REVIEWS

Plant protein improvement by genetic engineering: use of synthetic genes

J. M. Jaynes, M. S. Yang, N. Espinoza and J. H. Dodds

Methods now exist to construct genes coding for synthetic proteins enriched in essential amino acid content. The production of these synthetic proteins in potato tubers can improve the nutritive value of the potato and increase its importance as a basic food crop.

Protein nutrition

The biosynthesis of amino acids from simpler precursors is a process vital to all forms of life as these amino acids are the building blocks of proteins[1]. Organisms differ markedly with respect to their ability to synthesize amino acids. In fact, virtually all members of the animal kingdom are incapable of manufacturing some amino acids. There are twenty common amino acids which are utilized in the fabrication of proteins and essential amino acids are those protein building blocks which cannot be synthesized by the animal. It is generally agreed that humans require eight of the twenty common amino acids in their diet[2]. A nutritionally adequate diet must include a minimum daily consumption of these amino acids (Fig. 1).

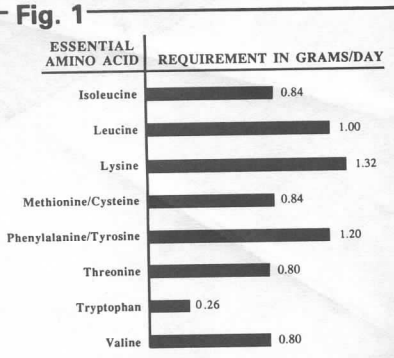
When diets are high in carbohydrates and low in protein over a protracted period, this results in deficiencies of essential amino acids and leads to a condition called 'Kwashiorkor' which is an African word meaning 'deposed child' (deposed from the mother's breast by a newborn sibling). This debilitating and malnourished state, characterized by a bloated stomach and reddish-orange discolored hair, is more often found in children than adults because of their great need for essential amino acids during growth and development. For normal physical and mental maturation, a balanced daily source of essential amino acids is a requisite. This is as important a feature of the diet as total protein quantity or total calorie intake.

Some foods, such as milk, eggs and meat, have very high nutritional values because they contain a high level of essential amino acids[3]. On the other hand, most foodstuffs obtained from plants possess a poor nutritional value because of their relatively low content of some or, in a few cases, all of the essential amino acids. Generally, the essential amino acids which are found to be most limiting in plants are isoleucine, lysine, methionine, threonine, and tryptophan (Fig. 2).

In the Western world few amino acid deficiency problems arise, primarily because the diet is composed of a mixture of a wide range of animal and plant proteins. In many developing countries, however, a single crop can be perhaps 80–90% of the total food intake. Rice, for example, is the major staple in Asia while potatoes are the staple in the Andean region of South America. With heavy dependence on plant protein from a single source, its essential amino acid composition becomes of critical importance. In these situations it would be highly beneficial to 'engineer' the plant to produce proteins with a balanced essential amino acid content. Several potential methods exist to achieve this objective. Firstly, utilization of conventional plant breeding techniques to improve protein quality; secondly, manipulation of the existing storage protein genes to increase levels of the essential amino acids; and thirdly, construction of synthetic genes which encode proteins enriched in essential amino acids for overall enhancement of the quality of total plant protein.

Although the first method has been applied with encouraging results in the case of maize[4] (i.e. the Opaque 2 mutant, so-called high lysine mutant) the screening techniques and effort involved make this an ineffective method. Besides, plant proteins are deficient in more than one of the essential amino acids which compounds the level of difficulty to achieve a nutritionally comple

Fig. 1

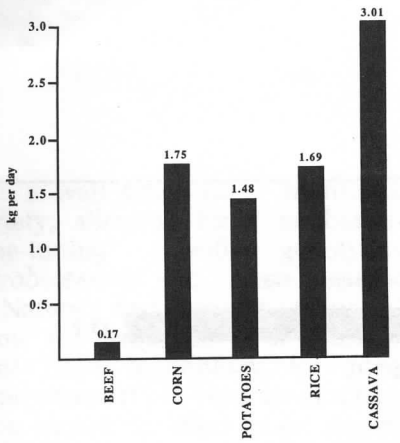


The daily essential amino acid requirements for a 20 kg child as listed are those necessary to maintain good health and normal growth and development as established by FAO. (A 90 kg adult male's daily essential amino acid requirement, to maintain good health, would be approximately the same – emphasizing the extreme importance of an adequate diet during early life.) The amino acid histidine has not been shown as it is generally regarded as being essential in humans for only the first three months of life. Also, the intake of cysteine and tyrosine can 'spare' (reduce) the requirements for methionine and phenylalanine, respectively.

