JAGIELLONIAN UNIVERSITY
IN KRAKÓW

Faculty of Biology
Neurobiology

Expression of the *Oprm1* gene in the mouse striatum and prefrontal cortex

Ekspresja genu *Oprm1* w prążkowiu oraz korze przedczołowej myszy

Magdalena Chrószcz

Master thesis prepared under the supervision of

Jan Rodriguez Parkitna, Ph.D.

at the Department of Molecular Neuropharmacology

Maj Institute of Pharmacology, Polish Academy of Sciences

Krakow 2022

Pragnę serdecznie podziękować Promotorowi,
Panu dr hab. Janowi Rodriguezowi Parkitnie,
za życzliwość, niewyczerpaną cierpliwość,
oraz za wszelkie wskazówki udzielone podczas tworzenia tej pracy

Dziękuję Panu mgr Łukaszowi Szumcowi oraz Panu dr Sławomirowi Gołdzie
za wprowadzenie w techniki biologii molekularnej
i pomoc podczas przeprowadzania doświadczeń

Pani mgr Klaudii Misiołek oraz Panu mgr Marcinowi Siwcowi
dziękuję za współpracę przy doświadczeniach z hybrydyzacją in situ

Dziękuję Panu mgr Jackowi Hajto oraz Panu dr Marcinowi Piechocie
z Pracowni Farmakogenomiki Instytutu Farmakologii im. Jerzego Maja PAN
za przeanalizowanie danych z sekwencjonowania Nanopore

Serdeczne podziękowania składam
wszystkim pracownikom Zakładu Neurofarmakologii Molekularnej
za miłą atmosferę pracy, otwartość i dzielenie się doświadczeniem

Table of contents

Abbreviations

bp – base pair

cAMP – 3', 5'-cyclic adenosine monophosphate

CREB – cAMP response element-binding protein

DAMGO – [D-Ala2-MePhe4-Gly5-ol] enkephalin

GDP – guanosine-5'-diphosphase

GIRK – G protein-gated inwardly rectifying potassium channels

GRK – G-protein-coupled receptor kinases

GTP – guanosine-5'-triphosphate

IBNtxA – 3-iodobenzoyl-6β-naltrexamide

IGV – Integrative Genome Viewer

ISH – *in situ* hybridization

LTR – long tandem repeat

MAPK – mitogen-activated protein kinases pathway

MOR-1 – μ opioid receptor

ONS – Oxford Nanopore Sequencing

PKC – protein kinase C

qPCR – quantitative PCR

RT-PCR – reverse transcription PCR

SINE – short interspersed nuclear elements

TM – transmembrane

UTR – untranslated region

VGCC – voltage-gated calcium channels

Abstract

The murine μ opioid receptor is encoded by the *Oprm1* gene located in the qA1 region of chromosome 10. The gene undergoes extensive alternative splicing, with approximately 30 transcript variants and multiple protein isoforms reported so far. Here, I investigated the expression of *Oprm1* transcripts in selected brain regions, specifically the striatum and the prefrontal cortex of adult mice bred on C57BL/6N background. In the experiments, several mRNA detection assays were used, among them, qPCR, RT-PCR, Oxford Nanopore Sequencing (ONS), and RNAscope *in situ* hybridization (ISH). The research confirmed the expression of variant types that contain either exon 4 or exon 7 at their 3' end. Furthermore, RT-PCR revealed the presence of *Oprm1* mRNA containing exon 19, similar to previously reported MOR-1U, MOR-1Eii, and MOR-1Eiv. No evidence for transcription of exon 11 was found. Interestingly, the sequencing data showed the existence of a previously unidentified 10 kbp-long contig downstream of exon 4 – a putative long 3'UTR of *Oprm1*. To confirm that the contig is a part of the μ opioid receptor's mRNA, I have performed a preliminary ISH analysis using the RNAscope method. Initial results confirmed the presence of the putative long *Oprm1* 3' UTR in the striatum and the prefrontal cortex, though due to technical issues transcript variant overlap could not be determined. In future research, I would like to repeat the ISH experiments with protocol modifications, as well as assess the distribution of the putative long *Oprm1*-3'UTR in various cell types.

Key words: *Oprm1*, alternative transcription, 3'UTR, μ-opioid receptor, sequencing

Streszczenie

Receptor opioidowy μ u myszy kodowany jest przez znajdujący się na chromosomie 10 gen *Oprm1*. Gen ulega silnej alternatywnej transkrypcji, w wyniku której powstaje ponad 30 wariantów mRNA, co skutkuje istnieniem kilkunastu form białkowych receptora. Celem pracy magisterskiej było określenie występowania wybranych wariantów transkrypcyjnych w prążkowiu oraz korze przedczołowej dorosłych myszy hodowanych na tle genetycznym C57BL/6N. W pracy wykorzystano szereg technik badania mRNA, w tym: reakcje RT-PCR, qPCR, sekwencjonowanie Oxford Nanopore oraz hybrydyzację *in situ* metodą RNAscope. Uzyskane dane pozwoliły na potwierdzenie obecności dwóch rodzajów transkryptów genu *Oprm1*: posiadających egzon 4 oraz posiadających egzon 7. Wyniki reakcji RT-PCR umożliwiły zaobserwowanie transkryptów posiadających egzon 19, podobnych do wcześniej zidentyfikowanych form MOR-1U, MOR-1Eii i MOR-1Eiv. Nie zaobserwowano ekspresji wariantów posiadających egzon 11 na końcu 5'. Co ciekawe, wyniki sekwencjonowania Oxford Nanopore wykazały występowanie rozciągającego się na długości około 10 tyś. par zasad regionu o wysokim pokryciu odczytami. Region zaobserwowano poniżej egzonu 4, co sugeruje, że może on stanowić obszar 3'UTRu dla wariantów transkrypcyjnych zawierających ten egzon. W celu potwierdzenia występowania długiego transkryptu genu *Oprm1*, wykonano hybrydyzację *in situ* metodą RNAscope. Wstępne wyniki potwierdziły obecność długiego 3'UTRu w prążkowiu i korze przedczołowej myszy, jednak z powodów technicznych nie pozwoliły na określenie kolokalizacji z standardową sondą dla genu *Oprm1*. Celem kolejnych doświadczeń będzie kontynuacja doświadczeń z użyciem metody RNAscope i dostosowanie jej protokołu, w celu określenia kolokalizacji mRNA dla *Oprm1* oraz wskazania typów komórek, w których obecne są długie formy transkrypcyjne.

Słowa kluczowe: *Oprm1*, alternatywna transkrypcja, 3'UTR, receptor opioidowy μ, sekwencjonowanie

1. Introduction

Murine μ opioid receptor is encoded by the *Oprm1* gene. The receptor is the main target of opioid analgesic drugs, including morphine, oxycodone, and fentanyl. The receptor is best known for its involvement in mediating analgesia, but it also plays a significant role in modulating affective states, reward, motivation, and stress response (Emery & Akil, 2020). Activation of the receptor is associated with numerous adverse effects of opioid analgesics, such as tolerance, physical dependence, sedation, respiratory depression, miosis, and constipation, that limit clinical applications of opioids. Despite extensive research, many gaps remain in our understanding of the μ opioid receptor activity, such as the molecular mechanisms behind adverse effects, unexpected differences in pharmacological properties between various agonists, or even the existence of multiple receptor subtypes (Pasternak & Pan, 2013). Due to the well-known clinical and physiological importance of the receptor, it remains necessary to expand our knowledge and try to fill the remaining gaps. In this thesis, I focus on the variability that arises from multiple variants of the μ opioid receptor transcripts.

1.1. Endogenous opioid system

The existence of common receptor binding sites for opioid analgesics was first postulated in 1954 by Beckett and Casy, based on the research on structural requirements for the activity of examined antinociceptive opiates (Beckett & Casy, 1954). In the 1960s the existence of multiple opioid receptors or different modes of interaction with ligands was proposed (Portoghese, 1965). Within a decade, the first specific opiate binding sites were identified in the mouse brain and guinea pig intestine using radioligand assays (Pert & Snyder, 1973). And the following studies allowed for experimental distinction of three putative types of opioid receptors, which were named after the pharmacological profiles of the compounds used in the studies or after their anatomical location: μ (after *m*orphine), δ (after vas *d*eferens), and κ (after *k*etocyclazocine) (Lord et al., 1977; Martin et al., 1976). This was followed by studies investigating potentially different subtypes of opioid receptors, which culminated in the identification of cDNAs of three distinct receptors (Chen et al., 1993; Evans et al., 1992; Meng et al., 1993). These discoveries paved the way for the generation of mice with 'knockout genes', providing conclusive evidence for the hypothesis of single genes encoding each type of opioid receptors and expanded our understanding of the functions of each receptor.

Opioid receptors are members of the metabotropic, rhodopsin-like G-protein coupled receptor (GPCR) superfamily. In the mouse, they are encoded by separate genes: *Oprm1, Oprd1, Oprk1* for μ, δ, and κ receptors, respectively. These genes, along with *Oprl1,* which encodes for the nociceptin/orphanin FQ receptor, are paralogous – they originate from a single ancestral gene that underwent quadruplication about 500 million years ago during tetraploidisation of the genome (Larhammar et al., 2015). The opioid receptors possess an extracellular N-terminus, seven transmembrane domains, 3 extracellular and 3 intracellular loops, and an intracellular C-terminus. The transmembrane regions are highly conserved, with 70-90% amino acid identity among the four receptors in the sequences encoding the intracellular loops and transmembrane domains, but only 20% identity in N-, C- terminal regions and extracellular loops (Law et al., 2000).

The opioid receptors are activated by endogenous ligands: enkephalins, dynorphins, neoendorphins, and β-endorphin. The peptides are derived from three precursors, preproenkephalin, preprodynorphin, and proopiomelanocortin, encoded in the mouse genome by single genes: *Penk, Pdyn,* and *Pomc,* respectively (Akil et al., 1984). To produce active opioid peptides, the protein precursors undergo extensive enzymatic cleavage and modifications. All the ligands share a common N-terminal signature sequence that consists of 4 amino acids, YGGF, and is critical for receptor binding. The endogenous ligands are characterized by limited selectivity towards their receptors, with short pentapeptides: Met- and Leu-enkephalins binding preferentially to δ opioid receptor, other members of the enkephalin family towards μ, but also κ and δ receptors, dynorphins binding mainly to κ opioid receptor, and β-endorphin displaying a similar affinity towards μ and δ opioid receptor (Akil et al., 1984; Mansour et al., 1995). The affinities for specific receptors are graphically summarised in Figure 1.

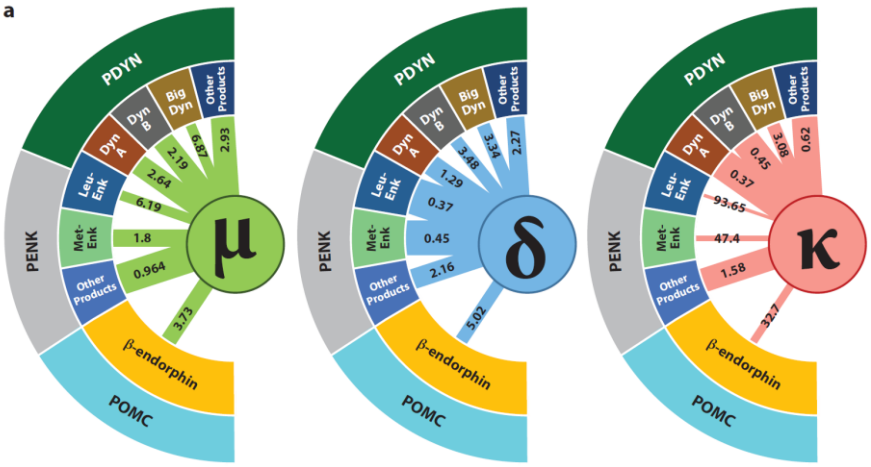

Figure 1. Endogenous opioid peptides and their receptors. The diagram summarizes the affinities of enkephalins, dynorphins, and beta-endorphin to the μ, δ, and κ receptors. The outer rings represent the peptide precursors, and the inner rings correspond to the processed, active peptides. The numbers in the spokes are reported K_d values (nM), the greater the spoke's thickness the higher the peptide's affinity for the receptor. Source: Emery and Akil, 2020, adapted from Mansour et al., 1995.

It should be noted that an additional group of peptides that specifically bind μ opioid receptor was reported, the endomorphins (Zadina et al., 1997). Endomorphin-1 and endomorphin-2 are tetrapeptides with amino acid sequences: YPDG and YPFF, respectively. They are characterized by high affinity (K_i = 0.36 nM and 0.69 nM) and selectivity to the μ opioid receptor (> 4 000-fold preference over the δ and κ receptors). They were suggested to participate in processing of nociceptive information; however, relatively little is known about this group, as their protein precursor remains unknown (Gu et al., 2017).

1.2. μ opioid receptor mechanisms of action

The opioid receptors are canonically coupled to the effector heterotrimeric $G_{i/o}$ protein that consists of membrane-anchored subunits: Gα and Gβγ. Studies using cryomicroscopy revealed that the main interaction sites between μ opioid receptor and Gα subunit occur between the intracellular loops 2, 3 and the transmembrane (TM) loops 3, 5, and 6 of the receptor and the α-helices of the Gα (Koehl et al., 2018).

Activation of the receptor induces changes in its conformation, which in turn facilitates conformational changes in the Gα protein and promotes the exchange of guanosine-5'-diphosphase (GDP) with guanosine-5'-triphosphate (GTP) (Figure 2). The exchange induces the dissociation of the $G_{o/i}$ from the receptor and its split into Gα and Gβγ subunit. Dissociated Gβγ subunit activates G protein-gated inwardly rectifying potassium channels (GIRK) (Al-Hasani & Bruchas, 2011). The outward potassium

current causes hyperpolarization and acts to inhibit tonic neuronal firing. Additionally, the Gβγ suppresses the opening of voltage-gated calcium channels (VGCC). The process blocks the release of neurotransmitters and also contributes to the inhibition of neuronal firing (Siegelbaum et al., 2013). Concurrently, the Gα subunit inhibits adenylyl cyclase attenuating production of the second messenger, 3', 5'-cyclic adenosine monophosphate (cAMP). A low level of cAMP yields a decrease in the cAMP-dependent influx of Ca^{2+} and downregulation of protein kinase A, which contributes to changes in the activity of transcription factors and regulatory proteins, for example, the cAMP response element-binding protein (CREB) (Al-Hasani & Bruchas, 2011; Law et al., 2000).

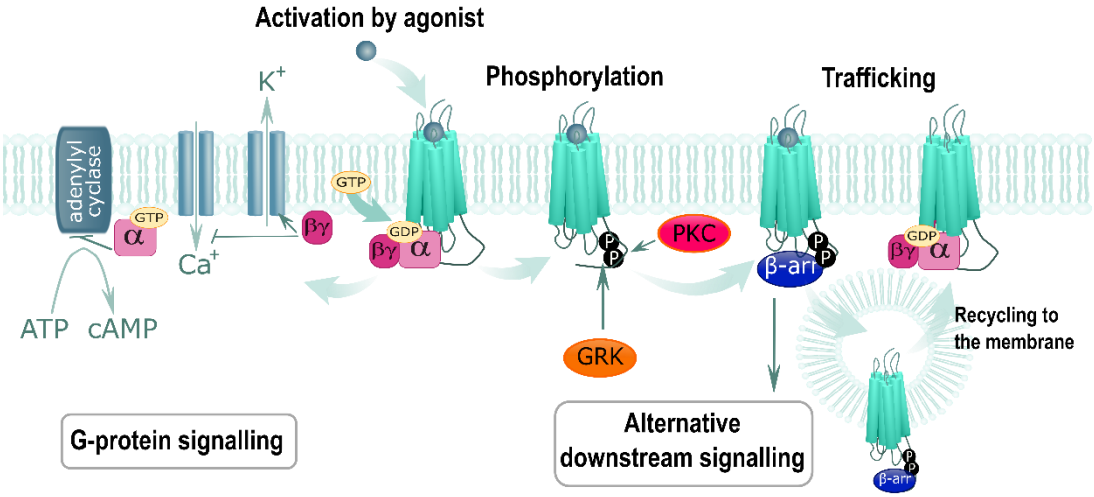

Figure 2. μ opioid receptor signalling pathways after activation by agonist. G-protein dependent and independent signalling pathways are presented. G-protein dependent pathway: activation of GIRC, inhibition of VGCC and adenylyl cyclase with G-protein subunits. G-protein independent pathway: alternative downstream signalling pathway with receptor phosphorylation and trafficking promoted by β-arrestins (β-arr).

Upon activation, the C-terminus of the μ opioid receptor becomes a target for phosphorylation by serine/threonine kinases: protein kinase C (PKC) and G-protein-coupled receptor kinases (GRK) (Figure 2) (Allouche et al., 2014). The phosphorylation pattern is strongly dependent on the ligand and determines β-arrestin binding. β-arrestins 1 and 2 are scaffold proteins that promote receptor uncoupling from the G-protein and enables the recruitment of alternate signal transduction cascades, including the mitogen-activated protein kinases pathway (MAPK): ERK1/2, JNK1-3 and p38. These processes further contribute to changes in the activity of numerous transcriptional factors and regulatory proteins, e.g., CREB (Al-Hasani & Bruchas, 2011; Allouche et al., 2014). β-arrestins are also crucial for receptor internalization via endocytosis. After endocytosis, μ opioid receptors are returned to the cellular membrane.

Earlier research suggested that signalling bias towards G-protein or β-arrestin-associated signalling pathways may explain discrepancies in the pharmacological profiles of various μ opioid compounds. It was postulated that ligands preferentially activating the 'classical' G-protein signalling pathway would be devoid of adverse effects, e.g. respiratory depression and tolerance, which were attributed to β-arrestin signalling (Bohn et al., 2000). Due to difficulties in replication of the research, the hypothesis is criticised (Gillis et al., 2020) and a more unequivocal picture of the role of downstream signalling pathways and receptor trafficking in the regulation of μ opioid receptor is supported (Allouche et al., 2014; Gillis et al., 2020; Williams et al., 2013).

Regardless, many other factors have been hypothesized to contribute to the complexity of the μ opioid activity, like allosteric modulation, receptor's dimerization or the existence of receptor isoforms. In the following subsection, I will focus on the transcriptional diversity of the *Oprm1* gene, that leads to generation of multiple isoforms of the μ opioid receptor.

1.3. The *Oprm1* gene

The μ opioid receptor is encoded by the *Oprm1* gene in rodents (*OPRM1* in humans). The gene is evolutionarily conserved and has been sequenced in more than 200 species including chicken, zebrafish, and sea lamprey. In humans the gene is located on chromosome 6 (6q25.2), in rat on chromosome 1 (1q11), and in mouse on chromosome 10 (10qA1). Its *locus* spans over 210, 264, and 270 kbp, respectively. Up to date, there are at least 11 exons reported in humans, 13 in rats (Rattus norvegicus), and 19 in mice (Mus musculus) that comprise for the *Oprm1* gene (Doyle et al., 2007; Pasternak & Pan, 2013). In these species, exon 1 encodes for N-terminal end and the first transmembrane domain. Exons 2 and 3 encode for the transmembrane regions 2-7 as well as intra- and extracellular loops. They are highly preserved and present in nearly all reported variants. Additionally, exon 3 also contains a part of the carboxy terminal encoding sequence. The downstream exons encode for the remaining parts of C-terminus and the 3'UTRs of the synthetised mRNA. Putative STOP codons are present in exons 4, 6, or 9 (Pasternak & Pan, 2013).

The *Oprm1* gene undergoes extensive alternative splicing. Over 30 transcript variants of the murine gene were reported (Figure 3). These variants generate multiple receptor isoforms that differ substantially in their N- and C- termini. Most of the transcript variants contain exons 1-3 and one or a few additional exons at the 3' end. Such variants

produce functional 7-TM receptors, . In addition to the 'classical' 7-TM isoforms, proteins with 6- or 1-TM domains were postulated (Pasternak, 2001). The truncated 6-TM receptors are derived from variants that contain exon 11 instead of exon 1, while the 1-TM isoforms arise from transcripts that lack exon 2 or 3 (Pasternak & Pan, 2013). According to evidence from mice, deletion of exon 1 results in the generation of 6-TM protein isoforms (Majumdar et al., 2011). Such proteins are functional receptors with region-specific distribution in the brain (Abbadie et al., 2004; Xu et al., 2014) and distinct affinity and selectivity towards μ opioid ligands compared to 7-TM proteins. For example, mice with exon-1 deletion did not produce morphine-induced analgesia, while the analgesia produced by another μ opioid agonist, 3-iodobenzoyl-6β-naltrexamide (IBNtxA) was preserved. On the contrary, deletion of exon 11 did not affect morphine analgesia, however, it strongly reduced IBNtxA-induced analgesia (Lu et al., 2018; Majumdar et al., 2011). The 1-TM isoforms were not shown to bind opioid ligands. They were suggested to have chaperon-like function and dimerise with 7-TM receptors, preventing them from misfolding and degradation in the endoplasmic reticulum (Xu et al., 2013).

The expression of exon 1- or exon 11-containing variants is dependent on promotor. There are two independent promoter sites in the *Oprm1* gene, the first is located upstream of exon 11 and encoding mainly 6-TM isoforms, while the second localize upstream exon 1, encoding mainly 7-TM isoforms (Pan, 2002). The exon 1-associated promotor consists of the proximal and the distal elements and lacks a TATAbox, which is important for the initiation of transcription. It possesses several transcription starts and includes a few GC-rich cis-acting motifs that bind transcriptional factors like STAT1/3, AP1, Ik, NF-κB. The structure of the promoter suggests an expression pattern more typical for housekeeping genes (Pasternak & Pan, 2013). On the other hand, exon 11-associated promoter is the only promoter among the opioid receptor gene family that contains a TATAbox sequence (Pasternak & Pan, 2013). It also includes cis-acting elements like NF-1 or cMyc/Max that modulate its activity, however, because the promoter requires the formation of a preinitiation complex and recruitment of RNA polymerase II, it is classified as eukaryote class II promoter (Pan, 2002; Pasternak & Pan, 2013).

