

Supplemental Materials and Methods: Details about the categorization process of the TOBEST clusters by a curator, in complementation to the automatic KOGnitor categorization.

It was observed that many of the genes previously described in *N. tabacum* as pistil-specific, involved in growth, development or pollen-pistil interactions, like the pistil-specific extensin-like protein PELP (Goldman et al., 1992), the transmitting tissue-specific TTS (Cheung et al., 1993) and the stigma-specific STIG1 (Goldman et al., 1994) were categorized as ‘no related KOG’ by the automated KOGnitor. In addition, several sequences with similarity to genes involved in the synthesis of hormones, as well as proteins regulated (induced or repressed) by hormones were also initially categorized as ‘no related KOG’. For this reason, a new category ‘Growth and Development’ was introduced to accommodate the sequences mentioned above. It is worth mentioning that not all the pistil-specific or hormone related genes are in this new category, as the sequences automatically categorized by KOGnitor were not moved from their original category. For example, the sequence TOBS042G05, encoding the pistil ABC transporter NtWBC1 (Otsu et al., 2004), was automatically associated to KOG0061 and categorized by KOGnitor as ‘Secondary metabolites biosynthesis, transport and catabolism’. The TOBS035C08, encoding a protein similar to an auxin transporter-like protein, was automatically associated to KOG1303 and to category ‘Amino acid transport and metabolism’. Additionally, flower or pistil-specific genes, for which a possible defense role has been suggested, like the (1-3)-beta-glucanase sp41 (Ori et al., 1990), the flower-specific gamma-thionin - FST (Gu et al., 1992), the flower-specific polyphenol oxidase tobP1 (Goldman et al., 1998) and the transmitting tissue-specific thaumatin (SE39b)-like protein (Kuboyama, 1998), and that were initially classified as ‘no related KOG’, were manually categorized as ‘Defense mechanisms’. Therefore, the ‘Growth and Development’ category may be under represented. In the ‘Defense mechanisms’ category were also included the sequences with similarity to stress-response proteins, as dehydration and hypoxia responsive proteins.

Kuboyama T (1998) A novel thaumatin-like protein gene of tobacco is specifically expressed in the transmitting tissue of stigma and style. *Sexual Plant Reproduction* **11**: 251-256