J. M. Jaynes and M. S. Yang are at the Department of Biochemistry, 322 Choppin Hall, Louisiana State University, Baton Rouge, Louisiana 70803, USA. N. Espinoza and J. H. Dodds are at the International Potato Center, P.O. Box 5969, Lima, Peru.

 0166 – 9430/86/$02.00

Fig. 2



The lower quality of plant-derived proteins is apparent when one looks at the total amount of consumption necessary for a 20 kg child to receive 100% of his essential amino acid requirement and compare this value with the amount required from beef protein. The calculations assumed that the proteins were 100% bioavailable, so these values should be viewed as lower limits.

maize utilizing this method. The second method has been discussed previously in *TIBTECH*[5].

The innovative methods of genetic engineering offer a novel approach to modifying the essential amino acid composition of plant proteins and can thus increase their nutritive value. Our approach is not to try and modify the existing plant proteins, but rather to supplement them with new synthetic proteins which have a high content of essential amino acids. To produce a synthetic protein of known amino acid composition, a DNA fragment with an appropriate codon sequence can be produced using modern chemical techniques and introduced into various microorganisms using recombinant DNA technology[6]. The cloning and expression of synthetic DNAs, containing repeated codons for a single amino acid, could be used as a means to supplement the essential amino acid content of plant protein[7]. However, because plant proteins are deficient in more than a single essential amino acid, the design of a useful synthetic protein must take into consideration the total amino acid composition of plant proteins.

Gene design and synthesis

It is now possible, utilizing machines, to synthesize DNA fragments with an individual length up to about 100 bases[8]. Using such a machine we have constructed, cloned and obtained expression of genes in bacteria which code for proteins with a high content of the essential amino acids found to be most deficient in plant-derived proteins[9]. This was done in bacteria first, in order to facilitate the analysis of the new genes and their protein products. The sequences for several of these synthetic gene fragments – the 'high essential amino acid encoding' gene (HEAAE-gene) – have been deduced and the particular encoded protein sequences were obtained by inspection of the genetic code (Fig. 3).

Fig. 3

Gene Fragment 1

***Reading direction of top strand* (A)->->->**

```
AATTCGGGGATCGTAAGAAATGGATGGATCGTCATCCATTTCTTCATCCATTTCTTAC
    GCCCCTAGCATTCTTTACCTACCTAGCAGTAGGTAAAGAAGTAGGTAAAGAATG

GATCCATCCATTTCTTAAGAAATGGATGAAGAAATGGATGACGATCCATCCATTTCTT
CTAGGTAGGTAAAGAATTCTTTACCTACTTCTTTACCTACTGCTAGGTAGGTAAAGAA

CATCCATTTCTTCATCCATTTCTTACGATCAAGAAATGGATGAAGAAATGGATGAAGA
GTAGGTAAAGAAGTAGGTAAAGAATGCTAGTTCTTTACCTACTTCTTTACCTACTTCT

AATGGATGAAGAAATGGATGCATCCATTTCTTAAGAAATGGATGAAGAAATGGATGAA
TTACCTACTTCTTTACCTACGTAGGTAAAGAATTCTTTACCTACTTCTTTACCTACTT

GAAATGGATGACGATCGATCGTAAGAAATGGATGACGATCCATCCATTTCTTACGATC
CTTTACCTACTGCTAGCTAGCATTCTTTACCTACTGCTAGGTAGGTAAAGAATGCTAG

CCCG
GGGCTTAA
```

***Reading direction of bottom strand* (B) <-<-<-**

***Sequence of protein* (A)->->->**

```
GlyAspArgLysLysTrpMetAspArgHisProPheLeuHisProPheLeuThrIleHisProPheLeu-
- LysLysTrpMetLysLysTrpMetThrIleHisProPheLeuHisProPheLeuHisProPheLeuThr-
- IleLysLysTrpMetLysLysTrpMetLysLysTrpMetLysLysTrpMetHisProPheLeuLysLys-
- TrpMetLysLysTrpMetLysLysTrpMetThrIleAspArgLysLysTrpMetThrIleHisProPhe-
- LeuThrIlePro
```

***Sequence of protein* (B)->->->**

```
GlyAspArgLysLysTrpMetAspArgHisProPheLeuThrIleAspArgHisProPheLeuHisPro-
- PheLeuHisProPheLeuLysLysTrpMetHisProPheLeuHisProPheLeuHisProPheLeuHis-
- ProPheLeuAspArgLysLysTrpMetLysLysTrpMetLysLysTrpMetAspArgHisProPheLeu-
- HisProPheLeuLysLysTrpMetAspArgLysLysTrpMetLysLysTrpMetThrIleHisProPhe-
- LeuThrIlePro
```

