisons (Testart, 1998). According to him, this would explain why some significant differences were found. This explanation is invalid; as mentioned in our paper, all data (excepted those evaluating pre-weaning development for which sexes were not separated) were analysed in a three-way analysis of variance, according to the experimental design. Paired comparisons were made only when the analysis of variance (ANOVA) showed a significant main effect or an interaction between cryopreservation and one of the other studied factors.

**No report of some negative results**

Testart also suggested that we intentionally omitted negative results, especially regarding body weight (Testart, 1998). The debate deserves better arguments than such insinuations. All the results that were not presented in tables or figures were mentioned in the text, and we confirm that body weight was only recorded at the ages that we mentioned.

**Discussion of the results**

**Maternal factor as a possible source of bias**

The use of outbred mice as recipient females has been questioned, as they could have provided an heterogeneous uterine environment that might have contributed to the observed differences. Although Naval Medical Research Institute (NMRI) mice are not inbred, they show high consanguinity and histocompatibility. Additionally, we carefully controlled their age (8 weeks) and weight (28–30 g) and we verified that offspring’s adult weights were not correlated with maternal ones. Moreover, this alternative explanation itself implies other hypotheses; the embryos from one category (for instance, the cryopreserved ones) should have been preferentially transferred to females having particular characteristics. This seems very unlikely, because the foster females were chosen at random and the transfers alternated embryos from the different groups. Furthermore, there should be specific relationships between this hypothetical particular type of females and the various differences that we observed. So, is such a hypothesis really less speculative than a relationship with cryopreservation?

**Variability of the observed differences**

Testart pointed out the variability of the effects of cryopreservation in relation to sex, genotype or age (Testart, 1998). According to him, this pattern revealed that there were only sampling variations giving rise, occasionally, to statistically significant differences. This view is based on a misunderstanding of our statistical procedures, as mentioned above. More fundamentally, it also seems to assume that true biological effects of cryopreservation should be similar whatever the genetic background. This is refuted by common experience with embryo freezing and, more generally, by numerous observations in biological research (Fraser and Fainstat, 1951; Tuchmann-Duplessis, 1975; Roubertoux et al., 1990, 1992). Of course, false significant differences due to random or to undetected biases cannot be completely ruled out but, in our study, it was very unlikely that all the effects found in ANOVA were due to random variations. Moreover, the experimental cautions taken, and the fact that clear differences were observed between sexes or strains made it difficult to identify the influence of an uncontrolled factor. Conversely, it is conceivable that both the primary cellular effects of cryopreservation and their later phenotypic expression could vary according to the genetic background, including sex.

**Size of the differences**

As we first stated in discussing our results, cryopreservation did not induce major anomalies, even in old animals. When observed, the differences between cryopreserved and control animals were similar in magnitude to other differences related to genotype or sex. Is this fact by itself sufficient to think that these differences were not due to cryopreservation but to sampling variations? In our opinion, given the hypotheses about the possible cellular effects of cryopreservation, finding differences that were equivalent or sometimes larger (e.g. body weight at 67 weeks) than genotype-related ones deserves attention. Variations of multifactorial quantitative traits like those we evaluated, whether spontaneous or experimental, are often moderate (Ramel, 1983; Vogel and Motulski, 1997). Thus, the pattern of the cryopreservation-related differences was compatible with the hypothesis of slight and probably heterogenous alterations of intrinsic developmental determinants. Of course, identifying these alterations by cellular and molecular investigations in cryopreserved embryos would be necessary to confirm a causal relationship between cryopreservation and later phenotypic changes.

**Selective bias**

Some embryos are destroyed by the freeze–thaw process, or fail to implant and proceed to further development. Therefore, cryopreservation operates a selection among the initial population of embryos. Testart proposes that such a selection might have caused the differences observed between the cryopreserved mice and the controls, which were not exposed to selection. Even if it supposes a relationship between tolerance to freezing–thawing (which may involve accidental features, like for instance the stage of cellular cycle at which the embryo is frozen) and later expressed phenotypic traits, this explanation could be pertinent when the embryos are genetically heterogeneous, as in humans. However, in our experiment, all the embryos from one strain were genetically identical, since they were F1 hybrids derived from inbred parents. There is no solid reason to suppose that some unknown factor had caused, before cryopreservation, a genetic heterogeneity of the embryos inside each strain.

**Evaluation of embryo cryopreservation and other assisted reproductive technologies, today and future**

In our paper, we concluded that this long-term study confirmed in the mouse that embryo cryopreservation is not severely
detrimental, but also indicated that it may not be absolutely free of long-term effects. What could this imply about the safety of embryo freezing in man? Of course, directly extrapolating from mice to humans was not conceivable. In humans, moreover, phenotypic changes like those we observed in highly controlled experimental conditions could be shadowed by the combined influence of genetic heterogeneity and highly diverse environmental conditions. Nevertheless, our results implied a possible susceptibility of development and post-natal phenotype to preimplantation manipulations or environmental factors. Such a notion had already been proposed (Reik et al., 1993). By itself, it justifies a more careful evaluation of embryo cryopreservation and other assisted reproductive technologies (ART).

