There are important classes of drugs, including analgesics, anticonvulsants and antidepressants, where not all patients benefit, and medicines are tried out sequentially or in combination. We now know that minor genetic variations in the drug target may also lead to interindividual variation in drug responses. A recent study has shown that an exploratory new drug differed by tenfold in affinity for its target, the P2X7 ion-channel, solely depending on two SNPs in the protein[19,20]. Polymorphisms may be unrelated to known disease but determine which patients do and do not respond to a drug. For some drug targets, there are hundreds of variants. Having genetic sequences available that cover human diversity tells us the frequency of allelic variation in proteins. *In vitro* experiments are needed to know whether those variants affect drug responses. We are now in the realm of needing thousands of hiPSC lines.

Given that hiPSC lines have the potential to aid these important areas of research, what kinds of difficulties are associated with their use in large collections? With many laboratories across the world making hiPSC lines, there will inevitably be substantial heterogeneity in the cells produced. Sources of variation including different tissue sources of hiPSCs (e.g., hair, skin or blood), the donor's age and state of health, and the conditions for making, selecting and maintaining the hiPSCs. Systematic understanding of the biological sources of such variation remains in its infancy. In such a fast-moving field, it will not be possible to standardize methodology in the near term, and a concerted effort will be required to assimilate best practices.

Rather than being too prescriptive, we should collect hiPSC lines with associated key information and learn what works and what doesn't from scientists using those lines. It is important to consolidate information on which lines prove most consistent and useful. Banks grow in value with the data deposited. Initially, some simple standard criteria should be applied to confirm that a cell is indeed an hiPSC, that it is free from mycoplasma or other contamination, and that its unique identity is verifiable, for example, by short tandem repeat fingerprinting. When using hiPSCs for experiments, three pieces of information should ideally be available: the clinical description of the patient, their genetic sequence and a differentiation protocol to produce the relevant cell type with all associated methodological data. Appropriate consent and donor anonymization are therefore critical.

To be effective and most useful, a bank should have the following attributes:

1. Fully-informed donor consent supporting the donation of tissue to generate hiPSCs together with genetic information and relevant medical history. The ethical considerations here are not insignificant.
2. A process to anonymize donors and maintain a robust database.
3. Where donated cells and associated information are to be used for research, we must recognize that the cell lines made are not restricted to one group of researchers but are made broadly available to all researchers who can contribute to the understanding of disease and its treatment, including those from academia, biotech and pharma.
4. Standardized protocols for storage, retrieval, culture and differentiation, where known.
5. A mechanism to collect knowledge on any phenotypic abnormalities arising after differentiation and characteristics unique to particular cell types.
6. A searchable electronic 'catalog' where cells can be requested based on specific gene sequence or medical background, and a quick, easy way of shipping cells to scientists globally.

A future can be envisaged in which thousands of hiPSC lines with some fundamental elements of quality control are broadly available. The challenge is substantial, not least in terms of ethical review, data management, cost and logistics. The only economically viable path forward is to generate such a bank (or network of banks) precompetitively and collaboratively. Generating, validating and expanding hiPSC lines is costly, with estimates of $10,000–20,000 per line. It is also time consuming, requiring 4–6 months from tissue harvest to robust characterization of the expanded line. Yet the costs are surely outweighed by the benefits, as ensuring that hiPSCs become standardized, readily accessible, high-quality reagents will enable scientists to optimize time spent in understanding human biology and disease and in generating new therapies.

ACKNOWLEDGMENTS

We acknowledge the valuable contribution of Dr. Devyin Smith.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper (*doi:10.1038/nbt.2710*).

Ruth McKernan[1] & Fiona M Watt[2]

[1]Neusentis, Pfizer, Granta Park, Cambridge, UK. [2]Centre for Stem Cells and Regenerative Medicine, King's College London, London, UK.

