DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
103

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
103

STRUCTURE OF THE 2,4-DICHLOROPHENOXYACETIC ACID-DEGRADATIVE PLASMID pEST4011

EVE VEDLER

TARTU UNIVERSITY
PRESS

Department of Genetics, Institute of Molecular and Cell Biology, University of Tartu, Estonia

Dissertation is accepted for the commencement of the degree of Doctor of Philosophy (in Genetics) on March 30th, 2005 by the council of the Institute of Molecular and Cell Biology, University of Tartu.

Opponent: Professor Jan Roelof van der Meer, PhD (Dept. of Fundamental Microbiology, University of Lausanne, Switzerland)

Commencement: May 20, 2005

Publication of this dissertation is granted by the University of Tartu

ISSN 1024–6479
ISBN 9949–11–031–9 (trükis)
ISBN 9949–11–032–7 (PDF)

Tartu Ülikooli Kirjastus
www.tyk.ee
Tellimus nr. 141

CONTENTS

LIST OF ORIGINAL PUBLICATIONS 7

ABBREVATIONS 8

1. INTRODUCTION 9

2. OVERVIEW OF LITERATURE 10
   2.1. Microorganisms and genes involved in aerobic 2,4-D degradation 10
   2.2. Common characteristics of plasmids 12
      2.2.1. Plasmid replication, stable inheritance and copy number control 13
      2.2.2. Plasmid incompatibility 14
      2.2.3. Plasmid transfer 14
   2.3. IncP1 plasmids 16
      2.3.1. Replication determinants 18
      2.3.2. Copy number control 20
      2.3.3. Conjugative transfer 20
         2.3.3.1. Mating pair formation complex 21
         2.3.3.2. DNA transfer and replication functions 22
      2.3.4. Stable inheritance 23
         2.3.4.1. The active partitioning mechanism 23
         2.3.4.2. Post-segregational killing and multimer resolution systems 26
      2.3.5. The regulatory network of IncP1 plasmids 27
   2.4. The 2,4-D-degradative plasmid pEST4011 isolated in Estonia 30

3. RESULTS AND DISCUSSION 32
   3.1. Determination of the role of TfdR in the expression of the *tfdCB* operon (Reference I) 32
   3.2. Mapping of the transcription initiation points of *tfdR* and *tfdCB* (Reference I) 33
   3.3. Species identification of strain EST4002 and isolation of pEST4011 from this strain (Reference II) 34
   3.4. Complete nucleotide sequence of and organization of pEST4011 (Reference III) 35
   3.5. Genes and IS elements of the pEST4011 catabolic transposon (References II and III) 36
   3.6. Comparison of the pEST4011 backbone with other IncP1 backbones (Reference III) 40

3.7. Multiple alignments of *trfA2*, *korA* and *traG* gene products of IncP1 plasmids (Reference III) 42
3.8. Comparison of pEST4011 and its laboratory ancestor pD2M4 (Reference III) 43

4. CONCLUSIONS 45

REFERENCES 46

SUMMARY IN ESTONIAN 58

ACKNOWLEDGEMENTS 60

PUBLICATIONS 63

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original papers which will be referred by their Roman numerals in the text.

I Vedler, E., Kõiv, V. and Heinaru, A. (2000) TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011. Gene 245:161–168.

II Vedler, E., Kõiv, V. and Heinaru, A. (2000) Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pEST4011 of *Achromobacter xylosoxidans* subsp. *denitrificans* strain EST4002. Gene 255:281–288.

III Vedler, E., Vahter, M. and Heinaru, A. (2004) The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring *tfd* genes for 2,4-dichlorophenoxyacetic acid degradation. Journal of Bacteriology 186:7161–7174.

ABBREVATIONS

aa	amino acid(s)
ADP	adenosine diphosphate
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
bp	basepair(s)
C-terminus	carboxyl-terminus
CC12O	chlorocatechol 1,2-dioxygenase
cco	central control operon
ccr	central control region
2,4-D	2,4-dichlorophenoxyacetic acid
3,5-DCC	3,5-dichlorocatechol
2,4-DCP	2,4-dichlorophenol
2,4-DCPH	2,4-dichlorophenol hydroxylase
Dtr	DNA transfer and replication
IS	insertion sequence
kb	kilobasepair(s)
kDa	kilodalton(s)
Mpf	mating pair formation
mrs	multimer resolution system
N-terminus	amino-terminus
ORF	open reading frame
psk	post-segregational killing
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gelelectrophoresis

1. INTRODUCTION

2,4-dichlorophenoxyacetic acid (2,4-D) was the first successful selective herbicide developed. It controls broad-leaf weeds in a crop without damaging that crop itself. After nearly 60 years of use, it is one of the most widely used herbicides due to its relatively moderate toxicity, when in concentrations resulting form adequate use in agriculture, and to its biodegradability in the soil (http://www.24d.org). 2,4-D is the most thoroughly researched herbicide in the world.

Microbial degradation of 2,4-D is a well-studied process. Various soil bacteria can use this compound as a carbon and energy source. Therefore, 2,4-D has become a model for studying the evolution and distribution of genes for the degradation of xenobiotic chloroaromatic compounds. A number of bacterial strains belonging to different phylogenetic groups have been found to possess genetically and enzymatically different genes and pathways for 2,4-D degradation (Fulthorpe *et al*., 1995; Fulthorpe *et al*., 1996; Itoh *et al*., 2002; Ka *et al*., 1994; Kamagata *et al*., 1997). Bacterial catabolic genes are often encoded by mobile genetic elements, including transposons and conjugative plasmids. Accumulating information from complete nucleotide sequences of these elements is helpful to understand the mechanisms leading to assembly and dissemination of different functional degradative pathways.

Nearly 20 years ago, several 2,4-D-metabolizing bacterial strains were isolated from different soil samples from Estonian agricultural enterprises. They all contained 2,4-D degradative plasmids that were the same size (Ausmees and Heinaru, 1990; V. Kõiv, unpublished data). One of these strains, *Achromobacter xylosoxidans* subsp. *denitrificans* EST4002 with a stable 2,4-D^+ phenotype contains the plasmid pEST4011.

The aim of the present thesis was to study the expression of two 2,4-D degradative genes found in pEST4011, namely *tfdB* coding for 2,4-dichlorophenol hydroxylase and *tfdC* coding for 3,5-dichlorocatechol 1,2-dioxygenase, and the role of TfdR, the LysR-type transcriptional activator, in the regulation of the *tfdCB* operon. The final goal was to determine the complete nucleotide sequence of the plasmid pEST4011, and to describe and analyze all open reading frames and other features found in this plasmid.

2. OVERVIEW OF LITERATURE

2.1. Microorganisms and genes involved in aerobic 2,4-D degradation

Many 2,4-D degrading bacteria have been isolated both from 2,4-D exposed (Don and Pemberton, 1981; Ka *et al*, 1994) and from nonexposed (pristine) (Itoh *et al*., 2000; Kamagata *et al*., 1997) soils. On evolutionary and physiological bases, 2,4-D degrading bacteria are divided into three groups (Kamagata *et al*., 1997). The first group consists of fast-growing copiotrophic bacteria belonging to β and γ subdivisions of *Proteobacteria*, including such genera as *Achromobacter*, *Alcaligenes*, *Burkholderia*, *Cupriavidus*, *Delftia*, *Halomonas*, *Pseudomonas*, *Rhodoferax* and *Variovorax*. These bacteria have been isolated from soils and sediments polluted with 2,4-D, and they all have *tfd* genes in common, usually present in a transmissible plasmid (Cavalca *et al*., 1999; Fulthorpe *et al*., 1995; McGowan *et al*., 1998; Vallaeys *et al*., 1999).

The best-studied 2,4-D degradative microorganism is the β-proteobaterium *Cupriavidus necator* JMP134 (formerly named as *Wautersia eutropha*, *Ralstonia eutropha* and *Alcaligenes eutrophus*). This bacterium was originally isolated from 2,4-D exposed soil in Australia (Don and Pemberton, 1981) and it contains the plasmid pJP4. The 87,688-bp complete nucleotide sequence of the latter has been determined recently (Trefault *et al*., 2004). The plasmid pJP4 contains among other genes the 2,4-D catabolic genes *tfdA* to *tfdF* as well as the isofunctional genes $tfdB_{II}$ to $tfdF_{II}$, two identical genes *tfdR* and *tfdS* both coding for the LysR-type transcriptional activator TfdR, and *tfdK* coding for transporter protein for active 2,4-D uptake (Fig. 1). All these genes are well-characterized together with the corresponding enzymes responsible for converting 2,4-D to central metabolic intermediate 3-oxoadipate (Laemmli *et al*., 2000; Laemmli *et al*., 2004; Leveau *et al*., 1999; Leveau *et al*., 1996; Plumeier *et al*., 2002; You and Ghosal, 1995) (Fig. 2). During the initial step in this pathway, the acetate side chain is cleaved from 2,4-D to produce 2,4-dichlorophenol (2,4-DCP). This reaction is mediated by an α-ketoglutarate-dependent dioxygenase, encoded by *tfdA* (Fukumori and Hausinger, 1993a,b). 2,4-DCP is subsequently hydroxylated to 3,5-dichlorocatechol (3,5-DCC) by 2,4-DCP hydroxylase encoded by *tfdB* and $tfdB_{II}$ (Farhana and New, 1997; Liu and Chapman, 1984; Perkins *et al*., 1990). 3,5-DCC is further degraded through a modified *ortho*-cleavage pathway encoded by *tfdCDEF* and $tfdD_{II}C_{II}E_{II}F_{II}$ (Laemmli *et al*., 2000; Perkins *et al*., 1990). In the presence of 2,4-D in the growth medium, expression of both *tfd* and tfd_{II} genes is induced by TfdR (Leveau *et al*., 1999). The presence of two isofunctional, but phylogenetically different *tfd* gene sets in pJP4 (amino acid (aa) identities range between 22 to 60% in cases of individual *tfd* and tfd_{II} gene products) may be required for efficient degradation of 2,4-D

and other related compounds (Laemmli *et al*., 2000; Pérez-Pantoja *et al*., 2000; Pérez-Pantoja *et al*., 2003; Plumeier *et al*, 2002).

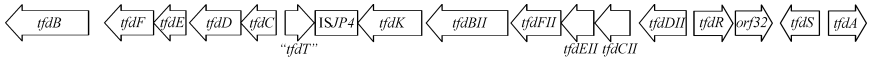

Figure 1. Arrangement of *tfd* genes in plasmid pJP4 (AY365053).

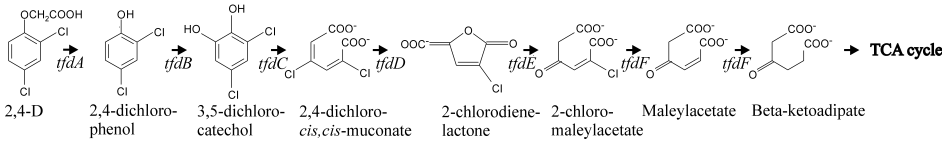

Figure 2. 2,4-D degradation pathway encoded by *tfd* genes. TCA, tricarboxylic acid.

Nucleotide sequence similarities of *tfdA*, *tfdB* and *tfdC* within this first group of 2,4-D degraders are more than 76, 34 and 55%, respectively (McGowan *et al*., 1998; Vallaeys *et al*., 1999). 2,4-D degraders of this group have supposedly obtained the respective catabolic activity by independent *tfd* gene recruitment through horizontal gene transfer (Fulthorpe *et al*., 1995; Ka and Tiedje, 1994; Ka *et al*., 1994; Neilson *et al*., 1994; Top *et al*, 1995; Vallaeys *et al*., 1996). The observed large diversity of gene sequences and gene assemblies suggests an ancient origin of *tfd* genes (Vallaeys *et al*., 1999). Besides pJP4, there are only two cases in which the DNA regions containing *tfd* genes for the whole 2,4-D degradation pathway have been sequenced, namely a chromosomal transposon-like structure (about 30 kb) from *Delftia acidovorans* P4a (Hoffmann *et al*., 2003) and Tn*5530* (41 kb) located in plasmid pIJB1 from *Burkholderia cepacia* 2a (Poh *et al*., 2002; Xia *et al*., 1998). In all three cases, the homologous *tfd* genes are organized in different ways.

The second group of 2,4-D degraders consists of α-*Proteobacteria* which are closely related to a *Bradyrhizobium* sp. They are slow-growing oligotrophic bacteria and have been isolated from pristine environments in Hawaii, Canada, Chili and Japan (Itoh *et al*., 2000; Kamagata *et al*., 1997). For 2,4-D degradation, these bacteria possess either the *cad* genes, which have nucleotide sequences not similar to the *tfd* genes (Kitagawa *et al*., 2002), or *tfd*-like genes less similar to those of the members of the first group of 2,4-D degraders (Itoh *et al*., 2004). As these bacteria have been isolated from environments with no exposure to 2,4-D, the enzymes performing 2,4-D degradation should carry functions other than 2,4-D degradation. *tfdA*-like genes were also shown to be present in all root nodule *Bradyrhizobium* strains tested, none of which showed 2,4-D-degrading activity (Itoh *et al*., 2004). Therefore, it is supposed that these

2,4-D degraders could provide useful information on the origin and evolutionary relationship of 2,4-D catabolic genes.

The third group of 2,4-D degraders have been isolated from polluted sites and they are copiotrophic bacteria belonging to the genus *Sphingomonas* of α-*Proteobacteria* (McGowan *et al*., 1998; Suwa *et al*., 1996; Vallaeys *et al*., 1999). They have been shown to possess *tfdB* and *tfdC* genes, similar to those of the first group, but the initial step in 2,4-D degradation is not likely performed by an α-ketoglutarate-dependent dioxygenase, instead the *cadA* gene homologs have been suggested to be involved in this step (Kitagawa *et al*., 2002).

The existence of three distinct ecological and genetic classes of 2,4-D degraders indicates a diversity of 2,4-D degradation genes and pathways among 2,4-D degraders. However, the 2,4-D degradation genes of the latter two groups of 2,4-D degraders have not been studied well.

2.2. Common characteristics of plasmids

Plasmids are extrachromosomal units of genetic inheritance that range in size from approximately 300 basepairs (bp) to 2400 kilobasepairs (kb). Small plasmids are usually present in the host cell in multiple copies, while large plasmids are present in one or two copies per cell. Plasmids can be found in all three domains of the living world, in *Archaea, Bacteria* and *Eukarya*. Plasmids may constitute up to 25% of the genetic material of the cell. The majority of gram-negative and gram-positive bacteria harbor plasmids. Plasmids usually reside as covalently closed circular molecules. The diversity of plasmids as to their size, copy number and genetic makeup show that these molecules have evolved to best achieve fitness and survival in a cellular environment. It has been argued that the addition of plasmids in the cell may be of greatest advantage to that cell in its survival in an existing environment or in adjustment to a new one. Assurance of cell survival as conferred by the plasmid also ensures the survival of the plasmid itself (Kado, 1998).

The traits favoring the survival of the cell, located on plasmids, are classified into four different major categories: 1. Resistance (to antibiotics, heavy metals, UV, phages, bacteriocins etc.); 2. Energy metabolism (catabolism of carbohydrates, organic acids, anabolism of amino acids, vitamins etc.); 3. Virulence, pathogeneicity and symbiosis (endo/exotoxins, colonization, serum resistance, nodulation etc.); and 4. Dissemination and perpetuation (sex pili synthesis, mobilization, chemotaxis etc.) (Kado, 1998).

2.2.1. Plasmid replication, stable inheritance and copy number control

In order to survive in the host cell, plasmids must be able to initiate its replication independently from chromosomal DNA and they must possess mechanisms like multimer resolution and active partitioning for adequate distribution of plasmid copies to daughter cells during cell division (Gerdes *et al*., 2000; Helinski *et al*., 1996; Hiraga, 1992; Møller-Jensen *et al*., 2000; Nordstrom and Austin, 1989). Active partitioning is important for low- and intermediate-copy-number plasmids. High-copy-number plasmids rely primarily on random partitioning. In addition, the former two plasmid-types use post-segregational killing mechanisms, also called as plasmid addiction systems (for example a plasmid encoded toxin and antitoxin system), which result in death of plasmid-free segregants (Gerdes *et al*., 1997; Jensen and Gerdes, 1995; Kobayashi, 2000).

Regardless of plasmid size, the basic replicon of a plasmid generally consists of a definable origin, where DNA replication initiates (*ori*), and one or more adjoining controlling elements. The *ori*, harbored within several hundred bp, contains recognition sites for plasmid and host proteins involved in replication initiation. In the majority of replicons a structural gene, often designated as *rep*, is present encoding a plasmid-specific protein required for the initiation of replication (Helinski *et al*., 1996).

Three general replication mechanisms are known for circular plasmids: theta type, strand displacement and rolling circle (del Solar *et al*., 1998). DNA replication through the theta mechanism begins with the melting of the parental strands at the *ori* with the help of the plasmid-encoded Rep and chromosome-encoded DnaA proteins. Host's primases, helicases, topoisomerases and DNA polymerases I and III are required for replication to proceed until termination occurs at *ter* sequences, which are the binding sites of the proteins that promote termination of plasmid replication. The DNA synthesis of both strands is coupled and occurs continuously on one strand (leading strand) and discontinuously on the other strand (lagging strand). Replication by the theta-type mechanism is widespread among plasmids from gram-negative bacteria, but it can be also found in plasmids of gram-positive bacteria. The DNA synthesis can start from one or several origins, and replication can be either uni- or bidirectional (del Solar *et al*., 1998).

The strand-displacement mechanism of replication was mostly studied on the promiscuous plasmids of the IncQ family. Members of this family require three plasmid-encoded proteins for initiation of DNA replication – RepA (DNA helicase), RepB (primase) and RepC (*ori* binding protein). The replication is bidirectional, and synthesis of both strands occurs continuously and results in the displacement of the complementary strand. Replication of the displaced strand is initiated at the exposed single-stranded *ori*. Due to the activities of the

plasmid replication proteins RepA, RepB and RepC, the replication is independent of host-encoded primases and helicases, and this may account for the broad-host-range character of plasmid harboring this kind of replicons (del Solar *et al*., 1998).

Replication by the rolling-circle mechanism is unidirectional and since the synthesis of the leading and lagging strand is uncoupled, it is assymmetric. The newly synthesised leading plus strand remains covalently bound to the parental plus strand. Replication is initiated by the plasmid-encoded Rep protein, which introduces a site-specific nick on the plus strand, at a region termed double-stranded origin (*dso*). The Rep protein becomes covalently attached to the 5' end at the nick site, the 3' end is used as a primer for leading strand synthesis. Host's DNA polymerase III, single-stranded DNA-binding proteins and a helicase are needed for the DNA synthesis. The end products of leading strand replication are a double-stranded DNA molecule constituted by the parental minus and the newly synthesised plus strand, and a single-stranded DNA intermediate, the parental plus strand. Finally, the latter is converted into double-stranded DNA by host proteins initiating the synthesis at the single-stranded origin (*sso*). Replication by rolling-circle is widespread among multicopy plasmids of a size that is usually smaller than 10 kb (del Solar *et al*., 1998).

Each plasmid has a characteristic copy number within a given host under fixed growth conditions. To define and maintain this copy number, plasmids use negative regulatory circuits consisting of plasmid-encoded control elements that regulate the inititation the replication (del Solar *et al*., 1998; del Solar and Espinosa, 2000).

2.2.2. Plasmid incompatibility

When more than one plasmid coexist in a cell, they are compatible; closely related plasmids compete for the same replication or maintenance sites which because of their limited numbers do not allow stable coexistence of these plasmids in the same cell. On that basis, plasmids are grouped into different incompatibility (Inc) groups (Helinski *et al*., 1996), designated as IncF, IncN, IncP, IncQ, IncW etc. The dominant incompatibility element, the one that determines the Inc group to which the plasmid is assigned, controls the copy number (Yarmolinsky, 2000).

2.2.3. Plasmid transfer

Some plasmids seem to have expanded their habitat by developing machineries that enable their spread to different hosts. An important mechanism in this

regard is conjugation, which involves DNA transfer following the establishment of direct contact between a donor and a recipient cell. Genes encoding these functions are commonly located on plasmids, termed conjugative or self-transmissible plasmids (Firth *et al*., 1996).