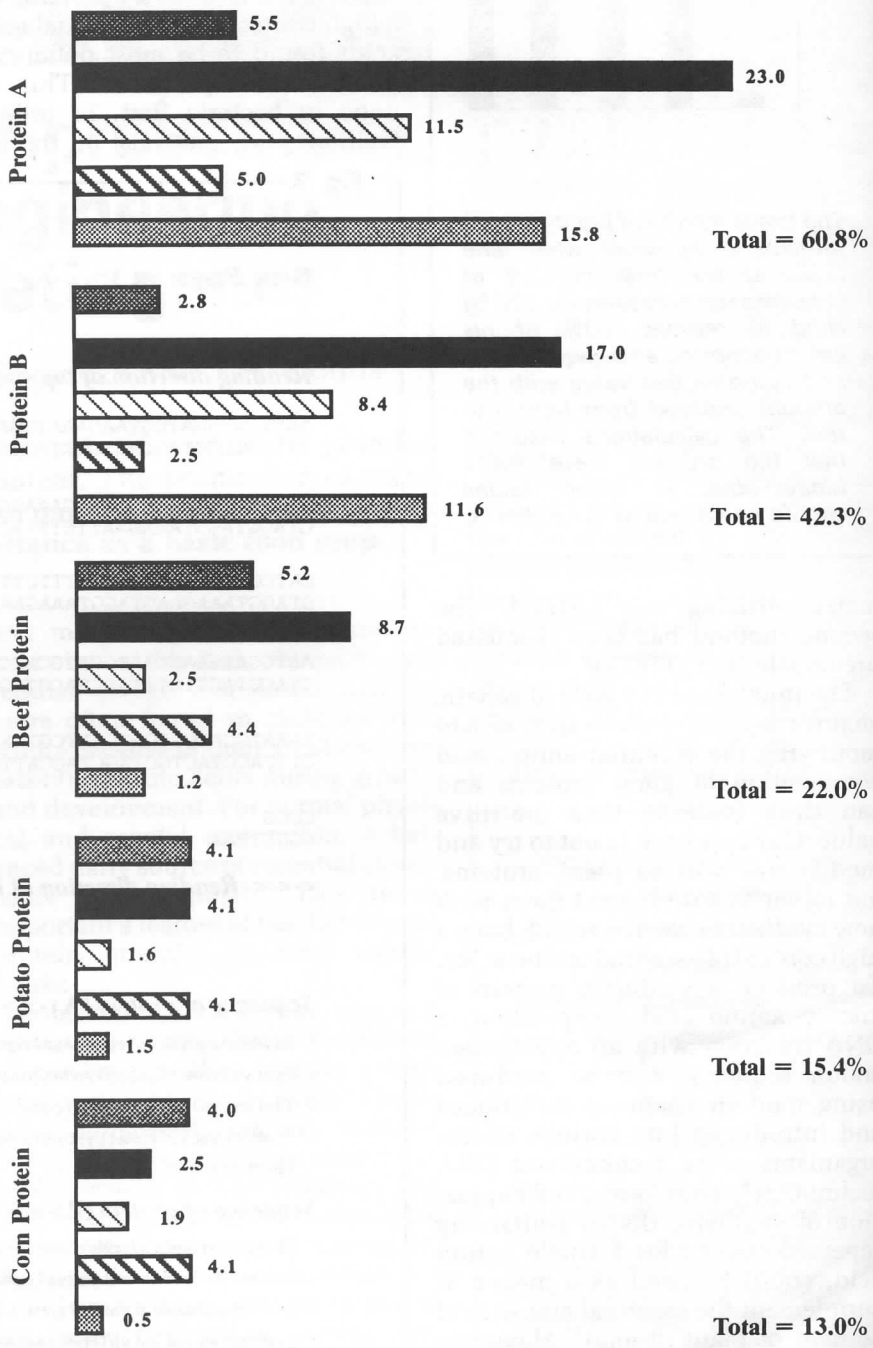
The nucleotide sequence of gene fragment 1 and the derived protein sequences. Gene fragment 1 yields two proteins designated A and B and one or the other will be produced depending upon the orientation of the gene when it is fused to the plant-gene promoter. In other words, the gene was constructed symmetrically so that a protein of high essential amino acid content would be produced no matter which DNA strand was read by the cell's protein synthesis machinery. Those amino acids highlighted are essential to the human and need to be consumed every day for good health. Protein A, the best so far analysed, is composed of about 80% essential amino acids.