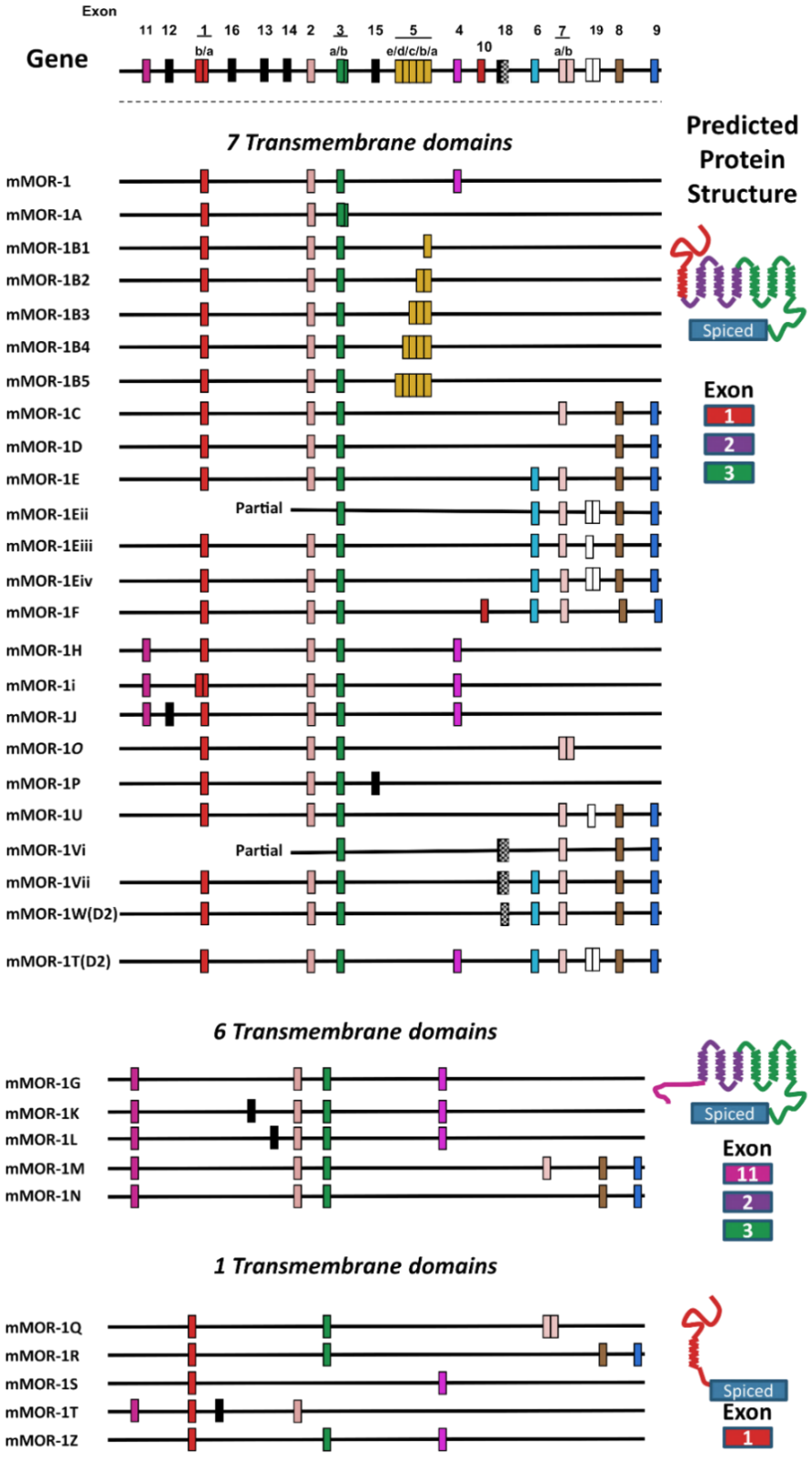

Figure 3. Schematic representation of the mouse *Oprm1* gene and reported transcript variants. Variants divided into three groups according the number of encoding transmembrane domains. Exons are depicted as rectangles. Schematic representation of putative translated proteins is shown on the rigth. Adapted from Pasternak & Pan, 2013

Oprm1 expression is also modulated by multiple transcriptional factors, repetitive elements and epigenetic factors like methylation of GpCs, histone methylation or histone acetylation (Cuitavi et al., 2021; Wu et al., 2019). For instance, it was observed that exon 1-associated promoter contains multiple GpC sites. High methylation in this region was shown to supress the expression of the *Oprm1* in the dorsal root ganglia (Sun et al., 2021). Also in humans, it was shown to contribute to various pathological conditions from increased risk of alcohol or opioid dependence (Ebrahimi et al., 2018; Zhang et al., 2012) to Alzheimer Disease diagnosis (Xu et al., 2018).

1.4. Expression of the μ opioid receptor in the central nervous system

The μ opioid receptor is abundant in the telencephalon, mesencephalon, pons, medulla, and spinal cord. Relatively high protein levels were observed in the laminar parts of the frontal cortex, piriform cortex and entorhinal cortex, the hippocampal formation and subiculum, as well as in the amygdala, especially the medial and lateral nuclei (Mansour et al., 1995). The high level of the receptors is present in the basal ganglia, dorsal and ventral striatum, ventral pallidum, globus pallidus, and the bed nucleus of the stria terminalis. The highest abundance of the receptor were observed in the thalamus and medial habenula. The hypothalamus is characterized by a relatively low receptor level in the dorsomedial nucleus and the lateral hypothalamic area. Furthermore, the receptors were observed in the mesencephalic structures, the interpeduncular nucleus, the substantia nigra, the ventral tegmental area, the superior and inferior colliculi, as well as the dorsal raphe nucleus. μ opioid receptor is also abundant in the pons, especially the parabrachial nucleus, the nucleus tractus solitarius, the spinal trigeminal nucleus, and the nucleus raphe magnus. High levels of μ opioid receptors were observed in the substantia gelatinosa of the spinal cord (Mansour et al., 1988; . The presence of μ opioid receptors in the radioligand-binding assays does not necessarily correlate with high levels of *Oprm1* mRNA. The apparent discrepancies in distribution occur in regions of the neocortex, olfactory bulb, superior colliculus, and spinal cord. The discrepancies may result from receptor mRNAs transport along the axons (Mansour et al., 1995).

The expression of the *Oprm1* mRNA in the striatum was confirmed by ISH assays, qPCRs, and northern blotting (Banghart et al., 2015; Minami et al., 1994; Oude Ophuis et al., 2014). Within the dorsal striatum, *Oprm1* transcripts co-localize with *Pdyn* mRNA. Expression of μ opioid receptors occurs in both D1- and D2-positive spiny projection neurons, with slight preference towards D1-positive cells (Banghart et al., 2015; Oude

Ophuis et al., 2014). The μ opioid receptor proteins create characteristic patches that converge with inputs from the limbic cortex (Banghart et al., 2015) Additionally, the proteins were also observed in 'giant' cholinergic interneurons and presynaptic terminals of neurons projecting from the cortex (Ponterio et al., 2013; H. Wang & Pickel, 1998). In the ventral striatum, *Oprm1* transcripts were observed in the nucleus accumbens core, especially the dorsomedial part referred to the 'hedonic hot-spot', the nucleus accumbens shell, and the olfactory tubercle. In the prefrontal cortex μ opioid reactivity was found in all layers I-V of the cortex, in layer I, the receptor was expressed together with Neuropeptide Y, whereas as in layers II-IV with enkephalin and vasoactive intestinal peptide (Férézou et al., 2007) In another study the μ opioid receptor reactivity was observed on small, bipolar, non-pyramidal neurons, located in layers II-IV (Taki et al., 2000).

The distribution of *Oprm1* mRNA in the striatum and the cortex can be observed in images of spatial transcriptomics (Ståhl et al., 2016), generated in previous research at the Department of Molecular Neuropharmacology (data not published). Spatial transcriptomics utilizes short-read sequencing that captures the poli-A tails of the mRNAs. Usually, the reads cover 150-300 nucleotides, and through specific barcodes, they can be assigned to the specific spots on the prepared slide, each spot 55 μm in diameter. The visualization of the *Oprm1* transcripts obtained from spatial transcriptomics is presented in Figure 4. The Figure depicts predominant expression of μ opioid receptor mRNA in the dorsal striatum, nucleus accumbens, ventral pallidum, islands of Calleja, lateral septal nucleus and dorsal endopiriform nucleus with lower expression in the cortex, especially the cingulate and piriform cortex, which remains in line with the literature (Mansour et al., 1995)

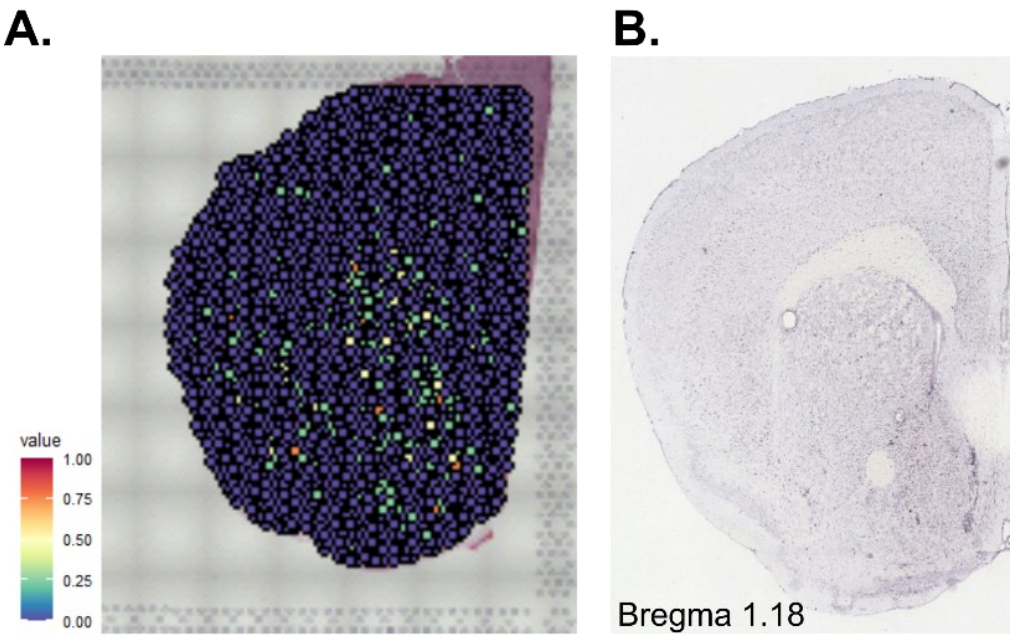

Figure 4. A. The visualization of *Oprm1* transcripts in spatial transcriptomics. The color of the dots reperesents expression levels according to the scale shown on the left.. B. the corresponding brain slice using ISH with probe targeting the *Oprm1* gene (*Gene :: Allen Brain Atlas: Mouse Brain*, 2011).

After discarding spatial information, the data revealed a few contigs suggesting the existence of several types of transcripts with diverse polyadenylation sides. The peaks of reads are schematically visualized in Figure 5. Whereas the contigs located in close proximity to exons 4 and 9 may be annotated to the *Oprm1* transcripts containing exon 4 and exon 9, the other peaks are difficult to annotate to the specific *Oprm1* exon.

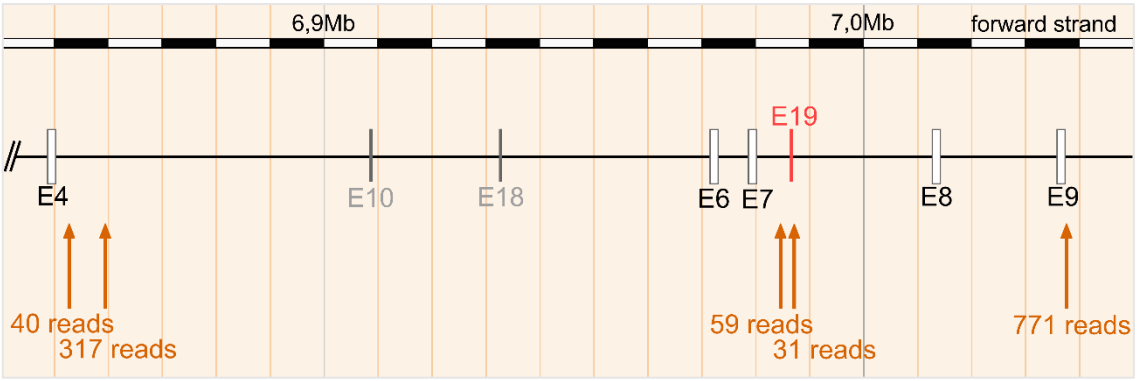

Figure 5 Polyadenylation sites of the *Oprm1* gene in the striatum and the prefrontal cortex revealed by the spatial transcriptomics method. Scheme of exons 4 to 9 of the *Oprm1* gene (GRCm38.v6). The double-slanted lines represent parts of the gene that are omitted in the scheme. Observed reads of peaks are indicated by arrows.

In summary, accumulated evidence indicates that the *Oprm1* gene has multiple transcripts that may generate receptor proteins of distinct functions, different affinities for ligands, and particularly high variability at the carboxy terminus, a region involved in

signal transduction pathways and internalization of the receptor. The existence of multiple types of the receptor implies that it could be possible to develop selective medications targeting specific subpopulations of μ opioid receptors and thus to develop improved analgesics. However, existing data on variants is largely based on just one analysis method (i.e., PCR) and there is limited data on the cell-specificity of expression of the reported variants.

2. Aim

The primary aim of the study was to determine the presence of selected transcripts of the *Oprm1* gene in the striatum of adult mice. The second aim was to determine the distribution of specific transcript variants in the striatum and the prefrontal cortex.

3. Materials and Methods

3.1. Materials

3.1.1. Chemicals

Name	Manufacturer	Application
10 × PBS Buffer	Serva, Germany	RNAscope ISH
50 × Wash Buffer	Advanced Cell Diagnostics, Inc., USA	RNAscope ISH
6 × DNA Loading Dye	VWR Life Science, USA	RT-PCR
Agarose Bio Standard	Prona Agarose, Spain	RT-PCR
Isopropanol	POCH, Poland	RNA isolation
Nuclease-free Water	Ambion Inc., USA	RNA isolation, RT-PCR
Oligo-dT Primers	Invitrogen, USA	RNA isolation
Optical Cutting Temperature	Cell Path, UK	RNAscope ISH
Paraformaldehyd Reinst, DAC	Carl Roth, Germany	RNAscope ISH
Prolong Gold Antifade Mountant	Invitrogen, USA	RNAscope ISH
RNAlater Stabilization Solution	Qiagen Inc., USA	Tissue preparation
RNAscope® Probes	Advanced Cell Diagnostics, Inc., USA	RNAscope ISH
RNAscope® Protease IV	Advanced Cell Diagnostics, Inc., USA	RNAscope ISH
TaqMan Probes	Thermo Fisher Scientific, UK	Quantitative PCR
TEA buffer	A&A Biotechnology, Poland	RT-PCR
Trizol Reagent	Invitrogen, USA	RNA isolation

3.1.2. Kits and ready-made solutions

Name	Manufacturer	Application
DreamTaq Green PCR Master Mix (2 ×)	Thermo Fisher Scientific, UK	PCR
Gel Purification Kit	Qiagen Inc., USA	Short-read sequencing
Omniscript RT Kit	Qiagen Inc., USA	Reverse transcription
PCR Product Purification Kit	Qiagen Inc., USA	Short-read sequencing
RNA 6000 Nano Methods LabChip	Agilent, USA	Nanopore sequencing
RNAscope Fluorescent Multiplex Detection Kit	Advanced Cell Diagnostics, Inc., USA	RNAscope ISH

RNAscope Intro Pack For Multiplex Fluorescent Reagent Kit	Advanced Cell Diagnostics, Inc., USA	RNAscope ISH
TaqMan Universal PCR Master Mix (2 ×)	Thermo Fisher Scientific, UK	Quantitative PCR

3.1.3. PCR primers' sequences

Name	Sequence 5'-3'	Targeted exon/ exon junction
PF1	GACCCATGCGGTCCTAACC	1
PF11	TCTGGGCCGATGATGGAAGC	11
PR2	TCTGCCAGAGCAAGGTTGAA	2
PR11/2	TCTGAGGTAACTCTTCCCCTCTTGA	11/2
PF3	TCCTCCACAATCGAACAGCAA	3
PF2/3	CAGGCAGGGGTCCATAGATTGC	2/3
PR_6860kb	GATCTGTGTAGGATGGAGCTATT	-
PR_6990kb	GCCACATTCACATGCCCAAG	19

3.1.4. RNAscope probes

Gene Name	Accession/region	Reactivity/species	Catalogue Number	Channel
Penk	NM_001002927.2	*Mus Musculus*	318761-C2	2
Oprm1	NM_001039652.1	*Mus Musculus*	315841-C3	3
Oprm1-O7	chr10:6858476-6861848 (GRCm38)	*Mus Musculus*	1178981-C1	1

3.2. Animals

In the experiment tissue from adult, C57BL/6 mice (*Mus musculus L.*) was used. The qPCR and Oxford Nanopore Sequencing (ONS) were performed on adult *wild type* mice from the *Penk*^{tm1Rpa} strain (Misiołek, 2020). The RNAscope experiments were conducted on 14-week old, *wild type* male mice from the *Pdyn*^{tm1Zim} strain (Zimmer et al., 2001). Genetically modified mice were congenic to the C57BL/6N strain. The animals were kept in the in-house breeding facility of the Maj Institute of Pharmacology of the Polish Academy of Sciences. Animal genotypes were confirmed by PCR on DNA

isolated from distal tail tissue biopsy. Primers as follows were used, for *Penk* KO: primer forward – CTACAGGCGCGTTCTTCTCT and primer reverse – GGATGGAGCAG GTAGTGG, for *Pdyn* KO: primer forward – CTTCAGAATAGGTATTGGGGTTC TCCTGGG and primer reverse – CGCACCGTCCATTTTAATGAGGAGGACTTG.

Mice were group-housed (2-6 animals per cage) in standard type II L cages, with *ad libitum* access to food and water. The breeding facility was kept in a 12/12h light/dark cycle with the light phase starting at 7:00 a.m. The temperature and humidity were controlled and maintained at 22 ±2 ºC, 55±10%, respectively. The ARRIVE guidelines were followed in planning of the experiments and reporting the results. A summary of animals used in the experiments is shown in Table 1.

Experiment	Number of Animals	Strain, Genotype	Sex	Age
RT-PCR and product sequencing	5	C57BL/6	males	adult
qPCR	3	*Penk*^{tm1Rpa} wt wt	males	14-18 weeks
ONS	2 2	*Penk*^{tm1Rpa} wt wt	males females	16-18 weeks
RNAscope ISH	2	*Pdyn* ^{tm1Zim} wt wt	males	14 weeks

Table 1. Number, strain (genotype), sex, and age of the animals used for specific experiments

Mice that were used for the RNAscope experiment underwent mild behavioral testing prior to sacrifice.

3.3. Quantitative PCR

3.3.1. Tissue preparation

Animals were euthanized by cervical dislocation. Brains, hearts, and spleens were removed immediately after decapitation, submerged in falcons with RNAlater Stabilization Solution (Qiagen Inc., USA) and placed at 4°C. The next day, the striatum (caudate putamen and nucleus accumbens) was isolated under stereo microscope (StemiDV4, Carl Zeiss, Germany) with 23G sterile needles. Small sections containing heart apexes and spleens were excised. All tissue samples were placed in 2.0 ml round bottom Eppendorf® tubes and kept at -20ºC.

3.3.2. RNA isolation

RNA was isolated using the single-step acid guanidinium thiocyanate-phenol-chloroform extraction (AGPC) method (Chomczynski, 1993). Samples were thawed at

room temperature, then the RNAlater solution was removed, and stainless steel balls were placed in the tubes. First, tissue samples were homogenized in 1 ml of TRIzol reagent (Invitrogen, USA) in the TissueLyser II apparatus (Qiagen Inc., USA). Homogenization was performed twice for 3 minutes at 25 Hz. To precipitate RNA, 200 µl of chloroform was added to each tube. Samples were incubated for 10 minutes on ice, mixed by vortexing, and centrifuged (12 000*g*, 20 minutes, 4ºC). The aqueous phase (approx. 450 µl) was transferred to the new 1.5 ml tubes, and an equal volume of isopropanol (POCH, Poland) was added to each tube. Samples were mixed, placed at -70ºC for 20 minutes and then centrifuged (12 000*g*, 30 minutes, 4ºC). Supernatant was discarded and the RNA-containing pellets were resuspended in 1 ml of 70% (v/v) ethanol and further centrifuged (12 000*g*, 10 minutes, 4ºC). After washing with ethanol, the pellets were air-died and dissolved in 30µl of sterile, nuclease-free water (Ambion Inc., USA). Finally, Samples were incubated in thermomixer (Thermomixer Comfort, Eppendorf, Germany) set at 65ºC for 5 minutes, then mixed and stored at -20 ºC. RNA concentration was measured using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA). Total RNA concentration was calculated based on the values obtained at A_{260}. The obtained results are presented in Table 2.