1. Carr, A.J. *et al. PLoS ONE* **4**, e8152 (2009).
2. Chambers, S.M. *et al. Nat. Biotechnol.* **30**, 715–720 (2012).
3. Ming, G.L. *et al. J. Neurosci.* **31**, 16070–16075 (2011).
4. Paşca, S.P. *et al. Nat. Med.* **17**, 1657–1662 (2011).
5. Israel, M.A. *et al. Nature* **482**, 216–220 (2012).
6. Brennand K.J. *et al. Nature* **473**, 221–225 (2011).
7. Shi, Y. *et al. Sci. Transl. Med.* **4**, 124ra29 (2012).
8. Nica, A.C. & Dermitzakis, E.T. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120362 (2013).
9. ENCODE Project Consortium. *Nature* **489**, 57–74 (2012).
10. Grskovic, M. *et al. Nat. Rev. Drug Discov.* **10**, 915–929 (2011).
11. McNeish, J. *et al. J. Biol. Chem.* **285**, 17209–17217 (2010).
12. Xu, X. *et al. Stem Cell Res. (Amst.)* **10**, 213–227 (2013).
13. Lee, G. *et al. Nat. Biotechnol.* **30**, 1244–1248 (2012).
14. Liang, P. *et al. Circulation* **127**, 1677–1691 (2013).
15. Kia, R. *et al. Br. J. Clin. Pharmacol.* **75**, 885–896 (2013).
16. Dib-Hajj, S.D. *et al. Nat. Rev. Neurosci.* **14**, 49–62 (2013).
17. Illing, P.T. *et al. Curr. Opin. Immunol.* **25**, 81–89 (2013).
18. Jorgensen, A.L. *et al. PLoS ONE* **7**, e44064 (2012).
19. McHugh, S.M. *et al. Br. J. Clin. Pharmacol.* **74**, 376–380 (2012).
20. Gregory, A.P. *et al. Nature* **488**, 508–511 (2012).

A proposal to use gamete cycling *in vitro* to improve crops and livestock

To the Editor:

The grand challenge of producing enough food, fiber and fuel for an ever-growing global population has benefited tremendously from genetic improvements in agriculturally important plants and animals over the past century[1]. These genetic modifications have enabled billions more people to meet basic needs while using less arable land and providing good returns on research investment[2,3] (**Supplementary Fig. 1**). Yet despite reductions in malnutrition-driven stunting and wasting, many humans remain undernourished[4]. Satisfying these basic needs becomes more challenging with climate variability, constraints on productive farmland and limited availability of off-farm inputs (e.g., water, pesticides, phosphorus)[5,6]. Here, we outline the potential implications of an *in vitro* approach (thus far demonstrated

in mice) for generating mature, fully functional female gametes from embryonic and induced pluripotent stem cells, as a means for rapidly breeding and introducing new traits into livestock and crops.

The slow pace of conventional breeding as a means of meeting food demands in the coming years is a substantial drawback. Generation times, ranging from months to decades, continue to be a bottleneck in the development of new animals and plants with valuable agronomic traits. For major crops, developing new cultivars can take 5–40 years; in the developed world, historic rates of genetic gain are only 0.8–2% per year[7,8]. Long-generation and perennial crops, in which the return on research investment has always been distant, have commonly received only paltry R&D investment. Advancements in breeding methodologies for annual crops that dominate modern agricultural productivity have generally involved increasing the number of progeny evaluated per unit time, the precision and speed of evaluation, and/or the breadth of genetic variation accessible to breeding programs[6,8] (**Supplementary Note**).

In this correspondence, we propose that a recently described approach[9]—the cycling of gametes *in vitro* (COGIV)—could enable genetic material to be recombined more rapidly than in conventional breeding. Furthermore, the approach could do so at a comparatively marginal cost because 'proto-organisms' generated by COGIV need not mature into adults, permitting selection among vast numbers of genotypes for rare, desirable allele combinations. Finally, the temporary by-passing of a sporophytic stage may reduce barriers to hybridization, providing access to alleles that are beyond the reach of conventional crosses.

The COGIV procedure for generating fully functional gametes from somatic cells *in vitro* greatly decreases the potential generation time for mice[9]. In the protocol, ordinary mouse somatic cells were reprogrammed into totipotent cells, then specialized into male and female gamete precursors, and matured into fully competent gametes by culturing in testicular and ovarian cell broths. These gametes were then united pairwise, and the zygotes grown *in vivo* to generate fully functional and fertile mouse offspring[9]; if the parental genotypes were known, the genome of the zygote could be determined from the genotype of the discarded polar body. In an alternative protocol, a mouse zygote could be briefly cultured *in vitro* to generate an implantation-ready cell mass—a blastula—that could source ≥100 somatic mouse cells for use as gamete precursors for the next breeding cycle.