Proteins A and B represent sequences derived from both possible reading directions of gene fragment 1. That is, gene fragment 1 was constructed symmetrically so that a protein, containing a high content of essential amino acids, would be produced no matter which strand of the synthetic DNA was ultimately read by the cell's protein synthesis machinery.

The composition of proteins A and B, for the five most deficient essential amino acids of plant-derived proteins, is compared with those found in proteins from several foods (Fig. 4). Beef protein has been shown for comparison purposes as it is composed of an unusually high content of these essential amino acids. Proper supplementation of plant-derived foods with proteins encoded by these synthetic gene fragments would markedly improve their essential amino acid balance.

This method of gene synthesis is flexible enough to produce proteins possessing any particular amino acid composition. Therefore, proteins could be specifically designed to supplement any desired animal feed or human food. It should be pointed out that the insertion of lysine at frequent intervals in these synthetic proteins provides numerous sites for proteolytic attack by trypsin (one of the main protein-degrading enzymes found in the digestive tract). This feature is important as it increases the bioavailability of the supplemental protein – the amount of amino acids absorbed from a particular dietary protein and used by the organism to make its own protein. The next major step has been the insertion of the synthetic genes into an important crop plant, the potato, using an *Agrobacterium*-based gene vector system.

Fig. 4



The percent composition found in the synthetic proteins A and B, for the five most deficient essential amino acids of plant-derived proteins, is compared with the content of proteins from some common foods. Synthetic protein A is composed of almost three times the content of these most important essential amino acids than is beef protein. ▩ Isoleucine: ■ Lysine: ▨ Methionine: ▨ Threonine: ▩ Tryptophan.

How the synthetic genes were introduced into potato plants

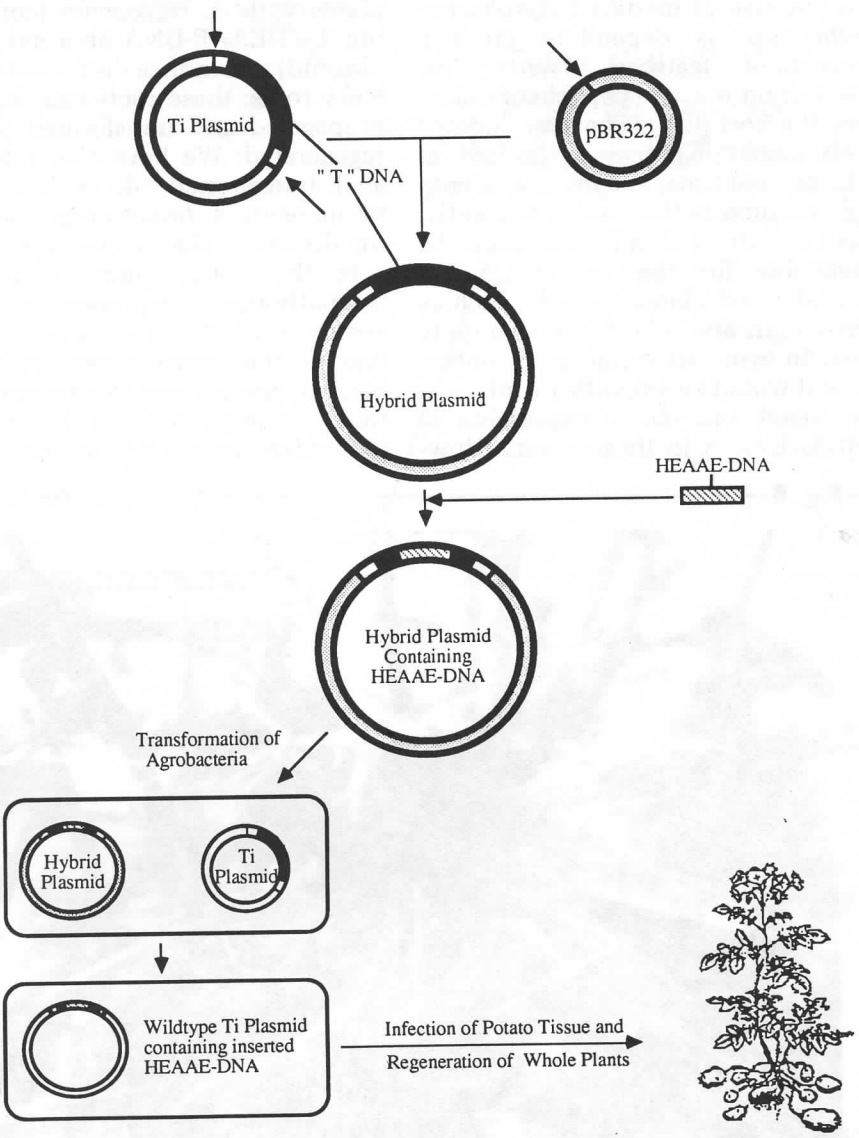
Plant breeding, the movement of desirable traits between plants by the use of the traditional techniques, has without question profoundly improved the important inheritable characters (yield, resistance to disease, etc.) of the world's major crop plants such as the potato. There are, however, certain limitations to these techniques. For example, it can take as long as ten years to introduce, select and establish a particular trait into a plant cultivar, and some traits are impossible to incorporate by these traditional techniques[10].