Sample ID	ng/µl	260/280	260/230
Blank	0	-	-
Heart1	689.92	1.97	0.34
Heart2	1016	2	0.76
Heart3	1585.76	2	1.31
Spleen1	2280.11	1.98	1.47
Spleen2	1562.77	2.01	1.05
Spleen3	3607.2	1.87	1.5
Striatum1	792.75	1.97	0.73
Striatum2	180.28	1.87	0.47
Striatum3	360.62	1.95	0.61

Table 2. mRNA concentrations (ng/µl) in the samples from the murine striata, hearts and spleens after reverse transcription. A260 – absorbance at 260 nm, A280 – absorbance at 280 nm. 260/280 – the ratio of absorbance at 260 nm and 280 nm. A ratio ~2.0 indicates 'pure' RNA . 260/230 – the ratio of absorbance at 260 nm and 230 nm. Expected 260/230 value for nuclei acids vary between 2.0-2.2. The low 260/230 ratios may indicate the presence of contaminants (*NanoDrop Spectrophotometers Nucleic Acid Purity Ratios T042 Technical Biuletin*, 2009)

3.3.3. Reverse transcription

Reverse transcription was performed using the Omniscript RT Kit (Qiagen Inc., USA). RNA samples were thawed on ice, mixed by vortexing and briefly spun down at 4ºC. 10 × Buffer RT, dNTP Mix, RNase-free water and oligo-dT primers (Invitrogen, USA) were thawed in RT, mixed, short spin centrifuged, and kept on ice. The volumes of

the components were calculated to obtain 2 µg of template RNA in 20 µl of the reaction mixture, as in Table 3. 10 × Buffer RT, dNTP Mix and oligo-dT primers and the Omniscript Reverse Transcriptase enzyme were mixed together and added to the tubes with diluted template RNA, the prepared mixtures were gently stirred with a pipette. The samples were incubated for 60 minutes at 37ºC and then diluted 20 times with sterile water.

Component	Volume/Reaction
10 × Buffer Rt	2 µl
dNTP Mix (5 mM each dNTP)	2 µl
Oligo-dT (10 µm)	2 µl
Omniscript Reverse Transcriptase	1 µl
RNase-free Water	variable
Template RNA	variable
Total Volume:	20 µL

Table 3. Volumes of the reverse-transcription reaction components

3.3.4. Quantitative PCR (qPCR)

The abundance of µ opioid receptor mRNA was analysed using the TaqMan Mm01188089_m1 probe (Thermo Fisher Scientific, UK). The probe target junction between exons 2 and 3 of the MOR-1C transcript of the murine *Oprm1* gene (NCBI reference sequence – NM_001039652.2). The *Hprt1* probe (NM_013556.2) was used as a control. Both probes were labelled with FAM dye (Carboxyfluorescein).

After thawing, the probes were mixed with TaqMan Universal PCR Master Mix (2 ×) (Thermo Fisher Scientific, UK) in the volumes indicated in Table 4. The mixture was gently stirred, vortexed, and aliquoted into a 6-well PCR plate (BIO-RAD, USA). Subsequently, the template cDNA was dispensed to the wells. The plate was sealed with an adhesive film (BIO-RAD, USA) and briefly centrifuged. The reaction was performed on the C1000 CFX96 Touch Real-Time PCR cycler (BIO-RAD, USA). The reaction parameters are presented in Table 5. Cycle threshold values (Ct) were automatically calculated by the PCR machine software, relative gene expression was determined using ΔCt method.

Component	Volume/Reaction
TaqMan Universal PCR Master Mix (AmpliTaq Gold® DNA Polymerase, Uracil-N glycosylase (UNG), dNTPs with dUTP, ROX™ Passive Reference, optimized buffer components) (2 ×)	10 µl
TaqMan Probe	1 µl
Template cDNA	9 µl
Total Volume:	20µl

Table 4. Volumes of the qPCR reaction components

Stage	Temperature	Time	Repeats
UNG Activation	50ºC	2 min	
Polymerase Activation, DNA Denaturation	95 ºC	10 min	
Amplification	95 ºC 60 ºC	15s 60s	40 ×

Table 5. Cycling parameters of the qPCR

3.4. RT-PCR and product sequencing

3.4.1. Polymerase chain reaction

To assess the presence of selected *Oprm1* transcripts in the mouse striatum RT-PCR was performed. The cDNA used as templates was prepared as described in Subsections 3.3.1 - 3.3.3. Primers targeting different variants of murine *Oprm1 splicing* were designed based on transcript sequences available in the Ensemble Genome Browser (https://www.ensembl.org/Mus_musculus/Gene/TranscriptComparison?db=core;g=ENSMUSG00000000766;r=10:6708506-6988198) using the NCBI Primer BLAST tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The primers were supplied by the DNA Sequencing and Oligonucleotide Synthesis Lab at the Institute of Biochemistry and Biophysics, PAS and prepared according to manufacturer's instructions.

Samples were thawed at room temperature. Reaction reagents: dH_2O, primers, and Master Mix (DreamTaq Green PCR Master Mix (2 ×), Thermo Scientific, USA) including optimized DreamTaq Green buffer, $MgCl_2$, and dNTPs were mixed together and parcelled into a 96-Well PCR Plate (BIO-RAD, USA). The samples were added to the wells. The volumes of specific components per well are shown in Table 6. The plate was then sealed with an adhesive film (BIO-RAD, USA), briefly centrifuged and placed in the Thermocycler (C100 Touch™ Thermal Cycler, BIO-RAD, USA). The temperature profiles and cycling parameters were adjusted to the expected lengths of the amplified

products and the thermodynamic properties of the primers. To increase the specificity of obtained products, the annealing temperature was gradually decreased with each cycle (touchdown PCR). Cycling parameters are summarised in Table 7.

Component	Volume/Reaction
DreamTaq Green PCR Master Mix (2 ×) (DreamTaq DNA polymerase, 2x DreamTaq Green buffer, dNTP, 0.4 mM each, 4 mM MgCl2)	10 µl
dH_2O	8 µl
Primer Forward (10 mM)	1 µl
Primer Reverse (10 mM)	1 µl
Template cDNA (50-150 ng/µl)	1 µl
Total Volume:	21 µl

Table 6. Volumes of the PCR reactions components

Stage	Temperature	Time	Repeats
Initial Denaturation	95ºC	3 min	
Denaturation	95 ºC	15s	
Annealing	65-60 ºC	30s	35-40 ×
Elongation	72 ºC	60s-90s	
Final Elongation	72 ºC	10 min	
Cooling	4 ºC	∞	

Table 7. Cycling parameters of the PCRs

3.4.2. Agarose gel electrophoresis

In order to identify the length of the products obtained in PCR reactions, electrophoretic separation was performed. Electrophoresis was carried out in 2% (w/v) agarose gel (Nu Micropore, Prona Agarose, Spain) with ethidium bromide and TAE buffer (A&A Biotechnology, Poland). The samples were loaded into gel wells mixed with a loading buffer (6 × DNA Loading Dye, VWR Life Science, USA). The gels were run for approx. 45 minutes at 120V. After separation, the gels were placed on a UV Transilluminator and photographed (BioDoc-It™ Imaging System, UVP, UK).

3.4.3. DNA purification and preparation for sequencing

To confirm the specificity of the products obtained in the PCR reaction, the bands were excised from the gel and prepared for sequencing. First, to extract DNA fragments from agarose gels, the Gel Extraction Kit (Qiagen Inc., USA) was used according to the manufacturer's instructions. 0.4 µl of each product was further used as a template cDNA in a second round of PCRs, the other reagents and primer pairs were added as in

Subsection 3.1.3. The cycling parameters were as in Table 7, except the annealing temperature, which was 60ºC (constant) and the number of cycle repeats (20 cycles). For each sample, five repeats were performed to obtain an adequate amount of product. After the PCR, 88 µl of the total volume for each sample was processed for further purification and 10 µl was used in electrophoresis. For purification, the PCR Purification Kit (Qiagen Inc., USA) was used following the manufacturer's instructions. After purification, the samples were stored in 1.5 ml Eppendorf tubes. Sanger sequencing was performed by the DNA Sequencing and Oligonucleotide Synthesis Lab at the Institute of Biochemistry and Biophysics, PAS.

3.5. Oxford Nanopore Sequencing

In order to reconstruct the entire length of splicing variants of the *Oprm1* gene, Oxford Nanopore Sequencing (ONS) was performed. The technology is a third generation sequencing that allows for obtaining long reads up to 2 Mb, which is not possible for most of available sequencing strategies, e.g. Illumina (Amarasinghe et al., 2020). In ONS, a nanometer-sized pore (nanopore) is inserted into an electrically resistant membrane created by synthetic polymers. The membrane is bathed in an electrochemical solution so that, when a potential is applied, an ionic current can be generated through the nanopore. Molecules can be identified by measuring the disruptions in the current, which are characteristic of a specific compound that enters the nanopore.

For this experiment, the RNA was prepared as described in Subsections 3.3.1 - 3.3.3. with few modifications. Brains immersed in RNAlater were cut on a vibratome VT1200 (Leica, Germany) into 200 µm sections. The entire slices containing the striatum and the prefrontal cortex (1.34-0.74 anterior to Bregma, Paxinos & Franklin, 2001) as well as the thalami isolated from the slices at -0.94 - -1.82 posterior to Bregma were selected to reverse transcription. After RNA isolation, its integrity was determined using capillary chip-based electrophoresis with an RNA 6000 Nano Methods 59 LabChip Kit and an Agilent Bioanalyzer 2100 (Agilent, USA) according to the manufacturer's instructions. The samples underwent reverse transcription and were prepared for shipment. The ONS was performed by Novogene UK using the Nanopore PromethION platform.

Sample Id	ng/µl	260/280	260/230	RIN
TH_7570	139.68	1.92	0.38	7.9
TH_7566	128.24	1.91	0.46	7.8
STR_7570	302.7	1.95	0.59	8.1
STR_7566	339.35	1.99	0.86	7.9

Table 8 Results of capillary chip-based electrophoresis for the mRNA samples for the ONS experiment. mRNA concentration (ng/µl), A260 – absorbance at 260 nm, A280 – absorbance at 280 nm. 260/280 – the ratio of absorbance at 260 nm and 280 nm. A ratio ~2.0 indicates 'pure' RNA . 260/230 – the ratio of absorbance at 260 nm and 230 nm. Expected 260/230 value for nuclei acids vary between 2.0-2.2. The low 260/230 ratios may indicate the presence of contaminants (*NanoDrop Spectrophotometers Nucleic Acid Purity Ratios T042 Technical Biuletin*, 2009). RIN (RNA integrity number) –the values vary between 10 (intact mRNA) to 1 (totally degraded mRNA) (Schroeder et al., 2006).

3.6. RNAscope *in situ* hybridization

The RNAscope assay is an advanced method of ISH that allows the analysis of RNA target expression within tissues and cell cultures (Wang et al., 2012). In the experiments, the RNAscope™ Fluorescent Multiplex Assay was performed according to the manufacturer's instructions.

The animals were killed by cervical dislocation. Brains were removed immediately from the skulls, fresh frozen on dry ice, and stored at -80ºC for up to 1 month. The day before sectioning, the brains were embedded in OCT Compound (Optimal Cutting Temperature, Cell Path, UK) and then stored at -80ºC. The frozen brains were sliced into 10 µm coronal sections on cryostat (CM 3050 S, Leica, Germany) with both the object and chamber temperature set at -20ºC. Slices were thaw-mounted on positively charged Superfrost® plus microscope slides and stored at -80ºC for up to one week. Slides representing the dorsal striatum, nucleus accumbens, and prefrontal cortex (the cingulate cortex and infralimbic cortex) were selected for ISH.

The RNAscope assay started with pre-fixing the slides with ice cold 4% paraformaldehyde (PFA), followed by dehydration at increasing concentrations of ethanol. After the dehydration, the slides were air-died and the hydrophobic barrier was drawn. To permeabilize the tissue, Protease IV was applied for 30 minutes at room temperature. The slides were then washed in 1 x PBS and hybridized with a specific probe for 2 h at 40ºC. After this step, four hybridizations with subsequent amplifiers were performed, in turns, with washing the slides in the buffer. In the final hybridization, fluorophores were attached to the targets. After hybridization, DAPI stain was applied to the slides for 30 seconds. The coverslips were mounted with ProLong Gold Antifade Mountant (Invitrogen, USA), the slides were stored at 4ºC in the opaque slide box.

3.7. Image acquisition and processing

The image acquisition was performed one month after the ISH. For imaging a fluorescent microscope Axio Imager.Z2 (Carl Zeiss, Germany) and Plan-Apochromat 63×/1.40 Oil M27 lens were used. Laser lengths equal to 644 (At647), 553 (At550), 493 (AF488), and 353 nm (DAPI) were set to excite the fluorophores. The laser power was adjusted to the brightest sample and remained the same throughout the experiment. For each stage, 5 μm-thick Z-stacks containing 21 slices were acquired. Image processing was carried out using ZEN lite software (3.5.093.00002 version, Carl Zeiss, Germany). For each Z-stack, the maximum orthogonal projection method was applied.

3.8. Statistical analysis

Statistical analysis was conducted in R version 4.0.3 using stats package. In order to compare the relative expression of *Oprm1* mRNA in qPCR experiment, Kruskal-Wallis test was performed. The statistical significance level was set at 0.05.

4. Results

4.1. *Oprm1* transcript abundance in the murine striatum, spleen, and heart

The expression level of *Oprm1* was evaluated by quantitative PCR. I have evaluated the abundance of gene transcripts in the mouse striatum and compared it with expression levels in peripheral tissues, spleen, and heart. I used the Mm01188089_m1 TaqMan probe that targeted the region within exons 2 and 3. The expression of *Oprm1* in the spleen and the heart was close to the detection limit, with one sample from the heart and two samples from the spleen not crossing the cycle threshold. The results showed that the relative expression level in the striatal samples was ~1000 times higher than in the peripheral organs (Figure 6).

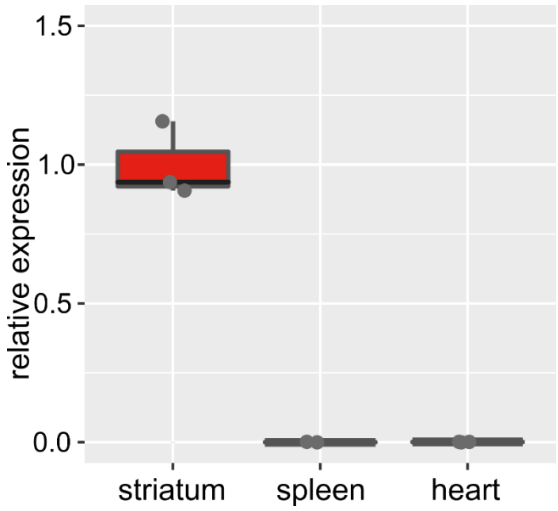

Figure 6. Relative expression of the *Oprm1* gene in the murine striatum, spleen, and heart. Relative expression was assessed by ΔCt method with *Hprt1* used as a reference gene and standardization to mean values obtained in the striatum. The box depicts the interquartile range with a black thick line representing median, and whiskers ranging between minimum and maximum values.

The discrepancies were not likely the result of technical problems, as the values of the cycle thresholds for the housekeeping gene were similar in all samples. The statistical test used to compare the relative *Oprm1* expression suggested that there were no significant differences between the organs (Kruskal-Wallis, χ^2= 5.6092, p=0.06053), though it should be noted that the sample size was small (n=3/group).

4.2. Expression of specific *Oprm1* transcripts in the murine striatum

In order to assess the expression of selected *Oprm1* transcripts RT-PCR followed by gel electrophoresis was performed. The identity of selected products was confirmed by sequencing.

Determination of 5'-terminal exons of *Oprm1*

First, we wanted to determine the existence of transcripts containing exon 1 or exon 11. A reaction with two primer pairs targeting these exons was performed on the material from 5 separate striatal samples obtained from adult male mice. The expected product sizes were 221 bp for the primer pair that flanked exon 1, and 90 bp for the pair that targeted exon 11. The results obtained are presented in Figure 7. Although the primers for exon 1 gave clear bands of expected size, the primers targeting exon 11 produced a weak, diffuse signal that did not allow for product identification.

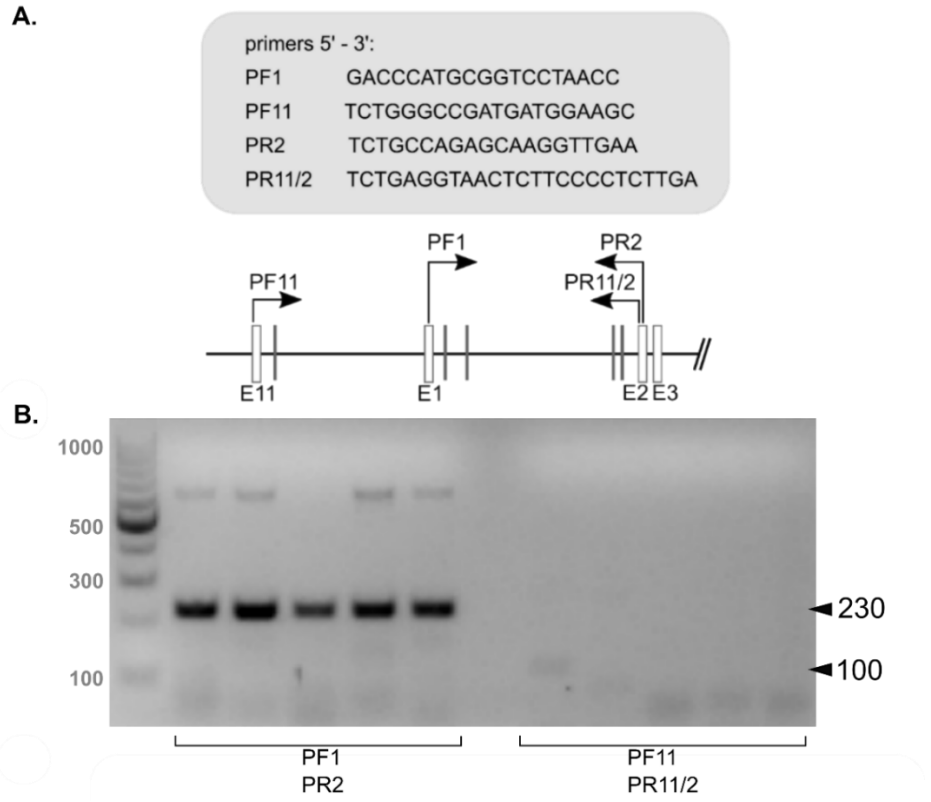

Figure 7. Determination of the 5'-terminal *Oprm1* exons in the striatum. A. The sequences of the primers used in the experiment with their localization in the *Oprm1* gene. The double-slanted lines represent parts of the gene that are omitted in the scheme. B. Electropherogram of the products of RT PCR. A 100 bp DNA ladder was run in the first lane from the left. Products of interest are indicated by arrows. The bands visible on the left side, at 230 bp, suggest the existence of variants containing exon 1. On the right, a weak band visible at ~100 bp in one of the samples may represent transcripts that contain exon 11.

Determination of 3'-terminal exons of *Oprm1*

Spatial transcriptomics sequencing revealed the existence of several potential 3' terminations of transcripts within the *Oprm1 locus.* Based on these results, we decided to design reverse primers targeting regions that were not previously annotated to *Oprm1* gene, at 6 860 000 and 6 986 000 bp (GRCm38.v6). To increase the reproducibility of the experiments, the previously-used forward primers were chosen. No potential amplicons of the *Oprm1* gene were produced using the PR_6860kb primer (Figure 8).

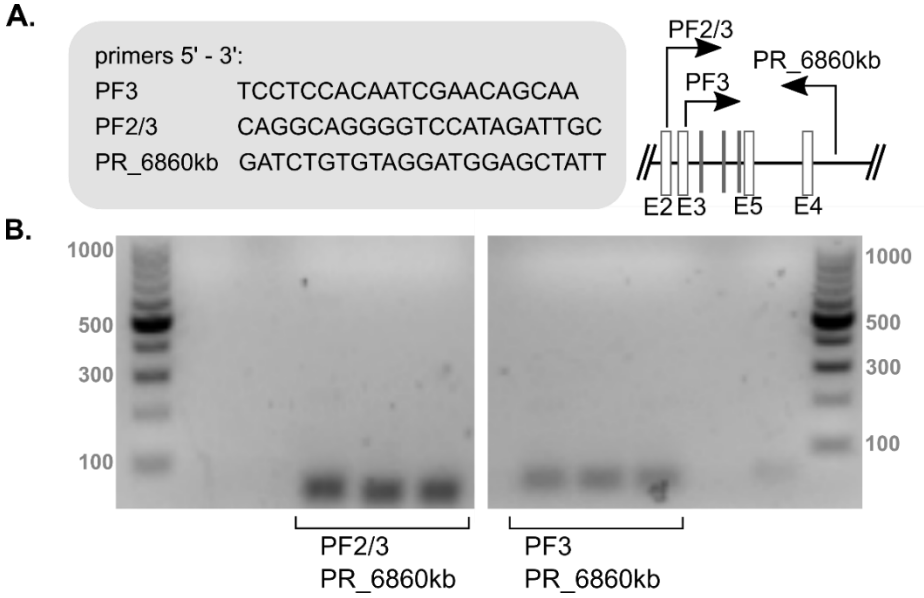

Figure 8. Evaluation of the possible exon at 6860 kbp (GRCm38). A. Sequences of the primers used in the reaction with their localization in the *Oprm1* gene. The double-slanted lines represent parts of the gene that are omitted in the scheme. B. Electropherograms of the products of the RT PCR reactions. A 100 bp DNA ladder was run in the peripheral lanes. The electropherograms suggest lack of the products of expected length.