The initial step in bacterial conjugation involves the formation of physical contact between the donor and recipient cells. This contact is established through thin, tube-like extracellular filaments, called conjugative pili, protruding from the donor cell. The tip of the pilus binds to a receptor on the recipient cell. A depolymerisation step is thought to pull the two cells together, thus allowing the cell envelopes to engage in intimate contact – a mating pair is formed. To establish the actual DNA transfer, the conjugative plasmid DNA in the donor cell is relaxed at the origin of transfer (*oriT*) by proteins belonging to the DNA transfer and relaxation system, and channeled into the periplasm through the lumen of a complex of mating pair formation (Mpf) proteins. The latter complex spans the cell membrane and envelope and is required to transfer the single-stranded plasmid DNA into the recipient cell (Kalkum *et al*., 2002). In both donor and recipient cell, recircularisation of plasmid DNA occurs, the DNA synthesis of the complementary strand is performed by the host enzymes (Firth *et al*., 1996). In plasmid F, the transfer region (*tra*) that encodes the conjugation machinery is 33.5 kb in size and consists of approximately 40 genes (Frost *et al*., 1994).

In addition to self-transfer, the transfer systems of conjugative plasmids often facilitate the transfer of other DNA sequences that are present in the donor cell. One mechanism involves the independent transfer of non-conjugative, mobilisable plasmids that are co-resident in the donor cell. These plasmids usually harbor an *oriT* and several adjacent genes that encode the proteins needed for nicking at *oriT*, strand separation and mobilisation of DNA (Meyer, 2000). Besides via plasmid mobilisation, DNA sequences can also be transferred after integration of sequences from the bacterial chromosome, transposons, and/or foreign plasmids into the conjugative plasmid (Firth *et al*., 1996). Since conjugative processes enable bacteria to transfer different DNA not only between members of their own kingdom, but also to fungi, plants and even mammalian cells, it can be stated that bacterial conjugation efficiently mediates horizontal gene transfer in a highly promiscuous manner (Kalkum *et al*., 2002). Therefore conjugation is a phenomenon of fundamental evolutionary, ecological, medical (spread of antibiotic resistance determinants) and environmental (spread of pollutant-degradative genes) importance.

2.3. IncP1 plasmids

Plasmids play a major role in the ability of bacterial populations to degrade a wide variety of chemical compounds, and also in their adaptation to novel, man-made (xenobiotic) substances. Rapid dissemination of antibiotic resistance in bacterial populations can also be greatly attributed to the plasmid-mediated horizontal gene transfer. Plasmids capable of being transferred and stably maintained in a wide range of bacteria, the so-called broad-host-range plasmids, are of special interest with respect to interspecies gene exchange. The incompatibility group P plasmids have been studied intensively for many years. Interestingly, among these, IncP2 and IncP9 catabolic plasmids often encode degradation of naturally occurring compounds, whereas genes that encode degradation of xenobiotic compounds (for example 3-chlorobenzoate, 2,4-D, *p*-toluenesulfonate), as well as genes conferring to antibiotic resistance, are often encoded by IncP1 plasmids. These plasmids are the most promiscuous self-transmissible plasmids characterised to date, with a host range that is much wider than that of IncP2 and IncP9 plasmids. IncP1 plasmids are able to transfer between and stably maintain themselves in almost all gram-negative bacterial species (Thomas and Smith, 1987). They are also highly prevalent in nature (Dröge *et al*., 2000; Heuer *et al*., 2002; Top *et al*., 2000). The wide host range enables these plasmids to act as shuttle vectors for locally adapted genes to facilitate rapid local adaptation of both phylogenetically related and distinct populations present in the same bacterial community (Top and Springael, 2003). More DNA sequence information for IncP1 plasmids and a better understanding of their dynamics in response to recent pollution will be necessary to understand their role in bacterial adaptation to xenobiotics.

IncP1 plasmids are divided into three, α, β and γ subgroups (Adamczyk and Jagura-Burdzy, 2003). Ten naturally occurring IncP1 plasmids have been (almost) completely sequenced: RK2, representing a number of indistinguishable Inp1 α antibiotic resistance plasmids (also RP1, RP4, R18, R68) (Pansegrau *et al*., 1994); and members of the phylogenetically distinct IncP1 β group, namely antibiotic resistance plasmids R751 (Thorsted *et al*., 1998), pB2, pB3 (Heuer *et al*., 2004), pB4 (Tauch *et al*., 2003), pB10 (Schlüter *et al*., 2003), and the catabolic plasmids pADP-1 (Martinez *et al*., 2001), pUO1 (Sota *et al*., 2003), pJP4 (Trefault *et al*., 2004) and pTSA (Tralau *et al*., 2001) coding for atrazine, haloacetate, 2,4-D and *p*-toluenesulfonate degradation, respectively. Although only the catabolic region of the 2,4-D degradative plasmid pIJB1 has been sequenced, it has been shown that this plasmid belongs to the IncP1 β subgroup, as well (Poh *et al*., 2002). Plasmid pQKH54 (Hill *et al*., 1992) have been assigned to the γ subgroup of IncP1 plasmids (personal communication with C. M. Thomas), but its nucleotide sequence is not available yet.

Even though IncP1 plasmids are infectiously transmitted with high rates, it has been assumed that horizontal transfer is not sufficient for plasmids to be

maintained as genetic parasites, as they would present the burden to their host. In order to be maintained in bacterial populations, they need to carry some advantageous traits (Bergstrom *et al*., 2000). This hypothesis is supported by the fact that no cryptic IncP1 plasmid (comprised solely of genes for plasmid replication, maintenance and transfer) have been found so far. Instead, all known IncP1 plasmids carry "loads" as "protective arsenal", i.e. large regions with aquired genes encoding various resistance or degradation traits. The latter may influence the host survival by augmenting its fitness, and thus benefit the plasmids indirectly. However, Heuer *et al*. (2004) provided evidence that IncP1 β plasmids without any accessory genes may exist in microbial communities. These plasmids occasionally acquire mobile genetic elements resulting in those plasmids that have been isolated so far based on selectable phenotypic markers. The cryptic plasmids have remained undiscovered because they do not confer the phenotypic traits that researchers are selecting for when screening for the presence of plasmids. Also, it has been proposed that some plasmid backbone genes might confer advantages to the host, for example by promoting biofilm formation (Ghigo, 2001), thus providing sufficient benefits to a bacterial community for them to be maintained over longer evolutionary times.

Comparison of the physical maps and DNA sequences of the sequenced IncP1 plasmids (Fig. 3) show that all these plasmids are very similar and share a common structural organisation (Adamczyk and Jagura-Burdzy, 2003; Heuer *et al*., 2004; Martinez *et al*., 2001; Pansegrau *et al*., 1994; Schlüter *et al*., 2003; Sota *et al*., 2003; Tauch *et al*., 2003; Thorsted *et al*., 1998; Tralau *et al*., 2001; Trefault *et al*., 2004). The regions coding for plasmid replication initiation, stable maintenance in the host cell, and the conjugation machinery constitute the plasmid backbone. The predicted gene products of the IncP1 β backbones are 65 to 100% identical with each other. The respective aa identities between RK2 and R751, the archetype plasmids of IncP1 α and IncP1 β subgroups, are 38 to 87%. Various additional functions such as resistance and catabolic determinants are inserted mainly into two regions in the IncP1 plasmid backbones (Fig. 3) – between the origin of vegetative replication, *oriV*, and the *trfA* gene coding for replication initiation protein; and/or between the *tra* and *trb* operons (Tra1 and Tra2 regions, respectively) coding for conjugative transfer apparatus. Most of the above mentioned plasmids contain their "loads" in both locations, except for pJP4 and pB10 which are interrupted only between *oriV* and *trfA*; and also pB2 and pB3 with interruptions only between Tra1 and Tra2. The only exceptions here are pB4 and RK2 which have also the third insertion in another location – between Tra1 and the region responsible for replication control and stability in case of pB4, and inside the *klcB* gene in case of RK2.

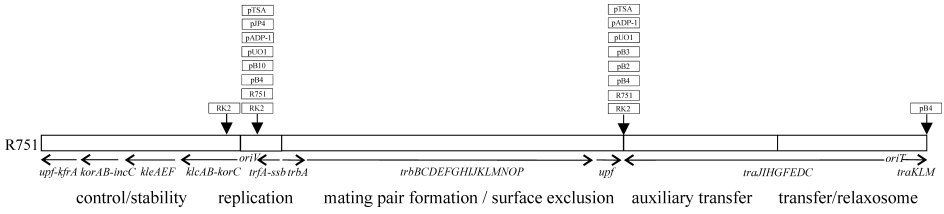

Figure 3. Genetic organization of IncP1 plasmids referred to R751 (NC_001735). Relevant backbone functions and backbone genes are indicated. Arrows show transcriptional units of R751 backbone. The boxes with plasmid names show the locations of "loads" in the respective plasmids.

2.3.1. Replication determinants

IncP1 plasmids replicate by theta mode, described in chapter 2.2.1. The replication of plasmid RK2 in *Escherichia coli* is quite well studied. In order to replicate, IncP1 plasmids require *oriV*, from which replication proceeds unidirectionally (at least in *E. coli*) and the *trfA* gene, whose two in-frame products, the larger TrfA-44 (TrfA1) and the smaller TrfA-33 (TrfA2), are necessary to activate replication from *oriV* (Pansegrau *et al*., 1994). It has been shown that the smaller protein is sufficient for replication in many hosts (Zhong *et al*., 2003).

Analysis of the *oriV* nucleotide sequences from RK2 and R751 revealed conserved features in this region (Pansegrau *et al*., 1994). The minimal DNA segment providing *oriV* activity is approximately 400 bp long and consists of five direct repeats called iterons (17-mers) to which monomeric TrfA binds, DnaA boxes, A+T-rich and G+C-rich regions (Fig. 4). The A+T-rich repeats (13-mers) designated L, M1, M2 and R are the initial sites of helix destabilization. The carboxyl (C)-terminus of TrfA constitutes the dimerization domain. The latter is highly conserved in both RK2 and R751 TrfA proteins and shows homology to the substrates of *E. coli* ClpX chaperone. It has been shown that ClpX is involved in TrfA monomerization (Konieczny and Helinski, 1997b) (Fig. 4). TrfA dimers are believed to play a key role in the negative control of replication ("handcuffing model") (Toukdarian and Helinski, 1998): when present in excess, TrfA monomers bound to the iterons of the replication origins of two plasmids interact with each other and form a sandwich-like structure. This makes the *oriV* inaccessible for other components of the replication complex. Thus, the TrfA proteins and iterons act as positive initiators of replication and negative regulators of plasmid copy number (Adamczyk and Jagura-Burdzy, 2003), which makes such control system very economical.

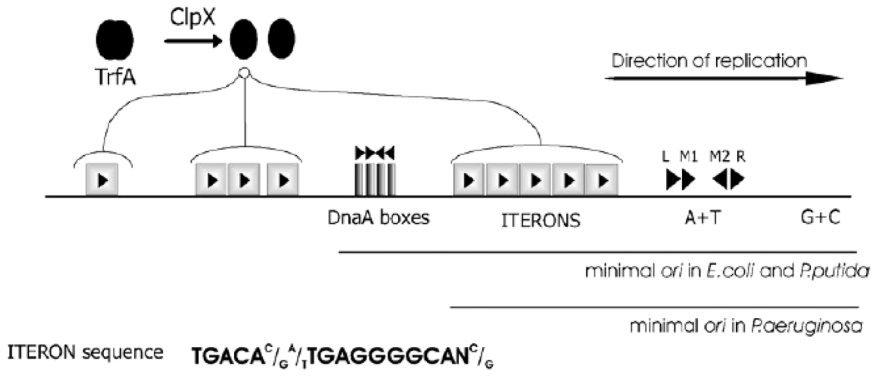

Figure 4. The *oriV* region of plasmid RK2 (Adamczyk and Jagura-Burdzy, 2003). The orientation of iterons, the DnaA boxes and 13-mers in A+T rich region (designated L, M1, M2 and R) are indicated by arrowheads.

The positive action of TrfA as an initiator of replication is modulated by the host's DnaA protein. This factor binds to four DnaA boxes arranged as inverted pairs of a 9 bp consensus sequence separated by few base pairs. These boxes are located immediately upstream of the block of five iterons. In *E. coli*, DnaA is the key replication initiator protein which binds to the chromosomal origin of replication, *oriC*. This binding promotes destabilization of nearby A+T-rich sequences, resulting in unwinding of the DNA double helix and the formation of an open complex (Bramhill and Kornberg, 1988; Krause *et al*., 1997). DnaA protein cannot by itself form an open complex in *oriV*, the binding of DnaA to DnaA boxes rather enhances or stabilizes the formation of the TrfA-mediated open complex (Konieczny *et al*., 1997). DnaA bound to its recognition sites interacts with the DnaB helicase to form a DnaABC complex in which DnaB is responsible for interaction between the TrfA protein and a prepriming form of the DnaBC complex (Konieczny and Helinski, 1997a). DnaC is an adenosine triphosphatase (ATPase) (Lanka and Schuster, 1983). In case of RK2, one molecule of the DnaB helicase is loaded at the A+T-rich region and replication proceeds toward the G+C-rich region (Adamczyk and Jagura-Burdzy, 2003).

The requirements for replication initiation are not universal in different hosts. In *E. coli* and *Pseudomonas putida*, loading and activation of DnaB helicase at the RK2 *oriV* occur with either TrfA-33 or TrfA-44 proteins, while in *Pseudomonas aeruginosa* these functions are dependent on the TrfA-44 protein (Caspi *et al*., 2001). In addition, the DnaB proteins of the two *Pseudomonas* species could be loaded and activated at the RK2 *oriV* in the absence of a DnaC-like accessory ATPase protein. Jiang *et al*. (2003) showed that TrfA-44 is unique among plasmid initiation proteins as it can load and activate the DnaB helicase of *P. putida* and *P. aeruginosa* at the RK2 origin *in vitro* in the absence of the DnaA protein.

2.3.2. Copy number control

IncP1 plasmids exist in bacterial cells at a copy number of 5 to 7 per chromosome. The number of replication initiation events depends on the cellular level of the initiation protein TrfA. The control of copy number is achieved by tight regulation of the expression of *trfA* operon. The *trfA* promoter is a very strong promoter so it may control expression at a sufficient level in a wide spectrum of hosts. This promoter is regulated by at least two plasmid-encoded repressor proteins – KorA and KorB (described in chapter 2.3.5) that act cooperatively and decrease the activity of *trfA* promoter about 1000-fold (Adamczyk and Jagura-Burdzy, 2003; Smith *et al*., 1984).

The level of active TrfA is also regulated by its ability to form dimers, inactive in binding to iterons and initiating replication (Adamczyk and Jagura-Burdzy, 2003). These dimers, as mentioned above, are dissociated into active monomers by the action of the ClpX and ClpB chaperones (Konieczny and Helinski, 1997b; Konieczny and Liberek, 2002). It has also been demonstrated that the "extra" iterons which lie outside of minimal *ori* (Fig. 4) decrease the level of available TrfA molecules (Thomas *et al*., 1984). Additionally, as described in the previous chapter, TrfA acts as a negative regulator of copy number when present in excess.

2.3.3. Conjugative transfer

IncP1 plasmids have a broad host-range with respect to their vegetative replication, originating at *oriV*, as well as with respect to conjugation, originating at *oriT* (Adamczyk and Jagura-Burdzy, 2003). The conjugative replication functions are different from those required for vegetative replication. Conjugative replication proceeds by the rolling circle mode. The range of recipient bacteria is broader than the host range in which the plasmids may replicate. As the process is dependent on factors encoded by the plasmid and expressed in the donor cell, any type of cell can serve as a recipient. It has been demonstrated, that the plasmid RK2 mediates the transfer of DNA from bacteria to yeast (Heinemann and Sprague, 1989) and to mammalian cells (Waters, 2001), although it cannot replicate in those eukaryotic organisms.

All IncP1 plasmids are conjugationally highly promiscuous. As the transfer loci of IncP1 α and IncP1 β plasmids share a common arrangement (Pansegrau *et al*., 1994), all IncP1 plasmids probably rely on similar conjugational transfer mechanisms (Adamczyk and Jagura-Burdzy, 2003). The IncP1 conjugation system is related to the type IV secretion system involved in conjugative processes – the *vir* system of Ti plasmids which transfers T-DNA from *Agrobacterium* species to plant cells, as well as in the transfer of virulence factors from human pathogens such as *Bordetella pertussis*, *Legionella pneumophila*

and *Helicobacter pylori* to mammalian cells (Christie, 2001; Lessl *et al*., 1992). Thus, conjugation can be considered as the DNA rolling-circle replication system linked to the type IV secretion system.

2.3.3.1. Mating pair formation complex

Spread of plasmids among bacteria depends on intimate molecular contacts between the surfaces of two cells. After initial contact has been established, DNA is thought to be transported in a complex with associated proteins through a channel at the mating bridge between the donor and recipient cells. Formation of this channel as well as pilus assembly depends on gene products of the mating pair formation (Mpf) system (Adamczyk and Jagura-Burdzy, 2003; Daugelavicius et al, 1997; Eisenbrandt et al, 2000; Haase *et al*., 1995; Pansegrau *et al*., 1994). Mpf genes *trbBCDEFGHIJKL* are located in the Tra2 region except for *traF*. The latter is located in the Tra1 region. The products of these genes are essential for mating except for TrbK, the surface exclusion protein located in the inner membrane and probably preventing formation of the DNA entry pore in the recipient cell envelope (Giebelhaus *et al*., 1996; Haase *et al*., 1996). The Mpf gene products exhibit typical features of bacterial membrane proteins, such as signal sequences and membrane-spanning regions. This suggests that most, if not all of these proteins are membrane associated (Grahn *et al*., 2000; Lessl *et al*., 1992).

The first gene in the *trb* operon, *trbA*, encodes a transcriptional regulator which controls expression of the *tra* and *trb* operons (see chapter 2.3.5). TrbA is not essential for Mpf functions (Adamczyk and Jagura-Burdzy, 2003).

Six TrbB molecules form a ring-shaped hexameric complex which is loosely associated with the inner membrane (Adamczyk and Jagura-Burdzy, 2003). These proteins possess weak ATPase and kinase activities and they might play a role in a chaperone-like function during the pilus assembly process (Krause *et al*., 2000).

The pilus subunit is encoded by the gene *trbC*. IncP1 pilin is composed of a 78-aa TrbC polypeptide forming a ring structure via an intramolecular head-to-tail peptide bond (Eisenbrandt *et al*., 1999). Two chromosomally encoded proteases and the plasmid-encoded TraF protein process the prepilin, after that the latter protein conducts the intramolecular cyclization reaction of TrbC (Eisenbrandt *et al*., 2000). The resulting pili of RK2 and R751 seen in the electron microscope appear to be very rigid, inflexible structures measuring 10 nm in diameter (Eisenbrandt *et al*., 1999).

Functions of other *trb* gene products are not clear at present.

2.3.3.2. DNA transfer and replication functions

The first steps in the transfer of IncP1 plasmids are relaxosome formation and initiation of rolling circle replication. The proteins encoded by the Tra1 region are involved in DNA transfer and replication (Dtr) functions (Adamczyk and Jagura-Burdzy, 2003).

The Tra1 region of RK2 consists of three operons: the relaxase operon *traJIH*, the primase operon *traGFEDCBA* and the leader operon *traKLM* (Pansegrau *et al*., 1994). The plasmid R751 lacks the *traA* and *traB* genes and shows poor conservation of *traE* and *traD* (Thorsted *et al*., 1998).

The transfer origin (*oriT*) sequence of 250 bp, also located within Tra1, is the only *cis*-acting DNA element essential for transfer (Pansegrau *et al*., 1994). The nick region is highly conserved in *oriT* of RK2, R751, and also in IncI1, IncQ and Ti plasmids (Pansegrau and Lanka, 1996; Pansegrau *et al*., 1994). The *oriT* of both RK2 and R751 lies in the intergenic region of the divergently transcribed *traJ* and *traK* genes. It consists of *traJ* and *traK* promoter sequences located to the one side of the nick sequence (Fig. 5) where TraK binds, and 19 bp inverted repeats recognized by TraJ on the other side of the nick site. A TraI binding site overlaps the nick site (Adamczyk and Jagura-Burdzy, 2003). Binding of TraJ has been proposed as the first step in functional relaxosome assembly (Pansegrau *et al*., 1994). In the second step, the relaxase protein, TraI, interacts with the TraJ-*oriT* DNA complex. An altered DNA conformation is thought to allow TraI to interact with the nick region. TraI access to its target site is also facilitated by TraK, which binds in the divergent promoter region and autoregulates the expression of both operons with the help of other relaxosome proteins (Pansegrau *et al*., 1994; Zatyka et al, 1994). It has been demonstrated that TraK wraps DNA, induces local changes in DNA super-helicity and helps to melt the nick region.