Genetic engineering offers the possibility of introducing a single trait, without altering other agronomically important characters of the plant,

into a well accepted or traditional variety, allowing for a method of 'fine-tuning' specific genotypes. *Agrobacterium* spp., often regarded as Nature's first 'genetic engineers', allow the exploitation of recombinant DNA technology for plant improvement[11]. Several species have been shown to affect plant growth and development. *Agrobacterium rhizogenes* causes a proliferation of roots at the point of infection on the host plant and this response has been termed 'hairy root'. The etiology (expression of symptoms) of hairy root is quite distinct from that of crown gall (a type of plant tumor), which is caused by *Agrobacterium tumefaciens.* The genes causing changes in the growth characteristics of infected plants have been shown to be located on large plasmids in both *A. rhizogenes* and *A. tumefaciens*[12]. Most importantly, a discrete portion of this plasmid DNA (called T-DNA) is found to be incorporated into the genome of the infected plant, a process known as 'plant transformation'. This T-DNA determines several fundamental characteristics of transformed plant tissue[13]. For example, these plant cells will grow without adding phytohormones to the culture medium, which are normally necessary for cell growth in culture. Also, the transformed cells will synthesize unusual compounds called 'opines'. In addition, the plasmids confer to the bacterium the ability to metabolize the particular opine produced by the transformed plant tissue. The evolutionary significance of transformation is that *Agrobacterium* causes the plant to produce special substances (the opines) which only the bacteria can utilize as a food source; thus, the bacteria have built a secure niche to promote their perpetuation at the expense of the plant.

One strategy for introducing foreign DNA into a plant, by means of *Agrobacterium* plasmids, is to use a small recombinant plasmid into which a known fragment of T-DNA has been inserted[14] (Fig. 5). This protocol can give rise to modified plant cells growing in culture. However, it is difficult to obtain healthy plants from these transformed cells. We have used a similar experimental strategy but employing an *A. rhizogenes* plasmid vector system in order to introduce the HEAAE-DNA into potato plant (Yang, M. S., Jaynes, J. M., Dodds, J. H. and White, F. F., unpublished). This procedure incites hairy or 'rooty' tumors and the infected root tissue can then be propagated. This appears to be a good method as the roots transformed with the modified T-DNA gave rise to plant cells which could be regenerated into intact plants (Fig. 6). The presence of HEAAE-DNA, its transcription into mRNA and translation into HEAAE-protein has been demonstrated in the potato by Southern transfer, 'northern blotting' and 'western' blotting, respectively. This method has proven to be very successful for the introduction and expression of HEAAE-DNA in transformed, regenerated potato plants (Espinoza, N., Yang, M. S., Jaynes, E. M., Dodds, J. H. and Schnorr, K. unpublished). A newer technique has emerged recently where *A. tumefaciens* plasmid

Fig. 5



Through the use of restriction enzymes, modified Ti plasmids can be constructed which contain foreign genes (in this case, the HEAAE-gene). These new plasmids can be introduced into* Agrobacteria *and these strains can then be used to infect potato plants. The resultant transformed tissue can be regenerated into whole, reproductively viable plants.

vectors have been constructed with the tumor-forming genes removed[15]. These so-called 'disarmed' vectors have allowed for the regeneration of plants from sections of the transformed leaf tissue. Transformed potato tissue employing this newer method has also been obtained.

Tissue culture of transformed potato plants

It is clear that techniques exist to insert the HEAAE-DNA into *Agrobacterium* plasmids; however, the controlled transformation of plants by the use of modified *Agrobacterium* spp. is dependent on two important features. Firstly, the bacterium must be capable of infecting the host plant. The usefulness of this infection process is limited, at the present time, to a few major crop plants such as the potato. Secondly, tissue culture techniques must be available for the regeneration of transformed plants. Cereals, such as rice, corn and wheat, are also quite low in essential amino acid content and it would be a significant advance to insert and obtain expression of HEAAE-DNA in these plants. However, cereals remain refractory to transformation by *Agrobacterium* and techniques do not exist to routinely regenerate whole cereal plants from single cells, a process which is extremely simple in the case of the potato[16]. Therefore, in order to genetically engineer cereals we must exploit new and different technologies.