In contrast, multiple bands were produced using PR_6990kb primer (Figure 9). Although the expected product lengths were not possible to predict, the span between forward primers suggested that the products of PF3 should be 433 bases longer than PF2/3, which might be applicable to bands appearing at 950 and 510 bp. To confirm the results, the selected PCR products were purified, amplified, and sequenced.

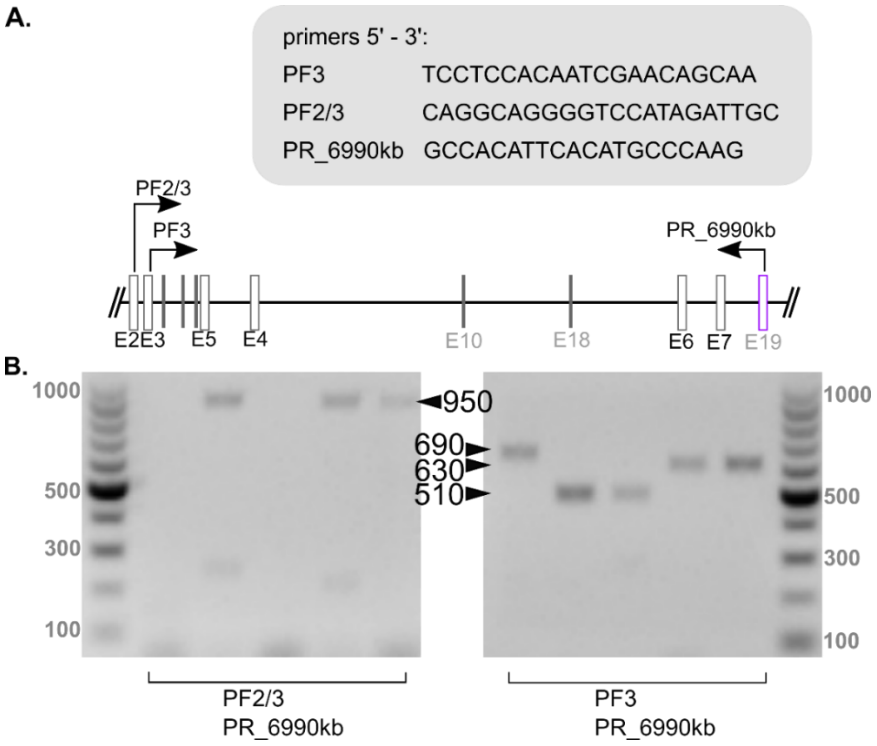

Figure 9. Evaluation of the possible exon at 6990 kbp (GRCm38). A. Sequences of the primers used in the reaction with their localization in the *Oprm1* gene. The double-slanted lines represent parts of the gene that are omitted in the scheme. B. Electropherograms of the products of the RT PCR reactions. A 100 bp DNA ladder was run in the peripheral lanes. Possible *Oprm1* transcripts are indicated by arrows at 510, 630, 690 and 950 bp.

4.3. Identification of RT PCR products

As the 630 bp and 950 bp bands provided most reproducible results, they were excised from the gel, cleaned up and then amplified in a secondary PCR using the same primers as previously. Small fractions of the amplicons were separated by electrophoresis to assess the accuracy of the procedures (Figure 10). The remaining material underwent sequencing. The cDNA concentrations in the samples were as follows: 75 ng/µl for PCR1, 88 ng/µl for PCR2, and 132 ng/µl for PCR3.

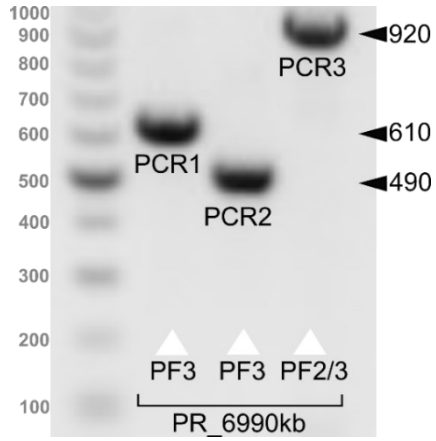

Figure 10. Evaluation of PCR clean-up and amplification procedures in samples prepared for Sanger sequencing. A 100 bp DNA ladder was run in the 1 lane from the left. Products of interest are indicated by arrows. The second lane corresponds to samples possessed from 630 bp bands of the previous experiment, and the other two lanes correspond to 950 bp bands. The discrepancy in length between the third and fourth lane stems from different forward primers used in the samples.

Sequences were aligned to the genome using the BLAT tool (University of California, Santa Cruz). In general, the results confirmed that the excised bands matched the specific

regions within the *Oprm1* gene. The sequences showed a relatively high alignment score and identity with the genomic sequence (≥ 96.0%) (Table 9). It is worth to note that the alignment did not include the terminal regions of the query sequences, due to the high noise level in these regions.

Query sequence	Query length	Alignment score	Query cover	Sequence identity	Chromo-some	Genome start	Genome end
PCR1-PF3	606	572	17-592	100.0%	chr10	6830444	6985996
PCR1-PR_6990kb	602	578	8-591	99.9%	chr10	6830411	6985971
PCR2-PF3	589	429	16-487	96.0%	chr10	6830444	6986007
PCR2-PR_6990kb	883	435	10-479	96.6%	chr10	6830411	6985971
PCR3-PF2/3	920	862	25-899	99.5%	chr10	6830017	6985988
PCR3-PR_6990kb	918	814	17-897	96.4%	chr10	6829988	6985965

Table 9. BLAT alignments of sequenced RT PCR samples. 'Alignment score' represents the number of nucleotides matched to the genome with a penalty for mismatches and gaps. 'Query cover' describes the location of the beginning and end of the alignment in the query sequence. 'Sequence identity' indicated the number of nucleotides and gaps in percent. 'Genome start' and 'Genome end' represent the beginning and end of the alignment in the genome (GRCm38.v6).

BLAT analysis showed that all sequenced fragments contained exon 3, 7, and part of exon 19, the PCR1 sample also contained exon 6. Reconstruction of the 'partial' transcripts is presented in Figure 11, while alignment with the genome and the whole sequences are shown in Appendix. As reads that were generated from forward primers (assigned as PCR_-PF_) had better quality, they were used for further analysis.

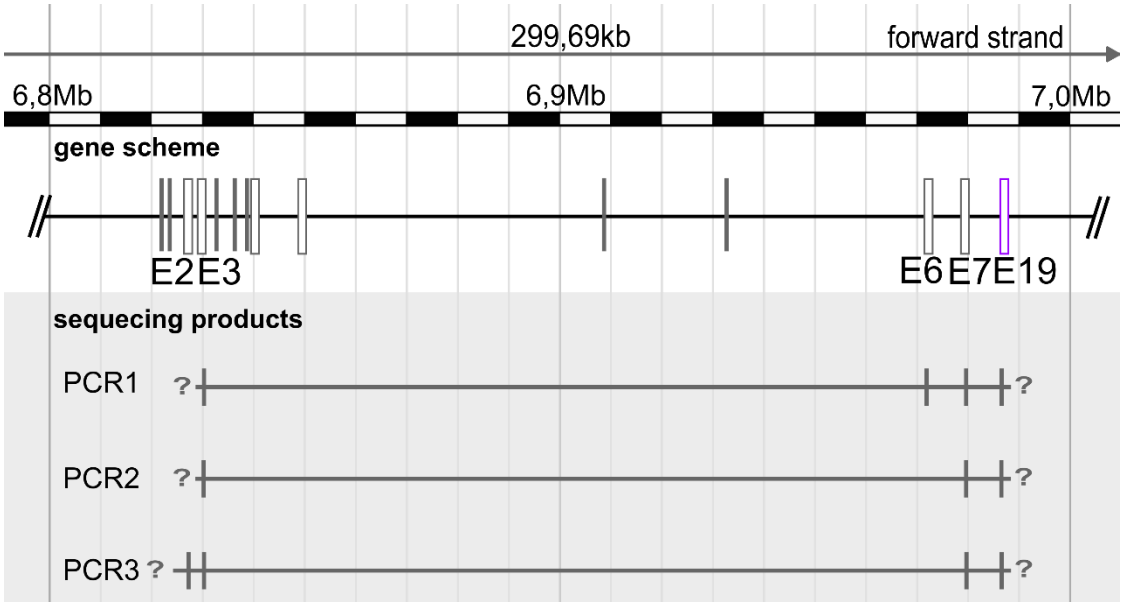

Figure 11. Schematic presentation of alignments for the sequenced PCR samples. On the top, *Oprm1* scheme is presented. On the bottom, the products of the PCR sequencing are presented, and recognized exons are marked with vertical lines.

In order to further characterize the transcripts, results were analysed with the blastn tool (National Center for Biotechnology Information) using the reference RNA sequences database (refseq_RNA). Sequences were matched with multiple transcripts, with the highest query coverage for MOR-1U, MOR-1M, and MOR-1C (Table 10). Detailed analysis revealed common exon 3, exon 7 between the samples and the aforementioned variants. A part of the sequenced exon 19 was identical to MOR-1U. After comparison with the literature, PCR1 was identified as MOR-1Eii or MOR-1Eiv. PCR2 and PCR3 showed the greatest similarity to MOR-1U; however, their exon 19 was longer (Doyle et al., 2007).

Query sequence	Transcript variant	NCBI accession	Query cover (%)	Query cover (bp)	Identity (%)
PCR1	MOR-1U	NM_001304955.1	40%	17-69 188-381	97.49%
	MOR-1M	NM_001302795.1	23%	17-69 188-279	100%
	MOR-1C	NM_001039652.2	23%	17-69 188-279	100%
PCR2	MOR-1U	NM_001304955.1	44%	2-265	94.05%
	MOR-1M	NM_001302795.1	26%	2-157	96.27%
	MOR-1C	NM_001039652.2	26%	2-157	96.27%
PCR3	MOR-1U	NM_001304955.1	73%	25-696	98.82%
	MOR-1M	NM_001302795.1	61%	25-594	99.82%
	MOR-1C	NM_001039652.2	61%	25-594	99.82%

Table 10. Comparison of RT-PCR sequences with existing transcripts.

Next, putative protein products were predicted using Expasy Translate Tool (SIB Swiss Institute of Bioinformatics) (Figure 12). The analysis revealed the presence of putative STOP codons in exons 6 and 19. PCR1 encoded 15 amino acids downstream exon 3, while PCR2 and PCR3 – 44 amino acids. A non-synonymous A125C mutation was observed in the PCR3 sample. The substitution was an artefact and did not appear in any other sample nor the reference genome.

PCR1

```
AAGAAAAAGCTGGACTCCCAGAGAGGGTGTGTACAGCATCCAGTGTGACCTGTCCCTTGT       60
K  K  K  L  D  S  Q  R  G  C  V  Q  H  P  V  -
                                                          6|
CTTTGAGCCTGGGGGCCATCTTCTTTCACAGCATACCATTTCCTTGTATCCTCTCTGAAG|      120
|7                          7|19
CCAACCCTGGCAGTCAGCGTGGCCCAGAT|CTTTACAGGATATCCTTCTCCGACTCATGTT      180
GAAAAACCCTGCAAGAGTTGCATGGACAGACTGCTACAACAGAAAACAACAGACTGGGTC      240
ACTCAGAAAAAATAAAAAGAAGAAGAAGAGGAGGAAGAACAAGCAAAATATCCTAGAAGC      300
TGGAATATCTAGGCAAGGTCAAGTTGCTAACAATGAAGATGACCTCTTCCTGACTTGCAG      360
ACATCTATGATCTTAATATAACCTCACATGGCAAATAGGAAATTCCTGTTCACCAGATAA      420
AGGCACAACTGCCATCGATGAGGACGCCACCCTCATGACCTATGACCTCTCAACACATCT      480
CTGAACAGCATCCCATTTTGAATTGAATTAGGACTTGGGCATG                       523
```

PCR2

```
CCAACCCTGGCAGTCAGCGTGGCCCAGATCTTTACAGGATATCCTTCTCCGACTCATGTT      60
P  T  L  A  V  S  V  A  Q  I  F  T  G  Y  P  S  P  T  H  V
                          7|19
GAAAAACCCTGCAAGAGTTGCATGGACAG|ACTGCTACAACAGAAAACAACAGACTGGGTC      120
E  K  P  C  K  S  C  M  D  R  L  L  Q  Q  K  T  T  D  W  V
ACTCAGAAAAAATAAAAAGAAGAAGAAGAGGAGGAAGAACAAGCAAAATATCCTAGAAGC      180
T  Q  K  K  -
TGGAATATCTAGGCAAGGTCAAGTTGCTAACAATGAAGATGACCTCTTCCTGACTTGCAG      240
ACATCTATGATCTTAATATAACCTCACATGGCAAATAGGAAATTCCTGTTCACCAGATAA      300
AGGCACAACTGCCATCGATGAGGACGCCACCCTCATGACCTATGACCTCTCAACACATCT      360
CTGAACAGCATCCCATTTTGAATTGAATTAGGACTTGGGCATGTGAATGTGGCA            414
```

PCR3

```
CCAACCCTGGCAGTCAGCGTGGCCCAGATCTTTACAGGATATCCTTCTCCGACTCATGTT      60
P  T  L  A  V  S  V  A  Q  I  F  T  G  Y  P  S  P  T  H  V
                          7|19
GAAAAACCCTGCAAGAGTTGCATGGACAG|ACTGCTACAACAGAAAACAACAGACTGGGTC      120
E  K  P  C  K  S  C  M  D  R  L  L  Q  Q  K  T  T  D  W  V
ACTCCGAAAAAATAAAAGAAGAAGAAGAGGAGGAAGAACAAGCAAAATATCCTATAAGC      180
T  P  K  K  -
TGGAATATCTAGGCAAGGTCAAGTTGCTAACAATGAAGATGACCTCTTCCTGACTTGCAG      240
ACATCTATGATCTTAATATAACCTCACATGGCAAATAGGATATTCCTGTTCACCACATAA      300
AGGCACAACTGCCATCGATGAGGACACCACCCTCATGACCTATGACCTCTCAACACATCT      360
CTGAAC AGCATCCCATTTTGAATTGAATTAGGACTGCTGGGGGGGGGGGGGGGAC          414
```

Figure 12. Partial sequences of PCR samples (downstream of exon 3) with their translation. To simplify the view, only sequences downstream exon 3 are included. Vertical lines mark exon junctions, and numbers indicate exons. The amino acid sequences are underlined in pink. STOP codons are marked in red.

Finally, I compared the putative amino-acid sequences with protein isoforms of mouse and human orthologues of *Oprm1* listed in the UniProt database (Figure 13). The alignment was prepared using ClustalW2 (EMBL's European Bioinformatics Institute) with the "Gap Opening Penalty" set at 90.00 and the "Delay divergent sequences" at 0%. The analysis revealed that the predicted amino acid sequence of PCR1 was identical with mMOR-1E, mMOR-1Eiii and mMOR-1Eiv, while the PCR2 and PCR3 appeared to be partially convergent with mMOR-1O, mMOR-1C, mMOR-1M, mMOR-1U, and hMOR-1C. Despite high similarity, the putative PCR1 and PCR2 protein sequences differ significantly in the carboxy terminus from already reported isoforms. This shows that if

the spotted variant encodes a functional protein, it produce an unique isoform of μ opioid receptor.

```
PRC1      ------------------------------KKKLDSQRGCVQHPV--------------
mMOR-1E   ------------------------------KKKLDSQRGCVQHPV--------------
mMOR-1Eiii------------------------------KKKLDSQRGCVQHPV--------------
mMOR-1Eiv ------------------------------KKKLDSQRGCVQHPV--------------
mMOR-1P   ----------------------------------------CV------------------
mMOR-1B5  ----------------------------------------CV------------------
mMOR-1F   -APCACVPGANRGQTKASDLLDLELETVGSHQADAETNPGPYEGSKCAEPLAISLVPLY-
hMOR-1Y2  ----------------------------IRDPISNLPRVSVF-----------------
hMOR-1Y3  ----------------------------IRDPISNLPRVSVF-----------------
hMOR-1Y   ----------------------------IRDPISNLPRVSVF-----------------
hMOR-1V   ----------------------------IRDPISNLPRVSVF-----------------
mMOR-1A   ---------------------------------------VCAF-----------------
hP35372-14--------------------------------NYYIIHRCCCNTPLISQKPVLLWFCD--
mMOR-1W   -------------------------------------AFGCCNEHHDQR-----------
PCR2      ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMDRLL
PCR3      ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMDRLL
mMOR-1O   ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMDR--
mMOR-1C   ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMD---
mMOR-1M   ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMD---
mMOR-1U   ----------------------------PTLAVSVAQIFTGYPSPTHVEKPCKSCMDSVD
hMOR-1C   ----------------------------PPLAVSMAQIFTRYPPPTHREKTCNDYMKR--
hMOR-1B3  ---------------------------------------GPPAKFVADQLAGSS-----
hMOR-1A   ---------------------------------------VRSL-----------------
hMOR-1W   ---------------------------------------VRSL-----------------
mMOR-1B4  ----------------------------AHQKPQECLKCRCLSLTILVICLHFQHQQFFI
mMOR-1R   ----------------------------AHQKPQECLKCRCLSLTILVICLHFQHQQFFI
hMOR-1B4  ------------------------------------------------------S------
mMOR-1B2  ---------------------------------KLLMWRAMPTFKRHLAIMLSLDN---
mMOR-1B3  ------------------------------------------------TSLTLQ----
mMOR-1Q   ------------------------------------------------TSLTLQ----
mMOR-1V   -------------------------KQEKTKTKSAWEIWEQKEHTLLLGETHLTIQHLS-
hMOR-1B1  -------------------------------------KIDLFQKSSLLNCEHTKG-----
mMOR-1B1  -------------------------------------KIDLF-----------------
hMOR-1B2  ---------------------------------------RERRQKSDW------------
hMOR-1B5  ----------------------------VELNLDCHCENAKPWPLSYNAGQSPFPFPGRV-
hMOR-1    ---------------------------------------LENLEAETAPLP----------
mMOR-1    ---------------------------------------LENLEAETAPLP----------
mMOR-1H   ---------------------------------------LENLEAETAPLP----------
mMOR-1J   ---------------------------------------LENLEAETAPLP----------
mMOR-1T   ---------------------------------------LENLEAETAPLP----------
hMOR-1R   CLPIPSLSCWALEQGCLVVYPGPLQGPLVRYDLPAILHSSCLRGNTAPSPSGGAFLLS--
hMOR-1X   CLPIPSLSCWALEQGCLVVYPGPLQGPLVRYDLPAILHSSCLRGNTAPSPSGGAFLLS--
mMOR-1D   ------------------------------------------------RNEEPSS----------
mMOR-1N   ------------------------------------------------RNEEPSS----------


PRC1      ------------------------------------------------------------
mMOR-1E   ------------------------------------------------------------
mMOR-1Eiii------------------------------------------------------------
mMOR-1Eiv ------------------------------------------------------------
mMOR-1P   ------------------------------------------------------------
mMOR-1B5  ------------------------------------------------------------
mMOR-1F   ------------------------------------------------------------
hMOR-1Y2  ------------------------------------------------------------
hMOR-1Y3  ------------------------------------------------------------
hMOR-1Y   ------------------------------------------------------------
hMOR-1V   ------------------------------------------------------------
```

```
mMOR-1A   ------------------------------------------------------------
hP35372-14------------------------------------------------------------
mMOR-1W   ------------------------------------------------------------
PCR2      QQKTTDWVTQKK--------------------------------------------
PCR3      QQKTTDWVTPKK--------------------------------------------
mMOR-1O   ------------------------------------------------------------
mMOR-1C   --------------------------------RGMRNLLPDDGPRQESGEGQLGR
mMOR-1M   --------------------------------RGMRNLLPDDGPRQESGEGQLGR
mMOR-1U   CYNRKQQTGSLRKNKKKKKRRKNKQNILEAGISRGMRNLLPDDGPRQESGEGQLGR
hMOR-1C   ------------------------------------------------------------
hMOR-1B3  ------------------------------------------------------------
hMOR-1A   ------------------------------------------------------------
hMOR-1W   ------------------------------------------------------------
mMOR-1B4  MIKKNVS---------------------------------------------------
mMOR-1R   MIKKNVS---------------------------------------------------
hMOR-1B4  ------------------------------------------------------------
mMOR-1B2  ------------------------------------------------------------
mMOR-1B3  ------------------------------------------------------------
mMOR-1Q   ------------------------------------------------------------
mMOR-1V   ------------------------------------------------------------
hMOR-1B1  ------------------------------------------------------------
mMOR-1B1  ------------------------------------------------------------
hMOR-1B2  ------------------------------------------------------------
hMOR-1B5  ------------------------------------------------------------
hMOR-1    ------------------------------------------------------------
mMOR-1    ------------------------------------------------------------
mMOR-1H   ------------------------------------------------------------
mMOR-1J   ------------------------------------------------------------
mMOR-1T   ------------------------------------------------------------
hMOR-1R   ------------------------------------------------------------
hMOR-1X   ------------------------------------------------------------
mMOR-1D   ------------------------------------------------------------
mMOR-1N   ------------------------------------------------------------
```

Figure 13. Comparison of human - mouse protein sequences for μ opioid receptor. Human isoforms are assigned with the prefix 'h', while mouse isoforms are assigned with the 'm'. Translation of the sequenced samples was presented in red.