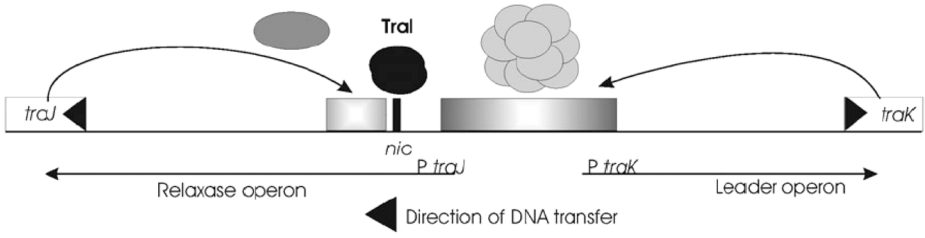

Figure 5. *oriT* of plasmid RK2 (Adamczyk and Jagura-Burdzy, 2003). TraI, TraJ and TraK proteins, involved in nick site (*nic*) recognition as well as their binding regions (represented by boxes) are shown.

TraH has been proposed to have chaperone-like activity, which might stabilize the relaxosomal nucleoprotein complex by specific interactions with both TraI and TraJ. After relaxosome formation has been completed, TraI generates single-strand nicking within the 6 bp *nic* site and becomes covalently bound to the 5' terminal guanosine (G) nucleotide of the nicked strand (Pansegrau and Lanka, 1996). Rolling circle replication may then start and the donor cell is ready to transfer DNA to the recipient cell (Adamczyk and Jagura-Burdzy, 2003).

TraG is known as the "coupling protein" linking the DNA processing and transfer reactions by delivering relaxosome-DNA complex to the Mpf apparatus (Adamczyk and Jagura-Burdzy, 2003). TraG was found to have a tendency to form oligomers and was shown to bind DNA without sequence specifity. Topology analysis revealed that TraG is a transmembrane protein with cytosolic amino (N)- and C-termini and a short periplasmic domain close to the N-terminus. This multimeric inner membrane protein has been predicted to form a pore which binds to the relaxosome via TraG-DNA and TraG-TraI interactions. Although experimental data are controversial, physical interactions between TraG and Mpf complex have been postulated (Carbezón *et al*., 1997; Grahn *et al*., 2000; Hamilton *et al*., 2000; Schröder *et al*., 2002).

No functions have yet been assigned to the products of *traA*, *traB*, *traD* and *traE* which are thought to be nonessential for conjugation (Pansegrau *et al*., 1994). TraE of R751 is predicted to be a DNA topoisomerase on the basis of sequence analysis only (Lin *et al*., 1997; Thorsted *et al*., 1998). Two products of the *traC* gene, TraC1 and TraC2, possess primase activity. During conjugation, TraC1 and Ssb (single-strand DNA binding protein encoded by *trfA* operon) are transferred to the recipient cells as part of the DNA-protein complex (Pansegrau *et al*., 1994).

2.3.4. Stable inheritance

Conjugative transfer is responsible for the spread of IncP1 plasmids across a wide range of bacteria and vegetative replication is the mean to establish these plasmids in new intracellular environments. In order to be stably maintained in the host populations, plasmids need to possess systems providing stable inheritance (Adamczyk and Jagura-Burdzy, 2003).

2.3.4.1. The active partitioning mechanism

Low-copy-number plasmids encode functions for stable maintenance that are very similar to the active partitioning process of bacterial chromosome segregation to daughter cells during cell division (Par system) (Bignell and Thomas,

2001). All Par systems (plasmid and chromosomal) which have been studied so far require three components: the two *trans*-acting factors ParA and ParB, and a *cis*-acting centromere-like site on which a nucleoprotein complex is formed. ParB is a DNA-binding protein that recognizes and binds to the centromere-like site and then interacts with ParA whose ATPase activity is essential to the segregation process. These proteins are almost always encoded by a single operon and are usually transcribed from an autoregulated promoter (Adamczyk and Jagura-Burdzy, 2003).

It has been suggested that a key driver of the partitioning cycle is a conformational change in ParA as it converts from the adenosine triphosphate (ATP)-bound to adenosine diphosphate (ADP)-bound form (Bignell and Thomas, 2001; Davey and Funnell, 1997). As the ParA proteins are known to be membrane bound, this model proposes that the ParA-ATP form promotes cell wall attachment of the partitioning machinery while ATP hydrolysis leads to its dissociation from the membrane (Lin and Mallavia, 1998).

In IncP1 plasmids, the IncC and KorB proteins are equivalent to ParA and ParB. These two proteins are highly conserved between the IncP1 α and IncP1 β subgroups. They are encoded by one *incC-korB* operon (Fig. 6), which is negatively autoregulated by KorB (Adamczyk and Jagura-Burdzy, 2003; Jagura-Burdzy *et al*., 1999a). The plasmids RK2 and R751 produce two forms of IncC – IncC-1 and IncC-2 (Fig. 6). The latter form is sufficient for IncP1 active partitioning (Williams *et al*., 1998) by directly interacting with KorB, the DNA-binding protein which is able to form dimers and multimers. The N-terminally extended form IncC-1 appears to play a regulatory role by enhancing KorB-mediated repression (Jagura-Burdzy *et al*., 1999a). The KorA protein, also encoded by the *incC* locus using a different reading frame (from a start codon located 4 bp upstream of the IncC-1 start) has an autoregulatory function, this protein recognizes the specific O_A operator sequence in *incC-korB* promoter (see chapter 2.3.5).

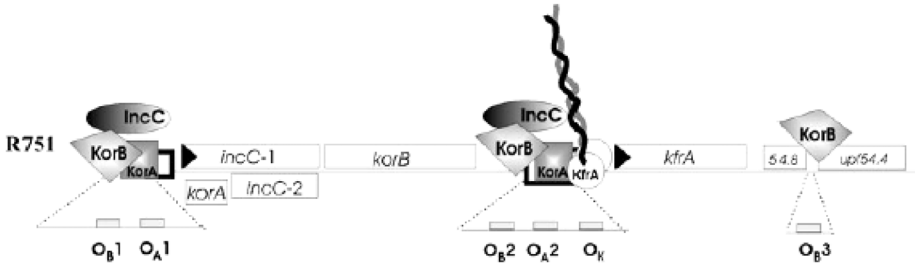

Figure 6. The *par* locus of plasmid R751 (Adamczyk and Jagura-Burdzy, 2003). The proteins KorA, KorB, IncC and KfrA are shown bound with promoters. Promoters are indicated by black open squares and direction of transcription by black arrows. O_A1, O_A2 – KorA binding sites; O_B1, O_B2, O_B3 – KorB binding sites; O_K – KfrA binding site; O_B3 – putative centromeric-like sequence.

The *cis*-acting centromere-like sequence has not yet been precisely identified for IncP1 plasmids. The 13 bp sequences (5'TTTAGCCGCTAAA3') recognized by KorB (O_B) are dispersed throughout the whole plasmid. There are twelwe such sequences (O_B1-O_B12) in RK2 and eleven in R751. O_B3 which is present in both plasmids (upstream of the *upf54.4* gene; Fig. 6) has been suggested to function as a centromere (Williams *et al*., 1998). It is also possible that all sites are equivalent and chosen for this function at random (Adamczyk and Jagura-Burdzy, 2003).

The cellular location of IncP1 plasmids has been studied by immunofluorescence microscopy using anti-KorB antibodies. The KorB foci reflected a symmetrical plasmid distribution pattern – the plasmids are replicated in the center of the cell after which they move to the ¼ and ¾ positions before division (Bignell *et al*., 1999). The mechanism by which plasmids are transferred to the desired locations within the cell is still unknown (Adamczyk and Jagura-Burdzy, 2003).

Studies on R751 have revealed that the *kfrA-upf54.8-upf54.4* region adjacent to the active partitioning operon (Fig. 6) plays an important role in the stability of plasmid (Adamczyk and Jagura-Burdzy, 2003). The *kfrA* gene encodes a protein of 308 aa in case of RK2 and 342 aa in case of R751, both having high alanine content. At the aa level, $KfrA_{RK2}$ and $KfrA_{R751}$ show only 39% identity/52% similarity. Despite that, both proteins have a conserved tertiary structure – an unusually long α-helical tail and globular head domain. The α-helical tails in KfrA dimers and higher order complexes might form coiled-coil structures (Fig. 6) organized in a filamentous network, which has been hypotesized to facilitate the plasmid partitioning process during cell division. Both KfrAs also repress their own transcription, this function is dependent on the N-terminal region (Jagura-Burdzy and Thomas, 1992; Williams and Thomas, 1992). The Upf54.4 aa sequence contains the purine/pyrimidine phosphoribosyl transferase signature (Pansegrau *et al*., 1994), whereas Upf54.8 shows the highest homology (58%) to the anthranilate synthase component II (Adamczyk and Jagura-Burdzy, 2003). Recent studies of R751 revealed that *upf54.8* and *upf54.4* form an operon together with *kfrA* and the genes have been renamed *kfrB* and *kfrC*. KfrA interacts with KfrB whereas KfrC has the ability to interact with KfrB implicating that all three proteins may form a multiprotein complex, important in the stable inheritance of IncP1 plasmids (Adamczyk and Jagura-Burdzy, 2003).

The efficiency of KorB/IncC active partitioning system is reportedly improved by functions encoded by the *klc/kle* region in at least two hosts – *P. aeruginosa* and *E. coli* (Bignell *et al*., 1999; Thorsted *et al*., 1998; Wilson *et al*., 1997). KlcA of both RK2 and R751 shows homology to the ArdB antirestriction proteins of the IncN plasmid pKM101, suggesting a protective role of this protein after conjugation to new hosts (Larsen and Figurski, 1994).

2.3.4.2. Post-segregational killing and multimer resolution systems

IncP1 β plasmids seem to rely on active partitioning as the stability function, while IncP1 α plasmids possess additional stability determinants that are also confusingly termed as *par* region (Gerlitz *et al*., 1990). The locus occurs between the Tra1 and Tra2 regions in RK2 and it encompasses two distinct stability functions: post-segregational killing (*psk*) and multimer resolution system (*mrs*), encoded by two divergently arranged operons, *parDE* and *parCBA*, respectively (Adamczyk and Jagura-Burdzy, 2003).

The presence of *psk* system results in the death of plasmid-free segregants (Roberts and Helinski, 1992). The system ensures that plasmids are maintained in their natural hosts for a number of generations in the absence of selection. The post-segregational killing effect is based on a toxin-antidote principle. The plasmid encodes two components – a stable toxin and an unstable antidote. The latter prevents the lethal action of the toxin by direct interaction (protein *psk*) or by posttranscriptional inhibition of its synthesis by antisense RNA. Both elements of *psk* remain in the cytoplasm of plasmid-free segregants and these cells are quickly deprived of the unstable antidote leaving the active toxin behind (Adamczyk and Jagura-Burdzy, 2003).

The *parDE* operon of RK2 codes for a proteic *psk* system. The homodimer of 12 kilodaltons (kDa) ParE is the toxin which inactivates DNA gyrase (Jiang *et al*., 2002) leading to inhibition of DNA replication and cell filamentation. ParD is a 9 kDa protein which also forms a homodimer and has two functions: inhibition of ParE toxicity and autoregulation of the *parDE* promoter (Roberts *et al*., 1993).

Similar killer gene systems are widely spread among plasmids and are also found in chromosomes of *E. coli* K-12 (Masuda *et al*., 1993), *Vibrio cholerae* (Heidelberg *et al*., 2000) and the plant pathogen *Xylella fastidiosa* (Frohme *et al*., 2000). It is still unclear how the chromosomally encoded *psk* could benefit bacteria but a possible role in programmed cell death under stress conditions has been discussed (Engelberg-Kulka and Glaser, 1999).

The *mrs* system present in almost all plasmids and chromosomes ensures that after replication or recombinational events each copy of the plasmid/sister chromosome will function as a separate unit of inheritance. In IncP1 α plasmids the *parCBA* operon seems to provide the *mrs* functions (Easter *et al*., 1998). ParA is a resolvase which acts on the plasmid multimer resolution site (*res*) to resolve multimers and has been shown to be absolutely essential for the stabilization of mini-RK2. The *res* site is located within a 100 bp region between the divergent *parCBA* and *parDE* promoters. ParB is an endo- and 5'-3' exonuclease (Johnson *et al*., 1999) with low sequence specifity. The function of ParC has not yet been determined. IncP1 β plasmids do not possess the *parABCDE* locus, however, their stability seems to be unaffected. The question of how the lack of this locus can be compensated is still unanswered

(Adamczyk and Jagura-Burdzy, 2003). It has been shown that the resolvase of a transposon acquired by the plasmid can double up for the *mrs* system, and may even be more effective than other systems (Tomalsky *et al*., 2000).

2.3.5. The regulatory network of IncP1 plasmids

In IncP1 plasmids, expression of each operon is controlled by more than one repressor. Otherwise these large plasmids could become an unacceptable metabolic load on its hosts. Such global regulatory network possibly enables broad-host-range plasmids to respond to variable concentrations of repressors in different hosts. In addition, cooperative action of different repressors has a better regulatory effect than a single repressor could have. Overlapping regulatory circuits provide coordination of regulatory response that is tight but sensitive. Besides the local control circuits, at least four global repressors KorA, KorB, KorC and TrbA are involved in the regulatory network of IncP1 plasmids (Adamczyk and Jagura-Burdzy, 2003). As described in section 2.3.4.1, KorB protein is important in partitioning functions. Interestingly, the same protein has been recruited to coordinate the global gene expression as well.

KorA and KorB are encoded by the same central control operon (*cco*), located in a region termed as the central control region (*ccr*) (Adamczyk and Jagura-Burdzy, 2003). KorA (102 aa in RK2) is a homodimer, it binds to O_A operators (5'TTTAGCTAAA3') and it co-operates with KorB (Jagura-Burdzy and Thomas, 1995; Kostelidou *et al*., 1999). There are seven KorA binding sites in RK2 and five in R751 (O_A1 to O_A5; Fig. 7). In RK2 the O_A operators have been divided into two classes – class I (O_A1 *korAp*, O_A3 *trfAp*, O_A7 *klaAp*) and class II (O_A2 *kfrAp*, O_A4 *klcAp*, O_A5 *kleAp*, O_A6 *kleCp*). This classification is based on the affinity of KorA binding which coincides with the localization of O_A within the promoter sequences – class I O_As overlap the –10 sequences whereas the low affinity class II O_As precede the —35 sequence of the regulated promoters. The affinity of binding depends on how perfect the palindromic sequence is and on the nature of flanking sequences. KorA has been demonstrated to bind to and repress all promoters in which O_A operators are present. The genes whose expression is regulated by KorA are involved in vegetative replication and stable maintenance. The strength of these regulated promoters is similar in both RK2 and R751. *korA* promoter, containing O_A1, is the most strongly KorA-repressed promoter in both plasmids (Kostelidou and Thomas, 2002). Through binding to the O_A3 operator, KorA mediates the switch in gene expression between vegetative replication and conjugation transfer that takes place at the *trfAp*/*trbAp* face-to-face divergent promoters. The strong transcriptional activity of *trfAp* inhibits RNA polymerase proceeding from *trbAp*. KorA plays a dominant role in the repression of the *trfAp*, causing elevation of the *trbAp* activity at the same time (Jagura-Burdzy and Thomas, 1994; Jagura-Burdzy and Thomas, 1995).

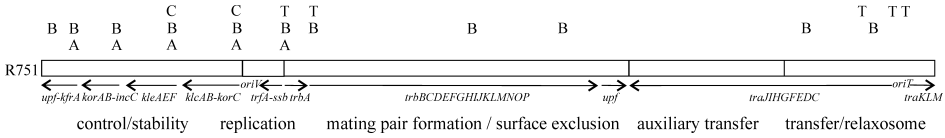

Figure 7. The global regulatory network of plasmid R751 (NC_001735). Relevant backbone functions and backbone genes are indicated. Arrows show transcriptional units of R751 backbone. The letters A, B, C and T stand for the binding sites of KorA (O_A), KorB(O_B), KorC(O_C) and TrbA(O_T) in R751, respectively.

KorB (358 aa in RK2) binds to O_Bs (5'TTTAGCG/CGCTAAA3') with different affinities (Adamczyk and Jagura-Burdzy, 2003). Among global regulators, KorB has the most general function in the IncP1 plasmid biology – it represses operons involved in the vegetative replication and stable maintenance, and also those required for conjugative transfer (Motallebi-Veshareh *et al*., 1992). There are eleven and twelwe highly conserved O_Bs in R751 and RK2, respectively (Fig. 7). The location of these sequences in respect to the different promoters is the basis for the classification of RK2 O_Bs into three classes. Class I sites (O_B1 *korAp*, O_B10 *trfAp*, and O_B12 *klaAp*) lie immediately upstream of the –35 region of each promoter. Operator O_B10 in *trfAp* is bound by KorB with the highest affinity, which implies that it is a key element in controlling replication. Class II sites (O_B2 *kfrAp*, O_B9 *trbBp*, O_B10 *trbAp*, and O_B11 *kleAp*) are located up to 189 bp upstream or downstream of the transcriptional start points. Class III is comprised of six intergenic/intragenic sites (O_B3 to O_B8) located at least 1 kb from the promoters (Balzer *et al*., 1992; Kostelidou and Thomas, 2000; Williams *et al*., 1993). Two RK2 O_Bs, namely O_B6 and O_B12 are not present in R751 while in the latter plasmid, one additional KorB binding site in the *klcA* promoter region exists (Fig. 7). It has been established that KorB represses transcription by binding to the class I and class II operators, but not to class III sites. In the case of class II O_Bs, KorB acts at a distance from where it binds, and formation of KorB dimers/multimers is required for the repressor function (Jagura-Burdzy *et al*., 1999b). KorB repression at a distance from the site of binding is facilitated by co-operation with other repressors KorA and TrbA (Kostelidou *et al*., 1999; Zatyka *et al*., 2001). Additionally, IncC-1 which is encoded by the same operon (*cco*) as KorA and KorB enhances KorB repression at class I and II operators. It also stimulates KorB binding and the formation of higher order complexes at all operators but O_B3 (Jagura-Burdy *et al*., 1999b). The O_B3 site was postulated to play a centromere-like role in the active partitioning process as described in the section 2.3.4.1.

A third global regulator, KorC (85 aa in both RK2 and R751) is encoded by a gene located in the region providing additional stability functions. This repressor acts at three promoters – *klcAp, kleAp* and *kleCp* – of RK2 and at two promoters – *klcAp* and *kleAp* – of R751 (Fig. 7), all located in the same region.

It recognizes the palindromic O_C sequence (5'TAGGGCATAATGCCCTA3'), located at the –10 region of the promoters (Kornacki *et al*., 1990; Thomas *et al*., 1988).

The last global repressor, TrbA (121 aa in RK2), has been shown to repress transcription from promoters of operons involved in conjugative transfer: *traGp, traJp, traKp*, *trbAp* and *trbBp* (Fig. 7) (Adamczyk and Jagura-Burdzy, 2003). The binding site for TrbA, O_T (5'CGATATATCG3') is poorly conserved but is located close to the –10 region of regulated promoters (Bingle and Zatyka, 2003; Zatyka *et al*., 1994). The C-terminal part of TrbA shares similarity with the C-terminal domain of KorA and is probably responsible for interaction with KorB. The co-operativity between TrbA and KorB in *trbAp* regulation takes place despite the significantly long distance (150 bp) between O_B and O_T operators (Zatyka *et al*., 2001; Zatyka *et al*., 1994).

Clustering of the regulatory functions into a single central control operon is unique among plasmids (Adamczyk and Jagura-Burdzy, 2003). Besides *korA, incC* and *korB*, the *cco* of RK2 also contains two additional regulatory genes *korF* and *korG*, not present in R751, which encode putative histone-like proteins. KorF (173 aa) and KorG (175 aa) repress transcription from *kfrAp* and *trfAp*, the two strongest promoters on plasmid RK2 (Jagura-Burdzy *et al*., 1991).