If the COGIV approach is translatable to species of agronomic importance, it offers many potential advantages for breeding. Above all, the method offers the possibility of reducing generation time while at the same time multiplying the number of nuclei that can be subjected to DNA marker-assisted selection (MAS), assuming all other factors remain equal. The time to complete each COGIV cycle, gamete to gamete, would likely vary by organism. For example, the mouse (a fast-generation higher animal) required about 5 weeks using the *in vitro* based method[9] versus about 9–10 weeks to go conventionally from fertilization of gametes to a sexually mature animal that could produce another generation of gametes. In contrast, algae (a lower plant more similar to proto-individual crop plants) can complete their life cycle from gamete to gamete in less than a week[10].

The speed advantage of COGIV would be most beneficial in agriculturally important livestock that currently take several years to develop from embryo to sexual maturation or in commercially important trees that require decades to mature[11]. With COGIV, long-generation crops and animals could advance through hundreds to thousands of generations of recombination in the time frame of a single conventional generation, if the MAS model did not need frequent updating[12,13]. In addition, in comparison to conventional breeding, the lower costs of propagating and maintaining proto-individuals *in vitro* by means of COGIV could enable selection with unprecedented stringency among vast numbers of genotypes at each step.

COGIV could also facilitate the process of trait introgression, often of a single gene, from an exotic donor genotype into an otherwise superior recurrent genotype. Conventionally, this is done by repeated backcrossing to the recurrent genotype, iteratively reducing the donor's genetic contribution while maintaining the target

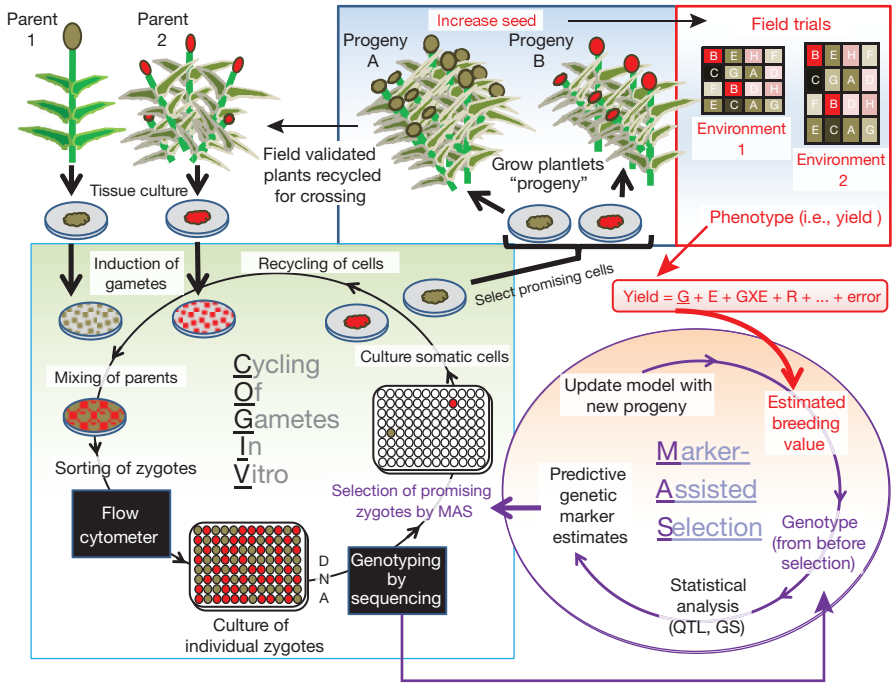

Figure 1 Schematic of COGIV. Somatic cells from two parents having desirable genetic traits are cultured, reprogrammed to totipotent cells and differentiated into gametes. *In vitro* fertilization methods are used to obtain zygotes[9,21]. Individual zygotes are sorted by flow cytometry and deposited in culture plates. DNA is extracted from a subset of these cells and used for genotyping. Cells with desirable genotypes are selected and either recycled into a new round of COGIV, or cultured into plantlets or embryos (two progeny are shown for simplicity but many would typically be employed). From seed produced from each plantlet or embryo, replicated multienvironmental field trials could be used to phenotype for traits of interest. The resulting variation would be parsed through a statistical model into genotype (G), environment (E), replicate (R), other variables (such as spatial variation) and error and/or noise. The phenotypic information could be combined with genotype and further partitioned into individual locus effects to map quantitative trait locus (QTL)[10] or used in a genomic selection (GS) model[11] to make selections in future generations; by applying this process iteratively, the marker-assisted selection (MAS) model would remain relevant as evolution progresses[12].