4.4. Long-read transcriptome sequencing

In order to perform an unbiased analysis of *Oprm1* transcripts, I have used Oxford Nanopore Sequencing (ONS). ONS is one of the third-generation sequencing methods. The technology enables sequencing of long DNA or unamplified RNA molecules. The ONS utilizes transmembrane proteins – nanopores, that are introduced into a synthetic membrane characterized by high electrical resistance. When an electric current is applied on both sides of the membrane, it creates a potential difference that attracts negatively charged nucleic acid polymers. Nucleotides passing through the nanopores create changes in the electric potential between both sides of the membrane. Such a variation is specific to each type of nucleotide and therefore can be recognized by the sequencing machine. In the experiment, samples from two regions of the brain were analysed, the whole slices containing the anterior striatum and prefrontal cortex (1.34-0.74 anterior to Bregma) and the isolated thalamus (at the -0.94 - -1.82 posterior to Bregma). Sequencing was performed using the PromethION platform, which is a high-throughput system. Each

sample was sequenced three times. The raw statistics that describe the quality and characteristics of the procedure are presented in Table 11. The number of sample bases varied between 5.2 and 11.5 Gbp per sample, while the number of reads fluctuated between 3.8 and 8.2 × 10^6. The median read length in all samples was relatively consistent and oscillated between 1 002 and 1 165 bp. The N50 read length, so the length of the shortest read in the subset of the longest sequences that together represent ≥50% of the nucleotides in the sample varied from 1 875 to 2 172 bp with higher values obtained in the samples containing the striatum and prefrontal cortex than those from the thalamus (Table 11). The median read quality was high and consistent between samples, reaching values between 12.1-13.0. The percentage of reads with relatively low quality (<7) did not exceed 3.5% in any of the samples.

Sample	Sample Bases (bp)	Number of reads	Median read length	Read length N50	Median read quality	>Q7 number(%)
TH_7566	11 489 387 734.0	8 233 012.0	1 062.0	1 957.0	12.2	7987049(97.0%)
TH_7566	9 993 495 631.0	6 851 893.0	1 117.0	2 055.0	13.0	6633785(96.8%)
TH_7566	11 489 387 734.0	8 233 012.0	1 062.0	1 957.0	12.2	7987049(97.0%)
TH_7570	6 057 697 607.0	4 577 351.0	1 002.0	1 875.0	12.1	4436861(96.9%)
TH_7570	5 218 273 416.0	3 785 962.0	1 047.0	1 955.0	12.9	3658243(96.6%)
TH_7570	6 057 697 607.0	4 577 351.0	1 002.0	1 875.0	12.1	4436861(96.9%)
STR_7566	6 057 697 607.0	4 750 245.0	1 019.0	1 972.0	12.1	4598738(96.8%)
STR_7566	6 535 407 003.0	3 933 810.0	1 067.0	2 072.0	12.9	3797426(96.5%)
STR_7566	5 649 055 794.0	4 750 245.0	1 019.0	1 972.0	12.1	4598738(96.8%)
STR_7570	6 535 407 003.0	6 142 736.0	1 106.0	2 068.0	12.1	5970938(97.2%)
STR_7570	8 968 497 314.0	5 110 330.0	1 165.0	2 172.0	12.9	4951231(96.9%)
STR_7570	8 968 497 314.0	6 142 736.0	1 106.0	2 068.0	12.1	5970938(97.2%)
TH_mean	8 384 323 288.0	6 043 096.8	1048.7	1945.7	12.4	5856641(96,9%)
STR_mean	7 119 093 672.5	5 138 350.3	1 080.3	2 054.0	12.4	4981335(96,9%)

Table 11. MinION Quality Control statistics for all reads. Read length and the average quality for all samples, each sample was sequenced 3 times. The means and t tests for each parameter were presented in the bottom rows. Samples containing the striatum and the prefrontal cortex are annotated as STR_, and samples containing the thalamus are annotated as TH_.

All FASTQ files for each brain region were merged and the data were aligned to the GRCm38_102/mm10 genome with the Minimap2 program (v. 2.24) by Jacek Hajto and Dr. Marcin Piechota from the Laboratory of Pharmacogenomics, Maj Institute of Pharmacology, PAS. Alignment results were filtered for strand mismatches and invalid 3' ends. The final number of valid alignments was ~35 million.

The output files were visualized using the Integrative Genome Viewer v2.12.3 (IGV; Thorvaldsdóttir et al., 2013; Figure 14). Initially, only the reads that contained exons 2 and 3 were classified as *Oprm1* transcripts. The analysis revealed that the vast majority of the transcripts (25 reads in the STR_ samples and 72 in the TH_ samples) contained exons 1, 2, 3, and 4. Four transcripts were observed that contained exon 7 instead of exon 4 (one in the striatum and prefrontal cortex and three in the thalamus). Three of them expanded through exon 8 to exon 9 and one terminated in exon 7. Additionally, in the thalamus, a variant with exon 6 was detected. No *Oprm1* transcripts comprised of exon 11 were observed.

In addition to the reads aligned to both exons 2 and 3, and defined as the *Oprm1* transcripts, multiple ‘shorter’ reads were detected. More than 100 reads of 150-250 kb length were reported in the region of exon 9 in both of the analysed areas. Notably, between 8850-8860 kbp, in the region adjacent to exon 4, a long contig was observed. The reads formed a shape described here as a ‘cat with a tail’. The genome coverage of the contig was similar to those observed in exons 2 and 3, except the region between 6 851 700 and 6 852 700 bp, where the number of reads was five times higher. Due to the localization in genome and high coverage with reads, the ‘cat with a tail’ region was hypothesised to be a part of the *Oprm1* gene. To verify the hypothesis, RNAscope ISH with probe targeting the region was performed.

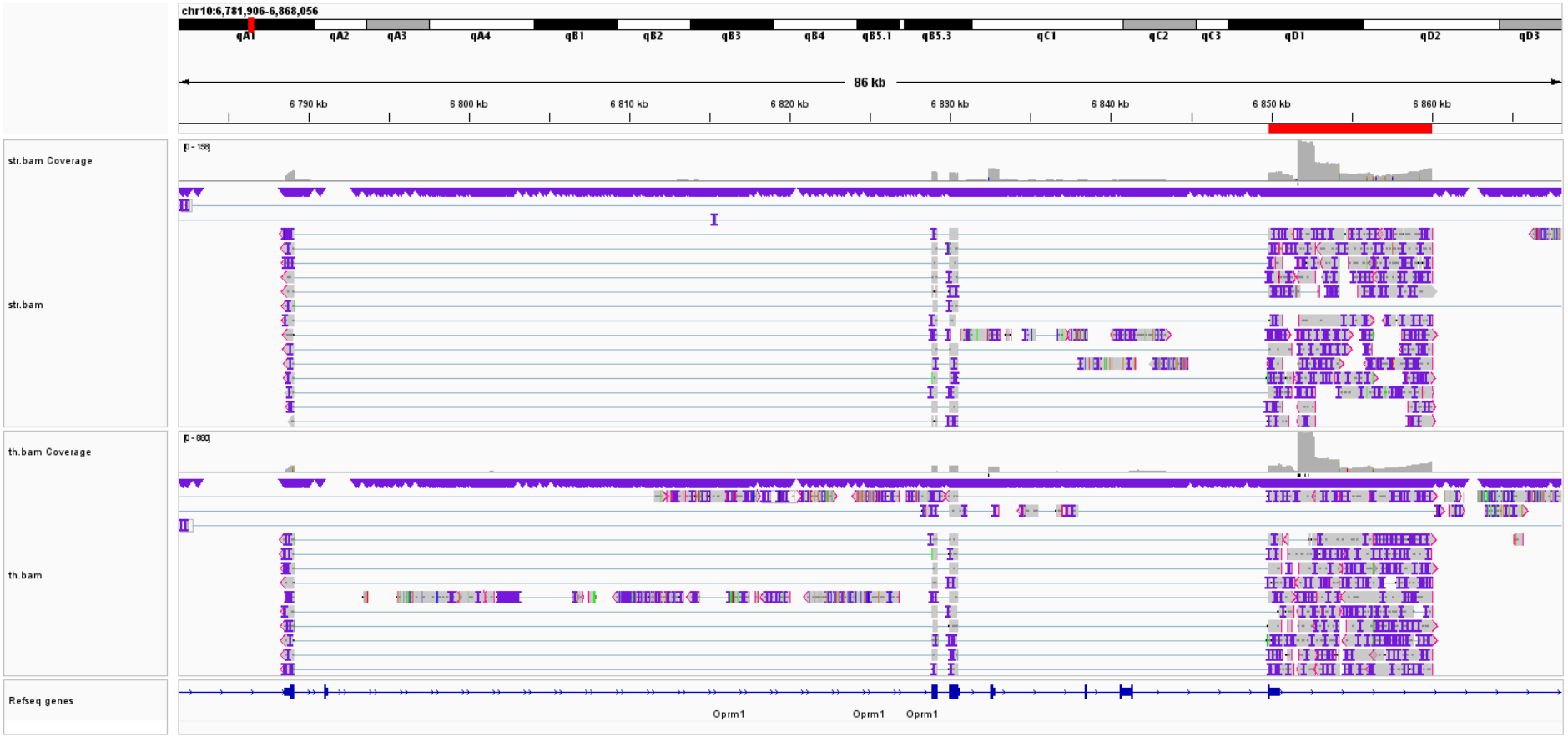

Figure 14. Representation of BAM files for the *Oprm1* gene in the striatum and the prefrontal cortex (upper panel) and in the thalamus (lower panel), as depicted in IGV. The overall coverage and reads are displayed as independent tracks. The top track represents the genome *locus*, the lowest track represents the reference *Oprm1* sequence. Small nucleotide deletions are marked in purple, clipped parts of the sequences are marked with red lines and angle brackets. In the picture, 4 exons (from 1 to 4) are captured. A 10 kb-long fragment of 'a cat with a tail' was marked in red.

4.5. Visualization of the *Oprm1* 3'UTR in RNAscope

The method is characterized by high sensitivity and specificity, which is provided by characteristic probe design and signal amplification strategy. Typically, for each target RNA, about 20 pairs of Z-shaped probes are designed that specifically hybridize with the target molecule. Both probes from single ZZ pair have to hybridize to the target sequence in order for signal amplification to occur. Hybridized ZZ probes pairs can bind a preamplifier which possesses numerous binding sites for amplifiers; such design enables for visualization of single RNA molecules and provides a high signal-to-noise ratio.

The aim of the experiment was to observe whether region spotted in the ONS experiment is a part of *Oprm1* transcripts and to compare the distribution of mRNA targeted by a standard probe (Mm-Oprm1-C3) with a custom-made probe (Mm-Oprm1-O7) hybridizing to the region between 6 858 476 and 6 861 848 bp (GRCm38.v6). Additionally, expression of the proenkephalin gene was assessed as a marker of the striato-pallidal medium spiny neurons. Three areas of the brain were analysed, the prefrontal cortex (the infralimbic cortex together with the cingulate cortex), the dorsal striatum and the nucleus accumbens core (Figure 15B, 16A) experiment revealed robust proenkephalin expression in many of the dorsostriatal and accumbal cells but not in the prefrontal cortex. There were also characteristic signal dots from the custom-made probe observed on slices. Unfortunately, the standard probe fluorescence exhibited low signal-to-noise ratio, what precluded localization of *Oprm1* transcripts (Figure 15C, D). Even though localization of *Oprm1* with a standard probe was not possible, the characteristics of the signal produced by the custom-made probe may suggest targeting the receptor's transcripts. The signal from the probe enveloped the nuclei, with lower expression in the background. The *Oprm1* signal was observed in both *Penk*-positive and *Penk*-negative cells within the dorsal and ventral striatum. The expression in the cortex was rather low in the outer layers, and slightly higher within inner layers (Figure 16B).

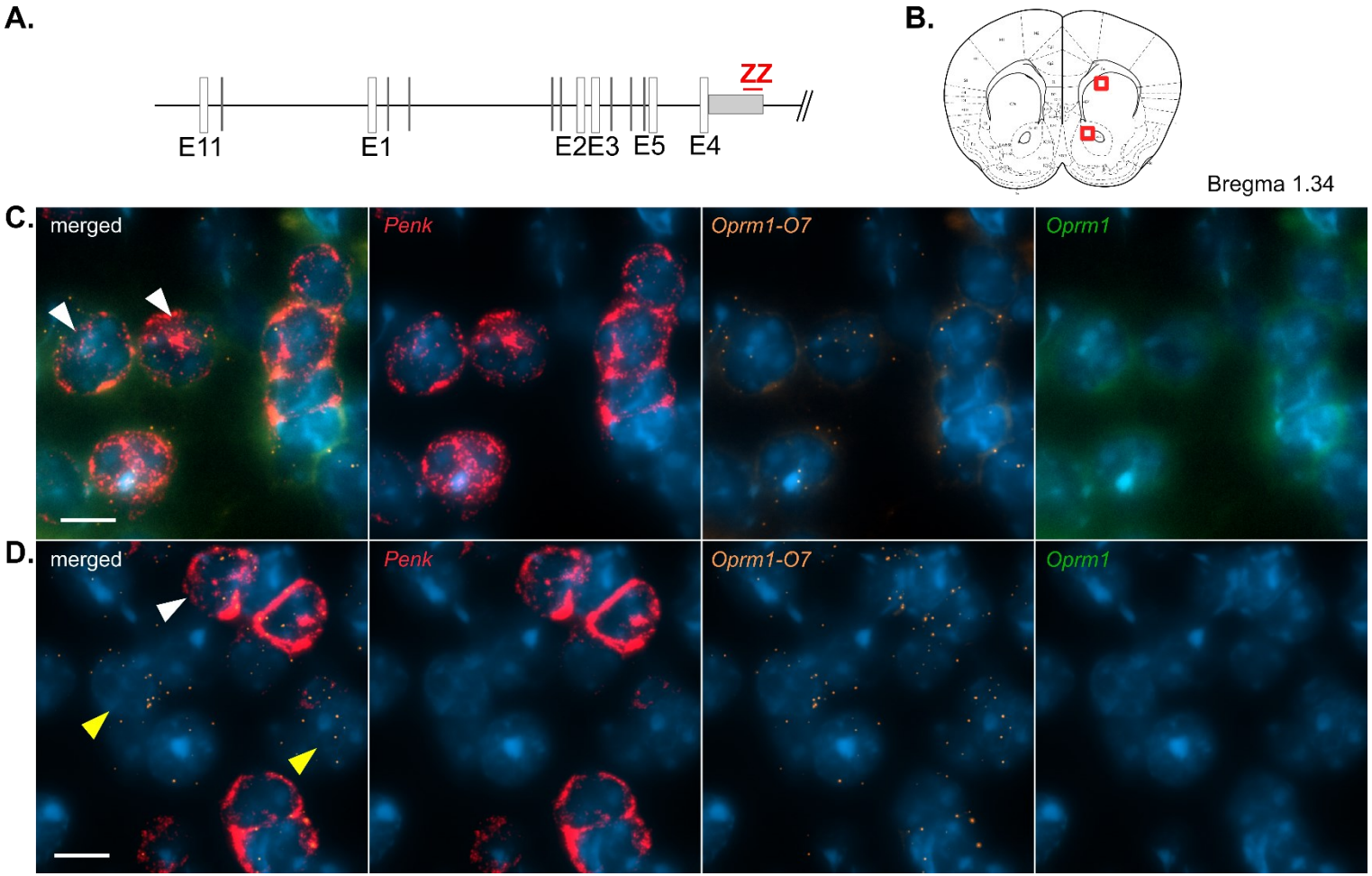

Figure 15. *Oprm1* transcript detection in the Dorsal Striatum and the Nucleus Accumbens as assessed by RNAscope ISH. A. Region targeted by the *Oprm1-O7* customized probe B. Reference section 1.34 mm from bregma, based on *The Mouse Brain in Stereotaxic Coordinates* (Paxinos & Franklin, 2001), area matching the regions presented on the photos delineated in red. C., D. Maximum intensity projections for *Penk* (red), *Oprm1-O7* (orange), *Oprm1* (green) in the Dorsal Striatum (C) and the Nucleus Accumbens (D), with corresponding photos for each transcript shown separately. *Oprm1* mRNA-expressing cells (white arrowheads), *Penk/Oprm1* mRNA expressing cells (yellow arrowheads). Scale bars 10 µ.

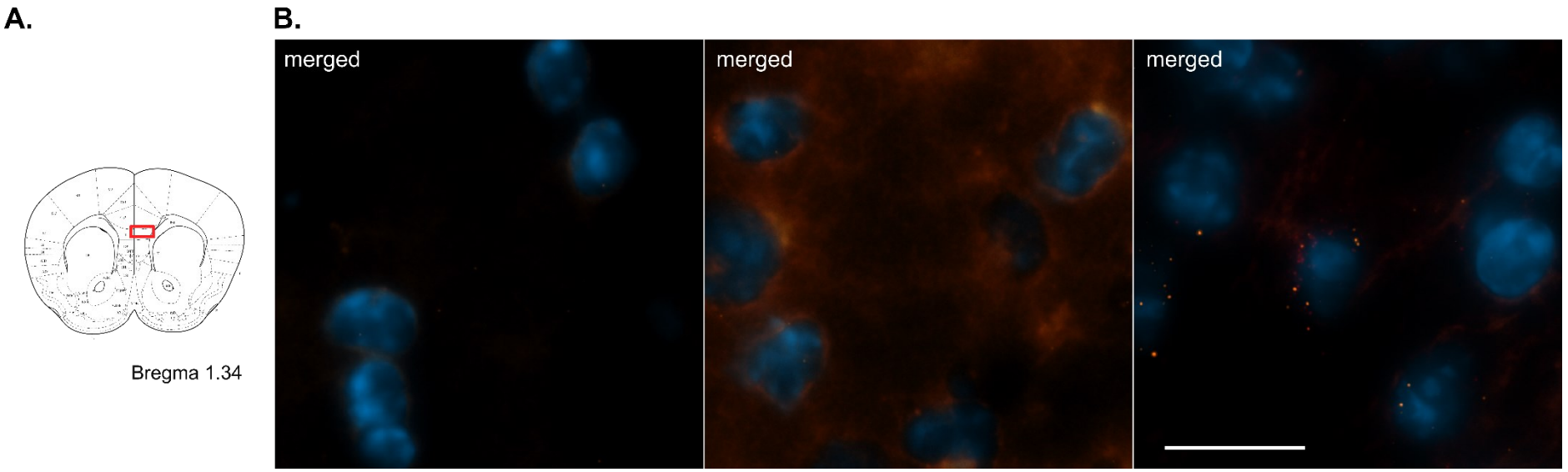

Figure 16. *Oprm1* transcript detection in the Cortex as assessed by RNAscope ISH. A. Reference section 1.34 mm from bregma, based on *The Mouse Brain in Stereotaxic Coordinates* (Paxinos & Franklin, 2001), area matching the region presented on the photos delineated in red. B. Maximum intensity projections for *Penk* (red), *Oprm1-O7* (orange), *Oprm1* (green) merged photos from the intersection of prefrontal cortex. Scale bar 20 μm.

5. Graphical summary

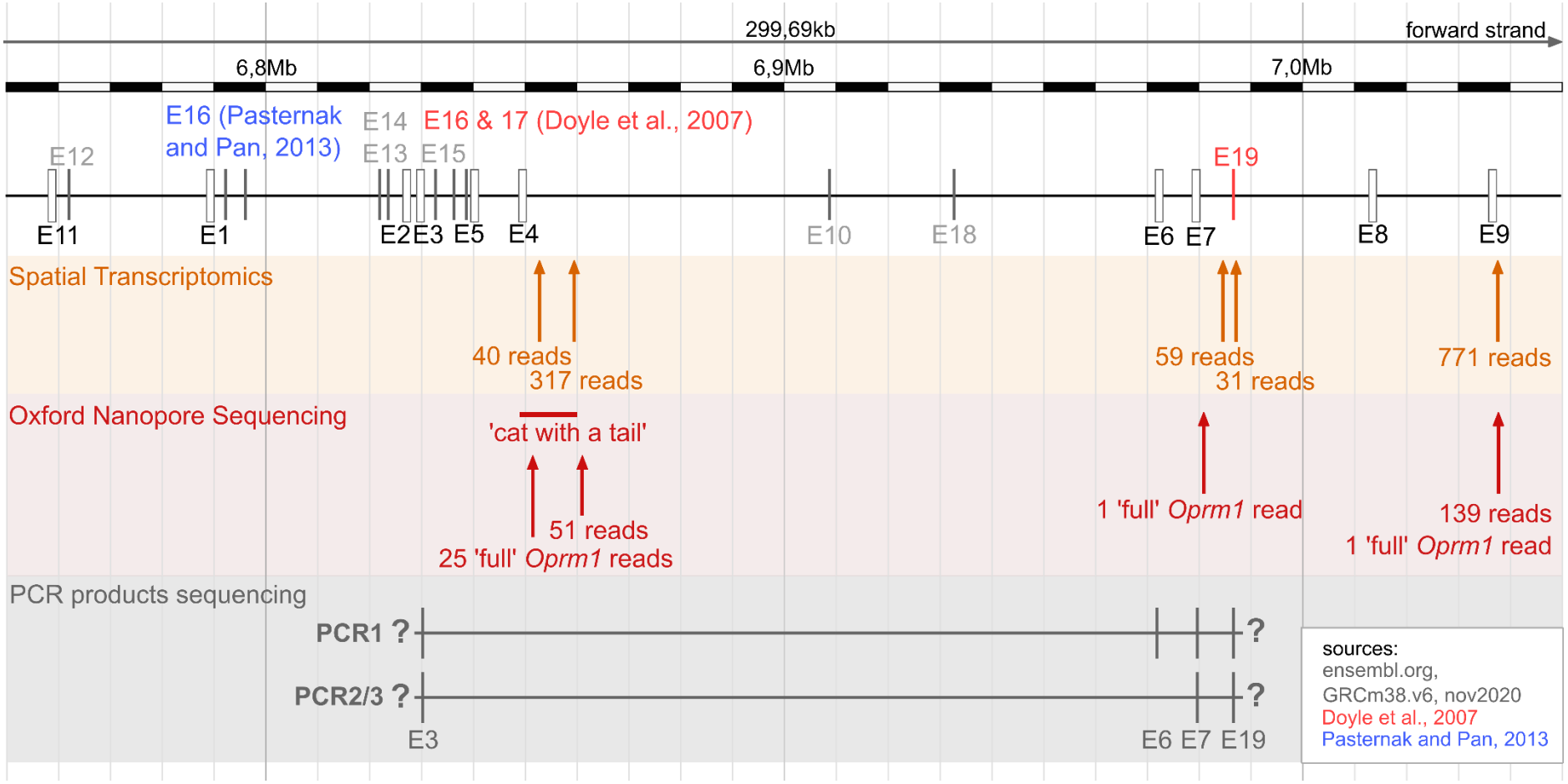

Figure 17. Graphical summary of the experiments. In the upper row - schematic visualization of the *Oprm1* gene (based on Ensembl.org). Summary of the results obtained from Spatial Transcriptomics in orange, peaks of reads are indicated by arrows. The ONS data from the striatum and the prefrontal cortex samples are indicated in red, the distal ends of the reads are indicated by arrows, long 3'UTR is labelled by as 'cat with a tail'. PCR product sequencing

visualization is presented in grey. Exons are indicated with vertical lines or rectangles. Sources: (Ensembl.org archive (GRCm38.v6, nov2020; Doyle et al., 2007; Pasternak & Pan, 2013)

6. Discussion

In this thesis I aimed to determine the presence of various transcripts of the *Oprm1* gene in the striatum and the prefrontal cortex of adult mice. The RT-PCR and ONS experiments enabled me for a detailed analysis of the *Oprm1* gene structure and provided evidence for the expression of specific transcript variants of the μ opioid receptor. Next, I tried to confirm the hypothesis of the long *Oprm1* transcript in the forebrain using RNAscope ISH. The results obtained are preliminary and further experiments should permit evaluating the specificity of distribution of the *Oprm1* transcripts.