Besides the central control operon *cco*, the *ccr* regions of RK2 and R751 both contain three genes transcribed in the same direction: *kfrA, upf54.8* and *upf54.4*. The possible role of the latter three genes is discussed in chapter 2.3.4.1.

Such multivalent regulatory system – global regulation network supported by autoregulatory circuits – provides unique coordination of all plasmid functions, i.e. vegetative replication, stable maintenance and conjugative transfer, despite that the regions coding for these "survival" functions may be physically separated by (large) mobile genetic elements coding for "additional" functions. It also leads to the fine-tuning of gene expression and minimizes the effects of genetic changes or cell-cycle fluctuations. The flexibility in the replication initiation, multiple stability mechanisms and coordinate regulation of all plasmid backbone functions possibly guarantees the enormous adaptability and stable maintenance of IncP1 plasmids in such a wide spectrum of hosts (Adamczyk and Jagura-Burdzy, 2003).

2.4. The 2,4-D-degradative plasmid pEST4011 isolated in Estonia

Nearly 20 years ago, several 2,4-D metabolizing bacterial strains were isolated from different 2,4-D pretreated soil samples of Estonian agricultural enterprises. Restriction analysis revealed that they all contained identical 2,4-D degradative plasmids (Ausmees and Heinaru, 1990). One strain, D2M4, containing plasmid pD2M4 of about 95 kb, was selected for further investigation as it showed the best growth characteristics on 2,4-D as a sole source of carbon and energy. However, the 2,4-D^+ phenotype of this strain was very unstable – 99% of cells lost this phenotype when grown in LB medium and all studied 2,4-D^- clones did not contain any plasmid. Subsequently, series of selections were performed – in each step D2M4 cells were cultivated in nonselective medium followed by selection on the medium containing 2,4-D as a sole source of carbon and energy. The resulting 2,4-D^+ clone with the best growth on 2,4-D was selected for the next step. Finally, a strain with improved stability of the 2,4-D^+ phenotype was obtained. The selected strain was named EST4002 and it contained a smaller plasmid pEST4002 (estimated size 78 kb) (Ausmees and Heinaru, 1990; V. Kõiv, unpublished data). In order to increase the stability of the plasmid and to overcome the difficulties with DNA extraction, pEST4002 was transferred by conjugation into a plasmid-free recipient strain *P. putida* PaW340. The resulting 2,4-D^+ transconjugants contained a plasmid which was named pEST4011 (on the basis of restriction analysis calculated to be of 70 kb). Subsequently, *P. putida* strain PaW85 was transformed with pEST4011. As a result, the 2,4-D^+ strain EST4021 was obtained.

Due to problems arising from instability of the 2,4-D^+ phenotype and the difficulties with DNA extraction, either the original strain D2M4 containing pD2M4 or the plasmid pEST4002 have not been maintained. As Mäe *et al*. have shown (1993), the plasmid pEST4002 had one additional 7.8 kb EcoRI restriction fragment as compared with pEST4011. However, the exact location and size of the deletion was not determined.

Hybridization experiments with pJP4 fragments containing the *tfd* genes located the pEST4011 catabolic genes for 2,4-D degradation in the 10.5 kb BamHI-B fragment. Subsequently, this fragment was cloned into the BamHI site of the vector plasmid pKT240. *P. putida* strain PaW85, containing the derived plasmid pEST4013 showed inducible expression of *tfdA* (α-ketoglutarate-dependent 2,4-D dioxygenase), *tfdB* (2,4-DCP hydroxylase, 2,4-DCPH) and *tfdC* (3,5-DCC 1,2-dioxygenase, CC12O) when 2,4-D and 2,4-DCP were used as inducers (Mäe *et al*., 1993). As the next step, the 6.1 kb PstI fragment of pEST4013 was subcloned into pKT240 to give pEST4017 (Fig. 1, reference I). The 4.2 kb XhoI-StuI fragment of pEST4017 was sequenced and three ORFs were detected (Kõiv *et al*., 1996). The aa sequences of the predicted products of ORF1 and ORF3 were found to be 62.3 and 61% identical to pJP4

TfdC and TfdB, respectively, and these ORFs were named as *tfdC* and *tfdB*. Three regions of the predicted product of ORF2 aa sequence were found to have 30–60% similarities to dienelactone hydrolase aa sequence encoded by *clcD* from *P. putida* plasmid pAC27 (Neidle *et al*., 1991). The protein products of *tfdC* and *tfdB* genes were analyzed on SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gelelectrophoresis). The occurrence of 32 kDa and 65 kDa proteins, corresponding to TfdC and TfdB, but no protein that could correspond to ORF2, were detected (Kõiv *et al*., 1996).

The aim of the present thesis was to study the expression of *tfdB* and *tfdC*, and the role of TfdR, the LysR-type transcriptional activator, in the regulation of the *tfdCB* operon. The final goal was to determine the complete nucleotide sequence of the plasmid pEST4011, and to describe and analyze all open reading frames and other features found in this plasmid.

3. RESULTS AND DISCUSSION

Previous results have shown that the plasmid pEST4011 enables its host a stable 2,4-D^+ phenotype. At least three genes necessary for 2,4-D degradation, namely *tfdA*, *tfdB* and *tfdC*, are coded by this plasmid. The expression of these genes has been shown to be inducible by 2,4-D and its degradation intermediate, 2,4-DCP. The aim of the present thesis was to study the expression of the previously sequenced 2,4-D catabolic genes *tfdB* and *tfdC*. Finally, in order to determine all genetic units which determine the stable 2,4-D^+ phenotype coded by pEST4011, the complete nucleotide sequence of this plasmid was determined.

3.1. Determination of the role of TfdR in the expression of the *tfdCB* operon (Reference I)

According to the hybridization analyses, Mäe *et al*. (1993) have mapped a regulatory gene very similar to *tfdR* of pJP4 upstream of the pEST4011 catabolic genes *tfdC* and *tfdB*. The first aim of this study was to find out whether this region is responsible for the induced expression of 2,4-DCPH and CC12O, coded by *tfdB* and *tfdC*, respectively. Initially, the 885 bp XhoI fragment (Fig. 1, reference I), adjacent to the 4.2 kb XhoI-StuI fragment described in section 2.4, was sequenced. Indeed, a 915 bp ORF was revealed. This ORF is transcribed divergently from *tfdC* and it predicts a polypeptide of 32.99 kDa. In order to detect the product of the *tfdR* gene on SDS-PAGE, the 1.6 kb PstI-StuI fragment of pEST4017 was cloned under the control of *lac* promoter into the plasmid pBluescript(SK). SDS-PAGE analysis showed that *E. coli* cells containing the derived plasmid pBLR (Fig. 1, reference I) indeed produced a protein of about 33 kDa (Fig. 2, reference I).

The size and aa sequence of TfdR was found to be similar to the members of the catechol-subgroup of the LysR family of activator proteins (Hennikoff *et al*., 1988; Schell, 1993) – TfdR of pJP4 (Matrubutham and Harker, 1994; Leveau and van der Meer, 1996), ClcR of *P. putida* plasmid pAC27 (Coco *et al*., 1994), TcbR of *Pseudomonas* sp. P51 plasmid pP51 (van der Meer *et al*., 1991b), CatM of *Acinetobacter calcoaceticus* (Neidle *et al*., 1989) and CatR of *P. putida* (Rothmel *et al*., 1990). All these regulators activate transcription from homologous downsteam divergent promoters of homologous genes coding for degradation of chlorinated or non-chlorinated catechols. The aa sequence of pEST4011 TfdR was 90% identical to TfdR of pJP4 (over 292 N-terminal aa-s). Thus, the former protein could act as a positive regulator of the downstream *tfdCB* operon. By analogy with other LysR family regulators, the effector mole-

cule for TfdR might be the product of CC12O activity, 2,4-dichloro-*cis,cis*-muconate.

The role of TfdR in the expression of *tfdCB* operon and the effector molecule for TfdR was determined by measuring the specific activities of 2,4-DCPH and CC12O in *P. putida* PaW85 carrying different plasmids (Fig. 1, reference I). The cells were grown 30 min with and without 2,4-DCP and 3,5-DCC as the substrates of 2,4-DCPH and CC12O, respectively. The results are shown in Fig. 3 of reference I. *P. putida* PaW85 harboring plasmids pEST4014 and pEST4017 containing the intact *tfdR* gene and *tfdCB* operon had basal 2,4-DCPH and CC12O activities without 2,4-DCP or 3,5-DCC, and induced 2,4-DCPH and CC12O activities in the presence of these compounds (however, in the case of 3,5-DCC induction occurred to a lesser extent).

In order to determine whether TfdR is responsible for *tfdCB* activation, plasmids without the intact *tfdR* were used: pEST4016 (contains only the intact *tfdC*) (Mäe *et al*., 1993) and pEST4018 (contains only the intact *tfdB*) (Fig. 1, reference I). *P. putida* PaW85 [pEST4018] showed no detectable activities of either 2,4-DCPH or CC12O with or without 2,4-DCP or 3,5-DCC. The cells harboring plasmid pEST4016 had no detectable 2,4-DCPH and only low basal CC12O activities in all growth conditions. These results show that without TfdR no induction of *tfdCB* operon was achieved.

At the determination of the effector molecule for TfdR the *tfdC* or *tfdB* genes were eliminated in order to achieve the situation where the generation of the possible effector molecule would be impossible. For this purpose we used plasmids pEST4015 (contains only the intact *tfdR* and *tfdC*) (Mäe *et al*., 1993) and pEST4017S (contains only the intact *tfdR* and *tfdB*) (Fig. 1, reference I). *P. putida* PaW85 [pEST4015] showed induced CC12O activity only in the presence of 3,5-DCC. *P. putida* PaW85 [pEST4017S] had only low constitutive levels of 2,4-DCPH in all growth conditions. These results strongly suggest that, as in case of other LysR family regulators, the activator molecule for TfdR is the product of CC12O activity, 2,4-dichloro-*cis,cis*-muconate.

3.2. Mapping of the transcription initiation points of *tfdR* and *tfdCB* (Reference I)

It has been shown that the intervening sequences of gene pairs *tfdS-tfdA, tfdT-tfdCDEF, tfdR-tfdD$_{II}$* of pJP4, *clcR-clcABD* of pAC27, *tcbR-tcbCDEF* of pP51 and *catR-catBC* of *P. putida* chromosome have regions with high conservation of nucleotides (Matrubutham and Harker, 1994). These regions are shown or proposed to contain the respective promoter sequences and binding sites for the regulatory proteins in case of TcbR (Leveau *et al*., 1994) and ClcR (Coco *et al*., 1994) with the T-N$_{11}$-A interrupted dyad binding consensus (Goethals *et al*., 1992). Database search done in this study revealed that the same

regions are also conserved in pEST4011 *tfdR-tfdC* (Fig. 5, reference I). This probably indicates similar molecular mechanisms of regulation in the case of pEST4011 and the well studied cases of pP51 and pAC27.

In order to find out whether the above mentioned conserved regions in pEST4011 also contain the promoter sequences of the *tfdR* gene and the *tfdCB* operon, the respective transcription initiation points were determined. Primer extension analysis of mRNA-s of *tfdR* and *tfdCB* demonstrated that identical *tfdR* extension products were detected when RNA was isolated from cultures grown with or without 2,4-DCP and the transcription starting point of *tfdR* mapped to the positions –85 and –84 relative to translation starting point of TfdR (Fig. 4A, reference I). The nucleotide sequence $CAGTCTN_{16}TGGTCT$ was found 6 (7) bp upstream from the transcription initiation site (Fig. 5, reference I). This sequence only very slightly resembles the *E. coli* RNA polymerase $sigma^{70}$-recognized promoter $TTGACAN_{16–18}TATAAT$ (two matches in both –35 and –10 boxes). In contrast, the transcript of *tfdCB* was observed with RNA isolated only from cultures grown on 2,4-DCP. The transcription starting point of *tfdCB* mapped to the position –292 relative to translation starting point of TfdC (Fig. 4B, reference I). The nucleotide sequence $TTAGACN_{18}TAGACT$, found 6 bp upstream from the transcription initiation site of this operon (Fig. 5, reference I), resembles the *E. coli* RNA polymerase $sigma^{70}$-recognized promoter more than the *tfdR* promoter does (two matches in –35 and four matches in –10 box). These results also indicated that transcription of *tfdR* gene is constitutive, while in the case of *tfdCB* operon transcription is induced by growth on 2,4-DCP. Primer extension analyses were also performed to detect a separate transcription starting point in front of *tfdB*, but no cDNA bands were detected in these cases.

Indeed, all these promoter sequences fall into the above mentioned highly conserved regions (Fig. 5, referenec I) supporting the idea that the molecular regulatory mechanisms are similar in the case of pEST4011 TfdR and other members of the catechol-subgroup of the LysR family activator proteins.

3.3. Species identification of strain EST4002 and isolation of pEST4011 from this strain (Reference II)

In 1990, the 2,4-D degradative strain EST4002 was named *Pseudomonas* sp. (Ausmees and Heinaru, 1990). In order to identify this strain more precisely, EST4002 was analyzed at a physiological level by microtiter plate-bound substrate utilization assay (Biolog; Biolog Inc., Hayward, Calif.) and at the phylogenetic level by sequencing and comparison of the nearly complete, PCR-amplified 16S rRNA gene. Surprisingly, the closest relative by Biolog identification appeared to be *Achromobacter xylosoxidans* subsp. *denitrificans*. The comparison of the 16S rRNA gene of strain EST4002 with data bank strains

supported the result of Biolog identification. Accordingly, we named the 2,4-D degradative strain *Achromobacter xylosoxidans* subsp. *denitrificans* EST4002. This discrepancy in strain identification might be due to older methods used more than 15 years ago.

As described in section 2.4, the strain EST4002 contained the plasmid pEST4002 (78 kb) which was transferred by conjugation into a plasmid-free recipient strain PaW340 (Mäe *et al*., 1993). The resulting 2,4-D$^+$ trans-conjugants contained the plasmid pEST4011 (70 kb). In 1999, we found that the strain EST4002 also contained the plasmid which had the same size and restrictional pattern than pEST4011 did. This plasmid was named pEST4011, as well. We found that pEST4011 is very stable in EST4002, growth of EST4002 in LB medium for 24 h resulted only in 1–2% cells with irreversible 2,4-D$^-$ phenotype, which all contained a deletant plasmid of about 30 kb. The latter strain was named EST4003. The corresponding deletant plasmid pEST4012 (from which the whole catabolic region has been deleted) was also achieved by Mäe *et al*. (1993) from 2,4-D$^-$ *P. putida* strain EST4022 when they cultivated *P. putida* EST4021 [pEST4011] cells in LB for 30 generations. The correlated restrictional maps of pEST4011 and pEST4012 are given in Figure 2 of reference II.

3.4. Complete nucleotide sequence and organization of pEST4011 (Reference III)

As the final step of this study, the complete nucleotide sequence of the plasmid pEST4011 was determined. The 2,4-D degradative plasmid pEST4011 is circular, 76958 bp long and has overall G+C content of 62.03%. Thus, pEST4011 is 7 kb larger than we have previously calculated on the basis of restriction analysis. A circular physical map of pEST4011 is shown in Fig. 8. The recognition site positions of 10 different restrictional endonucleases are also shown in this figure. Interestingly, 101 out of 115 of these sites are located in the catabolic transposon of pEST4011. It has been proposed that selection pressure would favor a broad-host-range plasmid with a minimum number of restriction enzyme sites. Regions of the plasmid acquired relatively recently (for example transposons) according to this proposal would contain the usual number of restriction sites (Thomas *et al*., 1980).

The plasmid pEST4011 consists of the 29 kb IncP1 backbone loaded with the 48 kb catabolic transposon containing among other genes the *tfd*-like genes for complete 2,4-D degradation, bordered by identical hybrid insertion elements IS*1071*::*1471*. The nucleotide sequence of this catabolic transposon is 99% identical to Tn*5530* (41 kb) of pIJB1, with the exception of duplication of a 6991 bp region in the case of pEST4011. As the result, the pEST4011 region 27334 to 34324 is identical to the adjacent region from 34325 to 41315 with

only 1 mismatch. In addition, the pEST4011 region 20101 to 31555 is 99% identical to the sequenced region of *Variovorax paradoxus* TV1 2,4-D degradative plasmid pTV1 (Vallaeys, 1992; Vallaeys *et al*., 1998).

According to the NCBI ORF Finder program and BLAST similarity searches, 68 ORFs were identified in the pEST4011 sequence (Fig. 8; Table 1 of reference III), 58 of them were named after the closest relative found in the GenBank. 22 ORFs are orientated in the same direction whereas the rest (46) are orientated in the opposite direction. Equally 34 ORFs are located both in the backbone and in the catabolic transposon of pEST4011 (9 of which are the *tfd* genes for 2,4-D catabolism). 6 ORFs in the catabolic region (including one *tfd* gene) have identical copies due to the duplication of a 7 kb DNA segment. 5 ORFs code for putative transposases, 2 of them are the identical *tnp1471* genes of IS*1071*::*1471*. 10 ORFs encode putative transcriptional regulators. 4 ORFs start with GTG, the rest with ATG translation start codon. Translation stop codons are TGA (42 ORFs), TAA (15 ORFs) and TAG (12 ORFs). A ribosome-binding site with low similarity (2 to 4 matches) to the Shine-Dalgarno consensus sequence (5'-TAAGGAGGT-3') precedes 24 ORFs while in case of 44 ORFs there are 5 to 7 matches to the consensus.

Restrictional map of the deletant plasmid pEST4012 indicates that this plasmid contains only the backbone and one insertion element IS*1071*::*1471*. Thus it is very likely that when grown in nonselective conditions the catabolic region of pEST4011 is deleted as the result of homological recombination between the two direct copies of IS*1071*::*1471*.

3.5. Genes and IS elements of the pEST4011 catabolic transposon (References II and III)

The *tfd* genes in the catabolic region of pEST4011 are organized as following (Fig. 9): a cluster of genes comprising *tfdC*, *tfdE*, *tfdB*, *tfdK* and *tfdA*; *tfdR* is located immediately upstream and is transcribed divergently from this cluster; almost 2 kb downstream of this cluster lies a gene *tfdF*; and finally, the two identical copies of *tfdD* are located 2.6 kb and 9.6 kb upstream of *tfdR*, respectively.

Earlier sequence analyses (Kõiv *et al*., 1996; References I and II) indicated that the ORF coding for *tfdE* in pEST4011 contains a GC dinucleotide deletion which leads to frameshifting. SDS-PAGE analysis also showed that no protein band was translated from this region (Kõiv *et al*., 1996). As all these experiments were performed using subclones of pEST4011 and the complete sequencing of this plasmid was done using the plasmid library, it is reasonable to believe that this GC deletion took place during subcloning procedures. Thus, the ORF between *tfdB* and *tfdC* in pEST4011 should code for an active TfdE protein.