gene by selection. This process remains critical for the generation of transgenic plants and animals (genetically modified organisms) because transgenes are often introduced into cultivars that are favorable for transformation but not commercially competitive; only subsequently are transgenes moved into elite cultivars using MAS. COGIV could dramatically accelerate this process and/or permit greater precision (i.e., reduced donor contributions) or speed. For example, gene conversions, where one allele directly and specifically replaces another without reciprocal recombination, are surprisingly frequent in polyploids, such as cotton[14], and might be harnessed as a practical and naturally occurring means of allele replacement.

COGIV could also reduce constraints to hybridization for some crop species. For example, in Brassica vegetables (e.g., rape seed, cabbage and mustard), COGIV may be used to mitigate sporophytic incompatibilities (the mechanism by which the self-incompatibility phenotype of pollen prevents self-fertilization) by postponing the sporophytic stage until after recombination has separated genes of interest from those conferring incompatibilities (**Supplementary Note**). It is reasonable to anticipate extensions of present breeding practices, in which basic advances occur primarily in transparent academic settings while applications to scale occur in the commercial sector.

All of these advantages notwithstanding, one may anticipate several challenges to translation of the COGIV approach to large animal and crop species. One drawback is that COGIV does not provide a means for meaningful phenotypic evaluation (few improved plant cultivars have ever been developed based solely on phenotyping *in vitro*). In this respect, DNA MAS is readily extended to applications *in vitro* and may provide a means to select putatively favorable genotypes. Indeed, DNA markers identified a priori, or through approaches such as quantitative trait locus mapping[15], genomic selection[16–18] and high-resolution chromosomal structure-and-content analyses[19], are already routinely used in the breeding of many agricultural species (**Supplementary Note**). Ongoing improvements in nucleic acid sequencing, robotically automated cell sorting and information management systems may allow genotyping of many-fold more proto-individuals than is possible today.

Breeding of staple crops and animals generally involves crosses between two elite parents and identification of progeny combining the best alleles from each parent. In the COGIV paradigm (**Fig. 1**), germ-cell precursors derived from each parent's somatic cells would be cultured *in vitro*, leading to fertilization-ready gametes. Following fertilization, zygotes would be sorted and grown *in vitro* to generate arrays of proto-organisms. MAS could then be used to identify potentially superior genotypes to advance into the next COGIV cycle and/or to grow into embryos for phenotypic validation. Periodic phenotypic evaluation combined with genotyping would enable MAS models to be continually updated to improve accuracy for future generations of selection[12].

Several technical hurdles stand in the way of wide implementation of COGIV. First, as yet, we have only a proof-of-principle study in mice; for wider agricultural application, it will be necessary to show reprogramming and differentiation into fully competent gametes for cells derived from all the relevant livestock and crop species. In this respect, *in vitro* fertilization with single gametes has been conducted in some plants[20], as has research into plant tissue differentiation[21] and microspore culture. Even so, we are not aware of any work aiming to produce plant gametes from somatic cells *in vitro*, perhaps because there has been no motivating application.

Second, it may not be possible to culture desired genotypes *in vitro*, to demonstrate normal recombination and development, or to avoid excessive somatic mutation during COGIV. For example, in some species, certain genotypes are more easily cultured than others (and yet the mechanisms underlying these differences remain poorly understood[22]). In addition, somatic mutations leading to somaclonal variation in phenotype are often observed in older tissue cultures. This variation may be either a function of the number of generations of somatic cells (in COGIV, differentiated from reprogrammed totipotent cells) or a function of time in culture. Thus, it is not clear how many mutations would accumulate in the rapid cycling system of COGIV[23]. Although selection among gametes may help mitigate errors in the genesis process, improve traits and assist in removing deleterious recessive alleles causing genetic load (**Supplementary Note**), it is conceivable that aspects of COGIV might sacrifice some thus far unknown benefits of ordinary sexual reproduction and full phenotype development. In addition, COGIV might impose arcane selection pressures that cumulatively have large and undesirable impacts[24].

Taking all of the above factors into consideration, we have presented a rationale for exploring COGIV as a complement to empirical breeding and phenotyping. We believe it could offer real advantages in the genetic improvement of agricultural traits in species that currently take years or decades to develop using traditional approaches.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper (doi:10.1038/nbt.2707).