Expression of *Oprm1* exon 4 and the long 3'UTR

The ONS results strongly suggest the existence of a long 3'UTR downstream exon 4. The observed contig expands beyond 10 kbp, and has not been mapped so far. The postulated length of exon 4 in the *Oprm1* transcripts varies between 53 and 808 bp, as reported in the Ensembl database. The length does not affect the produced protein size since the open reading frame is conserved, and a STOP codon occurs 39 bp downstream the beginning of the exon. There are several transcript variants reported that contain exon 4, among them those producing the canonical MOR-1 isoform. In the ONS experiment, the 5' end of exon 4 is clearly visible and converges with the Ensembl data. The 3' terminus, however, seems more diverse. There are multiple reads that contain exons 1 to 4 and then gradually diminish in counts between 6 850 050 kb and 6 851 400 kb. Some of them may be recognized as the reported variants *Oprm1-215, Oprm1-208,* and *Oprm1-203.* However, an early termination of exon 4 is not consistent with the data provided by Spatial Transcriptomics, where peaks of reads indicating 3'ends of mRNAs appear at 6 852 900 bp (40 reads) and at 6 860 000 bp (314 reads). Interestingly, despite the fact that full-length *Oprm1* transcripts terminate between 6 850 - 6 851 kbp, the high coverage of reads generated by the ONS expands up to 6 860 050 bp, creating specific contig described here as a 'cat with a tail'. The distal end of the contig matches the Spatial Transcriptomics results and is also consistent with a single-cell transcriptomics data acquired from neurons in the Tabula Muris Project (Schaum et al., 2018, Figure 18). Therefore, it is hypothesised that the visualized contig comprises for a 3'UTR of *Oprm1* transcripts that possess exons 1-4.

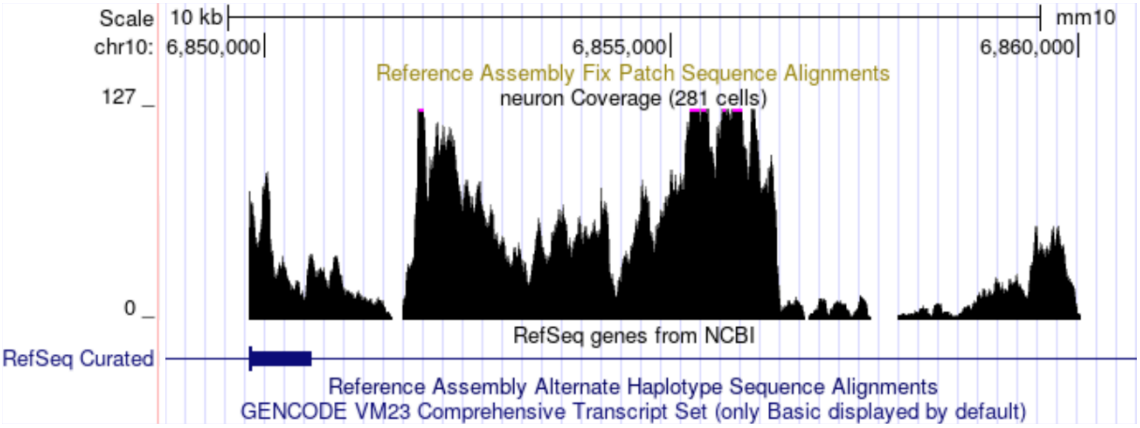

Figure 18. Visualization of Tabula Muris single cell sequencing on neuronal cells at 10:6 850 000 – 10:6 860 000 bp of the mouse genome (GRCm38). Picture obtained from the UCSC genome browser. 'RefSeq Curated' is a reference sequence that includes annotations whose accessions start with NM, NR, NP or YP prefixes in the NCBI database.

Several factors may contribute to the lack of 'full' *Oprm1* reads within the contig. First, the region analysed contains several repetitive elements, including one long-tandem repeat (LTR), three short interspersed nuclear elements (SINE; one mammalian-wide interspersed repeat and two Alu sequences), and 8 simple repeats of various sequences, as indicated by the RepeatMasker Open 4.0 tool (Smit et al., 2013). Both, LTR and SINE are transposable elements that can contribute to improper genome mapping and 'breakage' of a continuous sequence assembly (Tørresen et al., 2019). Additionally, even though, the long-read sequencing techniques should be able to span the repetitive sequences, they generate a relatively high error rate, caused mainly by frequent indels (insertion and deletions), which can negatively affect the correctness of genome alignment (Amarasinghe et al., 2020; Tørresen et al., 2019).

It should be noted that the discontinuity of reads in the ONS experiment was also observed, to some extent, in other genes, e.g. *Oprk1* and *Oprd1*. Although the genes were confirmed to contain 3 347 bp and 2 801 bp long 3 'UTRs, in both cases the last exons were not equally covered with reads. This may indicate that numerous transcripts were not mapped or sequenced as one read in the current experiment. While the ONS is known to generate 10-30 kb genomic libraries (with the longest reported length of 2.3 Mb), in our experiment only 4 826 out of ~68 million raw reads exceeded the length of 10 kbp. The possible explanation for the low long-read abundance includes fragmentation of the transcripts during the library preparation or nanopore blocking. The ONS is based on a protocol in which transcripts are sequenced towards the 5' end of the nucleic acid. This results in generation of truncated

reads may also result from a coverage bias towards the 3 'end' observed in this type of sequencing (Amarasinghe et al., 2020). During library preparation, transcription is carried out by a reverse transcriptase, which displays 3'-5' exonuclease activity. The enzyme types vary in their specifications and eficiency. Moreover, transcriptases display limted capability to deal with secondary structure mRNA, and stop their activity while recognizing such structure. Omniscript reverse transcriptase used to prepare the library in my experiment was found to have the highest transcription efficiency for mRNA fragments up to 1.2 kb (Martín-Alonso et al., 2021). Therefore, to improve the quality of cDNA, it would be beneficial to use enzyme that is dedicated to longer fragments and with higher detection sensitivity. An additional factor that may contibute to a vague picture of the 'cat with a tail' contig is a phenomenon of internal poli(A)s. It was shown that more than one poli(A) sequence can be included into a single mRNA molecule. During the library preparation, oligo(dT) primers used for trancription of mRNAs into cDNA can hybridize with the internal poli(A) and contribute to generation of truncated sequences (Nam et al., 2002). The phenomenon of internal poli(A) may be associated with the existence of alternative 3'UTRs, which were observed in neurons and will be described in the last section of the discussion.

Although the ONS results provide a massive amount of data, the aforementioned factors may account for a blurred overview of the real structure of the *Oprm1* gene. To finally confirm the ONS observations another suitable mRNA assay should be applied.

Confirmation of the long *Oprm1* 3'UTR in RNAscope ISH

To confirm the existence of the long 3'UTR, RNAscope ISH was performed. The main aim of the experiment was to colocalize the standard *Oprm1* probe with the custom-made probe targeting putative 3'UTR. The results did not provide satisfactory colocalization evidence, as the standard *Oprm1* probe produced a very low signal-to-noise ratio, impossible to distinguish by manual or automatic approaches. This may be partially explained by the properties of the fluorophore paired with the probe. The 3 'UTR targeting probe was conjugated with Alexa488, which is known to produce relatively high autofluorescence (Abuduwali et al., 2013). However, it cannot explain the complete lack of puncta. The other factor that can affect the results is the much longer time elapsed between the RNAscope assay and slides imaging than suggested (*RNAscope Fluorescent Multiplex Reagent Kit Part 2*, 2019).

The signal from the new, 3'UTR-targeting probe has nuclear localization, suggesting that the probe properly targets mRNAs. Furthermore, some characteristics of *Oprm1* expression were detected. Firstly, the puncta were observed to be more abundant in the dorsal striatum and the nucleus accumbens than in the cortical regions; secondly, the signal was present in both *Penk*-positive and *Penk*-negative cells, which is consistent with the literature on *Oprm1* gene expression (Banghart et al., 2015; Mansour et al., 1995; Xu et al., 2014). To unequivocally confirm the existence of a long *Oprm1* 3'UTR and analyse its distribution, it is necessary to repeat the RNAscope experiments with another set of fluorophores, following the suggested image acquisition time.

Expression of *Oprm1* exons 7-9

It should be noted that the discontinuity of reads does not neccessarily affect the 'cat with a tail' region only. In the ONS experiment few 'full' transcripts that contain exons 7-9 were observed. This may be surprising, as the variants with exon 7 were reported to account for a considrable fraction of transcribed *Oprm1* mRNA in the analysed brain regions (Xu et al., 2014). Although few 'full' transcripts were observed, there were many 'short' reads aligned to exon 9 in the ONS. Also, the Spatial Transcriptomics data show a robust genome coverage in this region (771 reads at 7 038 800 bp). This suggests a high abundance of *Oprm1* transcripts that possess exon 9.

In addition, PCR reactions provide evidence for transcription of several variants that expand beyond exon 4 and contain either exon 7 or the preceding exon 6. Noteworthily, the PCR revealed an unique transcriptional variant related to MOR-1U, but with a longer exon 19. The variant was not reported so far in the NCBI or Ensembl databases. If the revealed variant is transcribed into functional protein, then the carboxy terminal domain of the receptor would differ considerably from those present in the previously reported isoform, which, in turn, may affect its properties, for example, phospohorylation sites (Narayan et al., 2021). A caveat of this interpretation is that only 'partial' cDNA was sequenced. To assess the entire transcript sequence, the use of the RACE PCR (Rapid amplification of cDNA ends PCR) method would be beneficial.

Expression of *Oprm1* 5' terminal exons

According to the literature, exon 11 is predominantly expressed in transcripts encoding 6-TM isoforms – MOR-1G, MOR-1K, MOR-1L, MOR-1M and MOR-1N. Previous research showed robust exon 11-like immunoreactivity in the murine striatum (Abbadie et al., 2004), whereas mRNA levels for these isoforms in the striatum and

prefrontal cortex were scant (Xu et al., 2014). Here, neither the RT-PCR nor the ONS method provided evidence for exon 11 transcription in any of the regions examined. It may be due to the insufficient sensitivity of the methods used. However, the results do not exclude the expression of such transcripts in other regions of the central nervous system, e.g., the perigqueductal gray or the dorsal horn of a spinal cord.

Putative role of long 3'UTR in neuronal transcripts

In recent years, thousands of genes in both rodents and humans have been updated to have novel distant 3'UTRs. Substantially longer 3'UTRs were observed in neurons. Many of such transcripts were shown to possess >10kbp 3'UTR, and some of them exceeded >18kbp. The function of such a long untranslated region has not been fully established yet. It is suggested that the region contains sequences responsible for regulation of gene expression, stability and cellular localization of the mRNAs (Bae & Miura, 2020). For instance, research on slides from rodent hippocampus showed that the 3'UTRs found in the neuropil influenced longer mRNAs half-lives than the 3'UTRs in the somatic compartment (Tushev et al., 2018). Conversly, a knock-out of this region in the mTOR gene significantly limited the distribution of mRNAs to specific neuronal compartments (Terenzio et al., 2018).

Neuronal transcripts commonly contain several alternate 3'UTRs. Diversity is caused by alternative polyadenylation sites used during the generation of mRNAs. Up to 70% of protein-coding genes are estimated to have multiple poli(A) sites. Multiple poli(A)s may be associated with alternative splicing and variability in the last exon included in the transcript. However, in most events, alternative polyadenylation occurs in variants with the same open reading frame, but different 3' UTRs, the so-called 'tandem' 3'UTRs (Bae & Miura, 2020). 'Tandem' 3'UTRs were shown to generate mRNAs of different lengths, with a particular preference for longer transcripts observed in the central nervous system (Miura et al., 2013).

Such long 3'UTRs usually possess several microRNAs (miRNAs) binding sites, which contribute to a more precise regulation of gene expression, especially important in neurons. If the 'cat with a tail' is included in *Oprm1* transcripts, then the expression of μ opioid receptor can be modulated by miRNA targeting the region. According to the miRNA database, mirbase.org (Griffiths-Jones et al., 2006), the region contains two possible miRNA binding sites, targeted by mmu-mir-3470b and mmu-mir-3470a. The

miRNAs were identified in the ovary of newborn mice; their existence in the central nervous system was not evaluated (Ahn et al., 2010).

Conclusion

To conclude, my results provide evidence for a long 3'UTR downstream exon 4 in the murine *Oprm1* gene, which according to the data obtained in the ONS experiment represents for a considerable fraction of the transcribed μ opioid receptor. Furthermore, the data enabled me to determine the existence of splicing variants containing exon 19, including 'partial' cDNA of a form that was not previously reported. Finally, the results show a predominant expression of mRNAs containing exon 1, and no evidence of exon 11 transcription in the striatum and the prefrontal cortex. The data obtained so far confirm a highly complex architecture of the *Oprm1* gene. In future research, I would like to focus on assessing the cellular distribution of the long *Oprm1* 3'UTR in the brain, by continuing the experiments utilizing RNAscope assay. As long as the significance of 3'UTR sequences in the μ opioid receptor has not been fully elucidated yet, I would also like to investigate the putative role of the sequence in the context of the receptor function and potential role in its pathologies.

List of figures

List of tables

References

Abbadie, C., Pan, Y.-X., & Pasternak, G. W. (2004). Immunohistochemical study of the expression of EXON11-containing µ opioid receptor variants in mouse brain. *Neuroscience*, *127*(2), 419–430. https://doi.org/10.1016/j.neuroscience.2004.03.033

Abuduwali, N., Lossdörfer, S., Winter, J., Wolf, M., Götz, W., & Jäger, A. (2013). Autofluorescent characteristics of human periodontal ligament cells in vitro. *Annals of Anatomy - Anatomischer Anzeiger*, *195*(5), 449–454. https://doi.org/10.1016/j.aanat.2013.03.007

Ahn, H. W., Morin, R. D., Zhao, H., Harris, R. A., Coarfa, C., Chen, Z.-J., Milosavljevic, A., Marra, M. A., & Rajkovic, A. (2010). MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing. *Molecular Human Reproduction*, *16*(7), 463–471. https://doi.org/10.1093/molehr/gaq017

Akil, H., Watson, S. J., Young, E., Lewis, M. E., Khachaturian, H., & Walker, J. M. (1984). Endogenous opioids: Biology and function. *Annual Review of Neuroscience*, *7*, 223–255. https://doi.org/10.1146/annurev.ne.07.030184.001255

Al-Hasani, R., & Bruchas, M. R. (2011). Molecular Mechanisms of Opioid Receptor-dependent Signaling and Behavior: *Anesthesiology*, 1. https://doi.org/10.1097/ALN.0b013e318238bba6

Allouche, S., Noble, F., & Marie, N. (2014). Opioid receptor desensitization: Mechanisms and its link to tolerance. *Frontiers in Pharmacology*, *5*. https://www.frontiersin.org/article/10.3389/fphar.2014.00280

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., & Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biology*, *21*(1), 30. https://doi.org/10.1186/s13059-020-1935-5

Bae, B., & Miura, P. (2020). Emerging Roles for 3' UTRs in Neurons. *International Journal of Molecular Sciences*, *21*(10), E3413. https://doi.org/10.3390/ijms21103413

Banghart, M. R., Neufeld, S. Q., Wong, N. C., & Sabatini, B. L. (2015). Enkephalin disinhibits mu opioid receptor rich striatal patches via delta opioid receptors. *Neuron*, *88*(6), 1227–1239. https://doi.org/10.1016/j.neuron.2015.11.010

Beckett, A. H., & Casy, A. F. (1954). Synthetic analgesics: Stereochemical considerations. *The Journal of Pharmacy and Pharmacology*, *6*(12), 986–1001. https://doi.org/10.1111/j.2042-7158.1954.tb11033.x

Bohn, L. M., Gainetdinov, R. R., Lin, F.-T., Lefkowitz, R. J., & Caron, M. G. (2000). μ-Opioid receptor desensitization by β-arrestin-2 determines morphine tolerance but not dependence. *Nature*, *408*(6813), 720–723. https://doi.org/10.1038/35047086

Chen, Y., Mestek, A., Liu, J., Hurley, J. A., & Yu, L. (1993). Molecular cloning and functional expression of a mu-opioid receptor from rat brain. *Molecular Pharmacology*, *44*(1), 8–12.

Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques*, *15*(3), 532–534, 536–537.

Cuitavi, J., Hipólito, L., & Canals, M. (2021). The Life Cycle of the Mu-Opioid Receptor. *Trends in Biochemical Sciences*, *46*(4), 315–328. https://doi.org/10.1016/j.tibs.2020.10.002

Doyle, G. A., Sheng, X. R., Lin, S. S. J., Press, D. M., Grice, D. E., Buono, R. J., Ferraro, T. N., & Berrettini, W. H. (2007). Identification of five mouse mu-opioid receptor (MOR) gene (Oprm1) splice variants containing a newly identified alternatively spliced exon. *Gene*, *395*(1–2), 98–107. https://doi.org/10.1016/j.gene.2007.02.004

Ebrahimi, G., Asadikaram, G., Akbari, H., Nematollahi, M. H., Abolhassani, M., Shahabinejad, G., Khodadadnejad, L., & Hashemi, M. (2018). Elevated levels of DNA methylation at the OPRM1 promoter region in men with opioid use disorder. *The American Journal of Drug and Alcohol Abuse*, *44*(2), 193–199. https://doi.org/10.1080/00952990.2016.1275659

Emery, M. A., & Akil, H. (2020). Endogenous Opioids at the Intersection of Opioid Addiction, Pain, and Depression: The Search for A Precision Medicine Approach. *Annual Review of Neuroscience*, *43*, 355–374. https://doi.org/10.1146/annurev-neuro-110719-095912

Evans, C. J., Keith, D. E., Morrison, H., Magendzo, K., & Edwards, R. H. (1992). *Cloning of a Delta Opioid Receptor by Functional Expression*. https://www.science.org/doi/10.1126/science.1335167

Férézou, I., Hill, E. L., Cauli, B., Gibelin, N., Kaneko, T., Rossier, J., & Lambolez, B. (2007). Extensive overlap of mu-opioid and nicotinic sensitivity in cortical interneurons. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(8), 1948–1957. https://doi.org/10.1093/cercor/bhl104

Gene Detail: Allen Brain Atlas: Mouse Brain. (2011). http://mouse.brain-map.org/gene/show/18157

Gillis, A., Kliewer, A., Kelly, E., Henderson, G., Christie, M. J., Schulz, S., & Canals, M. (2020). Critical Assessment of G Protein-Biased Agonism at the μ-Opioid Receptor. *Trends in Pharmacological Sciences*, *41*(12), 947–959. https://doi.org/10.1016/j.tips.2020.09.009

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., & Enright, A. J. (2006). miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Research*, *34*(Database issue), D140–D144. https://doi.org/10.1093/nar/gkj112

Gu, Z.-H., Wang, B., Kou, Z.-Z., Bai, Y., Chen, T., Dong, Y.-L., Li, H., & Li, Y.-Q. (2017). Endomorphins: Promising Endogenous Opioid Peptides for the Development of Novel Analgesics. *Neurosignals*, *25*(1), 98–116. https://doi.org/10.1159/000484909

Koehl, A., Hu, H., Maeda, S., Zhang, Y., Qu, Q., Paggi, J. M., Latorraca, N. R., Hilger, D., Dawson, R., Matile, H., Schertler, G. F. X., Granier, S., Weis, W. I., Dror, R. O., Manglik, A., Skiniotis, G., & Kobilka, B. K. (2018). Structure of the μ-opioid receptor–Gi protein complex. *Nature*, *558*(7711), 547–552. https://doi.org/10.1038/s41586-018-0219-7

Larhammar, D., Bergqvist, C., & Sundström, G. (2015). Ancestral Vertebrate Complexity of the Opioid System. In *Vitamins & Hormones* (Vol. 97, pp. 95–122). Elsevier. https://doi.org/10.1016/bs.vh.2014.11.001

Law, P.-Y., Wong, Y. H., & Loh, H. H. (2000). Molecular Mechanisms and Regulation of Opioid Receptor Signaling. *Annual Review of Pharmacology and Toxicology*, *40*(1), 389–430. https://doi.org/10.1146/annurev.pharmtox.40.1.389

Lord, J. A. H., Waterfield, A. A., Hughes, J., & Kosterlitz, H. W. (1977). Endogenous opioid peptides: Multiple agonists and receptors. *Nature*, *267*(5611), 495–499. https://doi.org/10.1038/267495a0

Lu, Z., Xu, Jin., Xu, M., Rossi, G. C., Majumdar, S., Pasternak, G. W., & Pan, Y.-X. (2018). Truncated mu opioid receptors with six transmembrane domains are essential for opioid analgesia. *Anesthesia and Analgesia*, *126*(3), 1050–1057. https://doi.org/10.1213/ANE.0000000000002538

Majumdar, S., Grinnell, S., Le Rouzic, V., Burgman, M., Polikar, L., Ansonoff, M., Pintar, J., Pan, Y.-X., & Pasternak, G. W. (2011). Truncated G protein-coupled mu opioid receptor MOR-1 splice variants are targets for highly potent opioid analgesics lacking side effects. *Proceedings of the National Academy of*

Sciences of the United States of America, *108*(49), 19778–19783. https://doi.org/10.1073/pnas.1115231108

Mansour, A., Fox, C. A., Akil, H., & Watson, S. J. (1995). Opioid-receptor mRNA expression in the rat CNS: Anatomical and functional implications. *Trends in Neurosciences*, *18*(1), 22–29. https://doi.org/10.1016/0166-2236(95)93946-U

Mansour, A., Khachaturian, H., Lewis, M. E., Akil, H., & Watson, S. J. (1988). Anatomy of CNS opioid receptors. *Trends in Neurosciences*, *11*(7), 308–314. https://doi.org/10.1016/0166-2236(88)90093-8

Martin, W. R., Eades, C. G., Thompson, J. A., Huppler, R. E., & Gilbert, P. E. (1976). The effects of morphine- and nalorphine- like drugs in the nondependent and morphine-dependent chronic spinal dog. *The Journal of Pharmacology and Experimental Therapeutics*, *197*(3), 517–532.