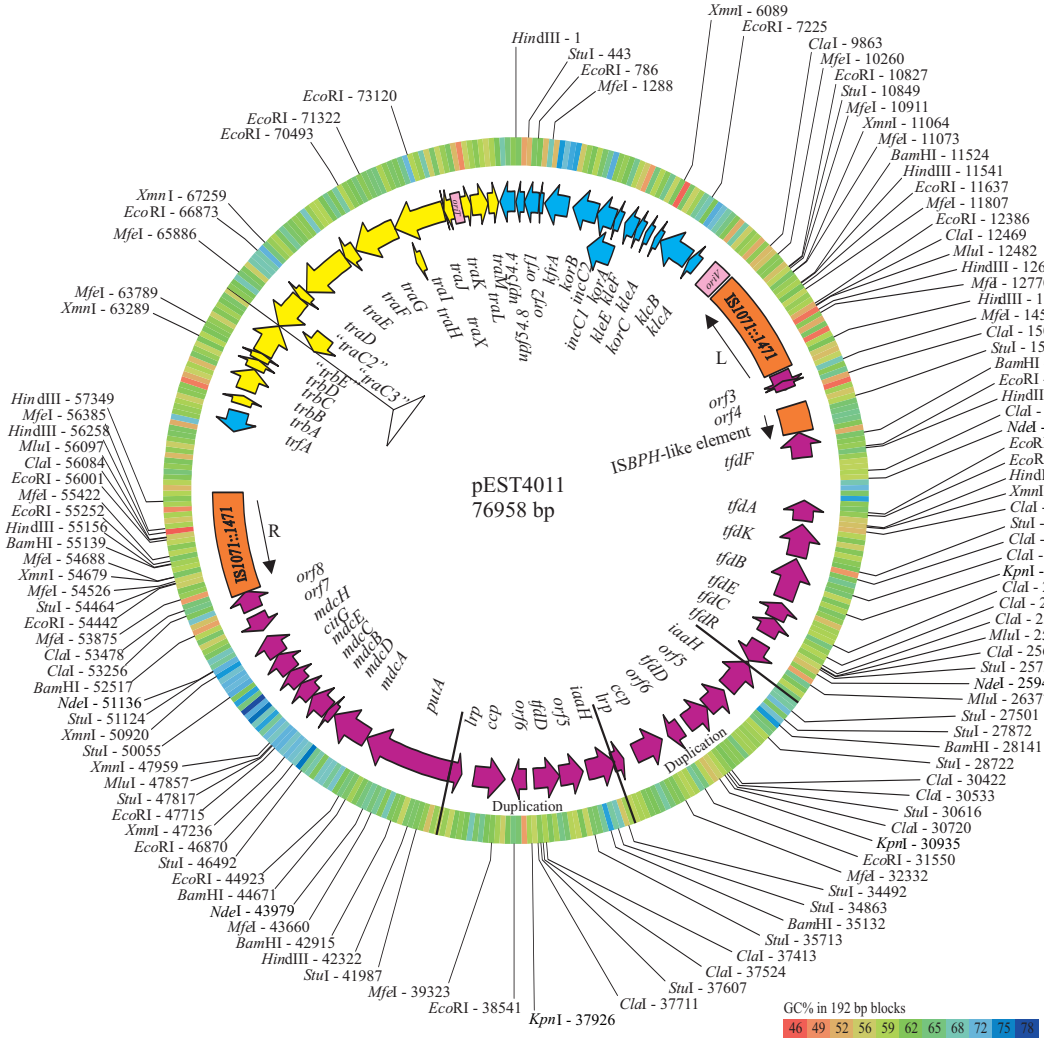

Figure 8. The circular physical map of the 76958 bp plasmid pEST4011. The identified ORFs, the IS elements and the vegetative (*oriV*) and transfer origins (*oriT*) are presented inside the circle. The blue genes are necessary for plasmid autonomous replication initiation, copy number control and stable maintenance in the host cell; the yellow genes are involved in plasmid conjugation; all genes in the pEST4011 catabolic transposon are purple. The 7 kb duplicated regions are shown by thick lines; the site from where the essential transfer genes *trbE* to *trbL* are missing, is shown by thin line and triangle. R stands for the right-hand and L for the left-hand copy of IS*1071*::*1471*. The recognition site positions of 10 restrictional endonucleases, used to verify the assembly of pEST4011, are presented outside the circle. The colored circle indicates G+C percentage in 192 bp blocks starting from the position 1 (the first nucleotide in the respective HindIII recognition site). The products of the genes in quotation marks are truncated and thus likely not functional.

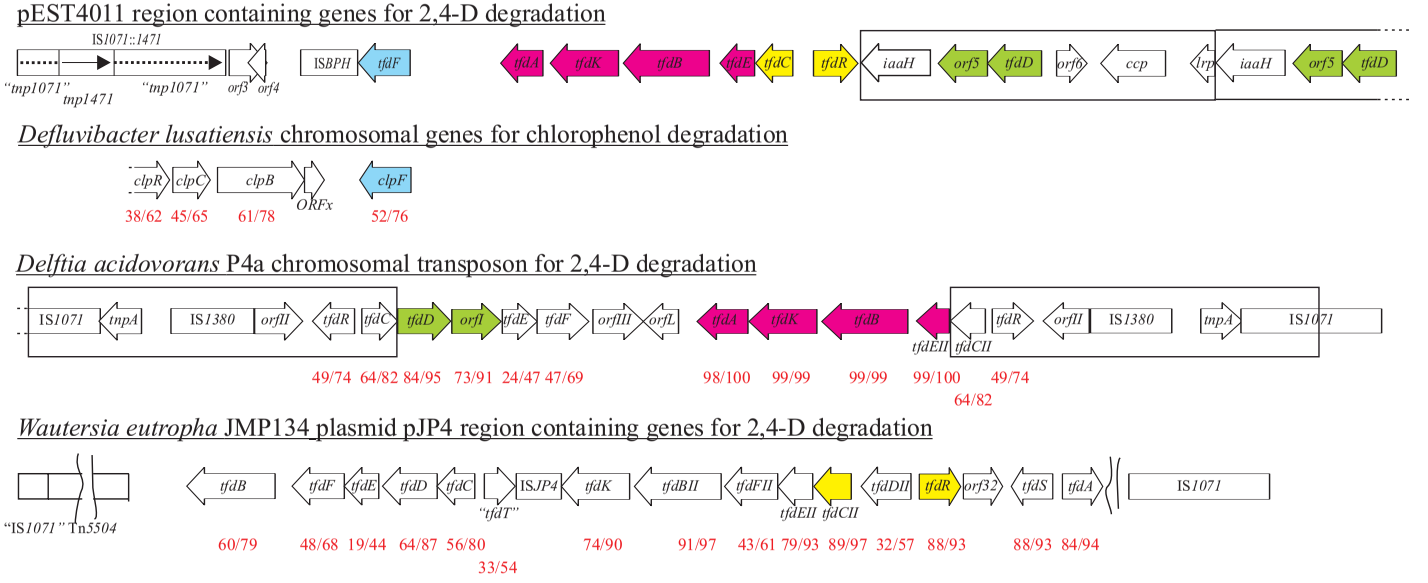

Figure 9. Comparison of the arrangement of the *tfd* genes in plasmid pEST4011 and the homologous genes in *Defluvibacter lusatiensis* (AJ536297), *Delftia acidovorans* P4a (AY078159) and *Cupriavidus necator* (*Wautersia eutropha*) JMP134 plasmid pJP4 (AY365053). A pEST4011 *tfd* gene and its most similar counterpart are shown with the same color. The red numbers below the genes stand for amino acid identity/similarity percentages in comparison with the respective pEST4011 *tfd* analogue. The boxes around the genes indicate duplicated regions.

The genes *tfdE*, *tfdB*, *tfdK* and *tfdA* are almost identical to the same genes of the *D. acidovorans* P4a chromosomal transposon for 2,4-D degradation and are also organized in the same order (Fig. 9). The genes *tfdC* and *tfdR* are most similar to respective genes of the *C. necator* JMP134 plasmid pJP4. The pEST4011 *tfdD* gene together with *orf5*, lying immediately downstream of *tfdD*, is most similar to *D. acidovorans* P4a *tfdD* and *orfI*. The function of the product coded by *orfI* is unknown, but this ORF has been found in many operons determining *ortho*-cleavage pathways of chlorocatechols. The strongly conserved aa sequences encoded by these ORFs suggest that the respective products might play an essential role in functioning of these operons (Ogawa and Miyashita, 1999). The chloromaleylacetate reductase coded by pEST4011 *tfdF* is quite distant from all other known (chloro)maleylacetate reductases having only 52% identical aa-s with the closest relative, the α-proteobacterial ClpF involved in chlorophenol degradation by *Defluvibacter lusatiensis* and *Bradyrhizobium japonicum* USDA 110 MacA (52/77% aa identity/similarity). Seibert *et al*. (2004) have presented dendrogram showing the relatedness of different maleylacetate reductases including TfdF of pEST4011. Their results also demonstrate that the latter protein occupies a separate branch of this phylogenetic tree.

Vallaeys *et al*. (1999) have shown that diverse gene cassettes are independently recruited during assemblage of 2,4-D catabolic pathways of different origins. All *tfd* gene products except for TfdD and TfdF are very similar to pJP4 TfdA (84/94% aa identity/similarity), $TfdB_{II}$ (91/97), $TfdC_{II}$ (89/97), $TfdE_{II}$ (79/93), TfdK (74/90) and TfdR (88/93) (Fig. 9). The ancestor of the pEST4011 *tfdRCEBKA* gene cluster has likely a common origin with the pJP4 $tfdRD_{II}C_{II}E_{II}F_{II}B_{II}K$ cluster. In the case of pEST4011, the regions containing *tfdD* and *tfdF* genes have been deleted. In the chromosomal transposon of *D. acidovorans* P4a there is only a portion of the *tfdRCEBKA* cluster present in pEST4011 (Fig. 9). As *tfdF* and *tfdD* are not pJP4-like genes and they are coded separately from the *tfdRCEBKA* gene cluster, they have probably been individually recruited in the precursor of pEST4011 catabolic transposon to replace the genes lost from the putative ancestor *tfdRDCEFBK(A)* cluster during evolution. As *tfdA* is present in the *tfdRCEBKA* cluster of pEST4011 and in the $tfdE_{II}BKA$ cluster of the chromosomal transposon of *D. acidovorans* P4a, it is reasonable to believe that it was also present in the putative ancestor *tfdRDCEFBK(A)* cluster. During the evolutionary events leading to the formation of pJP4 catabolic region, a model of which is proposed by Trefault *et al*. (2004), *tfdA* was probably lost from *tfdRDCEFBK(A)* cluster and replaced by phylogenetically different *tfdA* gene. McGowan *et al*. (1998) have sequenced internal conserved regions of *tfdA* genes of phylogenetically different 2,4-D degraders. When we aligned the corresponding region of pEST4011 *tfdA* (bases 453–765) to these sequences, we found that this fragment was identical to class III *tfdA* sequences, which by that time had been found only from phylo-

genetically distant, randomly isolated non-2,4-D degraders found in *Comamonas-Rhodoferax* group (β-*Proteobacteria*) and *Halomonas* group (γ-*Proteobacteria*). In these strains the *tfdA* gene may not be even expressed (Hogan *et al*., 1997). The pJP4 *tfdA* gene belongs to class I sequences, which have been found in *Burkholderia, Comamonas-Rhodoferax* and *Alcaligenes-Ralstonia* groups (all β-*Proteobacteria*).

1.2 kb upstream of *tfdD* there is a gene (*ccp*), the predicted product of which is most similar to *Burkholderia fungorum* EriC and *Xanthomonas axonopodis* pv. *citri* str. 306 (NC_00125171) YadQ (41/65% and 37/65% aa identity/similarity, respectively) – the chloride-channeling proteins (Purdy and Wiener, 2000). As chloride ions are liberated during 2,4-D dissimilation, this protein may have a role in removal of these ions from cells.

The proposed functions and aa identities/similarities of other predicted gene products coded by the pEST4011 catabolic transposon are given in Table 1 of reference III. Five ORFs code for (putative) transposases and four ORFs are translated into (putative) transcriptional regulators, functions of two ORFs are unknown. The remaining ten ORFs potentially code for different catabolic functions: indole acetamide hydrolase (two copies of *iaaH*), proline dehydrogenase/pyrroline-5-carboxylate dehydrogenase bifunctional protein (*putA*) and malonate decarboxylase (*mdc* operon of seven genes). As seen in Table 1 of reference III, the closest relatives of these catabolic genes are found in bacteria known as plant pathogens (*X. axonopodis* and *Pseudomonas syringae*) and in a nitrogen-fixing symbiontic bacterium (*B. japonicum*).

Indole acetamide hydrolase (IaaH) performs the second reaction in the two-step synthesis pathway of indole acetic acid (a plant hormone auxin) from tryptophan, found in associative plant growth-promoting rhizobacteria (Steenhoudt and Vanderleyden, 2000). Thus, even if the *iaaH* gene codes for a functional enzyme, the plasmid pEST4011 itself does not enable its host to synthesize auxin.

The PutA bifunctional protein converts L-proline to glutamate for its use as carbon and nitrogen source; in addition to that, it acts as a repressor of its own expression in response to proline supply (Ling *et al*., 1994). We analyzed the strains EST4002 [pEST4011] and EST4003 [pEST4012] by a microtiter plate-bound substrate utilization assay (Biolog). The scatterplot of substrate utilization activities for the two studied strains showed high correlation (R^2=97.8%, P<0.0001), which means that all 95 carbon substrates from Biolog GN2 microplate wells, including L-proline, were utilized by these strains with approximately the same rate (Fig. 4, reference III). Thus, the ability of the strain EST4002 to utilize L-proline as a growth substrate must rely on chromosomal genes.

Decarboxylation of malonate to acetate and CO_2 is a key reaction as it initiates decomposition of this compound in both aerobic and anaerobic bacteria (Dimroth and Hilbi, 1997; Kim, 2002). The predominant portion of malonate in

the environment originates from industrial production, therefore its bacterial degradation has become of great interest. In case of *B. cepacia* both strains 2a [pIJB1] and the mutant strain 2a-1 (contains the plasmid pIJB2 in which the whole catabolic region has been deleted) (Xia *et al*., 1998) were able to grow on malonate as a sole source of carbon and energy. The authors concluded that the ability to grow on malonate must reside on the chromosome, although the plasmid-borne analogues of *mdc* genes could also be active. According to the substrate utilization assay, the strains EST4002 and EST4003 do not use malonic acid as a growth substrate (Fig. 4, reference III). Therefore, as the nucleotide sequences of pEST4011 and pIJB1 *mdc* operons are 100% identical, this gene cluster alone is not adequate for malonate dissimilation in either host. Consequently, it seems that the only detectable catabolic property of EST4002 provided by the plasmid pEST4011 is the ability to degrade 2,4-D. However, it is possible that in other hosts this plasmid contributes to additional properties, for example auxin biosynthesis or malonate dissimilation.

The (putative) regulatory proteins coded by pEST4011, for example Lrp, may alter the expression pattern of both plasmid and chromosomal genes. Genome analyses have revealed that members of the Lrp family of transcriptional regulators are widely distributed among prokaryotes. The archetype Leucine-responsive Regulatory Protein (Lrp) from *E. coli* is a global regulator involved in modulating a variety of metabolic functions, including catabolism and anabolism of amino acids, as well as pili synthesis (Brinkman *et al*., 2003). Lrp has also been shown to be a positive modulator of conjugal transfer of F-like plasmids (Camacho and Casadesus, 2003; Starcic-Erjavec *et al*., 2003). However, the metabolic patterns of the strains EST4002 and EST4003 are not significantly different in case of the carbon substrates available on Biolog GN2 microplates (Fig. 4, reference III).

The 48 kb catabolic transposon of pEST4011 is bordered by identical (2 mismatches) copies of hybrid IS-element IS*1071*::*1471* (positions 9700 to 14019 and 53315 to 57634, respectively) (Fig. 8, Fig. 9). This element consists of 1.1 kb IS*1471* [belongs to IS*630* family by Mahillon and Chandler (1998)] inserted into a 3.2 kb class II transposable element IS*1071* (Xia *et al*., 1996). While IS*1071* flanks a variety of catabolic genes and operons (di Gioia *et al*., 1998), IS*1471* has been detected only in the hybrid IS-elements IS*1071*::*1471* present in pEST4011 and pIJB1. In case of pIJB1, both copies of IS*1071*::*1471* flanking Tn*5530* have been only partially sequenced, but Poh *et al.* (2002) have suggested that these copies are not identical.

1.5 kb upstream of the left-hand copy of IS*1071*::*1471* another IS element is found in pEST4011 (Fig. 8) – an IS*BPH*-like structure whose transposase gene is most similar to the *Achromobacter georgiopolitanum* KKS102 transposase gene *tnpBPH* of insertion sequence IS*BPH* (Table 1, reference III). In *A. georgiopolitanum* KKS102, IS*BPH* is located between *bphS* and *bphE*, the genes involved in biphenyl degradation (Ohtsubo *et al*., 2001).

Thus, only 11 of 34 ORFs (9 *tfd* genes plus two copies of *ccp*) coded by the catabolic transposon of pEST4011 are necessary for 2,4-D degradation. In addition, the presence of two copies of *orf5* could be bound to the same function (Ogawa and Miyashita, 1999). Significance of the other ORFs encoded by the catabolic transposon remains unclear at least in the strain EST4002. However, this transposon may confer different phenotypes to other hosts.

3.6. Comparison of the pEST4011 backbone with other IncP1 backbones (Reference III)

The 29 kb backbone of pEST4011 (positions 1 to 9699 and 57635 to 76958) contains genes for plasmid replication initiation, stable maintenance in host cell as well as conjugation machinery for plasmid transfer into new hosts (Fig. 8). The overall structure of the pEST4011 backbone is similar to other sequenced IncP1 plasmids, described in section 2.3 (Table 2 of reference III). However, the respective products of pEST4011 backbone show only 51–86% identity to either R751 or RK2 backbone gene products. The proposed functions and aa identities/similarities of all predicted pEST4011 backbone gene products are given in Table 1 of reference III.

Comparison of the plasmids pEST4011 and R751 revealed that pEST4011 contains only one catabolic transposon inserted between the core of *oriV* and the *trfA* gene (between the iterons 2 and 3) (Fig. 5, reference III). In the pEST4011 backbone two ORFs between the genes *upf54.8* and *kfrA*, namely *orf1* and *orf2*, were found whose predicted aa sequences are most similar to the *X. fastidiosa* 9a5c cell filamentation protein (probably involved in the regulation of cell division) and to an unknown protein, respectively (Table 1, reference III). These two ORFs are not present in any other IncP1 plasmid. In addition to that, two small genes, *kleB* and *kleG*, devoted to stable inheritance as well as two genes of unknown function, *kluA* and *kluB*, located in the R751 backbone were not identified in pEST4011.

In R751, RK2, pJP4 and pB4 there are two forms of replication initiation protein TrfA coded by genes *trfA1* and *trfA2* – the larger TrfA-44 (382–407 aa-s) and the smaller TrfA-33 (284–289 aa-s), respectively. These proteins are translated from the same reading frame, the translational starts spaced 97–122 aa-s apart. It has been shown that the smaller protein is sufficient for plasmid replication in many hosts (Zhong *et al*., 2003). In pUO1, pTSA and pEMT3 only the gene coding for TrfA-44 (406-411 aa-s) has been annotated, while in pADP-1 and pEST4011 there are only TrfA-33 proteins coded (303 and 289 aa-s, respectively). In pADP-1, the region corresponding to the N-terminal part of TrfA-44 is missing, in pEST4011 the deletion encompasses also the *ssb* gene coding for single-stranded-DNA-binding protein, involved in DNA replication,

recombination and repair and present in all other above-mentioned plasmids just upstream of *trfA*.

The nucleotide sequence of the 2360 bp region of pEST4011 (positions 57635 to 59995) between the catabolic transposon and the iteron 2 of *oriV* is not similar to any sequence in the GenBank.

In pEST4011, a large amount of DNA coding for the C-terminal half of TrbE together with the genes *trbF* to *trbN*, as well as the region coding for C-terminal majority of TraC2 and TraC3 and the whole *traC4* gene, present in all other sequenced IncP1 plasmids, is missing. In the latter plasmids, this region may also contain additional genes, but the important fact is that it includes seven *trb* genes (*trbE* to *trbL*) essential for mating pair formation during conjugation (Adamczyk and Jagura-Burdzy, 2003). Mating experiments using EST4002 [pEST4011] as the donor and *P. putida* PaW340 as the recipient resulted in no 2,4-D^+ transconjugants of PaW340 (data not presented) showing that pEST4011 is, as expected, not able to self-transmit.

As in other IncP1 plasmids, four global repressors are coded in the pEST4011 backbone – KorA, KorB, KorC and TrbA. The consensus binding sequences for these proteins (see section 2.3.5) are also present in pEST4011 except one TrbA binding site present in RK2 *traJ/traK* promoter-operator region (Table 3, reference III). Additionally, one KorB binding site upstream of *traJ* and one KorC binding site upstream of *klcA*, not occurring in R751 or RK2, were found in pEST4011. As seen in Table 3 of reference III, the binding sites of KorA and KorB are well conserved between the three plasmids, while the KorC binding sites of pEST4011 look like hybrids of the respective R751 and RK2 sites. Interestingly, the TrbA operator sequences are more heterogeneous and contain more mismatches from the proposed consensus binding site. In case of RK2, the individual TrbA operators contain one to three mismatches, while the respective sequences of pEST4011 and R751 have up to 6 mismatches. Only one TrbA binding site described in Table 3 of reference III is an exact match to the consensus, namely the one in the promoter region of the R751 *trbB* gene. Bingle and Zatyka (2003) have assumed in their paper that such suboptimal operators play an important role in the balanced regulation of plasmid transfer. Finally, the TraJ binding sites of pEST4011 and R751 are identical. This is in agreement with the fact that the aa sequence of pEST4011 TraJ protein is more similar to the respective sequence of R751 than to RK2, although the level of identity is only 55%.