The potato is a model plant for tissue culture and its plasticity of development allows intact plants to be regenerated from almost any plant part. It is possible to infect potato plants with *A. rhizogenes* (containing the HEAAE-DNA on a modified plasmid) and induce the formation of hairy roots; these roots can then be propagated and transformed plants regenerated. We have also regenerated transformed plants from leaf tissue using *A. tumefaciens* containing disarmed plasmid vectors.

In the potato plant, which is normally asexually propagated, there are no problems of segregation and loss of the inserted genetic information through meiotic recombination, although it is of interest to determine how this newly introduced gene (HEAAE-DNA) segregates and if it is transmitted sexually to the progeny.

Foreign DNA can be inserted at will into plants such as the potato. However, to ensure expression of a particular desirable gene, it must be inserted next to the proper control region of another already existing gene which has had its protein encoding part removed. These control regions, or promoters, usually found at the beginning of the gene, are vital switches attuned to the many influences which regulate the expression of the gene to its final protein product[17].

The potato, like many important plants, contains its food deposits in a specialized organ, in this case, a tuber. It would be of little import to improve the nutritional value of potato leaves as these are not consumed. Thus the ultimate goal of this research is to obtain high level expression of HEAAE-DNA only within the tuber (Fig. 7). This will increase the gene's value and importance as an inheritable genetic trait if this goal can be accomplished. It is important to have a method which can quickly analyse for the presence of the HEAAE-protein in the potato tuber. Techniques have now been developed which allow for the rapid (four weeks) induction of tubers *in vitro*[18]. These tubers are small, normally 3–5 mm in diameter but are morphologically and biochemically identical to field produced tubers (Fig. 8). This method will allow for speedy and efficient screening of the transformed regenerated plants for any increase in essential amino acid content of the tubers.