ACKNOWLEDGMENTS

We thank staff of the Agricultural Development Program of the Bill & Melinda Gates Foundation for its hospitality at the Perennial Crops Convening of October 2012, during which this collaboration commenced. Useful comments on the manuscript from B. Gates and D. Baltensperger are gratefully acknowledged.

AUTHOR CONTRIBUTIONS

L.W. and P.E. identified the application for the core mouse technology. All authors conceived applications, and co-wrote and edited the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper (*doi/10.1038/nbt.2707*).

Seth C Murray[1], Philip Eckhoff[2,3], Lowell Wood[2,4] & Andrew H Paterson[1,5]

[1]Department of Soil and Crop Sciences, Texas A&M University, College Station, Texas, USA. [2]Bill & Melinda Gates Foundation, Seattle, Washington, USA. [3]EMOD Project, Intellectual Ventures LLC, Bellevue, Washington, USA. [4]Hoover Institution, Stanford University, Stanford, California, USA. [5]Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia, USA.
e-mail: sethmurray@tamu.edu

1. Borlaug, N.E. *Plant Physiol.* **124**, 487–490 (2000).
2. Piesse, J. *et al. Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 3035–3047 (2010).
3. Fuglie, K. *et al. Science* **338**, 1031–1032 (2012).
4. Stevens, G.A. *et al. Lancet* **380**, 824–834 (2012).
5. Brummer, E.C. *et al. Front. Ecol. Environ*, **9**, 561–568 (2011).
6. Tester, M. *et al. Science* **327**, 818–822 (2010).
7. Cassman, K.G. *Proc. Natl. Acad. Sci. USA* **96**, 5952–5959 (1999).
8. Lee, E.A. *et al.* in *Handbook of Maize* (eds. J.L. Bennetzen, S.C. Hake,) 141–160 (Springer, 2009).
9. Hayashi, K. *et al. Science* **338**, 971–975 (2012).
10. Polle, J.E.W. *et al.* in *The Alga Dunaliella*, (eds. A. Ben-Amotz, J.E.W. Polle, D.V.S. Rao) 403–421 (Science Publishers, 2009).
11. Namkoong, G. *et al.* in *Tree Breeding: Principles and Strategies*, 177 (Springer, 1988).
12. Podlich, D.W. *et al. Crop Sci.* **44**, 1560–1571 (2004).
13. Wolc, A. *et al. Genet. Sel. Evol.* **43**, 23 (2011).
14. Paterson, A.H. *et al. Nature* **492**, 423–427 (2012).
15. Paterson, A.H. *et al. Nature* **335**, 721–726 (1988).
16. Heffner, E.L. *et al. Crop Sci.* **49**, 1–12 (2009).
17. Meuwissen, T.H.E. *et al. Genetics* **157**, 1819–1829 (2001).
18. Resende, M. Jr. *et al. New Phytol.* **193**, 617–624 (2012).

19. Dugoff, L. *N. Engl. J. Med.* **367**, 2249–2251 (2012).
20. Kranz, E. *Methods Mol. Biol.* **111**, 259–267 (1999).
21. Tran Thanh Van, K. *et al. Adv. Biochem. Eng.* **18**, 151–171 (1980).
22. Henry, Y. *et al. Euphytica* **79**, 45–58 (1994).
23. Bairu, M.W. *et al. Plant Growth Regul.* **63**, 147–173 (2011).
24. Mulcahy, D.L. *et al. Sex. Plant Reprod.* **9**, 353–356 (1996).

The regulatory landscape for actively personalized cancer immunotherapies

To the Editor:

Genetic heterogeneity is a hallmark of cancer. Distinct genotypes are found for individual tumors and even within single lesions[1,2]. This heterogeneity and the high diversity of the overall constitution of individual patients demand tailored approaches in tumor therapy. The promise of new approaches categorized as personalized medicine is that such drugs show increased efficacy and reduced adverse effects. Recent studies identified the therapeutic potential of immunogenic tumor mutations in mouse models[3,4]. For the first time, it has become technologically feasible to integrate data from high-throughput genome sequencing to identify immunogenic mutations and to design therapies tailored to the mutational composition of individual tumor genomes (mutanomes)[5,6]. Translation of genome-based vaccine approaches into human clinical trials is imminent[7]. However, regulatory challenges associated with such innovative personalized approaches are manifold, and specific regulatory guidance is not yet available. The choice of regulatory principles to be applied is facilitated when the level of personalization is appreciated. Here we propose a development strategy for actively personalized vaccines (APVACs) targeting multiple tumor mutations based on three distinct levels of personalization (**Fig. 1**). This strategy is based on the existing regulatory framework and thus should facilitate translation to human testing.