The pEST4011 origin of replication (*oriV*) consists of the same components as described for other IncP1 plasmids (section 2.3.1) – 10 binding sites (iterons) for TrfA protein – the iterons 2 to 11, plus one additional iteron (iteron 1) not present in R751 or RK2 (Table 3; Fig. 6A of reference III); 4 binding sites for the host protein DnaA, 3 palindromic sequences which are also proposed to bind TrfA, an A/T rich region and a G/C rich region. Table 3 of reference III shows that the conserved TGACA motif of the TrfA iterons of R751 and RK2 is

much more frequently AGACA in pEST4011. Also, the second part of the pEST4011 iterons is more frequently CTTGAG than it is in R751 and RK2. In RK2 this quite strictly alternates between CTTGAG and GATGAG.

The structure of the origin of transfer (*oriT*) is also similar to that of other IncP1 plasmids (section 2.3.3). It consists of a relaxation region containing TraJ and TraI binding sites (Table 3; Fig. 6B of reference III) and of the *nic* site GC (74414-74415 in pEST4011) where one strand of DNA is being nicked during conjugation (Pansegrau *et al*., 1993).

As presented above, despite similarities in their overall organization, the backbone gene products of pEST4011 are homologous but still quite different from the corresponding products of both RK2 and R751, the archetype plasmids of IncP1 α and IncP1 β subgroups, respectively.

3.7. Multiple alignments of *trfA2*, *korA* and *traG* gene products of IncP1 plasmids (Reference III)

In order to examine to which IncP1 subgroup pEST4011 belongs, we performed multiple alignments of available full-length aa sequences of the short form of TrfA (TrfA-33 or TrfA2), KorA and TraG of IncP1 plasmids (Table 2, reference III). In cases where the coding sequences for TrfA-33 were not annotated, we generated the necessary sequences from TrfA-44 by deleting the N-terminal 122 aa-s in case of pUO1 and pTSA, and 123 aa-s in case of pEMT3 (according to similarities with RK2 and R751). The bootstrapped neighbor-joining trees derived from these alignments (using the program CLUSTALX; Fig. 7, reference III) showed that pEST4011 TrfA, KorA and TraG are distinct from all other respective sequences and they occupy individual branches of each tree. Similar results were obtained with different PHYLIP programs used (data not presented). As C. M. Thomas and coworkers have assigned the plasmid pQKH54 (Hill *et al*., 1992) to the γ subgroup (personal communication with C. M. Thomas) and the aa sequence of TrfA of pQKH54 appeared to be distant from pEST4011 TrfA (unpublished data), we allocated pEST4011 to a new, IncP1 δ subgroup. The other four sequenced IncP1 catabolic plasmids – pJP4, pADP-1, pUO1 and pTSA – together with pIJB1 possess the well-conserved IncP1 β backbone. The TrfA2 phylogenetic tree (Fig. 7A, reference III) shows that the 2,4-D degradative plasmid pEMT3 is also distant from all other IncP1 plasmids and could form a separate subgroup, too. Unfortunately, the incompatibility group of the 2,4-D degradative plasmid pTV1 is unknown (personal communication with Tatiana Vallaeys).

3.8. Comparison of pEST4011 and its laboratory ancestor pD2M4 (Reference III)

As mentioned in section 2.4, we were not able to maintain either the original strain D2M4 containing pD2M4 or the plasmid pEST4002. However, the EcoRI and HindIII restriction analyses of the plasmid pD2M4 were performed by V. Kõiv in 1987 (unpublished data). These results made it possible to compare the present stable plasmid pEST4011 with its highly unstable laboratory ancestor pD2M4. As seen in Fig. 8 of reference III, pEST4011 had lost the HindIII restriction fragments of about 23.5 and 20.5 kb; and the EcoRI fragments of about 26, 9.5 and 4.6 kb as compared with pD2M4. In addition, the 19609 bp HindIII and the 6991 and 10872 bp EcoRI fragments of pEST4011 are not present in pD2M4. Instead of the 21849 bp pEST4011 HindIII fragment, pD2M4 has the respective band of about 15 kb.

These results show that first, pEST4011 has lost approximately 25–29 kb DNA from somewhere between the positions 57349 to 66873 (i.e. between the right-hand copy of IS*1071*::*1471* and the *traD* gene) and second, pD2M4 did not contain the 6991 bp duplication present in pEST4011, containing the genes *lrp* to *iaaH* (Fig. 8; Fig. 8C, reference III). Indeed, when pEST4011 was compared with R751, we noticed that in case of pEST4011 about 14 kb of backbone DNA containing among other genes a set of *trb* genes essential for conjugation (*trbE* to *trbL*) (Fig. 5, reference III), was missing. The remaining 11–15 kb could comprise additional DNA inserted between the Tra1 and Tra2 regions as reported for all other IncP1 plasmids discussed in this study except pJP4 and pB10.

The 6991 bp duplication in the pEST4011 catabolic region is not present in either pD2M4 or in the catabolic transposon Tn*5530* of pIJB1. At least two out of six ORFs coded by this region can be directly related to 2,4-D degradation. Namely, *tfdD* is responsible for the fourth step in the respective pathway (Fig. 2) and *ccp* can possibly function as a channeling protein for the removal of chloride, liberated during 2,4-D dissimilation, from cells. The predicted aa sequence of the pEST4011 TfdD protein is far more similar to TcbD of *Pseudomonas* sp. P51 plasmid pP51 (P27099) than to pJP4 $TfdD_I$ or $TfdD_{II}$ – 83/96, 64/87 and 32/57% aa identity/similarity, respectively. The *tcb* genes of pP51 are responsible for 1,2,4-trichlorobenzene degradation by its host strain (van der Meer *et al*., 1991a) and the substrate for TcbD is 2,3,5-trichloro-*cis,cis*-muconate. The respective 2,4-D degradation intermediate and substrate for pEST4011 and pJP4 *tfdD* gene products is 2,4-dichloro-*cis,cis*-muconate. It has been shown by Vollmer *et al.* (1999) that although they could not measure enzyme kinetics with 2,3,5-trichloro-*cis,cis*-muconate as the substrate, 2,4-dichloro-*cis,cis*-muconate was the best substrate for pP51 TcbD and pJP4 $TfdD_I$. However, the respective catalytic efficiency of TcbD was somewhat lower than that of TfdD (k_{cat}/K_m values 53 and 120 μM^{-1} min^{-1}, respectively). Therefore,

it is tempting to speculate that the 7 kb pEST4011 region was duplicated in order to increase the *tfdD* copy number and consequently the amount of the respective enzyme in a cell in order to degrade 2,4-D more efficiently.

Mercury resistance is a common heavy-metal resistance in bacterial isolates from soil and water environments (Kelly and Reanney, 1984). As described in Table 2 of reference III, all catabolic IncP1 plasmids except the least studied pEMT3 bear mercury resistance genes. Although the plasmid pEST4011 does not carry any Hg^r determinants, it is probable that the precursor plasmid pD2M4 did. In pADP-1 and pUO1 the *mer* genes are located between Tra1 and Tra2 regions; if this was the case for pD2M4 as well, the respective genes were lost during the laboratory evolution. However, mercury resistance has not been determined in either strain EST4002 or EST4003.

Thus, on the one hand these rearrangements in plasmid structure somehow ensured its host a stable 2,4-D^+ phenotype, but in the other hand pEST4011 lost the ability to self-transmit.

4. CONCLUSIONS

The results of the present thesis can be summarized as follows:

1. The 76958 bp plasmid pEST4011 ensures its host, *Achromobacter xylosoxidans* subsp. *denitrificans* EST4002 a stable 2,4-D^+ phenotype. This plasmid consists of a novel IncP1 δ backbone and a catabolic transposon harboring *tfd*-like (pJP4-like) genes for complete 2,4-D degradation.
2. The plasmid pEST4011 is a deletion and duplication derivative of pD2M4, the 95 kb highly unstable laboratory ancestor of pEST4011, and was self-generated during laboratory manipulations performed to increase the stability of the 2,4-D^+ phenotype of the original strain D2M4 (pD2M4), isolated in Estonia nearly 20 years ago.
3. The expression of at least two pEST4011 *tfd* catabolic genes, namely *tfdB* and *tfdC* is activated by the product of the *tfdR* gene, transcribed divergently from the *tfdCB* operon. The inductor molecule for this LysR-type regulator TfdR is the 2,4-D degradation intermediate, 2,4-dichloro-*cis,cis*-muconate.

Also, some general conclusions can be drawn:

1. Despite similarities in their overall organization, the backbones of IncP1 plasmids are far more heterogeneous than previously considered. Also, very similar IncP1 backbones, although loaded with different additional functions, can be isolated from geographically distinct regions of the world.
2. Diverse genes and gene cassettes are assembled into different (2,4-D) catabolic pathways, but once a working set of necessary genes has been brought together for example in a composite catabolic transposon, the latter may spread horizontally as one unit and insert itself into different vehicles like conjugative (broad-host-range) plasmids. Other catabolic and/or regulatory genes, coded by these composite transposons may be just entrapped passengers with no significance in one host, while they can most probably confer an altered phenotype to the other host in certain conditions.
3. Serious attention should be paid to the fact that occurrence of laboratory evolution caused by continuous enrichment procedures during the isolation of natural catabolic plasmids may hinder studies and distort our understanding of evolution in nature.

REFERENCES

Adamczyk, M., and G. Jagura-Burdzy. 2003. Spread and survival of promiscuous IncP-1 plasmids. Acta Biochim. Polon. 50:425–453.

Ausmees, N., and A. Heinaru. 1990. New plasmids of herbicide 2,4-dichlorophenoxyacetic acid biodegradation. Genetika 26:770–772 (in Russian).

Balzer, D., G. Ziegelin, W. Pansegrau, V. Kruft, and E. Lanka. 1992. KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative plasmid transfer. Nucleic Acids Res. 20:1851–1858.

Bergstrom, C. T., M. Lipsitch, and B. Levin. 2000. Natural selection, infection transfer and the existence conditions for bacterial plasmids. Genetics 155:1505–1519.

Bignell, C. R., A. S. Haines, D. Khare, and C. M. Thomas. 1999. Effect of growth rate and *incC* mutation on symmetric plasmid distribution by the IncP1 partitioning apparatus. Mol. Microbiol. 34:205–216.

Bignell, C. R., and C. M. Thomas. 2001. The bacterial ParA-ParB partitioning proteins. J. Biotechnol. 91:1–34.

Bingle, L. E. H., and M. Zatyka. 2003. Co-operative interactions control conjugative transfer of broad host-range plasmid RK2: full effect of minor changes in TrbA operator depends on KorB. Mol. Microbiol. 49:1095–1108.

Bramhill, D., and A. Kornberg. 1988. Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. Cell 52:743–755.

Brinkman, A. B., T. J. G. Ettema, W. M. de Vos, and J. van der Oost. 2003. The Lrp family of transcriptional regulators. Mol. Microbiol. 48:287–294.

Camacho, E. M., and J. Casadesus. 2003. Conjugal transfer of the virulence plasmid of *Salmonella enterica* is regulated by the leucine-responsive regulatory protein and DNA adenine methylation. Mol. Microbiol. 44:1589–1598.

Carbezón, E., J. I. Sastre, and F. De la Cruz. 1997. Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. Mol. Gen. Genet. 254:400–406.

Cavalca, L., A. Hartmann, N. Rouard, and G. Soulas. 1999. Diversity of *tfdC* genes: distibution and polymorphism among 2,4-dichlorophenoxyacetic acid degrading soil bacteria. FEMS Microb. Ecol. 29:45–58.

Christie, P. J. 2001. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. 40:294–305.

Coco, W. M., Parsek, M. R., and A. M. Chakrabarty. 1994. Purification of the LysR family regulator, ClcR, and its interaction with the *Pseudomonas putida clcABD* chlorocatechol operon promoter. J. Bacteriol. 176:5530–5533.

Daugelavicius, R., J. K. Bamford, A. M. Grahn, E. Lanka, and D. H. Bamford. 1997. The IncP plasmid-encoded cell envelope-associated DNA transfer complex increases cell permeability. J. Bacteriol. 179:5195–5202.

Davey, M. J., and B. E. Funnell. 1997. Modulation of the P1 plasmid partition protein ParA by ATP, ADP, and P1 ParB. J. Biol. Chem. 272:15286–15292.

del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62:434–464.

del Solar, G., and M. Espinosa. 2000. Plasmid copy number control: an ever-growing story. Mol. Microbiol. 37:492–500.

di Gioia, D., M. Peel, F. Fava, and R. C. Campbell. 1998. Structures of homologous composite transposons carrying *cbaABC* genes from Europe and North America. Appl. Environ. Microbiol. 64:1940–1946.

Dimroth, P., and H. Hilbi. 1997. Enzymic and genetic basis for bacterial growth on malonate. Mol. Microbiol. 25:3–10.

Don, R. H., and J. M. Pemberton. 1981. Properties of six pesticide degradation plasmids isolated from *Alcaligenes paradoxus* and *Alcaligenes eutrophus*. J. Bacteriol. 145:681–686.

Dröge, M., A. Pühler, and W. Selbitschka. 2000. Phenotypic and molecular characterization of conjugative antibiotic resistance plasmids isolated from bacterial communities of activated sludge. Mol. Gen. Genet. 263:471–482.

Easter, C. L., H. Schwab, and D. R. Helinski. 1998. Role of the *parCBA* operon of the broad-host-range plasmid RK2 in stable plasmid maintenance. J. Bacteriol. 180:6023–6330.

Eisenbrandt, R., M. Kalkum, E. M. Lai, R. Lurz, C. I. Kado, and E. Lanka. 1999. Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. J. Biol. Chem. 274:22548–22555.

Eisenbrandt, R., M. Kalkum, R. Lurz, and E. Lanka. 2000. Maturation of IncP pilin precursors resembles the catalytic dyad-like mechanism of leader peptidases. J. Bacteriol. 182:6751–6761.

Engelberg-Kulka, H., and G. Glaser. 1999. Addiction modules and programmed cell death and antideath in bacterial cultures. Annu. Rev. Microbiol. 53:43–70.

Farhana, L., and P. B. New. 1997. The 2,4-dichlorophenol hydroxylase of *Alcaligenes eutrophus* JMP134 is a homotetramer. Can. J. Microbiol. 43:202–205.

Firth, N., K. Ippen-Ihler, and R. A. Skurray. 1996. Structure and function of the F factor and mechanism of conjugation, p. 2377–2401. *In* F. C. Neidhart, *et al*. (editors), *Escherichia coli* and *Salmonella*, Cellular and molecular biology, 2nd edition. ASM Press, Washington, D.C.

Frohme, M., A. A. Camargo, S. Heber, C. Czink, A. J. Simpson, J. D. Hoheisel, and A. P. de Souza. 2000. Mapping analysis of the *Xylella fastidiosa* genome. Nucleic Acids Res. 28:3100–3104.

Frost, L., K. Ippen-Ihler, and R. A. Skurray. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. 58:162–210.

Fukumori, F., and R. P. Hausinger. 1993a. *Alcaligenes eutrophus* JMP134 "2,4-dichlorophenoxyacetate monooxygenase" is an α-ketoglutarate-dependent dioxygenase. J. Bacteriol. 175:2083–2086.

Fukumori, F., and R. P. Hausinger. 1993b. Purification and characterization of 2,4-dichlorophenoxyacetate/α-ketoglutarate dioxygenase. J. Biol. Chem. 268:24311–24317.

Fulthorpe, R. R., C. McGowan, O. V. Maltseva, W. E. Holben, and J. M. Tiedje. 1995. 2,4-dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. 61:3274–3281.

Fulthorpe, R. R., A. N. Rhodes, and J. M. Tiedje. 1996. Pristine soils mineralize 3-chlorobenzoate and 2,4-dichlorophenoxyacetic acid via different microbial populations. Appl. Environ. Microbiol. 62:1159–1166.

Gerdes, K., A. P. Gultyaev, T. Franch, K. Pedersen, and N. D. Mikkelsen. 1997. Antisense RNA-regulated programmed cell death. Annu. Rev. Genet. 31:1–31.

Gerdes, K., J. Møller-Jensen, and R. B. Jensen. 2000. Plasmid and chromosome partitioning: surprises from phylogeny. Mol. Microbiol. 37:455–466.

Gerlitz, M., O. Hrabak, and H. Schwab. 1990. Partitioning of broad-host-range plasmid RP4 is a complex system involving site-specific recombination. J. Bacteriol. 172:6194–6203.

Ghigo, J. M. 2001. Natural conjugative plasmids induce bacterial biofilm formation. Nature 412:442–445.

Giebelhaus, L. A., L. Frost, E. Lanka, E. P. Gormlay, J. E. Davies, and B. Leskiw. 1996. The Tra2 core of the IncP(alpha) plasmid RP4 is required for intergeneric mating between *Escherichia coli* and *Streptomyces lividans*. J. Bacteriol. 178:6378–6381.

Goethals, K., van Montagu, M., and M. Holsters. 1992. Conserved motifs in a divergent *nod* box of *Azorhizobium caulinodans* ORS571 reveal a common structure in promoters regulated by LysR-type proteins. Proc. Natl. Acad. Sci. USA 89:1646–1650.

Grahn, A. M., J. Haase, D. H. Bamford, and E. Lanka. 2000. Components of the RP4 conjugative transfer apparatus form an envelope structure bridging inner and outer membranes of donor cells: implications for related macromolecule transport systems. J. Bacteriol. 182:1564–1574.

Haase, J., R. Lurz, A. M. Grahn, D. H. Bamford, and E. Lanka. 1995. Bacterial conjugation mediated by plasmid RP4: RS1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. J. Bacteriol. 177:4779–4791.

Haase, J., M. Kalkum, and E. Lanka. 1996. TrfK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. J. Bacteriol. 178:6720–6729.

Hamilton, C. M., H. Lee, P. L. Li, D. M. Cook, K. R. Piper, S. B. von Bodman, E. Lanka, W. Ream, and S. K. Farrand. 2000. TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specifity to the conjugal transfer system of pTiC58. J. Bacteriol. 182:1541–1548.

Heidelberg, J. F., J. A. Eisen, W. C. Nelson, R. A. Clayton, M. L. Gwinn, R. J. Dodson, D. H. Haft, E. K. Hickey, J. D. Peterson, L. Umayam, S. R. Gill, K. E. Nelson, T. D. Read, H. Tettelin, D. Richardson, M. D. Ermolaeva, J. Vamathevan, S. Bass, H. Qin, I. Dragoi, P. Sellers, L. McDonald, T. Utterback, R. D. Fleishmann, W. C. Nierman, and O. White. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature 406:477–483.

Heinemann, J. A., and G. F. Sprague Jr. 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. Nature 406:477–483.

Helinski, D. R., A. E Toukdarian, and R. P. Novick. 1996. Replication control and other stable maintenance mechanisms of plasmids, p. 2295–2324. *In* F. C. Neidhart, *et al*. (editors), *Escherichia coli* and *Salmonella*, Cellular and molecular biology, 2nd edition. ASM Press, Washington, D.C.

Henikoff, S., Haughn, G. W., Calvo, J. M., and J. C. Wallace. 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA 85:6602–6606.

Heuer, H., E. Krögerrecklenfort, E. M. H. Wellington, S. Egan, J. D. van Elsas, L. van Overbeek, J.-M. Collard, G. Guillaume, A. D. Karagouni, T. L. Nikolakopoulou, and K. Smalla. 2002. Gentamicin resistance genes in environmental bacteria: prevalence and transfer. FEMS Microbiol. Ecol. 42:289–302.

Heuer, H., R. Szczepanowski, S. Schneiker, A. Pühler, E. M. Top, and A. Schlüter. 2004. The complete sequences of plasmids pB2 and pB3 provide evidence for a recent ancestor of the IncP-1β group without any accessory genes. Microbiology 150:3591–3599.

Hill, K. E., A. J. Weightman, and J. C. Fry. 1992. Isolation and screening of plasmids from the epilithon which mobilize recombinant plasmid pD10. Appl. Environ. Microbiol. 58:1292–1300.