Martín-Alonso, S., Frutos-Beltrán, E., & Menéndez-Arias, L. (2021). Reverse Transcriptase: From Transcriptomics to Genome Editing. *Trends in Biotechnology*, *39*(2), 194–210. https://doi.org/10.1016/j.tibtech.2020.06.008

Meng, F., Xie, G. X., Thompson, R. C., Mansour, A., Goldstein, A., Watson, S. J., & Akil, H. (1993). Cloning and pharmacological characterization of a rat kappa opioid receptor. *Proceedings of the National Academy of Sciences*, *90*(21), 9954–9958. https://doi.org/10.1073/pnas.90.21.9954

Minami, M., Onogi, T., Toya, T., Katao, Y., Hosoi, Y., Maekawa, K., Katsumata, S., Yabuuchi, K., & Satoh, M. (1994). Molecular cloning and in situ hybridization histochemistry for rat mu-opioid receptor. *Neuroscience Research*, *18*(4), 315–322. https://doi.org/10.1016/0168-0102(94)90167-8

Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., & Lai, E. C. (2013). Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. *Genome Research*, *23*(5), 812–825. https://doi.org/10.1101/gr.146886.112

Nam, D. K., Lee, S., Zhou, G., Cao, X., Wang, C., Clark, T., Chen, J., Rowley, J. D., & Wang, S. M. (2002). Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proceedings of the National Academy of Sciences*, *99*(9), 6152–6156. https://doi.org/10.1073/pnas.092140899

NanoDrop Spectrophotometers Nucleic Acid Purity Ratios T042 Technical Biuletin. (2009). Thermo Scientific.

Narayan, A., Hunkele, A., Xu, J., Bassoni, D. L., Pasternak, G. W., & Pan, Y.-X. (2021). Mu Opioids Induce Biased Signaling at the Full-Length Seven Transmembrane C-Terminal Splice Variants of the mu Opioid Receptor Gene, Oprm1. *Cellular and Molecular Neurobiology*, *41*(5), 1059–1074. https://doi.org/10.1007/s10571-020-00973-5

Oude Ophuis, R. J. A., Boender, A. J., van Rozen, A. J., & Adan, R. A. H. (2014). Cannabinoid, melanocortin and opioid receptor expression on DRD1 and DRD2 subpopulations in rat striatum. *Frontiers in Neuroanatomy*, *8*, 14. https://doi.org/10.3389/fnana.2014.00014

Pan, Y.-X. (2002). Identification and characterization of a novel promoter of the mouse mu opioid receptor gene (Oprm) that generates eight splice variants. *Gene*, *295*(1), 97–108. https://doi.org/10.1016/s0378-1119(02)00825-9

Pasternak, G. W. (2001). Insights into mu opioid pharmacology: The role of mu opioid receptor subtypes. *Life Sciences*, *68*(19), 2213–2219. https://doi.org/10.1016/S0024-3205(01)01008-6

Pasternak, G. W., & Pan, Y.-X. (2013). Mu Opioids and Their Receptors: Evolution of a Concept. *Pharmacological Reviews*, *65*(4), 1257–1317. https://doi.org/10.1124/pr.112.007138

Paxinos, G., & Franklin, K. B. J. (2001). *The Mouse Brain in Stereotaxic Coordinates* (second edition). Academic Press.

Pert, C. B., & Snyder, S. H. (1973). Opiate receptor: Demonstration in nervous tissue. *Science (New York, N.Y.)*, *179*(4077), 1011–1014. https://doi.org/10.1126/science.179.4077.1011

Ponterio, G., Tassone, A., Sciamanna, G., Riahi, E., Vanni, V., Bonsi, P., & Pisani, A. (2013). Powerful inhibitory action of mu opioid receptors (MOR) on cholinergic interneuron excitability in the dorsal striatum. *Neuropharmacology*, *75*, 78–85. https://doi.org/10.1016/j.neuropharm.2013.07.006

Portoghese, P. S. (1965). A new concept on the mode of interaction of narcotic analgesics with receptors. *Journal of Medicinal Chemistry*, *8*(5), 609–616. https://doi.org/10.1021/jm00329a013

RNAscope Fluorescent Multiplex Reagent Kit part 2. (2019). Advenced Cell Diagnostics, Inc.

Schaum, N., Karkanias, J., Neff, N. F., May, A. P., Quake, S. R., Wyss-Coray, T., Darmanis, S., Batson, J., Botvinnik, O., Chen, M. B., Chen, S., Green, F., Jones, R. C., Maynard, A., Penland, L., Pisco, A. O., Sit, R. V., Stanley, G. M., Webber, J. T., … Principal investigators. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*(7727), 367–372. https://doi.org/10.1038/s41586-018-0590-4

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., & Ragg, T. (2006). The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology*, *7*(1), 3. https://doi.org/10.1186/1471-2199-7-3

Siegelbaum, S. A., Clapham, D. E., & Schwartz, J. H. (2013). Modulation of Synaptic Transmission: Second Messengers. In E. R. Kandel (Ed.), *Principles of Neural Science* (5th ed, pp. 236–259). McGraw-Hill.

Smit, A., Hubley, R., & Green, P. (2013). *RepeatMasker Open* (4.0). http://www.repeatmasker.org

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P. I., Sahlén, P., Mulder, J., Bergmann, O., … Frisén, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78–82. https://doi.org/10.1126/science.aaf2403

Sun, N., Yu, L., Gao, Y., Ma, L., Ren, J., Liu, Y., Gao, D. S., Xie, C., Wu, Y., Wang, L., Hong, J., & Yan, M. (2021). MeCP2 Epigenetic Silencing of Oprm1 Gene in Primary Sensory Neurons Under Neuropathic Pain Conditions. *Frontiers in Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnins.2021.743207

Taki, K., Kaneko, T., & Mizuno, N. (2000). A group of cortical interneurons expressing mu-opioid receptor-like immunoreactivity: A double immunofluorescence study in the rat cerebral cortex. *Neuroscience*, *98*(2), 221–231. https://doi.org/10.1016/s0306-4522(00)00124-x

Terenzio, M., Koley, S., Samra, N., Rishal, I., Zhao, Q., Sahoo, P. K., Urisman, A., Marvaldi, L., Oses-Prieto, J. A., Forester, C., Gomes, C., Kalinski, A. L., Di Pizio, A., Doron-Mandel, E., Perry, R. B.-T., Koppel, I., Twiss, J. L., Burlingame, A. L., & Fainzilber, M. (2018). Locally translated mTOR controls axonal local translation in nerve injury. *Science*, *359*(6382), 1416–1421. https://doi.org/10.1126/science.aan1053

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Tørresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., Gruca, A., Grynberg, M., Kajava, A. V., Promponas, V. J., Anisimova, M., Jakobsen, K. S., & Linke, D. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Research*, *47*(21), 10994–11006. https://doi.org/10.1093/nar/gkz841

Tushev, G., Glock, C., Heumüller, M., Biever, A., Jovanovic, M., & Schuman, E. M. (2018). Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron*, *98*(3), 495-511.e6. https://doi.org/10.1016/j.neuron.2018.03.030

Wang, F., Flanagan, J., Su, N., Wang, L.-C., Bui, S., Nielson, A., Wu, X., Vo, H.-T., Ma, X.-J., & Luo, Y. (2012). RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues. *The Journal of Molecular Diagnostics*, *14*(1), 22–29. https://doi.org/10.1016/j.jmoldx.2011.08.002

Wang, H., & Pickel, V. M. (1998). Dendritic spines containing μ-opioid receptors in rat striatal patches receive asymmetric synapses from prefrontal corticostriatal afferents. *Journal of Comparative Neurology*, *396*(2), 223–237. https://doi.org/10.1002/(SICI)1096-9861(19980629)396:2<223::AID-CNE7>3.0.CO;2-2

Williams, J. T., Ingram, S. L., Henderson, G., Chavkin, C., von Zastrow, M., Schulz, S., Koch, T., Evans, C. J., & Christie, M. J. (2013). Regulation of μ -Opioid Receptors: Desensitization, Phosphorylation, Internalization, and Tolerance. *Pharmacological Reviews*, *65*(1), 223–254. https://doi.org/10.1124/pr.112.005942

Wu, S., Bono, J., & Tao, Y.-X. (2019). Long noncoding RNA (lncRNA): A target in neuropathic pain. *Expert Opinion on Therapeutic Targets*, *23*(1), 15–20. https://doi.org/10.1080/14728222.2019.1550075

Xu, C., Liu, G., Ji, H., Chen, W., Dai, D., Chen, Z., Zhou, D., Xu, L., Hu, H., Cui, W., Chang, L., Zha, Q., Li, L., Duan, S., & Wang, Q. (2018). Elevated methylation of OPRM1 and OPRL1 genes in Alzheimer's disease. *Molecular Medicine Reports*, *18*(5), 4297–4302. https://doi.org/10.3892/mmr.2018.9424

Xu, J., Lu, Z., Xu, M., Rossi, G. C., Kest, B., Waxman, A. R., Pasternak, G. W., & Pan, Y.-X. (2014). Differential Expressions of the Alternatively Spliced Variant mRNAs of the μ Opioid Receptor Gene, OPRM1, in Brain Regions of Four Inbred Mouse Strains. *PLOS ONE*, *9*(10), e111267. https://doi.org/10.1371/journal.pone.0111267

Xu, J., Xu, M., Brown, T., Rossi, G. C., Hurd, Y. L., Inturrisi, C. E., Pasternak, G. W., & Pan, Y.-X. (2013). Stabilization of the μ-opioid receptor by truncated single transmembrane splice variants through a chaperone-like action. *The Journal of Biological Chemistry*, *288*(29), 21211–21227. https://doi.org/10.1074/jbc.M113.458687

Zadina, J. E., Hackler, L., Ge, L.-J., & Kastin, A. J. (1997). A potent and selective endogenous agonist for the μ-opiate receptor. *Nature*, *386*(6624), 499–502. https://doi.org/10.1038/386499a0

Zhang, H., Herman, A. I., Kranzler, H. R., Anton, R. F., Simen, A. A., & Gelernter, J. (2012). Hypermethylation of OPRM1 promoter region in European Americans with alcohol dependence. *Journal of Human Genetics*, *57*(10), 670–675. https://doi.org/10.1038/jhg.2012.98

Appendix

Raw data obtained in PCR product sequencing

>PCR1-PF3
GCAGTCCCTATCGGTCAACACTAGGGACACCCCTCCACGGCTAATACAGTGGATCGAACTAACCACCAGAAGAAAAAGCT
GGACTCCCAGAGAGGGTGTGTACAGCATCCAGTGTGACCTGTCCCTTGTCTTTGAGCCTGGGGGCCATCTTCTTTCACAG
CATACCATTTCCTTGTATCCTCTCTGAAGCCAACCCTGGCAGTCAGCGTGGCCCAGATCTTTACAGGATATCCTTCTCCG
ACTCATGTTGAAAAACCCTGCAAGAGTTGCATGGACAGACTGCTACAACAGAAAACAACAGACTGGGTCACTCAGAAAAA
ATAAAAAGAAGAAGAAGAGGAGGAAGAACAAGCAAAATATCCTAGAAGCTGGAATATCTAGGCAAGGTCAAGTTGCTAAC
AATGAAGATGACCTCTTCCTGACTTGCAGACATCTATGATCTTAATATAACCTCACATGGCAAATAGGAAATTCCTGTTC
ACCAGATAAAGGCACAACTGCCATCGATGAGGACGCCACCCTCATGACCTATGACCTCTCAACACATCTCTGAACAGCAT
CCCATTTTGAATTGAATTAGGACTTGGGCATGGAAAAGGGGGGCAA

>PCR1-PR_6990kb
GCAACTTAAATGGGATGCTGTTCGAGATGTGTTGAGAGGTCATAGGTCATGAGGGTGGCGTCCTCATCGATGGCAGTTGT
GCCTTTATCTGGTGAACAGGAATTTCCTATTTGCCATGTGAGGTTATATTAAGATCATAGATGTCTGCAAGTCAGGAAGA
GGTCATCTTCATTGTTAGCAACTTGACCTTGCCTAGATATTCCAGCTTCTAGGATATTTTGCTTGTTCTTCCTCCTCTTC
TTCTTCTTTTTATTTTTTCTGAGTGACCCAGTCTGTTGTTTTCTGTTGTAGCAGTCTGTCCATGCAACTCTTGCAGGGTT
TTTCAACATGAGTCGGAAAAGGATATCCTGTAAAGATCTGGGCCACGCTGACTGCCAGGGTTGGCTTCAGAGAGGATACA
AGGAAATGGTATGCTGTGAAAGAAGATGGCCCCCAGGCTCAAAGACAAGGGACAGGTCACACTGGATGCTGTACACACCC
TCTCTGGGAGTCCAGCTTTTTCTTCTGGTGGTTAGTTCGATCCACTGTATTAGCCGTGGAGGGGTGTTCCCTAGTGTTTT
GACGGATTCGAGCAGAGTTTTGCTGTTCGATGGGGGAAGAAA

>PCR2-PF3
CCTGACTCGATCGTCAACACTAGGGACACCCCTCCACGGCTAATACAGTGGATCGAACTAACCACCAGCCAACCCTGGCA
GTCAGCGTGGCCCAGATCTTTACAGGATATCCTTCTCCGACTCATGTTGAAAAACCCTGCAAGAGTTGCATGGACAGTGT
GCACTGCTACAACACAAAACAACGGACTGGGTCACTCAAAAAAAATAAAAAGAAAAAGAAGAAGAAGAAGAACAAGCAAA
ATATCCTATAAACTGGAATATCTAGGCAAGGTCAAGTTGCTAACAATGAACATGACCTCTTCCTGACTTGCACACTTCTA
TGATCTTAATATAACCTCACATGGCAAATAGGAAATTCCTGTTCACCAAATAAAGGCACAACTGCCATCGATGACGACCC
CCCCCTCATGACCTATGACCTCTCATCACTTCTCTGAACACCATCCCATTTTGAATTGAATTAGGACTTGGCATGTGGAA
TGTGGCAAGTTTTTGCCCCCCCCCCCCGTTGGAGCCTCAAAAATAACAAGATAGCGAGAGTGATCTAATATTTTCTCACCA
TCTTTATTAATATTGTATAGGTGGTGTTC

>PCR2-PR_6990kb
CCCATTTTCAAATGGGATGCTGTTCAGAGATGTGTTGAGAGGTCATAGGTCATGAGGGTGGCGTCCTCATCGATGGCAGT
TGTGCCTTTATCTGGTGAACAGGAATTTCCTATTTGCCATGTGAGGTTATATTAAGATCATAGATGTCTGCAAGTCAGGA
AGAGGTCATCTTCATTGTTAGCAACTTGACCTTGCCTAGATATTCCAGCTTCTAGGATATTTTGCTTGTTCTTCCTCCTC
TTCTTCTTCTTTTTATTTTTTCTGAGTGACCCAGTCTGTTGTTTTCTGTTGTAGCAGTCTACCCTGTCCATGCTACTCTT
GCTTGTTTTTTCAACTTGAATCGGAAAAGGATATCCTGAAAAGATCTGGGCCACCCTGACTGCCAGGGTGGGGTGGTGGT
TAATTCAATCCATTGTATTAGCCGGGGAGGGGTGTTCCCTATTGTTTTGACCGATTCGAGCATAGTTTTGCTGTTCGATA
GTAAAGGAATTTTTCCCCCCCCCCCCCCTGGTTTGTCATTATTACAAATAATAGGAGAAGTGATCAATTCCTTTCTCGCC
ATGATTAATATTTATTATAGGTCAGCATCATCCGGGTATCGATGGAAACACATCCACATTATATCCATTCGAGTCGATCA
ACTCTGACATCACATCCCATAATTTCTCTATGTGCCGATAGTTTGAGCACAGTCTGGCTGGAAACAAGACAGCACGATCC
GGCCCAGATGGAAAGGATCTGAACCGAAAATCATTTTCATTTCCGCAATTGCATTGCTATGATTCCCAACCCCCAAAACA
AAAATAAAAAGCCCCCTTCGACCTTCCCCCGGGGGGCGTGGATTTTTTCAATAATCCCCCCCCCCCAATCAGTCTAATGGA
TAT

>PCR3-PF2/3
CCAACTAAACTCATATGAGAGTGAGTACTGGGAGACCTGCTCAAAATCTGTGTCTTCATCTTCGCCTTCATCATGCCGGT
CCTCATCATCACTGTGTGTTATGGACTGATGATCTTACGACTCAAGAGTGTCCGCATGCTGTCGGGCTCCAAAGAAAAGG
ACAGGAACCTGCGCAGGATCACCCGGATGGTGCTGGTGGTCGTGGCTGTATTTATTGTCTGCTGGACCCCCATCCACATC
TATGTCATCATCAAAGCACTGATCACGATTCCAGAAACCACTTTCCAGACTGTTTCCTGGCACTTCTGCATTGCCTTGGG
TTACACAAACAGCTGCCTGAACCCAGTTCTTTATGCGTTCCTGGATGAAAACTTCAAACGATGTTTTAGAGAGTTCTGCA
TCCCAACTTCCTCCACAATCGAACAGCAAAACTCTGCTCGAATCCGTCAAAACACTAGGGAACACCCCTCCACGGCTAAT
ACAGTGGATCGAACTAACCACCAGCCAACCCTGGCAGTCAGCGTGGCCCAGATCTTTACAGGATATCCTTCTCCGACTCA
TGTTGAAAAACCCTGCAAGAGTTGCATGGACAGACTGCTACAACAGAAAACAACAGACTGGGTCACTCCGAAAAAATAAA
AAGAAGAAGAAGAGGAGGAAGAACAAGCAAAATATCCTATAAGCTGGAATATCTAGGCAAGGTCAAGTTGCTAACAATGA
AGATGACCTCTTCCTGACTTGCAGACATCTATGATCTTAATATAACCTCACATGGCAAATAGGATATTCCTGTTCACCAC
ATAAAGGCACAACTGCCATCGATGAGGACACCACCCTCATGACCTATGACCTCTCAACACATCTCTGAACAGCATCCCAT
TTTGAATTGAATTAGGACTGCTGGGGGGGGGGGGGGACCA