Figure 1 shows three types of approaches towards precision tumor immunotherapy. The first type (**Fig. 1a**) is characterized by biomarker-based stratification of patients to be treated with one invariant drug product. Typically, patients are identified who express predictive biomarkers such as *HER2* (also known as *ERBB2*), *KRAS* or epidermal growth factor receptor (*EGFR*) mutations. This kind of personalization is well covered by the existing regulatory framework originally intended for the development of chemical and biological drug products.

A second type of therapy (**Fig. 1b**) involves treatment with autologous medicines; these are mainly represented by cell therapy products administered to individual patients from whom cells have been retrieved as starting material for drug product manufacture. Of note, such products are applied irrespective of biomarker screening results. As this second type of product bears a high degree of heterogeneity intrinsic to the genetic background of the patient and the disease type, they can be considered as being 'passively' personalized (PPVACs). Examples of such products are Provenge (sipuleucel-T; autologous peripheral blood mononuclear cells/antigen-presenting cells activated *ex vivo* with recombinant prostatic acid phosphatase fused to granulocyte macrophage colony stimulating factor), which was approved by the US Food and Drug Administration in 2010 and by the European Commission in 2013, and Oncophage (autologous tumor–derived peptides in combination with gp96 heat shock protein; HSPPC-96), which was approved in Russia in 2008.

A third type of personalized immunotherapy (APVACs; **Fig. 1c**) mandates the use of molecular tests not only for patient selection but also to 'actively' personalize a drug product for every single patient. APVACs can be created in two ways. A first approach is 'warehousing' (e.g., using libraries of synthetic peptides for vaccination). Some examples of synthetic peptide vaccines have already reached the clinic. Depending on predictive biomarker signatures in individual patients, off-the-shelf peptides can be selected for use in individual patients[8,9]. The other approach to APVACs is *de novo* synthesis of drug product components that are then mixed in the appropriate manner for an individual. Examples of this APVAC concept are vaccines targeting tumor-specific mutations in one or two defined target genes known to be commonly mutated in patients, such as *TP53*, *KRAS* or the Von-Hippel-Lindau (*VHL*) gene[10,11].

Recent advances in whole-genome sequencing that make individual tumor genomes available within days at affordable cost now expand the aforementioned approach to a new level and reach far beyond a restricted set of predefined antigen libraries or a few known mutated genes[12]. Complete *de novo* synthesis of multitarget vaccines directed against the immunogenic mutanome of individual patients allows, for the first time, the comprehensive exploitation of highly immunogenic epitopes not covered by central immune tolerance. Identification of tumor-specific nonsynonymous mutations in individual tumors is achieved by novel, massively parallel, next-generation sequencing technology. Based on a list of tumor-specific mutations, a standardized selection process defines the set of mutations to be included in the vaccine. This could be achieved by applying computer-based prediction tools[13], determining naturally presented mutated peptides based on elution of human leukocyte antigen–restricted peptides and subsequent peptide sequencing by mass spectrometry[14] or testing patients for pre-existing immunity against identified mutated epitopes. Finally, a small fraction of tumor-specific mutated antigens are selected for the vaccine. Such actively personalized vaccines do not add new markers, but merely enrich antigens that are also present in vaccine formats, such as tumor-cell lysates that have frequently been used clinically in the past.

Addressing regulatory challenges associated with APVACs and their possible solutions requires a differentiated look at quality, preclinical and clinical aspects. Quality attributes of APVACs targeting the mutanome need to be prospectively defined as for any medicine, which will be feasible when using vaccine formats, such as synthetic peptides or RNA that can be well characterized. Because of the 'on demand' production of APVACs and the variability of APVACS that might be associated with varying stabilities, the generation of stability data as normally required for well-defined products will be challenging. Nevertheless, the same principles as for autologous cellular therapies might be applied, by performing real-time stability studies on a defined number of batches and then assigning the established shelf life to all subsequent batches manufactured by the same manufacturing process (extrapolation). Unfortunately animal models are not useful for establishing