Hiraga, S. 1992. Chromosome and plasmid partition in *Escherichia coli*. Annu. Rev. Biochem. 61:283–306.

Hoffmann, D., S. Kleinsteuber, R. H. Muller, and W. Babel. 2003. A transposon encoding the complete 2,4-dichlorophenoxyacetic acid degradation pathway in the alkalitolerant strain *Delftia acidovorans* P4a. Microbiology 149:2545–2556.

Hogan, D. A., D. H. Buckley, C. H. Nakatsu, T. M. Schmidt, and R. P. Hausinger. 1997. Distribution of the *tfdA* gene in soil bacteria that do not degrade 2,4-dichlorophenoxyacetic acid, 2,4-D. Microb. Ecol. 34:90–96.

Itoh, K., R. Kanda, Y. Momoda, Y. Sumita, Y. Kamagata, K. Suyama, and H. Yamamoto. 2000. Presence of 2,4-D-catabolizing bacteria in a Japanese arable soil that belong to BANA (*Bradyrhizobium-Agromonas-Nitrobacter-Afipia*) cluster in α-*Proteobacteria*. Microb. Environ. 15:113–117.

Itoh, K., R. Kanda, Y. Sumita, H. Kim, Y. Kamagata, K. Suyama, H. Yamamoto, R. P. Hausinger, and J. M. Tiedje. 2002. *tfdA*-like genes in 2,4-dichlorophenoxyacetic acid-degrading bacteria belonging to the *Bradyrhizobium-Agromonas-Nitrobacter-Afipia* cluster in *α-Proteobacteria*. Appl. Environ. Microbiol. 68:3449–3454.

Itoh, K., Y. Tashiro, K. Uobe, Y. Kamagata, K. Suyama, and H. Yamamoto. 2004. Root nodule *Bradyrhizobium* spp. Harbor *tfdAα* and *cadA*, homologous with genes encoding 2,4-dichlorophenoxyacetic acid-degrading proteins. Appl. Environ. Microbiol. 70:2110–2118.

Jagura-Burdzy, G., J. P. Ibbotson, and C. M. Thomas. 1991. The *korF* region of broad-host-range plasmid RK2 encodes two polypeptides with transcriptional repressor activity. J. Bacteriol. 173:826–833.

Jagura-Burdzy, G., K. Kostelidou, J. Pole, D. Khare, A. Jones, D. R. Williams, and C. M. Thomas. 1999a. IncC of broad-host-range plasmid RK2 modulates KorB transcriptional repressor acitivity *in vivo* and operator binding *in vitro*. J. Bacteriol. 181:2807–2815.

Jagura-Burdzy, G., D. P. Macartney, M. Zatyka, L. Cunliffe, D. Cooke, C. Huggins, L. Westblade, F. Kahnim, and C. M. Thomas. 1999b. Repression at a distance by the global regulator KorB of promiscuous IncP plasmids. Mol. Microbiol. 32:519–532.

Jagura-Burdzy, G., and C. M. Thomas. 1992. *kfrA* gene of broad host range plasmid RK2 encodes novel DNA-binding protein. J. Mol. Biol. 225:651–660.

Jagura-Burdzy, G., and C. M. Thomas. 1994. KorA protein of promiscuous plasmid RK2 controls a transcriptional switch between divergent operons for plasmid replication and conjugative transfer. Proc. Natl. Acad. Sci. USA 91:10571–10575.

Jagura-Burdzy, G., and C. M. Thomas. 1995. Purification of KorA protein from broad host range plasmid RK2: definition of a hierarchy of KorA operators. J. Mol. Biol. 253:39–50.

Jensen, R. B., and K. Gerdes. 1995. Programmed cell death in bacteria: proteic plasmid stabilization systems. Mol. Microbiol. 17:205–210.

Jiang, Y., M. Pacek, D. R. Helinski, I. Konieczny, and A. Toukdarian. 2003. A multifunctional plasmid-encoded replication initiation protein both recruits and positions an active helicase at the replication origin. Proc. Natl. Acad. Sci. USA 100:8692–8697.

Jiang, Y., J. Pogliano, D. R. Helinski, and I. Konieczny. 2002. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. Mol. Microbiol. 44:971–979.

Johnson, E. P., A. R. Strom, and D. R. Helinski. 1999. Plasmid RK2 ParB protein: purification and nuclease properties. J. Bacteriol. 181:6010–6018.

Ka, J. O., W. E. Holben, and J. M. Tiedje. 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D treated field soils. Appl. Environ. Microbiol. 60:1106–1115.

Ka, J. O., and J. M. Tiedje. 1994. Integration and excision of a 2,4-dichlorophenozyacetic acid-degradative plasmid in *Alcaligenes paradoxus* and evidence of its natural intergenetic transfer. J. Bacteriol. 176:5284–5289.

Kado, C. I. 1998. Origin and evolution of plasmids. Antonie van Leeuewnhoek 73:117–126.

Kalkum, M., R. Eisenbrandt, R. Lurz, and E. Lanka. 2002. Tying rings for sex. Trends Microbiol. 10:382–387.

Kamagata, Y., R. R. Fulthorpe, K. Tamura, H. Takami, L. J. Forney, and J. M. Tiedje. 1997. Pristine environments harbor a new group of oligotrophic 2,4-dichlorophenoxyacetic acid-degrading bacteria. Appl. Environ. Microbiol. 63:2266–2272.

Kelly, W. J., and D. C. Reanney. 1984. Mercury resistance among soil bacteria: ecology and transferability of genes encoding resistance. Soil Biol. Biochem. 16:1–8.

Kim, Y. S. 2002. Malonate metabolism: biochemistry, molecular biology, physiology and industrial application. J. Biochem. Mol. Biol. 35:443–451.

Kitagawa, W., S. Takami, K. Miyauchi, E. Masai, Y. Kamagata, J. M. Tiedje, and M. Fukuda. 2002. Novel 2,4-dichlorophenoxyacetic acid degradation genes from oligotrophic *Bradyrhizobium* sp. strain HW13 isolated from a pristine environment. J. Bacteriol. 184:209–518.

Kobayashi, I. 2000. Selfishness and death: raison d'être of restriction, recombination and mitochondria. Trends Genet. 14:368–374.

Kõiv, V., R. Marits, and A. Heinaru. 1996. Sequence analysis of the 2,4-dichlorophenol hydroxylase gene *tfdB* and 3,5-dichlorocatechol 1,2-dioxygenase gene *tfdC* of 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011. Gene 174:293–297.

Konieczny, I., K. S. Doran, D. R. Helinski, and A. Blasina. 1997. Role of TrfA and DnaA proteins in origin opening during initiation of DNA replication of the broad host range plasmid RK2. J. Biol. Chem. 272:20173–20178.

Konieczny, I., and D. R. Helinski. 1997a. Helicase delivery and activation by DnaA and TrfA proteins during the initiation of replication of the broad host range plasmid RK2. J. Biol. Chem. 272:33312–33318.

Konieczny, I., and D. R. Helinski. 1997b. The replication initiation protein of the broad-host-range plasmid RK2 is activated by the ClpX chaperone. Proc. Natl. Acad. Sci. USA 94:14378–14382.

Konieczny, I., and K. Liberek. 2002. Cooperative action of *Escherichia coli* ClpB protein and DnaK chaperone in the activation of replication initiation protein. J. Biol. Chem. 277:18483–18488.

Kornacki, J. A., R. S. Burlage, and D. H. Figurski. 1990. The *kil-kor* regulon of broad-host range plasmid RK2: nucleotide sequence, polypeptide product, and expression of regulatory gene *korC*. J. Bacteriol. 172:3040–3050.

Kostelidou, K., A. C. Jones, and C. M. Thomas. 1999. Conserved C-terminal region of global repressor KorA of broad-host-range plasmid RK2 is required for co-operativity between KorA and a second RK2 global regulator, KorB. J. Mol. Biol. 289:211–221.

Kostelidou, K., and C. M. Thomas. 2000. The hierarchy of KorB binding at its 12 binding sites on the broad-host-range plasmid RK2 and modulation of this binding by IncC1 protein. J. Mol. Biol. 295:411–422.

Kostelidou, K., and C. M. Thomas. 2002. DNA recognition by the KorA proteins of Incp-1 plasmids RK2 and R751. Biochim. Biophys. Acta 1576:110–118.

Krause, S., M. Barcena, W. Pansegrau, R. Lurz, J. M. Carazo, and E. Lanka. 2000. Sequence-related protein export NTPases encoded by the conjugative transfer region of RP4 and by the *cag* pathogenicity island of *Helicobacter pylori* share similar hexameric ring structures. Proc. Natl. Acad. Sci. USA 97:3067–3072.

Krause, M., B. Reuckert, R. Lurz, and W. Messer. 1997. Complexes at the replication origin of *Bacillus subtilis* with homologous and heterologous DnaA protein. J. Mol. Biol. 274:365–380.

Laemmli, C. M., J. H. J. Leveau, A. J. B. Zehnder, and J. R. van der Meer. 2000. Characterization of a second *tfd* gene cluster for chlorophenol and chlorocatechol metabolism on plasmid pJP4 in *Ralstonia eutropha* JMP134 (pJP4). J. Bacteriol. 182:4165–4172.

Laemmli, C. M., C. Werlen, and J. R. van der Meer. 2004. Mutation analysis of the different *tfd* genes for degradation of chloroaromatic compounds in *Ralstonia eutropha* JMP134. Arch. Microbiol. 181:112–121.

Lanka, E., and H. Schuster. 1983. The DnaC protein of *Escherichia coli*. Purification, physical properties and interaction with DnaB protein. Nucleic Acids Res. 11:987–997.

Larsen, M. H., and D. H. Figurski. 1994. Structure, expression, and regulation of the *kilC* operon of promiscuous IncP alpha plasmids. J. Bacteriol. 176:5022–5032.

Lessl, M., D. Balzer, W. Pansegrau, and E. Lanka. 1992. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. J. Biol. Chem. 267:20471–20480.

Leveau, J. H. J., de Vos, W. M., and J. R. van der Meer. 1994. Analysis of the binding site of the LysR-type transcriptional activator TcbR on the *tcbR* and *tcbC* divergent promoter sequences. J. Bacteriol. 176:1850–1856.

Leveau, J. H. J., F. König, H. Füchslin, C. Werlen, and J. R. van der Meer. 1999. Dynamics of multigene expression during catabolic adaptation to *Ralstonia eutropha* JMP134 (pJP4) to the herbicide 2,4-dichlorophenoxyacetate. Mol. Microbiol. 33:396–406.

Leveau, J. H. J., and J. R. van der Meer. 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134 (pJP4). J. Bacteriol. 178:6824–6832.

Lin, Z., H. Hiasa, U. Kumar, and R. J. DiGate. 1997. The *traE* gene of plasmid RP4 encodes a homologue of *Escherichia coli* DNA topoisomerase III. J. Biol. Chem. 272:19582–19587.

Lin, Z., and L. P. Mallavia. 1998. Membrane association of active plasmid partitioning protein A in *Escherichia coli*. J. Biol. Chem. 273:11302–11312.

Ling, M., S. W. Allen, and J. M. Wood. 1994. Sequence analysis identifies the proline dehydrogenase and delta 1-pyrroline-5-carboxylate dehydrogenase domains of the multifunctional *Escherichia coli* PutA protein. J. Mol. Biol. 243:950–956.

Liu, T., and P. J. Chapman. 1984. Purification and properties of a plasmid-encoded 2,4-dichlorophenol hydroxylase. FEBS Lett. 173:314–318.

Mäe, A., R. Marits, N. Ausmees, V. Kõiv, and A. Heinaru. 1993. Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. 139:3165–3170.

Mahillon J., and M. Chandler. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62:725–774.

Martinez, B., J. Tomkins, L. P. Wackett, R. Wing, and M. J. Sadowsky. 2001. Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. J. Bacteriol. 183:5684–5697.

Masuda, Y., K. Miyakawa, Y. Nishimura, and Y. Ohtsubo. 1993. *chpA* and *chpB*, *Escherichia coli* chromosomal homologues of the *pem* locus responsible for stable maintenance of plasmid R100. J. Bacteriol. 175:6850–6856.

Matrubutham, U., and A. R. Harker. 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J. Bacteriol. 176:2348–2353.

McGowan, C., R. Fulthorpe, A. Wright, and J. M. Tiedje. 1998. Evidence for interspecies gene transfer in the evolution of 2,4-dichlorophenoxyacetic acid degraders. Appl. Environ. Microbiol. 64:4089–4092.

Meyer, R. 2000. Identification of the *mob* genes of plasmid pSC101 and characterization of a hybrid pSC101-R1162 system for conjugal mobilisation. J. Bacteriol. 182:4875–4881.

Møller-Jensen, J., R. B. Jensen, and K. Gerdes. 2000. Plasmid and chromosomal segregation in prokaryotes. Trends Microbiol. 8:313–320.

Motallebi-Veshareh, M., D. Balzer, E. Lanka, G. Jagura-Burdzy, and C. M. Thomas. 1992. Conjugative transfer functions of broad-host-range plasmid RK2 are coregulated with vegetative replication. Mol. Microbiol. 6:907–920.

Neidle, E. L., C. Hartnett, N. Ornston, A. Bairoch, M. Rekik, and S. Harayama. 1991. Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. J. Bacteriol. 173:5385–5395.

Neidle, E. L., C. Hartnett, and L. N. Ornston. 1989. Characterization of *Acinetobacter calcoaceticus catM*, a repressor gene homologous in sequence to transcriptional activator genes. J. Bacteriol. 171:5410–5421.

Neilson, J. W., K. L. Josephson, I. L. Pepper, R. B. Arnold, G. D. DiGiovanni, and N. A. Sinclair. 1994. Frequency of horizontal gene transfer of a large catabolic plasmid (pJP4) in soil. Appl. Environ. Microbiol. 60:4053–4058.

Nordstrom, K., and S. J. Austin. 1989. Mechanisms that contribute to the stable segregation of plasmids. Annu. Rev. Genet. 23:37–69.

Ogawa, N., and K. Miyashita. 1999.The chlorocatechol-catabolic transposon Tn*5707* of *Alcaligenes eutrophus* NH9, carrying a gene cluster highly homologous to that in the 1,2,4-trichlorobenzene-degrading bacterium *Pseudomonas* sp. strain P51, confers the ability to grow on 3-chlorobenzoate. Appl. Environ. Microbiol. 65:724–731.

Ohtsubo, Y., M. Delawary, K. Kimbara, M. Takagi, A. Ohta, and Y. Nagata. 2001. BphS, a key transcriptional regulator of *bph* genes involved in polychlorinated biphenyl/biphenyl degradation in *Pseudomonas* sp. KKS102. J. Biol. Chem. 276:36146–36154.

Pansegrau, W., and E. Lanka. 1996. Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPα plasmid RP4. J. Biol. Chem. 271:13068–13076.

Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas. 1994. Complete nucleotide sequence of Birmingham IncP*α* plasmids. J. Mol. Biol. 239:623–663.

Pansegrau, W., W. Schröder, and E. Lanka. 1993. Relaxase (TraI) of IncP*α* plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. Proc. Natl. Acad. Sci. USA 90:2925–2929.

Pérez-Pantoja, D., L. Guzman, M. Manzano, D. H. Pieper, and B. Gonzalez. 2000. Role of $tfdC_ID_IE_IF_I$ and $tfdC_{II}D_{II}E_{II}F_{II}$ gene modules in catabolism of 3-chlorobenzoate by *Ralstonia eutropha* JMP134 (pJP4). Appl. Environ. Microbiol. 66:1602–1608.

Pérez-Pantoja, D., T. Ledger, D. H. Pieper, and B. Gonzalez. 2003. Efficient turnover of chlorocatechols is essential for growth of *Ralstonia eutropha* JMP134 (pJP4) in 3-chlorobenzoic acid. J. Bacteriol. 185:1534–1542.

Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin. 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172:2351–2359.

Plumeier, I., D. Perez-Pantoja, S. Heim, B. Gonzalez, and D. H. Pieper. 2002. Importance of different *tfd* genes for degradation of chloroaromatics by *Ralstonia eutropha* JMP134. J. Bacteriol. 184:4054–4064.

Poh, R. P.-C., A. R. W. Smith, and I. J. Bruce. 2002. Complete characterization of Tn*5530* from *Burkholderia cepacia* strain 2a (pIJB1) and studies of 2,4-dichlorophenoxyacetate uptake by the organism. Plasmid 48:1–12.

Purdy, M. D., and M. C. Wiener. 2000. Expression, purification, and initial structural characterization of YadQ, a bacterial homolog of mammalian ClC chloride channel proteins. FEBS Letters 466:26–28.

Roberts, R. C., C. Spangler, and D. R. Helinski. 1993. Characteristics and significance of DNA binding activity of plasmid stabilization protein ParD from the broad host-range plasmid RK2. J. Biol. Chem. 268:27109–27117.

Rothmel, R. K., T. L. Aldrich, J. E. Houghton, W. M. Coco, L. N. Ornston, and A. M. Chakrabarty. 1990. Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive reguator of the *catBC* operon is a member of the LysR family. J. Bacteriol. 172:922–931.

Schell, M. A. 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. 47:597–626.

Schlüter, A., H. Heuer, R. Szczepanowski, L. J. Forney, C. M. Thomas, A. Pühler, and E. M. Top. 2003. The 64508 bp Incp-1β antibiotic multiresistance plasmid pB10 isolated from a wastewater treatment plant provides evidence for recombination between members of different branches of the Incp-1β group. Microbiology 149:3139–3153.

Schröder, G., S. Krause, E. L. Zechner, B. Traxler, H. –J. Yeo, R. Lurz, G. Waksman, and E. Lanka. 2002. TraG-like proteins of DNA transfer systems and of the *Helicobacter pylori* type IV secretion system: inner membrane gate for exported substrates? J. Bacteriol. 184:2767–2779.

Seibert, V., M. Thiel, I.-S. Hinner, and M. Schlömann. 2004. Characterization of a gene cluster encoding the maleylacetate reductase from *Ralstonia eutropha* 335^{T}, an enzyme recruited for growth with 4-fluorobenzoate. Microbiology 150:463–472.

Smith, C. A., V. Shingler, and C. M Thomas. 1984. The *trfA* and *trfB* promoter regions of broad host range plasmid RK2 share common potential regulatory sequences. Nucleic Acids Res. 12:3619–3630.

Sota, M., H. Kawasaki, and M. Tsuda. 2003. Structure of haloacetate-catabolic IncP-1*β* plasmid pUO1 and genetic mobility of its residing haloacetate-catabolic transposon. J. Bacteriol. 185:6741–6745.

Starcic-Erjavec, M., J. P. van Putten, W. Gaastra, B. J. Jordi, M. Grabnar, and D. Zgur-Bertok. 2003. H-NS and Lrp serve as positive modulators of *traJ* expression from the *Escherichia coli* plasmid pRK100. Mol. Genet. Genomics 270:94–102.

Steenhoudt, O., and J. Vanderleyden. 2000. *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. FEMS Microbiol. Rev. 24:487–506.

Suwa, Y., A. D. Wright, F. Fukumori, K. A. Nummy, R. P. Hausinger, W. E. Holben, and L. J. Forney. 1996. Characterization of a chromosomally encoded 2,4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase from *Burkholderia* sp. strain RASC. Appl. Environ. Microbiol. 62:2464–2469.

Tauch, A., A. Schlüter, N. Bischoff, A. Goesmann, F. Meyer, and A. Pühler. 2003. The 79,370-bp conjugative plasmid pB4 consists of an IncP-1*β* backbone loaded with a chromate resistance transposon, the *strA-strB* streptomycin resistance gene pair, the oxacillinase gene bla_{NPS-1}, and a tripartite antibiotic efflux system of the resistance-nodulation-division family. Mol. Gen. Genomics 268:570–584.

Thomas, C. M., M. A. Cross, A. A. Hussain, and C. A. Smith. 1984. Analysis of copy number control elements in the region of the vegetative replication origin of broad host range plasmid RK2. EMBO J. 3:57–63.

Thomas, C. M., J. P. Ibbotson, N. Y. Wang, C. A. Smith, R. Tipping, and N. M. Loader. 1988. Gene regulation on broad host range plasmid RK2: identification of three novel operons whose transcription is repressed by both KorA and KorC. Nucleic Acids Res. 16:5345–5359.