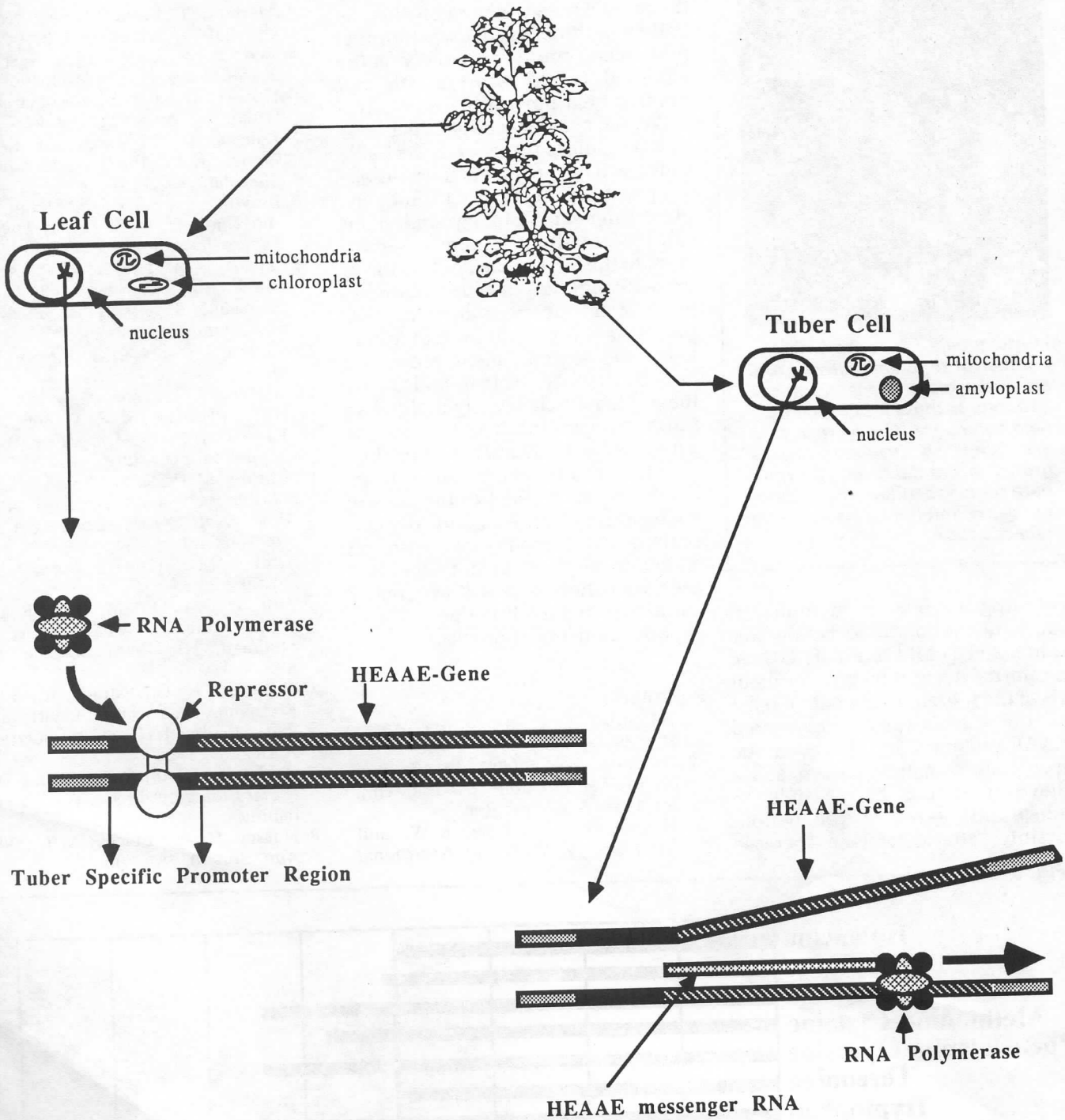
Fig. 6

Agrobacterium rhizogenes *incites hairy roots on* in vitro *propagated potato plants. Bacterial-free individual roots from these plants were placed on nutrient media to induce the formation of plantlets.*

Conclusions

The technology now exists to synthesize 'custom made' genes for proteins with an enriched essential amino acid content. We have been able to insert these genes into the potato and detect expression of the gene. An analysis of amino acid profiles of transformed and non-transformed plants is currently under study. Just how much HEAAE-protein will need to be produced within the tuber to make a significant impact on the essential amino acid content of the potato has yet to be ascertained.

Fig. 7



Modified Ti plasmids which contain foreign genes (e.g., the HEAAE-gene) can be constructed. The T-DNA region of the Ti plasmid was inserted into the E. coli *plasmid, pBR322. The pBR322 portion of this hybrid plasmid (i) permitted propagation of the hybrid in* E. coli *and (ii) bore an antibiotic resistance marker permitting subsequent selection of plasmid-containing clones. The HEAAE-DNA was then inserted into the T-DNA region of the hybrid plasmid. This new plasmid was then introduced into an* A. tumefaciens *strain carrying the unmodified* Agrobacterium *plasmid. Since these different plasmids contain homologous (similar) sequences, a rare double-recombination between the two plasmids will sometimes occur and result in a new plasmid in which the T-DNA harbors the insert of HEAAE-DNA. Strains of* Agrobacterium *containing the new plasmid can be identified and selected by their survival on media containing the antibiotic. These selected bacteria were then used to transfer the modified T-DNA (containing the HEAAE-DNA) into the plant genome.*

Fig. 8

In vitro *tubers can be grown from a number of different cultivars and are remarkably similar in shape and color to those produced under field conditions. On average they are 5 mm in diameter. Plants grown in the field, derived from both sources of tubers, yield about the same amount of consumable tubers.*