Since our report, other studies on children born from cryopreserved embryos have been published (Sutcliffe et al., 1995a,b; Olivennes et al., 1996; Wennenholm et al., 1997, 1998). They all concluded that cryopreservation induced no major pathological features. Some differences were found with children born after either in-vitro fertilization (IVF) without cryopreservation or natural conception, e.g. frequencies of major and minor congenital anomalies, parameters evaluating mental age (Sutcliffe et al., 1995a,b) and weight of girls (Wennenholm et al., 1997, 1998), but they were either not significant or significant but isolated and rather small. Obviously, they remained difficult to interpret in the context of these studies where, despite methodological precautions, several confounding factors could not be controlled.

In other mammalian species, it is generally considered that cryopreservation of the embryo does not affect its development after implantation. Most experiments have been done in the mouse and, as said in introduction, have shown that embryo cryopreservation is not strongly teratogenic or mutagenic in this species. However, several studies have suggested the possibility of detrimental effects beyond implantation, as evidenced by increased rates of post-implantation losses (Rall et al., 1987; Kono and Tsunoda, 1988; Trounson et al., 1988; Shaw and Trounson, 1989; Wilson and Quinn, 1989; Liu et al., 1993) or by reduced fetal weight (Shaw and Trounson, 1989). A lower post-implantation viability has also been reported after mouse oocyte cryopreservation (Van der Elst et al., 1993; Wood et al., 1993; George et al., 1994), and although the induction of aneuploidies due to spindle disorganization has been proposed as a cause of fetal loss, this remains controversial. Some years ago, we showed ourselves, in mouse embryos derived from frozen–thawed oocytes, an increased frequency of sister chromatid exchanges, suggesting a mutagenic effect of cryopreservation (Bouquet et al., 1993). To our knowledge, no recent study has focused on post-natal development of animals cryopreserved as embryos.

Interestingly, however, other experiments have evidenced an influence of preimplantation in-vitro culture conditions or embryo manipulations on post-implantation development. For example, lamb birth weight has been found to be affected by the medium in which the embryos were cultivated from the zygote stage to compaction (Thompson et al., 1995). Retarded fetal development and neural tube defects have been observed in mouse fetuses after pre-implantation exposure to ammonium ions (Lane and Gardner, 1994). More generally, it seems that IVF and embryo culture or manipulations can affect post-implantation viability, gestation length and developmental features (Wright and Ellington, 1995; Walker et al., 1996; Thompson, 1997). In humans, the results from different studies about the outcome of IVF pregnancies and the children’s health status are contradictory about the possibility of moderate differences with non-IVF pregnancies. Recently, the follow-up of children conceived by intracytoplasmic sperm injection (ICSI) has raised a similar debate about the frequency of malformations (Bonduelle et al., 1997; Kurinczuk and Bower, 1997) and the occurrence of functional disorders (Bonduelle et al., 1998; Bowen et al., 1998). In this case however, paternally inherited factors may be involved.

Thus, numerous observations indicate that post-implantation development and phenotypic characteristics at birth can be modified by diverse manipulations or even by exposure of the embryo, before implantation, to transitory environmental changes. It is noteworthy that in many cases, the effects were observed in only some of the animals, the other being apparently not affected. All these observations highlight the need for improving knowledge about the early embryonic stages’ susceptibility to environmental hazards especially (but not only) in the context of assisted procreation. If rather mild changes, such as variations in culture medium composition, can have delayed effects, then observing similar phenomena after cryopreservation, which induces much more dramatic disturbances in the embryo and its microenvironment should not be very surprising. In this respect, the recent finding of formaldehyde, a cytotoxic and mutagenic chemical, in cryoprotectant solutions (Karran and Legge, 1996; Mahadevan et al., 1998), might be a serious matter of concern. A recent meeting at the Jackson Laboratory (Maine, USA, 1997), which celebrated the 25th anniversary of the first birth of a mouse derived from a cryopreserved embryo, came to the similar conclusions. A consensus emerged, stressing the need for a long-term screening including detection of subtle modifications, in order to verify the innocuity of the technique.

**Transmission of information to the couples**

It is now generally recognized that ART have often been introduced in clinical practice ‘without sufficient prior animal experimentation’ (ISLAT Working Group, 1998). Increasing attention is being given to their potential unexpected consequences. The next years will probably bring new insights about this issue but, at the present time, there are still more uncertainties than certitudes. Should not the physicians explain these uncertainties to the infertile couples? It is quite feasible, given time and care, to provide full information covering the different aspects of ART, including the possibility of unknown risks, without being unduly alarmist. This attitude has been recommended recently by Van Steirteghem (Van Steirteghem, 1998) and for a long time by one of us (Auroux, 1987). The couples themselves are, almost always, firmly in demand of objective and comprehensive information. A young woman said to one of us ‘we are aware of the tendency of covering
our eyes to truth and we demand that doctors help us to act with responsibility, even if this means not acting at all’. With respect to the child, future parents and physicians share a responsibility which obliges them to face facts.

Conclusions

We have always been aware that our study did not give definitive conclusions, but could rather contribute to increase the level of vigilance and stimulate further investigations. It is true that follow-up of the children is very difficult and must not become, by itself, a source of disturbances. But adequate prospective studies are essential to be able to state whether, in humans, ART have delayed consequences or not and whether these consequences, if present, are detrimental or not. In parallel, experimental models remain necessary, not only to more accurately assess the possibility of delayed effects, but also to explore presumptive mechanisms. Thorough evaluation combining clinical and experimental investigations is more than a scientific necessity, it is an ethical obligation. Indeed, with ART, physicians do not repair an organism, they help to create a new being. Thus, their responsibility reaches another level: ‘Audaces fortuna juvat’ but, here especially, ‘primum non nocere’.

References


1144