Thomas, C. M., R. Meyer, and D. R. Helinski. 1980. Regions of broad-host-range plasmid RK2 which are essential for replication and maintenance. J. Bacteriol. 141:213–222.

Thomas, C. M., and C. A. Smith. 1987. Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. Annu. Rev. Microbiol. 41:77–101.

Thorsted, P. B., D. P. Macartney, P. Akhtar, A. S. Haines, N. Ali, P. Davidson, T. Stafford, M. J. Pocklington, W. Pansegrau, B. M. Wilkins, E. Lanka, and C. M. Thomas. 1998. Complete sequence of the IncP*β* plasmid R751: implications for evolution and organization of the IncP backbone. J. Mol. Biol. 282:969–990.

Tomalsky, M. E., S. Colloms, G. Blakely, and D. J. Sherratt. 2000. Stability by multimer resolution of pJHCMW1 is due to the Tn*1331* resolvase and not to the *Escherichia coli* Xer system. Microbiology 146:581–589.

Top, E. M., W. E. Holben, and L. J. Forney. 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. 61:1691–1698.

Top, E. M., Y. Moënne-Loccoz, T. Pembroke, and C. M. Thomas. 2000. Phenotypic traits conferred by plasmids, p. 249–286. *In* C. M. Thomas (editor), The horizontal gene pool. Amsterdam: Harwood Academic Publishers.

Top, E. M., and D. Springael. 2003. The role of mobile genetic elements in bacterial adaptation to xenobiotic organic compounds. Current Opinion in Biotech. 14:1–8.

Toukdarian, A. E., and D. R. Helinski. 1998. TrfA dimers play a role in copy-number control of RK2 replication. Gene 223:205–211.

Tralau, T., A. M. Cook, and J. Ruff. 2001. Map of the IncP1*β* plasmid pTSA encoding the widespread genes (*tsa*) for *p*-toluenesulfonate degradation in *Comamonas testosteroni* T-2. Appl. Environ. Microbiol. 67:1508–1516.

Trefault, N., R. De la Iglesia, A. M. Molina, M. Manzano, T. Ledger, D. Pérez-Pantoja, M. A. Sanchez, M. Stuardo, and B. Gonzalez. 2004. Genetic organization of the catabolic plasmid pJP4 from *Ralstonia eutropha* JMP134 (pJP4) reveals mechanisms of adaptation to chloroaromatic pollutants and evolution of specialized chloroaromatic degradation pathways. Environ. Microbiol. 6:655–668.

Vallaeys, T. 1992. Isolement d'une communauté microbienne degradant l'acide 2,4-dichlorophenoxyacetique a partir d'un sol de Dijon. Caracterisation cinetique et genetiques des souches impliques. Ph.D. thesis. University of Lille, France.

Vallaeys, T., L. Albino, G. Soulas, A. D. Wright, and A. J. Weightman. 1998. Isolation and charcterization of a stable 2,4-dichlorophenoxyacetic acid degrading bacterium, *Variovorax paradoxus* using chemostat culture. Biotechnol. Lett. 20:1073–1076.

Vallaeys, T., L. Courde, C. Mc Gowan, A. D. Wright, and R. R. Fulthorpe. 1999. Phylogenetic analyses indicate independent recruitment of diverse gene cassettes

during assemblage of the 2,4-D catabolic pathway. FEMS Microbiol. Ecol. 28:373–382.

Vallaeys, T., R. R. Fulthorpe, A. M. Wright, and G. Soulas. 1996. The metabolic pathway of 2,4-dichlorophenoxyacetic acid degradation involves different families of *tfdA* and *tfdB* genes according to PCR-RFLP analysis. FEMS Microbiol. Ecol. 20:163–172.

van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos. 1991a. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. J. Bacteriol. 173:2425–2434.

van der Meer, J. R., A. C. J. Frijters, J. H. J. Leveau, R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos. 1991b. Characterization of the *Pseudomonas* sp. strain P51 gene *tcbR*, a LysR-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon, and analysis of the regulatory region. J. Bacteriol. 173:3700–3708.

Vollmer, M. D., U. Schell, V. Seibert, S. Lakner, and M. Schlömann. 1999. Substrate specifities of the chloromuconate cycloisomerases from *Pseudomonas* sp. B13, *Ralstonia eutropha* JMP134 and *Pseudomonas* sp. P51. Appl. Microbiol. Biotechnol. 51:598–605.

Waters, V. L. 2001. Conjugation between bacterial and mammalian cells. Nat. Genet. 29:375–376.

Williams, D. R., D. P. Macartney, and C. M. Thomas. 1998. The partitioning activity of the RK2 central control region requires only *incC, korB* and KorB-binding site O(B)3 but other KorB-binding sites form destabilizing complexes in the absence of O(B)3. Microbiology 144:3369–3378.

Williams, D. R., M. Motallebi-Veshareh, and C. M. Thomas. 1993. Multifunctional repressor KorB can block transcription by preventing isomerization of RNA polymerase-promoter complexes. Nucleic Acids Res. 21:1141–1148.

Williams, D. R., and C. M. Thomas. 1992. Active partitioning of bacterial plasmids. J. Gen. Microbiol. 138:1–16.

Wilson, J. W., E. A. Sia, and D. H. Figurski. 1997. The *kilE* locus of promiscuous IncP alpha plasmid RK2 is required for stable maintenance in *Pseudomonas aeruginosa*. J. Bacteriol. 179:2339–2347.

Xia, X. S., S Aathithan, K. Oswiecimska, A. R. Smith, and I. J. Bruce. 1998. A novel plasmid pIJB1 possessing a putative 2,4-dichlorophenoxyacetate degradative transposon Tn*5530* in *Burkholderia cepacia* strain 2a. Plasmid 39:154–159.

Xia, X. S., A. R. Smith, and I. J. Bruce. 1996. Identification and sequencing of a novel insertion sequence, IS*1471*, in *Burkholderia cepacia* strain 2a. FEMS Microbiol. Lett. 144:203–206.

Yarmolinsky, M. B. 2000. A pot-pourri of plasmid paradoxes: effects of a second copy. Mol. Microbiol. 38:1–7.

You, I.-S., and D. Ghosal. 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetate catabolic plasmid pJP4. Mol. Microbiol. 16:321–331.

Zatyka, M., L. Bingle, A. C. Jones, and C. M. Thomas. 2001. Cooperativity between KorB and TrbA repressors of broad-host-range plasmid RK2. J. Bacteriol. 183:1022–1031.

Zatyka, M., G. Jagura-Burdzy, and C. M. Thomas. 1994. Regulation of transfer genes of promiscuous IncP alpha plasmid RK2: repression of Tra1 region transcription both by relaxosome proteins and by the Tra2 regulator TrbA. Microbiology 140:2981–2990.

Zhong, Z., D. Helinski, and A. Toukdarian. 2003. A specific region in the N-terminus of a replication initiation protein of plasmid RK2 is required for recruitment of *Pseudomonas aeruginosa* DnaB helicase to the plasmid origin. J. Biol. Chem. 278:45305–45310.

SUMMARY IN ESTONIAN

2,4-diklorofenoksüatsetaadi lagundamist tagava plasmiidi pEST4011 struktuur

Eestis isoleeritud bakter *Achromobacter xylosoxidans* subsp. *denitrificans* EST4002 on võimeline lagundama herbitsiid 2,4-diklorofenoksüatsetaati (2,4-D). Sellele tüvele tagab stabiilse 2,4-D^+ fenotüübi 76958 aluspaari (bp) suurune plasmiid pEST4011. Varem oli sellest plasmiidist sekveneeritud piirkond, kust leiti kaks geeni, *tfdB* ja *tfdC*, mis kodeerivad vastavalt 2,4-D lagundamisraja kahte ensüümi – 2,4-diklorofenooli hüdroksülaasi ja 3,5-diklorokatehhooli 1,2-dioksügenaasi. Käesolevas töös selgitati välja, et nende kahe geeni ekspressiooni aktiveerib LysR perekonda kuuluv regulaatorvalk, mida kodeeriv geen *tfdR* asub ülalpool *tfdCB* operoni. Induktormolekulina vajab TfdR valk 2,4-D lagundamisraja vaheühendit 2,4-dikloro-*cis*,*cis*-mukonaati.

Töö viimaseks eesmärgiks oli kindlaks teha plasmiidi pEST4011 täielik nukleotiidjärjestus, kirjeldada ja analüüsida kõiki sealt leitud avatud lugemisraame ning muid struktuuriosi.

Plasmiidi pEST4011 umbes 29 kiloaluspaari (kb) suuruses selgroos asuvad geenid kodeerivad plasmiidi replikatsiooniks, stabiilseks säilumiseks peremehes ning konjugatsiooniks vajalikke funktsioone. Siiski on plasmiidist pEST4011 osa konjugatsiooniks hädavajalikke geene puudu ja see plasmiid konjugatsioonivõimeline ei ole. Plasmiidi pEST4011 selgroos kodeeritud geenide produktid on kõige sarnasemad antibiootikumi resistentsusplasmiidide RK2 ja R751 sama funktsiooniga valkudega. Need kaks plasmiidi on vastavalt IncP1 α ja IncP1 β alagrupi arhetüüp-plasmiidid. IncP1 tüüpi plasmiidid on praegu teadaolevalt ühed kõige laiema peremeesteringiga konjugatiivsed plasmiidid, seega suuresti vastutavad neil kodeeritud lisafunktsioonide (kataboolsed või antibiootikumi resistentsusgeenid) kiire ja efektiivse levitamise eest mikroobipopulatsioonides. Peale pEST4011 on praeguseks täielikult või peaaegu täielikult sekveneeritud ainult neli kataboolset IncP1 plasmiidi – pJP4, pUO1, pADP-1 ning pTSA – mis kodeerivad vastavalt 2,4-D, haloatsetaadi, atrasiini ja *p*-tolueensulfonaadi lagundamist ning mille selgrood on peaaegu identsed R751 selgrooga. Plasmiidi pEST4011 selgroos kodeeritud geenide aminohappelised järjestused on aga kõigest 51–86% identsed nii R751 kui RK2 vastavate järjestustega. Seega, hoolimata plasmiidide üldise ülesehituse sarnasusest, on pEST4011 selgroog väga vähe sarnane IncP1 α ja IncP1 β alagruppi kuuluvate plasmiidide selgroogudega ja kuulub uude, IncP1 δ alagruppi. Seda väidet toetab ka antud töös teostatud kolme pEST4011 selgroos kodeeritud geeni, nimelt *trfA2*, *korA* ja *traG*, produkti fülogeneetiline analüüs. Kõik seni sekveneeritud IncP1 plasmiidid kuuluvad kas α või β alagruppi, samuti on C. M. Thomas kaastöötajatega isoleerinud plasmiidi pQKH54, mille nad on

paigutanud γ alagruppi, ent selle plasmiidi nukleotiidjärjestus ei ole veel avalikustatud. Käesolevas töös läbi viidud fülogeneetiline analüüs näitas selgelt, et IncP1 plasmiidid on järjestuse poolest palju heterogeensemad kui seni arvatud.

Plasmiidi pEST4011 47935 bp suurune kataboolne ala moodustab transposooni-taolise struktuuri, mille mõlemas otsas on hübriidne IS-element IS*1071*::IS*1471*. Plasmiidi pEST4011 ja arhetüüp-2,4-D-plasmiidi pJP4 kataboolsed alad sisaldavad mõlemad homoloogilisi *tfd* geene 2,4-D täielikuks lagundamiseks, kuid on muus osas väga erineva järjestusega. Samas on pEST4011 terve kataboolne ala praktiliselt identne ühe teise 2,4-D lagundamist kodeeriva plasmiidi pIJB1 kataboolse transposooniga, välja arvatud umbes 7 kb suurune duplikatsioon pEST4011 puhul. Plasmiidi pIJB1 selgroo nukleotiidjärjestus on täielikult määramata, kuid on näidatud, et see plasmiid kuulub samuti IncP1 β alagruppi. Need tulemused toetavad ka teiste teadlaste poolt tehtud järeldusi, et erineva päritoluga geenid ja geenikassetid assambleeritakse erinevate organismide genoomidesse erinevateks (2,4-D) lagundamisradadeks. Plasmiidide pEST4011 ja pIJB1 põhjal võib aga ka väita, et kui funktsioneeriv lagundamisrada on kord juba kokku pandud näiteks kataboolse transposooni koosseisu, siis see võib horisontaalselt levida ühe üksusena, inserteerudes näiteks erinevatesse konjugatiivsetesse plasmiididesse.

Käesolevas töös on ka võrreldud plasmiidi pEST4011 ja tema ebastabiilse eellase pD2M4 restriktsioonimustreid. Selgus, et plasmiidi pD2M4 2,4-D^+ fenotüübi stabiliseerimiseks laboris läbi viidud protseduuride käigus deleteerus sellest plasmiidist 25–29 kb suurune piirkond, mis sisaldas suure tõenäosusega nii kõiki konjugatsiooniks vajalikke geene kui ka veel mingeid lisafunktsioone, näiteks mõnda kataboolset transposooni. Viimases võisid analoogiliselt teiste kataboolsete IncP1 plasmiididega olla kodeeritud ka elavhõbeda resistentsusgeenid. Samuti puudus pD2M4-l ülalmainitud 7 kb suurune duplikatsioon kataboolses alas. Kuna plasmiid pD2M4 oli ülimalt ebastabiilne, siis ei õnnestunud meie laboris seda säilitada ja pole võimalik kindlaks teha, mida sisaldas see 25–29 kb suurune deleteerunud piirkond. Tõsist tähelepanu tuleks aga pöörata faktile, et suured looduslikud kataboolsed plasmiidid võivad ka laboris evolutsioneeruda rikastamisprotseduuride käigus, mida teostatakse nende plasmiidide looduslikust keskkonnast isoleerimisel. Taoline laboritingimustes toimuv evolutsioon võib tõsiselt moonutada meie arusaamu looduses reaalselt asetleidvatest evolutsiooniprotsessidest.

ACKNOWLEDGEMENTS

I thank my first supervisor, Viia Kõiv, who introduced me to the laboratory work and guided me to the exciting world of science. I would also like to thank my present supervisor, professor Ain Heinaru, for his support and fruitful ideas.

I wish to thank all my previous and present colleagues for help and advice. Special thanks go to Viia Kõiv, Merle Vahter, Eeva Heinaru and Jaak Truu for their contribution to the present study, and to Rita Hõrak and Andres Mäe for critical reading of the manuscript of this thesis.

I gratefully thank professor Christopher Thomas for fruitful discussion and sending me the amino acid sequence of TrfA of pQKH54. I also thank Tatiana Vallaeys for the information concerning the plasmid pTV1.

I thank Estonian Biocentre for providing the generous sequencing service, and also Jaan Lind who turned all my sequencing probes into nice colorful strings of letters A, C, G and T.

Finally, I want to thank my son Kristjan for being such a great fellow and for his plan to become a scientist, too.

PUBLICATIONS

Vedler, E., Kõiv, V. and Heinaru, A. (2000)
TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011. Gene 245:161–168.

Vedler, E., Kõiv, V. and Heinaru, A. (2000)
Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pEST4011 of *Achromobacter xylosoxidans* subsp. *denitrificans* strain EST4002. Gene 255:281–288.

Vedler, E., Vahter, M. and Heinaru, A. (2004)
The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring *tfd* genes for 2,4-dichlorophenoxyacetic acid degradation. Journal of Bacteriology 186:7161–7174.

CURRICULUM VITAE

Eve Vedler

Date of birth	12.06.1971
Nationality	Estonian
Family status	Married, one child
Place of employment	Department of Genetics, Institute of Molecular and Cell Biology, Tartu Univeristy, 23 Riia Street, Tartu 51010, Estonia
Phone	+372-7375014
Fax	+372-7420286
E-mail	eve.vedler@ut.ee

Education

1978–1989	Rakvere Secondary School No 1
1989–1994	Tartu University, Faculty of Biology and Geography, graduated as geneticist
1994–1999	Tartu University, Institute of Molecular and Cell Biology, MSc student
2004–2005	Tartu University, Institute of Molecular and Cell Biology, PhD student

Additional studies

1994	FEBS Advanced Course, Genetics and Cell Engineering (Riga, Latvia)

Professional employment

1999–2000	Tartu University, Centre of Strategic Competence, researcher
2000–present	Tartu University, Institute of Molecular and Cell Biology, researcher

Teaching

Lectures in Molecular Ecology of Microbes (BGMR.03.031) and Practical Courses in Genetics (BGMR.03.036)

Scientific organisations

Member of Estonian Society for Microbiology

CURRICULUM VITAE

Eve Vedler

Sünniaeg	12.06.1971
Kodakondsus	Eesti
Perekonnaseis	Abielus, üks laps
Aadress tööl	Geneetika õppetool, molekulaar- ja rakubioloogia instituut, Tartu Ülikool, Riia 23, Tartu 51010
Telefon	+372-7375014
Faks	+372-7420286
E-mail	eve.vedler@ut.ee

Haridus

1978–1989	Rakvere I Keskkool
1989–1994	Tartu Ülikool, bioloogia-geograafiateaduskond, geneetika eriala
1994–1999	Tartu Ülikool, molekulaar- ja rakubioloogia instituut, magistrant
2004–2005	Tartu Ülikool, molekulaar- ja rakubioloogia instituut, doktorant

Erialane enesetäiendamine

1994	FEBS Advanced Course, Genetics and Cell Engineering (Riia, Läti)

Teenistuskäik

1999–2000	Tartu Ülikool, tehnoloogiakeskus, teadur
Alates 2000	Tartu Ülikool, molekulaar-ja rakubioloogia instituut, teadur

Õppetöö

Leongud molekulaarses mikroobiökoloogias (BGMR.03.031) ja geneetika praktikum (BGMR.03.036)

Teadusorganisatsioonid

Eesti Mikrobioloogide Ühenduse liige

LIST OF PUBLICATIONS

Articles in peer-reviewed journals

1. **Vedler, E.**, Kõiv, V. and Heinaru, A. "TfdR, the LysR-type transcriptional activator, is responsible for the activation of the *tfdCB* operon of *Pseudomonas putida* 2,4-dichlorophenoxyacetic acid degradative plasmid pEST4011". Gene 245 (2000), 161–168. (Current Contents, Medline)
2. **Vedler, E.**, Kõiv, V. and Heinaru, A. "Analysis of the 2,4-dichlorophenoxyacetic acid-degradative plasmid pEST4011 of *Achromobacter xylosoxidans* subsp. *denitrificans* strain EST4002". Gene 255 (2000), 281–288. (Current Contents, Medline)
3. Heinaru, E., Viggor, S., **Vedler, E.**, Truu, J., Merimaa, M. and Heinaru, A. "Reversible accumulation of *p*-hydroxybenzoate and catechol determines the sequential decomposition of phenolic compounds in mixed substrate cultivations in pseudomonads". FEMS Microbiology Ecology 37 (2001), 79–89. (Current Contents, Medline)
4. Truu, J., Kärme, L., Talpsep, E., Heinaru, E., **Vedler, E.,** and Heinaru, A. "Phytoremediation of solid oil shale waste from the chemical industry". Acta Biotechnologica 23 (2003) 2–3, 301–307. (Current Contents, Medline)
5. Truu, J., Talpsep, E., **Vedler, E**., Heinaru, E., and Heinaru, A. "Enhanced biodegradation of oil shale chemical industry solid wastes by phytoremediation and bioaugmentation". Oil Shale 20 (2003) 3, 421–428.
6. **Vedler, E.,** Vahter, M, and Heinaru, A. "The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring *tfd* genes for 2,4-dichlorophenoxyacetic acid degradation". Journal of Bacteriology, 186 (2004), 7161–7174.

Articles in proceedings

1. Heinaru, E., Viggor, S., **Vedler, E.**, Merimaa, M. and Heinaru, A. "Biodegradation potential of three different catabolic types of phenol- and *p*-cresol-degrading bacteria and their efficiency in natural and synthetic wastewater treatment". Proceedings of the Fifth International Symposium on Environmental Biotechnolgy, July 9–13, 2000, Kyoto, Japan, 436–439.
2. Truu, J., Heinaru, E., **Vedler, E.**, Viirmäe, M., Juhanson, J., Talpsep, E., and Heinaru, A. "Remediation of oil shale chemical industry solid wastes using phytoremediation and bioaugmentation". Proceedings of the European Symposium on Environmental Biotechnology, April 25–28, 2004, Oostende, Belgium, 897–899.

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.
102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.