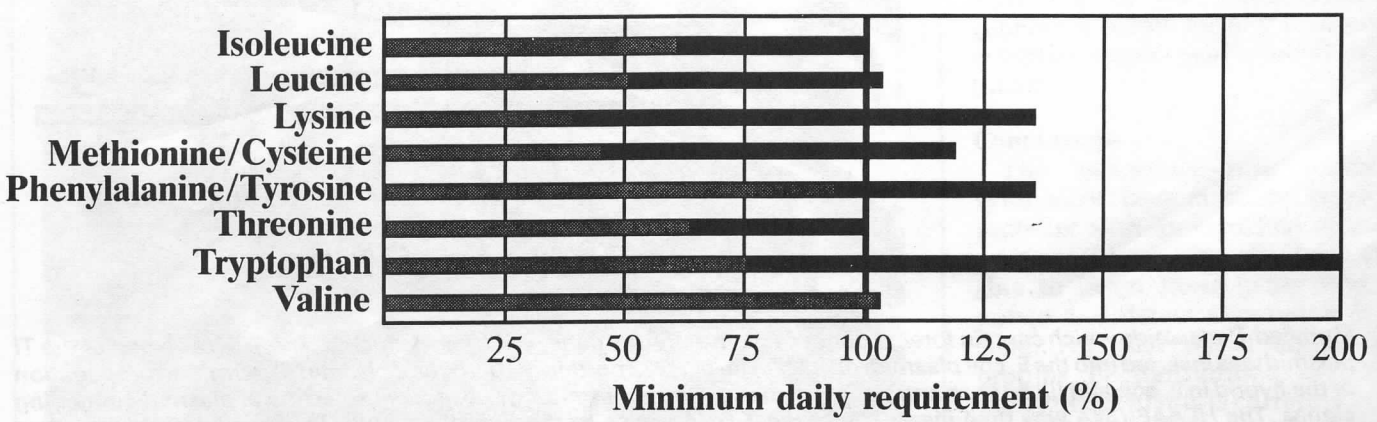
According to our calculations, if enough of the modified potato was fed to a 20 kg child to satisfy 20% of his calorific needs (600 g), then about 28% of the protein in the potato must be the genetically engineered HEAAE-protein (Fig. 9). Given the appropriate promoter sequence, we believe it should be possible to achieve this level of production. Certainly, any detectable increase would be significant. However, if we can achieve this level of expression of the protein then it is possible that the child would not need any additional protein and could exist solely on the modified potato with growth and development proceeding normally.

We have demonstrated that it is now possible to incorporate desirable genes into the potato via the methods of recombinant DNA technology, plant molecular biology and plant cell tissue culture. The major limitation to the large-scale application of these techniques for the improvement of the potato and other crop plants, is the general lack of knowledge about useful plant genes and how to identify, isolate, and purify them. This article has highlighted an innovative approach – if desirable genes cannot be located and purified, then they can be constructed synthetically. When applied to the genetic engineering of an economically important basic food crop with an agronomically significant synthetic gene we believe this will eventually culminate in a modified potato with superior nutritional qualities.

References

1 Doolittle, D. A. (1985) *Sci. Amer.* 253, 88–99
2 Arroyave, G. (1975) in *Protein-Calorie Malnutrition* (Olson, R., ed.), pp. 1–17, Academic Press
3 Hansen, R. G., Wyse, B. W. and Sorenson, A. W. (1979) *Nutritional Quality Index of Foods*, AVI Publishing
4 Mertz, E. T., Bates, L. S. and Nelson, O. E. (1964) *Science* 145, 279–280
5 Rao, A. S. and Singh, R. (1986) *Trends Biotechnol.* 4, 108–109
6 Gilbert, W. and Villa-Komaroff, L. (1980) *Sci. Amer.* 242, 74–94
7 Kangas, T. T., Cooney, C. L. and Gomez, R. F. (1982) *Appl. Environ. Microbiol.* 43, 629–636
8 Brown, E. L., Belagaje, R., Ryan, M. J. and Khorana, H. G. (1979) *Methods Enzymol.* 68, 109–151
9 Jaynes, J. M., Langridge, P., Anderson, K., Bond, C., Newman, C. W. and Newman, R. (1985) *Appl. Micro. Biotechnol.* 21, 200–205
10 Harlan, J. R. (1976) *Sci. Amer.* 235, 88–97
11 Albinger, G. and Beiderbeck, R. (1977) *Phytopathol. Z.* 90, 306–310
12 Kado, C. I. (1976) *Annu. Rev. Phytopathol.* 14, 265–308
13 Chilton, M. D., Saiki, R. K., Yadar, N., Gordon, M. P. and Quentier, F. (1980) *Proc. Natl Acad. Sci. USA* 77, 4060–4064
14 Chilton, M. D. (1983) *Sci. Amer.* 248, 50–59
15 Horsch, R. B., Fraley, R. T., Rogers, S. G., Sanders, P. R., Lloyd, A. and Hoffman, N. (1984) *Science* 223, 496–498
16 Espinoza, N. O., Estrada, R., Silva-Rodriguez, D., Tovar, P., Lizarraga, R. and Dodds, J. H. (1986) *Outlook Agric.* 15, 21–26
17 Schleif, R. (1985) *Genetics and Molecular Biology*, Addison-Wesley Publishing
18 Hussey, G. and Stacey, N. J. (1981) *Ann. Bot.* 48, 787–796

Fig. 9



The essential amino acid profile of the putative nutritionally complete modified potato tuber is shown above. The data presented assumes the consumption of 20% of the caloric requirement (600 g) of the modified potato tuber for a 20 kg child. The black rectangles are those essential amino acids provided by the HEAAE-protein showing that with the correct level of expression the essential amino acid deficiencies of the potato can be overcome. The potato would thus become a complete food.