>PCR3-PR_6990kb
AAACAGGCAAATGCGTGATGCTGTTCGAGATGTGTTGAGAGGTCATAGGTCATGAGGGTGGCGTCCTCATCGATGGCAGT
TGTGCCTTTATCTGGTGAACAGGAATTTCCTATTTGCCATGTGAGGTTATATTAAGATCATAGATGTCTGCAAGTCAGGA
AGAGGTCATCTTCATTGTTAGCAACTTGACCTTGCCTAGATATTCCAGCTTCTAGGATATTTTGCTTGTTCTTCCTCCTC
TTCTTCTTCTTTTTATTTTTTCTGAGTGACCCAGTCTGTTGTTTTCTGTTGTAGCAGTCTGTCCTTGCAACTCTTGCAGG
GTTTTTCAACATGAGTCGGAAAAGGATATCCTGTAAAGATCTGGGCCACGCTGACTGCCAGGGTTGGGTGGTGGTTAGTT
CAATCCACTGTATTAGCCGTGGAGGGGTGTTCCCTACTGTTTTGACGGATTCGAGCAGAGTTTTGCTGTTCTATTGTGGA
GGAAGTTGGGTTGCAAACCTCTCTAAAACATCGTTTGAATTTTTCATCCAAGAACGCAAAAAAAACTGGGTTCAGGCAGC
TGTTTGTGTAACCCAAGGCAATGCAAAAGTGCCAGGAAACAGTCTGGAAAGTGGTTTCTGGAATCGTGATCAGTGCTTTG
ATGATGACATAAATGTGGAGGGGGGTCCAGCACACAATAAATACAGCCACGACCACCACCCCCATCCGGGTGATCCTGCG

```
CAGGTTCCTGTCCTTTTCTTTGGAGCGCGACAGCATGCGGACACTCTTGAGTCGTAAGATCATCTCTCCATACCACACAG
TGATGATGAGGACCGGCATGATGAGAGCAAAGATGAAAACACAGATTTTGAGCAGGTTCTCCCATCACCATGTGATATGA
GAGAACATGAGGGTGCATCATGCCCCCCCCCCCCCCAA
```

The PCR product sequencing aligment to the genome (GRCm38). Analysis generated using the BLAT tool, data presented for both forward and reverse primers for each reaction:

PCR1-PF3 (+ strand)

```
0000017 aacactaggg.acacccctccacggctaatacagtggatcgaactaacca 0000065
>>>>>>> |||||||||| |||||||||||||||||||||||||||||||||||||| >>>>>>>
6830444 aacactagggaacacccctccacggctaatacagtggatcgaactaacca 6830493

0000066 ccag 0000069
>>>>>>> |||| >>>>>>>
6830494 ccag 6830497
```

```
0000070 aagaaaaagctggactcccagagagggtgtgtacagcatccagtgtgacc 0000119
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6972199 aagaaaaagctggactcccagagagggtgtgtacagcatccagtgtgacc 6972248

0000120 tgtcccttgtctttgagcctgggggccatcttctttcacagcataccatt 0000169
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6972249 tgtcccttgtctttgagcctgggggccatcttctttcacagcataccatt 6972298

0000170 tccttgtatcctctctgaag 0000189
>>>>>>> |||||||||||||||||||| >>>>>>>
6972299 tccttgtatcctctctgaag 6972318
```

```
0000190 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 0000239
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6979616 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 6979665

0000240 gactcatgttgaaaaaccctgcaagagttgcatggacag 0000278
>>>>>>> |||||||||||||||||||||||||||||||||||||| >>>>>>>
6979666 gactcatgttgaaaaaccctgcaagagttgcatggacag 6979704
```

```
0000279 actgctacaacagaaaacaacagactgggtcactcagaaaaaataaaaag 0000328
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985683 actgctacaacagaaaacaacagactgggtcactcagaaaaaataaaaag 6985732

0000329 aagaagaagaggaggaagaacaagcaaaatatcctagaagctggaatatc 0000378
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985733 aagaagaagaggaggaagaacaagcaaaatatcctagaagctggaatatc 6985782

0000379 taggcaaggtcaagttgctaacaatgaagatgacctcttcctgacttgca 0000428
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985783 taggcaaggtcaagttgctaacaatgaagatgacctcttcctgacttgca 6985832

0000429 gacatctatgatcttaatataacctcacatggcaaataggaaattcctgt 0000478
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985833 gacatctatgatcttaatataacctcacatggcaaataggaaattcctgt 6985882

0000479 tcaccagataaaggcacaactgccatcgatgaggacgccaccctcatgac 0000528
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985883 tcaccagataaaggcacaactgccatcgatgaggacgccaccctcatgac 6985932

0000529 ctatgacctctcaacacatctctgaacagcatcccattttgaattgaatt 0000578
```

```
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985933 ctatgacctctcaacacatctctgaacagcatcccattttgaattgaatt 6985982

0000579 aggacttgggcatg 0000592
>>>>>>> |||||||||||||| >>>>>>>
6985983 aggacttgggcatg 6985996
```

PCR1-PR_6990kb (- strand)

```
0000008 aaatgggatgctgttc.gagatgtgttgagaggtcataggtcatgagggt 0000056
<<<<<<< |||||||||||||||| ||||||||||||||||||||||||||||||||| <<<<<<<
6985971 aaatgggatgctgttcagagatgtgttgagaggtcataggtcatgagggt 6985922

0000057 ggcgtcctcatcgatggcagttgtgcctttatctggtgaacaggaatttc 0000106
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985921 ggcgtcctcatcgatggcagttgtgcctttatctggtgaacaggaatttc 6985872

0000107 ctatttgccatgtgaggttatattaagatcatagatgtctgcaagtcagg 0000156
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985871 ctatttgccatgtgaggttatattaagatcatagatgtctgcaagtcagg 6985822

0000157 aagaggtcatcttcattgttagcaacttgaccttgcctagatattccagc 0000206
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985821 aagaggtcatcttcattgttagcaacttgaccttgcctagatattccagc 6985772

0000207 ttctaggatattttgcttgttcttcctcctcttcttcttcttttatttt 0000256
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985771 ttctaggatattttgcttgttcttcctcctcttcttcttcttttatttt 6985722

0000257 ttctgagtgacccagtctgttgttttctgttgtagcagt 0000295
<<<<<<< ||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985721 ttctgagtgacccagtctgttgttttctgttgtagcagt 6985683
```

```
0000296 ctgtccatgcaactcttgcagggttttcaacatgagtcggaaaaggata 0000345
<<<<<<< ||||||||||||||||||||||||||||||||||||||||| |||||||| <<<<<<<
6979704 ctgtccatgcaactcttgcagggttttcaacatgagtcggagaaggata 6979655

0000346 tcctgtaaagatctgggccacgctgactgccagggttgg 0000384
<<<<<<< ||||||||||||||||||||||||||||||||||||||| <<<<<<<
6979654 tcctgtaaagatctgggccacgctgactgccagggttgg 6979616
```

```
0000385 cttcagagaggatacaaggaaatggtatgctgtgaaagaagatggccccc 0000434
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6972318 cttcagagaggatacaaggaaatggtatgctgtgaaagaagatggccccc 6972269

0000435 aggctcaaagacaagggacaggtcacactggatgctgtacacaccctctc 0000484
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6972268 aggctcaaagacaagggacaggtcacactggatgctgtacacaccctctc 6972219

0000485 tgggagtccagctttttctt 0000504
<<<<<<< |||||||||||||||||||| <<<<<<<
6972218 tgggagtccagctttttctt 6972199
```

```
0000505 ctggtggttagttcgatccactgtattagccgtggaggggtgttccctag 0000554
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6830497 ctggtggttagttcgatccactgtattagccgtggaggggtgttccctag 6830448

0000555 tgttttgacggattcgagcagagttttgctgttcgat 0000591
<<<<<<< ||||||||||||||||||||||||||||||||||||| <<<<<<<
6830447 tgttttgacggattcgagcagagttttgctgttcgat 6830411
```

PCR2-PF3 (+ strand)

```
0000016 aacactaggg.acacccctccacggctaatacagtggatcgaactaacca 0000064
>>>>>>> |||||||||| |||||||||||||||||||||||||||||||||||||| >>>>>>>
6830444 aacactagggaacacccctccacggctaatacagtggatcgaactaacca 6830493

0000065 ccag 0000068
>>>>>>> |||| >>>>>>>
6830494 ccag 6830497
```

```
0000069 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 0000118
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6979616 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 6979665

0000119 gactcatgttgaaaaaccctgcaagagttgcatggacag 0000157
>>>>>>> ||||||||||||||||||||||||||||||||||||| >>>>>>>
6979666 gactcatgttgaaaaaccctgcaagagttgcatggacag 6979704
```

```
0000158 tgtgc 0000162
>>>>>>> ||||| >>>>>>>
6980456 tgtgc 6980460
```

```
0000163 actgctacaacacaaaacaacggactgggtcactcaaaaaaaataaaaag 0000212
>>>>>>> |||||||||||| |||||||| |||||||||||||| |||||||||||||| >>>>>>>
6985683 actgctacaacagaaaacaacagactgggtcactcagaaaaaataaaaag 6985732

0000213 aaaaagaagaagaagaagaacaagcaaaatatcctataaactggaatatc 0000262
>>>>>>> || ||||||| || |||||||||||||||||||||| || |||||||||| >>>>>>>
6985733 aagaagaagaggaggaagaacaagcaaaatatcctagaagctggaatatc 6985782

0000263 taggcaaggtcaagttgctaacaatgaacatgacctcttcctgacttgca 0000312
>>>>>>> |||||||||||||||||||||||||| |||||||||||||||||||| >>>>>>>
6985783 taggcaaggtcaagttgctaacaatgaagatgacctcttcctgacttgca 6985832

0000313 cacttctatgatcttaatataacctcacatggcaaataggaaattcctgt 0000362
>>>>>>>  || |||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985833 gacatctatgatcttaatataacctcacatggcaaataggaaattcctgt 6985882

0000363 tcaccaaataaaggcacaactgccatcgatgacgacccccccctcatgac 0000412
>>>>>>> |||||| |||||||||||||||||||||||| ||| || |||||||||| >>>>>>>
6985883 tcaccagataaaggcacaactgccatcgatgaggacgccaccctcatgac 6985932

0000413 ctatgacctctcatcacttctctgaacaccatcccattttgaattgaatt 0000462
>>>>>>> |||||||||||||| ||| |||||||||| |||||||||||||||||||| >>>>>>>
6985933 ctatgacctctcaacacatctctgaacagcatcccattttgaattgaatt 6985982

0000463 aggacttgg.catgtggaatgtggca 0000487
>>>>>>> ||||||||| |||||| ||||||||| >>>>>>>
6985983 aggacttgggcatgtg.aatgtggca 6986007
```

PCR2-PR_6990kb (- strand)

```
0000010 aaatgggatgctgttcagagatgtgttgagaggtcataggtcatgagggt 0000059
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985971 aaatgggatgctgttcagagatgtgttgagaggtcataggtcatgagggt 6985922

0000060 ggcgtcctcatcgatggcagttgtgcctttatctggtgaacaggaatttc 0000109
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985921 ggcgtcctcatcgatggcagttgtgcctttatctggtgaacaggaatttc 6985872

0000110 ctatttgccatgtgaggttatattaagatcatagatgtctgcaagtcagg 0000159
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
```

```
6985871 ctatttgccatgtgaggttatattaagatcatagatgtctgcaagtcagg 6985822

0000160 aagaggtcatcttcattgttagcaacttgaccttgcctagatattccagc 0000209
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985821 aagaggtcatcttcattgttagcaacttgaccttgcctagatattccagc 6985772

0000210 ttctaggatattttgcttgttcttcctcctcttcttcttcttttttatttt 0000259
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985771 ttctaggatattttgcttgttcttcctcctcttcttcttcttttttatttt 6985722

0000260 ttctgagtgacccagtctgttgttttctgttgtagcagtctac 0000302
<<<<<<< ||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985721 ttctgagtgacccagtctgttgttttctgttgtagcagtctac 6985679

---

0000303 cctgtccatgctactcttgcttgtttttttcaacttgaatcggaaaaggat 0000352
<<<<<<< |||||||||||| ||||||||  | ||||||||| ||| ||||| |||||| <<<<<<<
6979705 cctgtccatgcaactcttgcagggtttttcaacatgagtcggagaaggat 6979656

0000353 atcctgaaaagatctgggccaccctgactgccagggt 0000389
<<<<<<< |||||| |||||||||||||| |||||||||||||| <<<<<<<
6979655 atcctgtaaagatctgggccacgctgactgccagggt 6979619

---

0000390 ggggtg 0000395
<<<<<<< |||||| <<<<<<<
6831001 ggggtg 6830996

---

0000396 gtggttaattcaatccattgtattagccggggaggggtgttccctattgt 0000445
<<<<<<< ||||||| ||| ||||| |||||||||||| |||||||||||||||| ||| <<<<<<<
6830494 gtggttagttcgatccactgtattagccgtggaggggtgttccctagtgt 6830445

0000446 tttgaccgattcgagcatagttttgctgttcgat 0000479
<<<<<<< |||||| |||||||||| |||||||||||||| <<<<<<<
6830444 tttgacggattcgagcagagttttgctgttcgat 6830411

PCR3-PF2/3 (+ strand)
0000025 gtactgggag.acctgctcaaaatctgtgtcttcatcttcgccttcatca 0000073
>>>>>>> |||||||||| ||||||||||||||||||||||||||||||||||||| >>>>>>>
6830017 gtactgggagaacctgctcaaaatctgtgtcttcatcttcgccttcatca 6830066

0000074 tgccggtcctcatcatcactgtgtgttatggactgatgatcttacgactc 0000123
>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830067 tgccggtcctcatcatcactgtgtgttatggactgatgatcttacgactc 6830116

0000124 aagagtgtccgcatgctgtcgggctccaaagaaaaggacaggaacctgcg 0000173
>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830117 aagagtgtccgcatgctgtcgggctccaaagaaaaggacaggaacctgcg 6830166

0000174 caggatcacccggatggtgctggtggtcgtggctgtatttattgtctgct 0000223
>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830167 caggatcacccggatggtgctggtggtcgtggctgtatttattgtctgct 6830216

0000224 ggacccccatccacatctatgtcatcatcaaagcactgatcacgattcca 0000273
>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830217 ggacccccatccacatctatgtcatcatcaaagcactgatcacgattcca 6830266

0000274 gaaaccactttccagactgtttcctggcacttctgcattgccttgggtta 0000323
>>>>>>> ||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830267 gaaaccactttccagactgtttcctggcacttctgcattgccttgggtta 6830316

0000324 cacaaacagctgcctgaacccagttctttatgcgttcctggatgaaaact 0000373
```

```
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830317 cacaaacagctgcctgaacccagttctttatgcgttcctggatgaaaact 6830366

0000374 tcaaacgatgttttagagagttctgcatcccaacttcctccacaatcgaa 0000423
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830367 tcaaacgatgttttagagagttctgcatcccaacttcctccacaatcgaa 6830416

0000424 cagcaaaactctgctcgaatccgtcaaaacactagggaacacccctccac 0000473
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6830417 cagcaaaactctgctcgaatccgtcaaaacactagggaacacccctccac 6830466

0000474 ggctaatacagtggatcgaactaaccaccag 0000504
>>>>>>> |||||||||||||||||||||||||||||| >>>>>>>
6830467 ggctaatacagtggatcgaactaaccaccag 6830497
```

```
0000505 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 0000554
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6979616 ccaaccctggcagtcagcgtggcccagatctttacaggatatccttctcc 6979665

0000555 gactcatgttgaaaaaccctgcaagagttgcatggacag 0000593
>>>>>>> |||||||||||||||||||||||||||||||||||||||| >>>>>>>
6979666 gactcatgttgaaaaaccctgcaagagttgcatggacag 6979704
```

```
0000594 actgctacaacagaaaacaacagactgggtcactccgaaaaaataaaaag 0000643
>>>>>>> |||||||||||||||||||||||||||||||||| ||||||||||||||| >>>>>>>
6985683 actgctacaacagaaaacaacagactgggtcactcagaaaaaataaaaag 6985732

0000644 aagaagaagaggaggaagaacaagcaaaatatcctataagctggaatatc 0000693
>>>>>>> |||||||||||||||||||||||||||||||||||||||| |||||||||||||| >>>>>>>
6985733 aagaagaagaggaggaagaacaagcaaaatatcctagaagctggaatatc 6985782

0000694 taggcaaggtcaagttgctaacaatgaagatgacctcttcctgacttgca 0000743
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985783 taggcaaggtcaagttgctaacaatgaagatgacctcttcctgacttgca 6985832

0000744 gacatctatgatcttaatataacctcacatggcaaataggatattcctgt 0000793
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||| |||||||| >>>>>>>
6985833 gacatctatgatcttaatataacctcacatggcaaataggaaattcctgt 6985882

0000794 tcaccacataaaggcacaactgccatcgatgaggacaccaccctcatgac 0000843
>>>>>>> |||||| |||||||||||||||||||||||||||||| |||||||||||||| >>>>>>>
6985883 tcaccagataaaggcacaactgccatcgatgaggacgccaccctcatgac 6985932

0000844 ctatgacctctcaacacatctctgaacagcatcccattttgaattgaatt 0000893
>>>>>>> |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| >>>>>>>
6985933 ctatgacctctcaacacatctctgaacagcatcccattttgaattgaatt 6985982

0000894 aggact 0000899
>>>>>>> |||||| >>>>>>>
6985983 aggact 6985988
```

PCR3-PR_6990kb (- strand)

```
0000017 gatgctgttc.gagatgtgttgagaggtcataggtcatgagggtggcgtc 0000065
<<<<<<< |||||||||| |||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985965 gatgctgttcagagatgtgttgagaggtcataggtcatgagggtggcgtc 6985916

0000066 ctcatcgatggcagttgtgcctttatctggtgaacaggaatttcctattt 0000115
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985915 ctcatcgatggcagttgtgcctttatctggtgaacaggaatttcctattt 6985866
```

```
0000116 gccatgtgaggttatattaagatcatagatgtctgcaagtcaggaagagg 0000165
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985865 gccatgtgaggttatattaagatcatagatgtctgcaagtcaggaagagg 6985816

0000166 tcatcttcattgttagcaacttgaccttgcctagatattccagcttctag 0000215
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985815 tcatcttcattgttagcaacttgaccttgcctagatattccagcttctag 6985766

0000216 gatattttgcttgttcttcctcctcttcttcttcttttattttttctga 0000265
<<<<<<< ||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6985765 gatattttgcttgttcttcctcctcttcttcttcttttattttttctga 6985716

0000266 gtgacccagtctgttgttttctgttgtagcagt 0000298
<<<<<<< |||||||||||||||||||||||||||||||| <<<<<<<
6985715 gtgacccagtctgttgttttctgttgtagcagt 6985683
```

```
0000299 ctgtccttgcaactcttgcagggtttttcaacatgagtcggaaaaggata 0000348
<<<<<<< |||||| |||||||||||||||||||||||||||||||| ||||||| <<<<<<<
6979704 ctgtccatgcaactcttgcagggtttttcaacatgagtcggagaaggata 6979655

0000349 tcctgtaaagatctgggccacgctgactgccagggttgg 0000387
<<<<<<< |||||||||||||||||||||||||||||||||||| <<<<<<<
6979654 tcctgtaaagatctgggccacgctgactgccagggttgg 6979616
```

```
0000389 tggtggttagttcaatccactgtattagccgtggaggggtgttccctact 0000438
<<<<<<< |||||||||||||| |||||||||||||||||||||||||||||||| | <<<<<<<
6830496 tggtggttagttcgatccactgtattagccgtggaggggtgttccctagt 6830447

0000439 gttttgacggattcgagcagagttttgctgttctattgtggaggaagttg 0000488
<<<<<<< |||||||||||||||||||||||||||||||| ||||||||||||||| <<<<<<<
6830446 gttttgacggattcgagcagagttttgctgttcgattgtggaggaagttg 6830397

0000489 ggttgcaaacctctctaaaacatcgtttgaatttttcatccaagaacgca 0000538
<<<<<<< || |||| | |||||||||||||||||||| |||||||||| ||||||| <<<<<<<
6830396 ggatgcagaactctctaaaacatcgtttgaagttttcatccaggaacgca 6830347

0000539 aaaaaaactgggttcaggcagctgtttgtgtaacccaaggcaatgcaaaa 0000588
<<<<<<<  ||| |||||||||||||||||||||||||||||||||||||||||| || <<<<<<<
6830346 taaagaactgggttcaggcagctgtttgtgtaacccaaggcaatgcagaa 6830297

0000589 gtgccaggaaacagtctggaaagtggtttctggaatcgtgatcagtgctt 0000638
<<<<<<< |||||||||||||||||||||||||||||||||||||||||||||||||| <<<<<<<
6830296 gtgccaggaaacagtctggaaagtggtttctggaatcgtgatcagtgctt 6830247

0000639 tgatgatgacataaatgtggaggggggtccagcacacaataaatacagcc 0000688
<<<<<<< |||||||||||||| ||||||| |||||||||||| |||||||||||||||| <<<<<<<
6830246 tgatgatgacatagatgtggatgggggtccagcagacaataaatacagcc 6830197

0000689 acgaccaccacccccatccgggtgatcctgcgcaggttcctgtccttttc 0000738
<<<<<<< |||||||||| | |||||||||||||||||||||||||||||||||||| <<<<<<<
6830196 acgaccaccagcaccatccgggtgatcctgcgcaggttcctgtccttttc 6830147

0000739 tttggagcgcgacagcatgcggacactcttgagtcgtaagatcatctctc 0000788
<<<<<<< |||||||| |||||||||||||||||||||||||||||||||  || <<<<<<<
6830146 tttggagcccgacagcatgcggacactcttgagtcgtaagatcatcagtc 6830097

0000789 cataccacacagtgatgatgaggaccggcatgatgagagcaaagatgaaa 0000838
<<<<<<< |||| |||||||||||||||||||||||||||||||  || ||||||||  <<<<<<<
6830096 cataacacacagtgatgatgaggaccggcatgatgaaggcgaagatgaag 6830047
```

```
0000839 acacagattttgagcaggttctcccatcaccatgtgatatgagagaacat 0000888
<<<<<<< |||||||||||||||||||||||||  ||||||||  |||||||||| | <<<<<<<
6830046 acacagattttgagcaggttctcccagtaccatgtgggatgagagaacgt 6829997

0000889 gagggtgca 0000897
<<<<<<< ||||||||| <<<<<<<
6829996 gagggtgca 6829988
```